

Product Design in Selection Markets*

André Veiga[†] E. Glen Weyl[‡]

October 2015

Abstract

In selection markets, where the cost of serving consumers is heterogeneous and non-contractible, non-price product features allow a firm to sort profitable from unprofitable consumers. An example of this “sorting by quality” is the use of downpayments to dissuade borrowers unlikely to repay. We study a model in which consumers have multidimensional types and a firm offers a single product of endogenous quality, as in [Spence \(1975\)](#). These two ingredients generate a novel sorting incentive in a firm’s first-order condition for quality, which is a simple ratio. The denominator is marginal consumer surplus, a measure of market power. The numerator is the covariance, among marginal consumers, between marginal willingness to pay for quality and cost to the firm. We provide conditions under which this term is signed, and contrast the sorting incentives of a profit maximizer and a social planner. We then use this characterization to quantify the importance of sorting empirically in subprime auto lending, analytically sign its impact in a model of add-on pricing, and calibrate optimal competition policy in health insurance markets.

Word count: 13428

JEL Classification Codes: D41, D42, D43, D86, G21, I13

*This paper replaces earlier working papers “Multidimensional Heterogeneity and Platform Design” and “Multidimensional Product Design” and incorporates some material from “The Leibniz Rule for Multidimensional Heterogeneity.” We thank Eduardo Azevedo, Bruno Jullien, Renato Gomes, Jacques Crémer, Matthew Gentzkow, Josh Mollner, Jesse Shapiro, Mike Whinston, and many workshop and conference participants for helpful comments. We thank Jonathan Baker, Liran Einav, Jon Levin, Jean Tirole, and Preston McAfee for excellent formal discussions of the paper. We acknowledge the financial support of the NET Institute, the research program on intellectual property at the Institut d’Économie Industrielle (IDEI), the Marion Ewing Kauffman Foundation, which funded the research assistance of Kevin Qian and especially Jack Mountjoy, and the Fundação para a Ciência e a Tecnologia (FCT). We are especially grateful to Ben Handel, Igal Hendel, and Mike Whinston for the summary statistics that allow our calibration, and to Liran Einav, Mark Jenkins, and Jon Levin for assistance in using their calibrated model. We thank the editor, Andrei Shleifer, and three anonymous referees for detailed feedback.

[†]Department of Economics and Nuffield College, University of Oxford, Manor Road Building, Oxford OX1 3UQ, UK; andre.veiga@economics.ox.ac.uk. In loving memory of António Luis Veiga.

[‡]Microsoft Research New England, One Memorial Drive, Cambridge, MA 02142 and Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637 USA; glenweyl@microsoft.com.

1 Introduction

When subprime auto lenders help finance a car purchase, they require that borrowers first make a payment to the lender upfront. These downpayments seem ineffective as revenue or collateral. Instead, they seem to address the most important challenge in subprime lending: dissuading those borrowers least likely to repay. In this industry, documentation is absent and credit ratings are noisy, so a borrower’s ability to raise even a small downpayment is a powerful signal of her ability to repay the loan. Charging higher interest rates (viz. raising the total amount owed) also discourages marginal borrowers, but tends to discourage precisely those marginal borrowers who intend to repay, because, for them, the debt burden is most onerous. Thus, raising downpayments while lowering interest rates allows the firm to maintain the same total number of borrowers while improving their composition, since it advantageously sorts out the riskiest borrowers. Although the low-information environment makes subprime lending an extreme example, this type of sorting is common in many other *selection markets*, such as finance and insurance, where consumer profitability to the firm is heterogeneous and non-contractible. This paper develops a quantitative theory of the use of non-price product features to sort consumers.

Firms can partly control which consumers they attract by changing price (e.g., the interest rate). This is the *selection* (by quantity) considered by [Akerlof \(1970\)](#) and [Einav and Finkelstein](#). However, the use of non-price features (e.g., a downpayment) greatly expands the ability to sort profitable from unprofitable consumers. As another example, [Rothschild and Stiglitz \(1976\)](#) suggest insurers attract profitable low-risk consumers (i.e., “cream-skim”) by offering cheap low-coverage plans. We label this effect *sorting* (by quality) to distinguish it from selection.

Regulation of non-price product features is a central policy concern in many selection markets. Several rules in the US Patient Protection and Affordable Care Act were explicitly designed to prevent the cream-skimming emphasized by [Rothschild and Stiglitz](#). The Dodd-Frank reforms impose similar restrictions on mortgages. Therefore, a general characterization of sorting has implications for regulation, as we discuss throughout the paper.

We propose a model of sorting in which consumers have multidimensional non-contractible types. We consider a general non-price product feature, which we refer to as “quality” for brevity, allowing each consumer’s preference for this quality to be arbitrary. This general specification remains tractable because of an important and restrictive simplification. Building off [Einav, Finkelstein and Cullen \(2010\)](#) and [Einav and Finkelstein \(2011\)](#), we assume firms offer a single contract, rather than a menu of contracts as has been common in the literature since [Stiglitz \(1977\)](#) and [Mussa and Rosen \(1978\)](#). This assumption allows us to

follow [Spence \(1975\)](#) and [Sheshinski \(1976\)](#)’s analysis of a monopolist’s optimal choice of *quality*, holding fixed the choice of *quantity*. In their analyses, concerns about direct costs and price discrimination entirely drive a firm’s choice of quality. Section 2.2 shows that, in selection markets, an additional sorting incentive appears in the firm’s first-order condition for quality. The per-customer sorting effect on profit of a marginal increase in quality is the ratio of two terms. The denominator is marginal consumers’ surplus. The numerator is the covariance, among marginal consumers, between consumers’ marginal willingness to pay for quality and consumers’ cost to the firm.

This covariance captures the extent to which, among marginal individuals, those most attracted by quality (high marginal willingness to pay for quality) are also those who are particularly costly. When this is the case, raising quality changes the composition of buyers (i.e., sorts) in a way that raises cost. This covariance is zero in typical models of selection markets, in which consumers have unidimensional types and utility is monotonic in type. In these settings, a unique marginal type is present, so the covariance necessarily vanishes. Intuitively, sorting is absent because changing quality does not allow the firm to achieve compositions of buyers beyond those already achievable by simply changing quantity. Sorting requires both multidimensional types and endogenous quality, hence its absence in the existing literature.¹

The denominator of the sorting incentive is marginal consumer surplus, a common measure of market power. This term captures the fact that, for firms with elastic demand, a small change in quality can induce a significant change in buyer composition and thus have a large effect on profit. This denominator also suggests competition increases the relative importance of sorting incentives, because it increases the elasticity of firms’ residual demands, as we discuss in Sections 2.2 and 3.3.

In Section 2.3, we use this characterization to contrast a firm’s and a social planner’s incentives to sort. When consumers are fully rational, a monopoly’s sorting incentive is socially optimal because marginal consumers are indifferent to purchasing the good, so the monopolist internalizes all relevant social welfare. The optimality of private sorting incentives fails for competing firms, who do not internalize the effect of sorting on their rival’s profit. As in [Rothschild and Stiglitz](#), the inefficient sorting incentives of competing firms lead to inefficient quality provision.

Section 2.4 gives conditions under which the sorting incentive can be signed. Its sign is often a function of the structure of preferences and cost alone. It is independent of the

¹As [Einav and Finkelstein](#) write, “We currently lack clear characterizations of...equilibrium...in which firms compete over contract dimensions as well as price, and in which consumers have multiple dimensions of private information.”

statistical distribution of consumer types, because, for a given contract, only consumer preferences determine which types are marginal. When the sorting term is positive, increasing quality raises average cost. We call this “adverse sorting,” by analogy to [Einav and Finkelstein](#)’s “adverse selection” whereby raising the price increases average cost. We also provide conditions under which the distortion highlighted by [Spence](#) and [Sheshinski](#) can be signed under multidimensional heterogeneity.

We illustrate the flexibility and usefulness of these results with three applications. First, [Einav, Jenkins and Levin \(2012\)](#) study the design of a single contract offered to (observationally equivalent) potential borrowers in subprime auto lending. The authors find higher downpayments decrease overall default probabilities, and attribute most of this effect to *selection*: high downpayments dissuade marginal consumers, and marginals are generally riskier than infra-marginals. In [Section 3.1](#), we use their calibrated model to show that about one third of the total effect they measure stems from heterogeneity within marginal consumers (i.e., sorting) rather than the difference between marginals and infra-marginals (i.e., selection). An increase in car price also dissuades marginal consumers. However, a higher downpayment reduces default rates significantly more because it dissuades precisely those marginal borrowers who are at greatest risk of default. We show that one-third of the total value of what [Einav, Jenkins and Levin](#) call selection is due to advantageous sorting within marginal borrowers by the downpayment requirement, rather than advantageous selection of inframarginal over marginal borrowers. The distinction between sorting and selection is particularly important in this context, as we discuss, as they have opposite interactions with financial stability externalities.

Our second application, in [Section 3.2](#), considers analytically a monopolist’s tendency to charge excessive add-on fees. If all consumers are perfectly rational, as in [Ellison \(2005\)](#), sorting reduces this inefficiency. Marginal consumers who frequently pay fees are more profitable, but they are also those more strongly dissuaded by higher fees. Therefore, sorting is adverse: higher fees dissuade the most profitable marginal consumers, which gives the firm an incentive to lower the fees. The opposite can occur if unsophisticated low-income consumers underestimate the fees they will incur, as in [Gabaix and Laibson \(2006\)](#). When consumer sophistication is elastic with respect to income, higher fees dissuade the least profitable marginal consumers, thereby further incenting the firm to increase add-on fees. Because sorting is the only deterrent to setting arbitrarily high add-on prices, when sophistication is income elastic all restraint on add-on pricing is removed. Existing empirical evidence suggests sophistication is large but less than unit elastic to income ([Stango and Zinman, Forthcoming](#); [Rees-Jones and Taubinsky, 2015](#)), so heterogeneous sophistication is likely to exacerbate, but not entirely remove constraints on, excessive add-on fees.

Our final application, in Section 3.3, calibrates optimal competition policy in an insurance market. In Rothschild and Stiglitz, a pooling contract that gives full insurance to all consumers maximizes welfare, but competitive cream-skimming leads to insufficient insurance in any equilibrium. We confirm the conjecture, given in that paper, that market power may mitigate these harmful incentives and thereby increase welfare. At relatively low levels, market power blunts inefficient cream-skimming more than it distorts quantity downward (Mahoney and Weyl, 2014). Social welfare first increases and then decreases with market power, so an interior optimal degree of competition exists.

We calibrate our competitive insurance model to Handel, Hendel and Whinston (2015)’s health insurance data. First, we consider an otherwise unregulated market where welfare is derived from revealed preferences. Allowing firms to charge prices 110% above average cost achieves nearly 97% of the maximum of welfare, whereas markups 91% above cost achieve 80% of the consumer surplus achieved at the maximum of welfare. We then consider a market where behavioral consumers overestimate their risk aversion, calibrated to measurements of this phenomenon by Handel and Kolstad (2015). In this case, lower insurance quality is socially desirable. Therefore, a lower (88%) markup achieves 99.7% of the maximum of (behavioral) welfare. In sum, our framework consistently shows the role of market power in mitigating cream-skimming, but can flexibly accommodate a wide range of assumptions regarding product features, preferences, costs, consumer heterogeneity, rationality, and welfare standards.

Despite the wide range of applications of our approach, we emphasize that assuming a single contract is restrictive. It implies we cannot analyze how broad a range of contracts will be made available, as in Lester et al. (2015), or which types of consumers will adopt what type of contract, as in Azevedo and Gottlieb (2015). This weakness is particularly acute in extremely competitive markets where the sort of symmetric equilibrium we study in Section 3.3 often fails to exist, but asymmetric equilibria often do.

However, in many empirical settings, consumers are effectively offered a single, but carefully designed, contract. Examples include the Einav, Jenkins and Levin subprime lending market and the health insurance market studied by Einav, Finkelstein and Cullen. Such “pooling” of many consumer types into a few contracts is often more socially desirable than the outcomes of an uninhibited competitive process, as in Rothschild and Stiglitz. As we discuss in Section 2.3, the desirability of pooling may explain why regulations often limit the range of products that firms can offer in selection markets. Our framework may therefore be particularly relevant in such partly regulated environments. Furthermore, as we highlight in Section 3.3, the implications of our approach are in line with the received wisdom of the existing literature when the two are comparable. Thus, our model offers a pragmatic and

flexible way to derive novel and policy-relevant characterizations of quality in selection markets with multidimensional types and market power, where classical contract theory struggles to obtain sharp results.

We collect derivations, proofs, and computational details into appendices following the main text.

2 Theory

We propose a simple but general model of a monopolistic firm that chooses the price and quality of its product, while facing customers heterogeneous in many dimensions. The model combines the central elements of selection markets analyzed by [Einav, Finkelstein and Cullen \(2010\)](#) and [Einav and Finkelstein \(2011\)](#) with the the analysis of monopolistic quality choice of [Spence \(1975\)](#) and [Sheshinski \(1976\)](#).²

2.1 Model

A unit mass of potential consumers exists that is characterized by a T -dimensional type $\theta \in \Theta$. We assume Θ is a (possibly unbounded) hyper-rectangle in \mathbb{R}^T and θ is distributed according to a continuously differentiable, full-support probability density function $f(\theta) > 0$. A monopoly offers a single product characterized by a price $p \in \mathbb{R}_{++}$ and “quality” x , chosen from an interval subset of \mathbb{R} . For brevity, we use “quality” to refer to any non-price product feature. The strong assumption of a single contract, made for tractability, is discussed in the introduction and conclusion.

A consumer of type θ is willing to pay $u(x, \theta)$ for a product of quality x and carries an expected cost to the firm of $c(x, \theta)$. Cost depends directly on the non-contractible consumer type θ , which is the defining feature of selection markets. Type θ captures any consumer characteristics that affect willingness to pay (WTP) or cost, such as health, risk aversion, sophistication, or income. We assume $u = u(x, \theta)$ and $c = c(x, \theta)$ are twice continuously differentiable in all arguments. Notice u and c can increase or decrease with “quality” x for each consumer.

It is convenient to decompose $\theta = (\zeta, \tau)$, where ζ ’s domain is a hyper-rectangle within \mathbb{R}^{T-1} , and $\tau \in (\underline{\tau}, \bar{\tau}) \subseteq \mathbb{R}$. We assume a dimension of type (τ) exists such that $\frac{\partial u(x, \zeta, \tau)}{\partial \tau} > 0$ and

$$\forall (x, \zeta) : \lim_{\tau \rightarrow \bar{\tau}} u(x, \zeta, \tau) = \infty \text{ and } \lim_{\tau \rightarrow \underline{\tau}} u(x, \zeta, \tau) \leq 0.$$

²[Weyl \(2010\)](#) extends the [Spence](#) and [Sheshinski](#) analyses to multidimensional heterogeneity in a context without selection.

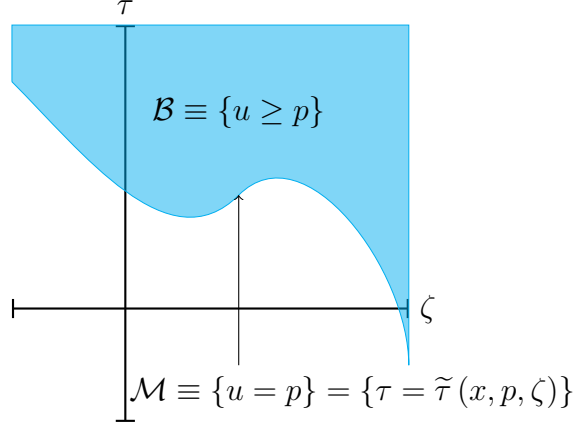


Figure 1: An example of the set of buyers, \mathcal{B} , and the set of marginal consumers, \mathcal{M} , for $T = 2$. The set of marginal consumers is described by $\tau = \tilde{\tau}(x, p, \zeta)$. For each ζ , individuals buy the product if τ is sufficiently large and do not buy if τ is sufficiently small.

This assumption implies that, for any (x, ζ) , values of τ exist for which a consumer of type ζ buys and does not buy the product, which simplifies the exposition. It relaxes the monotonicity assumptions (Spence, 1973; Mirrlees, 1971) typically made in models with a single-dimensional type.³ In particular, we assume only that *some* dimension of type exists that monotonically increases u , whereas existing models typically assume u is increasing in the unique dimension of type.

The set of buyers of the product is $\mathcal{B} = \{u \geq p\}$. Given the assumptions above, a unique function $\tilde{\tau} = \tilde{\tau}(x, p, \zeta)$ exists such that $u(x, \zeta, \tilde{\tau}(x, p, \zeta)) = p$. We use this function to provide a more convenient definition of the set of buyers as $\mathcal{B} \equiv \{\theta : \tau \geq \tilde{\tau}(x, p, \zeta)\}$.⁴ Similarly, the set of consumers who are indifferent between purchasing or exiting the market (“marginals”) is $\mathcal{M} \equiv \{\tau = \tilde{\tau}\} = \{u = p\}$. Note \mathcal{M} is a $T - 1$ dimensional surface bounding the T dimensional subspace \mathcal{B} , as illustrated in Figure 1 for $T = 2$.

These assumptions allow us to express demand as a simple iterated integral:

$$Q = Q(x, p) = \int_{\zeta} \int_{\tilde{\tau}(x, p, \zeta)}^{\bar{\tau}} f(\zeta, \tau) d\tau d\zeta \equiv \int_{\mathcal{B}} f(\theta) d\theta.$$

This definition of an integral over the (volume of the) set of buyers \mathcal{B} will be used throughout.

³A previous version of this paper, Veiga and Weyl (2012), considers a setup that is more general but without additional economic intuition. We refer interested readers to that earlier version. For technical details on the differentiation of more general integrals, see Flanders (1973) and Uryasev (1995).

⁴Throughout the paper, we use the symbol \equiv to note definitions being made.

Again using the function $\tilde{\tau}$, the derivative of demand can be obtained straightforwardly as

$$\frac{\partial Q}{\partial p} = \int_{\zeta} \left(-\frac{\partial \tilde{\tau}}{\partial p} \right) f(\zeta, \tilde{\tau}) d\zeta = - \int_{\zeta} \frac{f(\zeta, \tilde{\tau})}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}} d\zeta \equiv -M(x, p) < 0.$$

This follows from applying the implicit function theorem to the equation that defines the margin, $u(x, \zeta, \tilde{\tau}(x, p, \zeta)) = p$, to obtain $\frac{\partial \tilde{\tau}}{\partial p} = \frac{1}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}}$. Then, $\frac{\partial \tilde{\tau}}{\partial p} > 0$ by the assumptions above. The term $M = M(x, p) > 0$ is defined by a surface integral over the set of marginal types $\mathcal{M} \equiv \{\tau = \tilde{\tau}\}$ and captures the density of this set. In our multidimensional setting, the density M plays a role analogous to that played, in a typical unidimensional-types model, by the density of types evaluated at the unique marginal type. Importantly, in our setting, multiple types are marginal, as illustrated by Figure 1.

Because $f(\theta) > 0$ and $\frac{\partial \tilde{\tau}}{\partial p} > 0$, demand $Q(x, p)$ is strictly decreasing in p . Therefore, a differentiable inverse demand $P(x, q)$ exists such that $Q(x, P(x, q)) = q$ for any x . Using this relationship, we follow Spence in thinking of the firm as directly choosing the number of buyers q , and thereby inducing a price $p = P(x, q)$.⁵

Two expectation operators will be crucial for our analysis. For an arbitrary smooth function $z(x, \theta)$, the expectation conditional on the set of buyers \mathcal{B} is standard:

$$\mathbb{E}[z(x, \theta) \mid \mathcal{B}] \equiv \frac{1}{q} \int_{\mathcal{B}} z(x, \theta) f(\theta) d\theta.$$

The expectation of any $z(x, \theta)$ conditional on the set of marginal consumers \mathcal{M} is less standard. It is *defined* as

$$\mathbb{E}[z(x, \theta) \mid \mathcal{M}] \equiv \frac{1}{M} \int_{\zeta} z(x, \zeta, \tilde{\tau}) \frac{f(\zeta, \tilde{\tau})}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}} d\zeta,$$

where we use the definition of M provided above. This defines an expectation on a set of measure 0 in the type space Θ , and thus is an important non-trivial definition. This expectation uses the modified density $\frac{f(\zeta, \tilde{\tau})}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}}$, rather than simply $f(\zeta, \tilde{\tau})$. This modified density was also used in the definition of M , because it performs the integral in the economically meaningful space of utility rather than in an arbitrary mathematical space of types.⁶

⁵Our analysis holds if the firm chooses (x, p) , although considering the choice of (x, q) allows for a clearer intuition and is more in line with the existing literature.

⁶Suppose $\tau \in \mathbb{R}_+$ and the density of types is $f(\zeta, \tau)$. Now consider the change of variables $a = e^{\tau}$. The new density is $g(\zeta, a) = \frac{f(\zeta, \log(a))}{a}$ because the change of variables requires including the gradient of the transformation and $[\log(a)]' = \frac{1}{a}$. As a result, absent the modification of the density, a mechanical integral over the set of marginal a values does not yield the same solution as an integral over marginal τ , even though a is simply an equivalent re-parameterization of τ . However, such a change of variables would have no effect on the economically meaningful modified density, because the integral with changed variables would use the density

Other statistical operators, such as conditional variances and covariances, are defined analogously from the definitions above. We use Newton's notation to denote partial derivatives with respect to x , namely, $z'(x, \theta) \equiv \frac{\partial z(x, \theta)}{\partial x}$.

2.2 Profit Maximization

Recalling that $p = P(x, q)$, a profit-maximizing monopoly chooses (x, q) to maximize

$$\pi(x, q) \equiv \int_{\mathcal{B}} (P(x, q) - c(x, \theta)) f(\theta) d\theta,$$

where $\mathcal{B} = \{u(x, \theta) \geq P(x, q)\}$. The familiar first-order condition (FOC) with respect to quantity q is derived in Appendix A: $p - \mathbb{E}[c \mid \mathcal{M}] = MS$. The markup is equated to marginal consumer surplus $MS = \frac{q}{M} = q \frac{\partial P(x, q)}{\partial q}$, as in the standard monopoly model. Marginal cost $\mathbb{E}[c \mid \mathcal{M}]$ is the average cost among consumers who are marginal. Incentives for quantity choice are not the focus of this paper, and we refer the reader to [Einav and Finkelstein](#) and [Mahoney and Weyl \(2014\)](#) for additional discussion.

However, marginal surplus plays an important role in our analysis below, so we briefly relate it to other concepts in the literature. Marginal surplus is equal to the monopoly markup above marginal cost, as shown above. However, previous literature has also shown it equates to several other quantities: the ratio of price to the elasticity of demand, the inverse semi-elasticity of demand, the inverse of the hazard rate of the marginal distribution of WTP, and the gap between valuation and virtual valuation in auction theory ([Myerson, 1981](#)). In this paper, we emphasize the role of MS as a measure of market power.

We now characterize optimal choice of quality, which is the focus of our analysis.

Proposition 1. *The per-consumer marginal effect of quality x on profit is*

$$\frac{\partial \pi}{\partial x} \frac{1}{q} = - \underbrace{\mathbb{E}[c' \mid \mathcal{B}]}_{\text{direct cost}} + \underbrace{\mathbb{E}[u' \mid \mathcal{M}]}_{\text{private Spence term}} - \underbrace{\frac{\text{Cov}[u', c \mid \mathcal{M}]}{MS}}_{\text{private sorting incentive}}.$$

At the privately optimal quality x^* , this expression vanishes. First, when a firm increases quality x , it loses the average increase in the cost of all buyers, which results mechanically from the increase in provided quality, $\mathbb{E}[c' \mid \mathcal{B}]$. Second, as identified by [Spence](#) and [Sheshinski](#), a monopolist raises price by $\mathbb{E}[u' \mid \mathcal{M}]$ when it increases x , because, to hold fixed q , price must offset the average benefit that marginal consumers derive from the additional quality.

$$\frac{g(\zeta, a) \frac{1}{a}}{\frac{\partial u(x, \zeta, \log(a))}{\partial a}} = \frac{f(\zeta, \log(a)) \frac{1}{a}}{\frac{\partial u(x, \zeta, \tau) \frac{1}{a}}{\partial \tau}} = \frac{f(\zeta, \tau)}{\frac{\partial u(x, \zeta, \tau)}{\partial \tau}}.$$

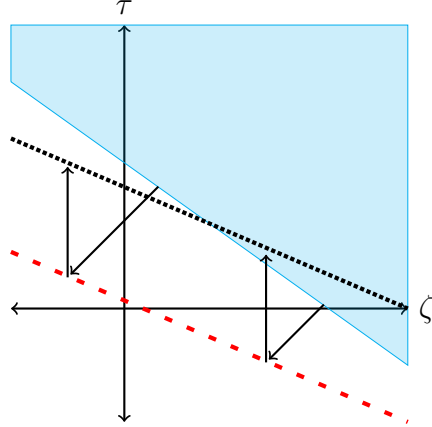


Figure 2: A visual illustration of sorting with $T = 2$, $u(x, \zeta, \tau) = \frac{17-2\zeta}{7}x + \frac{3}{7}\zeta + \tau$ and initially $(x, P(q)) = (0, \frac{8}{7})$. The downward shift results from increasing x by a unit and the upward shift from a compensating rise in p by 2 units.

Following the convention in the literature, we refer to this component of private incentives as the *private Spence term*.⁷

The third effect is the novel element in our model. In [Spence](#)'s analysis, the firm cares only about the impact of quality on cost and on the price it can charge, not on *which* consumers it serves. However, in a selection market, buyer composition is crucial. The impact of x on profit depends on whether the marginal consumers most strongly attracted by x (those with the largest u') are also those with relatively high or low cost c . It is therefore natural that the sorting effect of x on profit is proportional to the covariance, among marginal consumers, between marginal WTP for quality and cost, $\text{Cov}[u', c \mid \mathcal{M}]$. If this covariance is positive, “adverse sorting” occurs: the change in buyer composition induced by increasing x , together with the compensating change in price, increases cost. If the covariance is negative, sorting is “advantageous.”⁸

The left panel of [Figure 2](#) provides a visual illustration of sorting in two dimensions ($T = 2$). In this example, utility takes the simple form $u(x, \zeta, \tau) = \frac{17-2\zeta}{7}x + \frac{3}{7}\zeta + \tau$, for $(\zeta, \tau) \in \mathbb{R}^2$. For the pictured range, and for values $x = 0$ and $P(q) = \frac{8}{7}$, buyers are those with sufficiently high type in either dimension (the shaded blue area in [Figure 2](#)), so the

⁷[Spence](#) and [Sheshinski](#)'s analysis did not explicitly discuss foundations in a consumer type space and informally discussed their results in terms of a single marginal type. [Weyl \(2010\)](#) considers an explicit model with multidimensional types.

⁸Although, to our knowledge, this effect is new to the analysis of optimal contracting, it follows a logic similar to that of [Diamond \(1973\)](#) and [Green and Sheshinski \(1976\)](#)'s analysis of optimal Pigouvian taxation with imperfect targeting. The characterization of the optimal structure of rewards for innovation by [Weyl and Tirole \(2012\)](#) is also a special case of our analysis.

curve that defines the set of buyers slopes down in the (ζ, τ) plane.⁹ Raising x by a unit lowers this curve (causing more individuals in the pictured range and on average to buy) but also rotates it (attracts some types more intensely than others). The resulting set of buyers are those above the dashed red curve. In this case, marginal types with high τ have higher marginal WTP (u') than marginal types with high ζ , so the difference between the shaded area and the dashed red curve is greater for high values of τ . This first effect of an increase in x is illustrated by the diagonal arrows.

However, to hold quantity q fixed, price rises by 2 units, because most buyers have low ζ and thus gain strongly from the increase in x . This rise in price corresponds visually to a uniform upward shift of the dashed red line to the dotted black line, illustrated by the vertical arrows. The private Spence term corresponds visually to the vertical difference between the dashed and dotted lines. The rotation caused by the original increase in x implies that, after the price adjustment, this process ultimately results in attracting some consumers (those above the dotted line and below the shaded area) and dissuading others (those within the shaded area and below the dotted line). The sorting term $\text{Cov}[u', c | \mathcal{M}]$ captures the effect on profit of this change in buyer composition that holds the total mass of buyers (q) fixed. In this case, sorting shifts buyer composition toward high- τ buyers, and away from high- ζ buyers.

The importance of sorting relative to the Spencian terms is determined by the coefficient on the sorting term, $\frac{1}{MS}$, which increases with the elasticity of demand. The more elastic demand is, the greater the importance of sorting, because a small change in x can induce substantial changes in buyer composition. Moreover, because competition typically increases the elasticity of a firm's residual demand, the sorting term plays a crucial role in determining the effect of competition on quality. In general, MS is endogenous, determined by the demand system and competitive conditions that also determine \mathcal{B} and \mathcal{M} . Therefore, determining the effect of competition on quality analytically may not be straightforward. In Section 3.3, we consider a model with explicit Hotelling competition, and obtain further results about the link between sorting and competition.

The sorting term vanishes in three scenarios. First, $\text{Cov}[u', c | \mathcal{M}] = 0$ when consumers are homogeneous in their cost. Thus, sorting is relevant in selection markets, where consumers differ in their non-contractible profitability to the firm. Second, sorting vanishes if marginal consumers are homogeneous in their marginal preferences for quality. More generally, if quality is interchangeable with price at the same rate for all consumers, there is no scope for differential sorting by quality. Third, sorting is absent when a unique type of

⁹Because WTP is increasing in all arguments, if types (ζ, τ) and $(\hat{\zeta}, \hat{\tau})$ are both marginal, $\hat{\zeta} > \zeta$ must imply $\hat{\tau} < \tau$. This need not always be the case, as illustrated by Figure 1.

marginal consumer is present, which is usually the case in standard models with unidimensional types, because utility is typically monotonic in type. In such models, all types above a given type threshold $\tilde{\tau}$ buy the product, so the choice of quantity q immediately determines the composition of buyers. In other words, in these models, the ability to choose quality does not expand the set of buyer compositions attainable by the firm.¹⁰

We refer to this novel effect as “sorting” (by quality) to emphasize its distinction from what is commonly referred to as “selection” (by quantity). For instance, in [Akerlof \(1970\)](#) and [Einav and Finkelstein](#), adverse selection occurs when raising quantity lowers cost, for fixed quality, which occurs when marginal consumers (low $u = p$) have low cost c while infra-marginal consumers have high $u > p$ and high c . By contrast, adverse sorting occurs when raising quality raises cost, for fixed quantity (beyond the mechanical effect of quality on the cost of infra-marginals), which occurs when, among *marginal* consumers \mathcal{M} , those with higher *marginal WTP for quality* (high u') are more costly (high c) than other marginal consumers with low u' and low c . For instance, selection can occur when consumers have homogeneous marginal WTP with respect to a product feature, whereas sorting cannot. In [Section 3.1](#), we explore some policy implications of the distinction between selection and sorting.

2.3 Welfare Maximization

We now turn to welfare analysis. We define the gross contribution to welfare of a consumer of type θ purchasing the good as $w = w(x, \theta)$. Then, welfare is

$$W \equiv \int_{\mathcal{B}} (w(x, \theta) - c(x, \theta)) f(\theta) d\theta.$$

[Appendix A](#) shows the necessary condition for welfare maximization with respect to q is that additional (necessarily marginal) buyers generate as much social benefit as cost: $\mathbb{E}[w - c \mid \mathcal{M}] = 0$. If, as a baseline, we identify welfare w with private WTP u , this condition collapses to marginal cost pricing $p = \mathbb{E}[c \mid \mathcal{M}]$, because $\mathbb{E}[w \mid \mathcal{M}] = \mathbb{E}[u \mid u = p] = p$. However, this conclusion ignores the possibility of externalities, such as business stealing between firms or “externalities within individuals” if consumers are not perfectly rational, so we allow $w \neq u$ before specializing further.

Proposition 2. *The per-consumer marginal effect of quality on welfare is*

¹⁰[Araujo and Moreira \(2010\)](#) show this simple cut-off structure need not hold in unidimensional models if standard monotonicity properties, such as the [Spence \(1973\)-Mirrlees \(1971\)](#) condition, fails. In that case, sorting may occur even in one-dimensional models. However, as [Araujo and Moreira](#) note, such relaxations introduce most of the complexities of multi-dimensional type models, so the literature has almost never considered such models.

$$\frac{dW}{dx} \frac{1}{q} = - \underbrace{\mathbb{E}[c' | \mathcal{B}]}_{\text{cost}} + \underbrace{\mathbb{E}[w' | \mathcal{B}]}_{\text{social Spence term}} + \underbrace{\frac{\text{Cov}[u', w - c | \mathcal{M}]}{MS}}_{\text{social sorting incentive}}.$$

At the socially optimal quality x^{**} , this expression vanishes. To interpret this condition, begin by assuming no externalities ($w = u$). Then, $\mathbb{E}[w' | \mathcal{B}] = \mathbb{E}[u' | \mathcal{B}]$ and the difference between the social and private optima, emphasized by [Spence](#) and [Sheshinski](#), becomes apparent: a profit-maximizer only internalizes the preferences of marginal consumers $\mathbb{E}[u' | \mathcal{M}]$, whereas a social planner internalizes the preferences of all buyers $\mathbb{E}[u' | \mathcal{B}]$. The difference between the private and social Spence terms is the *Spence distortion*.

In the absence of externalities, the Spence distortion is the *only* distortion of monopoly profit maximization. When $\mathbb{E}[w | \mathcal{M}] = \mathbb{E}[u | \mathcal{M}] = p$ constant, the social sorting incentive becomes $-\frac{1}{MS} \text{Cov}[u', c | \mathcal{M}]$, equal to the private sorting incentive in [Proposition 1](#). Absent externalities, the contribution of marginal consumers to welfare is their contribution to profit. Sorting concerns only marginal consumers, so a monopolist's sorting incentive is socially optimal.

This stark conclusion depends on the typically unrealistic assumption of no externalities, which is easily relaxed by changing the definition of w . One natural relaxation, which we explore in [Sections 3.2 and 3.3](#), is to allow externalities within imperfectly rational individuals.¹¹ In this section, we instead consider an externality typical of industrial organization and contract theory: business stealing between firms.

Suppose consumers gain a gross utility of $v = v(x, \theta)$ from purchasing the focal good but have two other alternatives: buying nothing or buying another good for which WTP is $v_0 = v_0(\theta)$. The competing good is supplied at price p_0 and produced at cost $c_0 = c_0(\theta)$. In this setting, WTP for the focal firm's product is $u(x, \theta) = v(x, \theta) - \max\{0, v_0 - p_0\}$. The focal firm's buyers are still the set $\mathcal{B} = \{u \geq p\}$, those who prefer the focal product to the competitor and to purchasing nothing. The focal firm's margin now has two disjoint components. *Exiting* consumers are those indifferent between the focal product and not purchasing at all, $\mathcal{E} = \{v - p = 0 > v_0 - p_0\}$. *Switching* consumers are those indifferent between the focal product and the competitor, $\mathcal{S} = \{v - p = v_0 - p_0 > 0\}$. Let the density of these sets be E and S , respectively. The set of all marginals is the disjoint union $\mathcal{M} = \mathcal{E} \cup \mathcal{S}$, and its density is $M = S + E$. [Figure 3](#) provides an illustration.

This distinction is immaterial to a profit-maximizing firm, which ignores its rival's profit, so our analysis of profit maximization still applies to a duopolist. Furthermore, the cost and

¹¹We use “externalities” in the sense of [Marshall \(1890\)](#) and [Pigou \(1920\)](#) to denote any effects not captured in the price system. It is also common to use “externality” to refer to effects exterior to agents and “internality” to refer to imperfect rationality (i.e., effects within each agent). We refer to imperfect rationality as “externalities within individuals.”

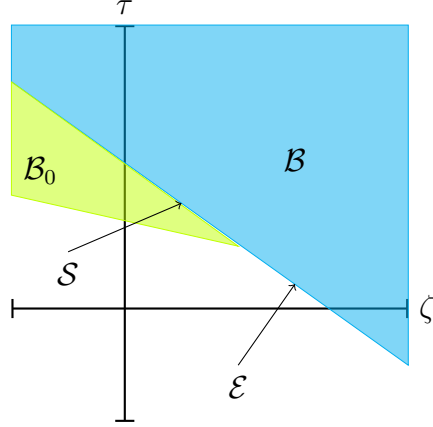


Figure 3: The sets of buyers \mathcal{B} , exiting consumers \mathcal{E} , switching consumers \mathcal{S} , and buyers of the outside product \mathcal{B}_0 .

Spence terms of Proposition 2 remain unchanged in the welfare maximizer's FOC. However, a social planner is concerned that, when a firm chooses quality to improve its own composition of consumers, it may leave the most costly consumers to its rival. This inefficient sorting incentive is the “cream-skimming” effect emphasized by [Rothschild and Stiglitz \(1976\)](#). The focal firm ignores the externality of an individual's decision to purchase its good, namely, the contribution to welfare of $v_0 - c_0$ that could have occurred if the consumer had purchased the rival good.

One could study this cream-skimming incentive by explicitly solving for the analog of Proposition 2 in this new duopolistic environment. A simpler approach, given our results so far, is to alter the definition of w so that Proposition 2 applies unchanged. This approach allows us to study welfare-maximizing incentives in a multi-product market while still considering only each product's buyers (\mathcal{B}) in turn. By the reasoning above, for switching consumers in \mathcal{S} , their net contribution to welfare from purchasing the focal good is $w = v - (v_0 - c_0) = p - p_0 + c_0$, where we used the definition of \mathcal{S} . Because p and p_0 are constant across θ , the social sorting incentive becomes

$$\underbrace{\frac{\text{Cov}[u', w - c \mid \mathcal{M}]}{MS}}_{\text{social sorting incentive}} = \underbrace{\frac{D \text{Cov}[u', c_0 \mid \mathcal{S}]}{MS}}_{\text{cream-skimming distortion}} + \underbrace{\frac{-\text{Cov}[u', c \mid \mathcal{M}]}{MS}}_{\text{private sorting incentive}}, \quad (1)$$

where $D = \frac{S}{M}$ is the *diversion ratio* ([Shapiro, 1996](#)).¹² This expression shows the divergence between private and social incentives is entirely due to inefficient cream-skimming incentives

¹²The diversion ratio captures the share of a firm's lost business captured by another firm following a price increase.

regarding switching consumers in \mathcal{S} , because no externality is associated with the decision of exiting consumers in \mathcal{E} .

This expression further suggests inefficient cream-skimming becomes more severe when the market is competitive, because, in this case, $MS \rightarrow 0$ (the firm loses market power) and $D \rightarrow 1$ (a higher density of consumers becomes indifferent between the firms). In the competitive limit as $\frac{D}{MS} \rightarrow \infty$, the firm's first-order condition becomes impossible to satisfy unless $\text{Cov}[u', c | \mathcal{S}] \rightarrow 0$. When sorting is adverse on the switching margin ($\text{Cov}[u', c | \mathcal{S}] > 0$), it is natural that this covariance is driven to 0 by reducing cost c , and that this is done by reducing quality x to 0. This extreme result is in line with [Rothschild and Stiglitz](#)'s observation that, under perfect competition, cream-skimming creates an overwhelming incentive to reduce insurance quality. This intuition underlies our analysis of an insurance market in [Section 3.3](#), where market power increases product quality and welfare by mitigating harmful cream-skimming incentives.

Equation 1 also suggests the following conjecture regarding optimal product diversity. Suppose costs (c and c_0) increase in quality (x and x_0). Suppose also that the competitor product is, in equilibrium, of lower quality ($x_0^* < x^*$) and lower cost ($c_0 = \alpha c$ for $\alpha < 1$). From equation 1, the cream-skimming distortion for the focal firm is proportional to α , but it is proportional to $\frac{1}{\alpha} > 1 > \alpha$ for the low-quality competitor who faces the symmetric condition. Thus, x_0^* is more severely distorted away from its socially optimal level x_0^{**} , and correcting x_0^* increases welfare more than correcting x^* . If sorting is adverse ($\text{Cov}[u', c_0 | \mathcal{S}] > 0$), both firms under-provide quality, but x_0 is more undersupplied than x . A social planner would incent both firms to raise quality, but would prioritize increasing x_0 , leading to more homogenous products. Therefore, in such a setting, a social planner would desire a market with more homogeneous products than what is provided in equilibrium, moving toward the “pooling” contracts that [Rothschild and Stiglitz](#) identify as desirable but undermined by cream-skimming.¹³ Governments may therefore wish to limit the range of contracts offered, making the single-contract case we study more relevant in partially regulated environments.¹⁴ A rigorous formulation of this conjecture is left for future research.

¹³If advantageous sorting exists, both firms oversupply quality, but x_0 is more oversupplied than x . A social planner would incent both firms to decrease their quality, but would prioritize a decrease in x_0 , causing *greater* product diversity than is supplied in equilibrium.

¹⁴This characterization relies heavily on a duopoly model. With multiple firms with differentiated products, switchers to high-quality products may substitute into other costly, high-quality products. In this case, the downward distortion of quality could actually be greater for high-quality products.

2.4 Signing terms

The magnitude of the sorting term is endogenous, but its sign can often be ascertained directly, independently of (x, p) and of the distribution of types $f(\theta)$, because \mathcal{M} depends on the preferences of consumers, so WTP u and the set \mathcal{M} are closely related.

Proposition 3. *Suppose exists a function $g = g(x, \theta)$ exists such that $c \equiv \hat{c}(x, g)$ and $u' \equiv \hat{u}(x, g)$. Suppose $\hat{c}(x, g)$ and $\hat{u}(x, g)$ are monotone in g . Then $\text{Cov}[u', c \mid \mathcal{M}]$ has the same sign as $\frac{\partial \hat{c}(x, g)}{\partial g} \frac{\partial \hat{u}(x, g)}{\partial g}$.*

Thus, sorting can be signed whenever u' and c , conditional on \mathcal{M} , can be expressed as monotone univariate functions with the same input “index.” This condition may seem special, but it applies naturally in several contexts, especially when x is a feature of pricing and consumers are rational. One such context, as we illustrate in Section 3.2, is when x captures the slope of a two-part tariff.

Proposition 3 also applies when types are two-dimensional and $u(x, \theta)$ increases in both dimensions of θ . In this case, \mathcal{M} is a one-dimensional curve and, within the margin, τ is a decreasing function of ζ , as illustrated by Figure ???. The function $g(x, \theta)$ required by Proposition 3 can therefore be constructed, as formalized in Corollary 1.

Corollary 1. *Suppose $\theta = (\zeta, \tau) \in \mathbb{R}^2$, and $u(x, \theta), c(x, \theta)$ are increasing in both dimensions of θ . Then $\text{Cov}[u', c \mid \mathcal{M}]$ has the same sign as*

$$S \equiv \left(\frac{\partial u' / \partial \zeta}{\partial u / \partial \zeta} - \frac{\partial u' / \partial \tau}{\partial u / \partial \tau} \right) \left(\frac{\partial c / \partial \zeta}{\partial u / \partial \zeta} - \frac{\partial c / \partial \tau}{\partial u / \partial \tau} \right) \Big|_{\mathcal{M}}$$

when S is signed.

Our primary focus is on the sorting term, but the Spence distortion $\mathbb{E}[u' \mid \mathcal{M}] - \mathbb{E}[u' \mid \mathcal{B}]$ is an important determinant of whether quality is over- or under-provided. The following result provides, to our knowledge, the first conditions for signing it in a context with multidimensional heterogeneity.

Proposition 4. *Suppose $u(x, \theta)$ and $u'(x, \theta)$ are increasing in each dimension of θ and the dimensions of θ are affiliated in the sense of [Milgrom \(1981\)](#). Then $\mathbb{E}[u' \mid \mathcal{M}] \leq \mathbb{E}[u' \mid \mathcal{B}]$. If $u(x, \theta)$ and $-u'(x, \theta)$ are increasing in θ , the inequality is reversed.*

Intuitively, if all dimensions of θ increase both u and u' , those with high u (who are in \mathcal{B}) will have higher u' than those with only moderate u (who are in \mathcal{M}). Affiliation, which need only hold within the set \mathcal{B} , ensures this effect is sufficiently monotone throughout \mathcal{B} to

make this intuition rigorous.¹⁵ Proposition 4 can also be used to sign the Spence distortion when $u \neq w$, based on the sign of $u' - w'$, as we illustrate in Section 3.2.

3 Applications

3.1 Selection vs. sorting in consumer credit markets

Our first application compares the empirical relevance of sorting and selection. Einav, Jenkins and Levin (2012) (henceforth EJJ) study an integrated used car vendor and sub-prime lender that chooses a downpayment x and demand q , implicitly choosing the price ($p = P(x, q)$).¹⁶ Selection arises because buyers differ in default risk, which determines $c(x, \theta)$. EJJ consider consumers heterogeneous in two dimensions: initial liquidity and value for the car. WTP is increasing in each of these dimensions. The authors estimate the joint distribution of buyer types and default probabilities $d = d(x, \theta)$, and determine the effects of changing x and p on the likelihood of default of the pool of buyers.

In the theoretical model that motivates their estimation, EJJ assume unidimensional types, which rules out sorting as discussed in Section 2.2. EJJ interpret their results in terms of “treatment effect” (the impact of x on the default rates of those already purchasing, captured by c' in Proposition 1) and “selection effect” (the difference between default rates of marginal and infra-marginal purchasers, $\mathbb{E}[c|\mathcal{B}] - \mathbb{E}[c|\mathcal{M}]$). However, they do not explicitly discuss sorting. To gauge the quantitative importance of sorting, we decompose the “selection effect” they measure into (i) the selection by q of marginal versus infra-marginal consumers, and (ii) the sorting by x of heterogeneous marginal consumers.

By simulating draws from EJJ’s calibrated model, we can determine \mathcal{B} for any (x, p) . The left panel of Figure 4 shows the difference between marginal consumers whose behavior is affected by a change in p or x . It plots the set of marginal borrowers, with infra-marginals (not shown) lying to the northeast. An increase in price p from \$11,000 by \$500 affects mostly consumers with low initial car value (“ p -marginals,” shown as red points). An increase in downpayment x from \$1000 by \$10 dissuades mostly consumers with low liquidity (“ x -marginals,” shown in blue x’s).

The right panel of Figure 4 graphs the variation in default probabilities among marginal individuals, with individuals more likely to default represented by larger and more saturated circles. EJJ find that default probability is driven almost entirely by liquidity. p -marginals have a wide range of liquidities but low car values, whereas x -marginals have low and fairly homogeneous levels of liquidity, so x -marginals tend to have higher default probabilities

¹⁵Strict versions of this result also hold if the hypotheses are strengthened to be strict.

¹⁶In this application, higher x hurts consumers.

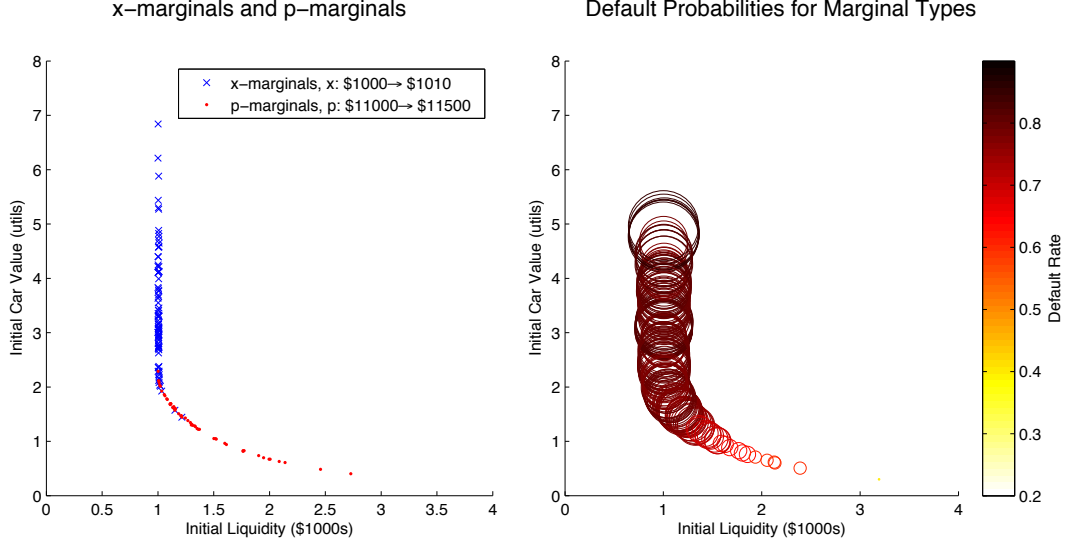


Figure 4: The set of marginal buyers; infra-marginal buyers are to the northeast (not shown). The left panel shows the types who are p -marginal (red points) and x -marginal (blue x's). The right panel shows the probability of default in the calibrated EJM model for marginal buyers, with larger and more saturated circles representing borrowers more likely to default. For further details, see Appendix B.

than p -marginals. Therefore, sorting by x is advantageous to the firm, because increasing x dissuades these costlier marginal borrowers.

Figure 5 quantifies the wedge between selection and sorting for values of (x, p) spanning the range in each dimension observed in their data. At each (x, p) , we compute average default rates among all buyers, x -marginal and p -marginals. Because the profits made off borrowers are closely related to their default probabilities, we can naturally proxy for selection ($\mathbb{E}[c|u \geq p] - \mathbb{E}[c|u = p]$) by the vertical gap between the default probabilities of p -marginals and all buyers, $\mathbb{E}[d|u \geq p] - \mathbb{E}[d|u = p]$.¹⁷ This selection effect captures the fact that the firm may benefit from dissuading marginal buyers by raising price. Increasing x , like increasing p , dissuades all marginal borrowers to some extent. However, sorting captures how an increase in x improves default rates beyond the selection effect achieved by quantity (through price), by disproportionately dissuading those who are particularly risky.

The vertical gap between the default rates of x -marginals and p -marginals is a measure of this sorting effectiveness of x relative to p . In particular, the default rate of x -marginals is $\frac{\mathbb{E}[u'd|u=p]}{\mathbb{E}[u'|u=p]}$ so the gap between the default rates of x - and p -marginals is

¹⁷The relative importance of selection and sorting is measured as the vertical distance between the lines at each point, rather than the slopes of the lines. These selection and sorting effects cannot be judged directly by the impact of changing p or x on average default rates (i.e., the slopes of the lines), because these reflect a mix of selection, sorting, and treatment effects. For instance, raising minimum downpayments directly reduces debt burdens and thus makes repayment more likely even for infra-marginal borrowers.

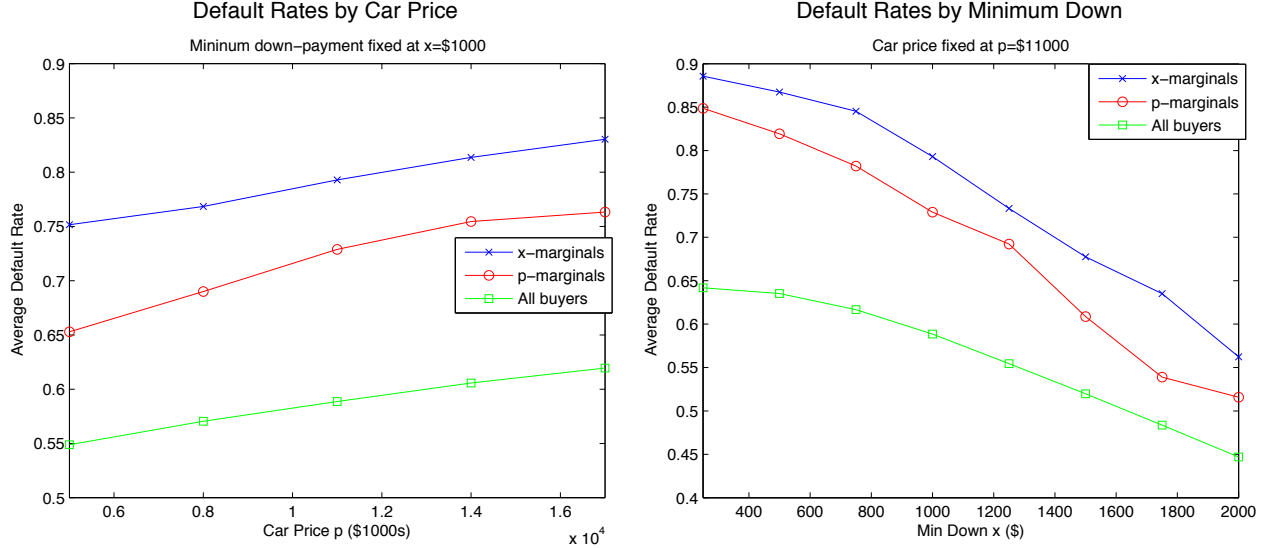


Figure 5: Average default rates as a function of price (left) and minimum downpayment (right). The graphs consider the set of all buyers, marginal buyers when price is raised (p -marginals), and marginal buyers when minimum downpayment is raised (x -marginals). The relative importance of selection and sorting is measured as the vertical distance between the lines for x -marginals and p -marginals at each point, rather than the slopes of the lines. For further details, see Appendix B.

$$\frac{\mathbb{E}[u'd|u=p]}{\mathbb{E}[u'|u=p]} - \mathbb{E}[d|u=p] = \frac{\mathbb{E}[u'd|u=p] - \mathbb{E}[u'|u=p]\mathbb{E}[d|u=p]}{\mathbb{E}[u'|u=p]} = \frac{\text{Cov}[u', d | \mathcal{M}]}{\mathbb{E}[u'|u=p]}.$$

It thus quantifies the importance of sorting in terms directly comparable to selection. Over the range of x and p shown in the figures, x -marginals default with probability about .1 higher than p -marginals, whereas p -marginals default with probability about .15 greater than average purchasers. Thus, the total effect of x on default rates is about $\frac{.15}{.1+.15} = 60\%$ due to selection, whereas the remaining 40% are due to sorting.

Our interest is primarily in establishing the relative magnitude of the two effects, but the distinction is also relevant for policy. Mahoney and Weyl (2014) show the advantageous selection present in the EJL data implies a large over-provision of loans.¹⁸ This distortion may compound with other externalities from excess lending related to systemic financial risk (Mian and Sufi, 2014). On the other hand, downpayment requirements sort advantageously, so lenders will have an incentive to raise downpayments, and excessively so if significant

¹⁸Advantageous selection tends to cause excess supply as described by de Meza and Webb (1987) and Einav and Finkelstein. Mahoney and Weyl estimate a subsidy of about 41% for reasonable assumptions about the market power of the firms.

competition is present as discussed in Section 2.2. Thus, sorting incentives could cause equilibrium lending standards to be excessively strict, mitigating the systemic risk caused by selection incentives. Thus, whether asymmetric information on net mitigates or exacerbates systemic risk depends crucially on the relative importance of selection and sorting.

3.2 Add-on pricing

Our second application illustrates how the generality of Section 2.4 can be combined with additional structure to derive novel applied theory results. We consider a monopolist's incentive to charge excessive add-on fees to imperfectly rational consumers, complementing the competitive analysis of add-on pricing by Ellison (2005) and Gabaix and Laibson (2006).

A bank chooses an account maintenance fee p and an (expected) overdraft fee x .¹⁹ Consumers differ in their income $\omega \geq 0$ and level of activity $a \geq 0$ (an exogenous number of transactions per period). Thus, types are $\theta = (\omega, a) \in \mathbb{R}_+^2$ with PDF $f(\omega, a) > 0$. We assume (ω, a) are positively affiliated, so those with higher income tend to perform more transactions.

Consumers perceive their WTP as $u = a[\alpha\omega - \lambda x]$. Each transaction brings $\alpha\omega > 0$ dollars of WTP, following the multiplicative specification of Mussa and Rosen (1978) and Shaked and Sutton (1982), but also has an independent probability of being an overdraft, resulting in expected overdraft fees x . However, consumers underestimate the overdrafting probability by a factor $\lambda \in [0, 1]$, which captures “sophistication” (e.g., due to overconfidence or limited attention). Thus, a consumer of type a and sophistication λ expects overdraft fees of $a\lambda x$, but in fact incurs expected fees of ax . Following Section 2.3, this imperfect rationality is accommodated by defining the contribution of consumers to welfare (w) to be their WTP without this behavioral bias: $w = a[\alpha\omega - x] < u$.

Strong evidence suggests limited attention is more common among the poor and less educated (Stango and Zinman, Forthcoming; Rees-Jones and Taubinsky, 2015), so we assume $\lambda = \lambda(\omega)$ is weakly increasing in income ($\lambda'(\omega) \geq 0$). If all consumers are neoclassical, as in Ellison (2005), then $\lambda(\omega) = 1, \forall \omega$. We also assume $\frac{\partial u}{\partial \omega} > 0 \Leftrightarrow \frac{\alpha}{x} > \lambda'$.²⁰ The set of buyers is $\mathcal{B} = \{u \geq p\}$ and the set of marginals is $\mathcal{M} = \{u = p\}$. Consumers pay fees to the bank, so the “cost” per consumer is $c = -ax$. The margin is defined by $\mathcal{M} = \{u = p\} = \{\omega = \frac{1}{\alpha}(\frac{p}{a} + \lambda x)\}$.

The overdraft fee is a transfer that reduces gains from trade, so $x^{**} = 0$ is socially optimal. Indeed, in the social planner's FOC, the social Spence and cost terms cancel out,

¹⁹The analysis extends beyond the banking industry to other markets with add-on pricing as discussed by Ellison (2005), Gabaix and Laibson (2006), and Grubb (2009). In this application, higher x hurts consumers.

²⁰This assumption is weak and is discussed further in Appendix C.

so Proposition 2 yields

$$-\underbrace{\mathbb{E}[-a \mid \mathcal{B}]}_{\text{cost}} + \underbrace{\mathbb{E}[-a \mid \mathcal{B}]}_{\text{social Spence term}} + \underbrace{\frac{\text{Cov}[-a\lambda, a\alpha\omega \mid \omega = \frac{1}{\alpha}(\frac{p}{a} + \lambda x^{**})]}{MS}}_{\text{social sorting incentive}} = 0 \Rightarrow \frac{x^{**}\alpha \text{Var}[a\lambda \mid \mathcal{M}]}{MS} = 0.$$

However, Proposition 1 shows the firm's FOC is

$$-\underbrace{\mathbb{E}[-a \mid \mathcal{B}]}_{\text{cost}} + \underbrace{\mathbb{E}[-a\lambda \mid \mathcal{M}]}_{\text{private Spence term}} - \underbrace{\frac{x^* \text{Cov}[-a\lambda, -a \mid \mathcal{M}]}{MS}}_{\text{private sorting incentive}}.$$

A profit-maximizing firm will choose $x^* > 0$ because, at $x^* = 0$, the firm's FOC is positive. To see this result, note the private sorting incentives vanish for $x^* = 0$. Moreover, $-u' = \lambda(\omega)a$ and u are increasing in both variables (ω, a) , which are affiliated. Then Proposition 4 and $\lambda \leq 1$ imply

$$\mathbb{E}[-a\lambda \mid \mathcal{M}] > \mathbb{E}[-a\lambda \mid \mathcal{B}] \geq \mathbb{E}[-a \mid \mathcal{B}].$$

Finally, it can easily be shown that any $x^* < 0$ violates second-order conditions.

Because the sum of Spence and cost terms is positive, satisfying the firm's FOC is only possible if $\text{Cov}[a\lambda, a \mid \mathcal{M}] > 0$. Otherwise, sorting by x is advantageous, so the firm will always gain from further raising x until it faces a regulatory or other constraint. However, as we now show, the sign of sorting depends on the nature of consumer sophistication.

If consumers are rational ($\lambda(\omega) = 1, \forall \omega$), sorting is adverse: $\text{Cov}[u', c \mid \mathcal{M}] = \lambda x^* \nabla[a \mid \mathcal{M}] > 0$.²¹ Intuitively, those who face the overdraft most often are the most profitable, but they are also those most deterred by an increase in the fee. Adverse sorting deters the firm from high x , mitigating the incentive to charge excessive fees. The importance of sorting grows with x^* , so under relatively weak conditions (omitted here for brevity), an interior value of x^* is optimal.

The same analysis applies if consumers are homogeneously unsophisticated, so $\lambda(\omega) = \bar{\lambda} < 1, \forall \omega$. In this case, rearranging Proposition 1 yields

$$x^* = \frac{MS (\mathbb{E}[a \mid \mathcal{B}] - \bar{\lambda} \mathbb{E}[a \mid \mathcal{M}])}{\bar{\lambda} \text{Var}[a \mid \mathcal{M}]}.$$

This expression suggests the fee x^* decreases with the amount of sophistication $\bar{\lambda}$. It also sug-

²¹This argument is an application of Proposition 3, where ω is the relevant index. The result holds also if $\gamma = \gamma(x, \omega, a)$ is endogenous to x and heterogeneous among consumers, in which case $\text{Cov}[u', c \mid \mathcal{M}] = \lambda x^* \nabla[a\gamma \mid \mathcal{M}] > 0$.

gests higher competition, by lowering MS , would tend to lower x^* . That is, when consumers are homogeneously sophisticated, competition should increase welfare. However, explicit comparative statics require stronger modeling assumptions, which we defer to our insurance application in Section 3.3.

However, when consumers differ in their sophistication, sorting by x can be advantageous. By Corollary 1, the sign of the sorting term is that of

$$\left(\frac{\partial u'/\partial \omega}{\partial u/\partial \omega} - \frac{\partial u'/\partial a}{\partial u/\partial a} \right) \left(\frac{\partial c/\partial \omega}{\partial u/\partial \omega} - \frac{\partial c/\partial a}{\partial u/\partial a} \right) = \left(\frac{-a\lambda'x}{a(\alpha - \lambda'x)} + \frac{\lambda x}{\alpha\omega - \lambda x} \right) \left(0 + \frac{x}{\alpha\omega - \lambda x} \right),$$

if this term is globally signed. Appendix C shows this expression has the sign of $1 - \eta$, where $\eta \equiv \omega \frac{\lambda'(\omega)}{\lambda(\omega)}$ is the income elasticity of sophistication. When sophistication is income inelastic ($1 > \eta$), adverse sorting ($\text{Cov}[u', c | \mathcal{M}] > 0$) incents the firm to lower x^* , mitigating excessive fees. However, when sophistication is income elastic ($1 < \eta$), advantageous sorting ($\text{Cov}[u', c | \mathcal{M}] < 0$) exacerbates socially inefficient overdraft fees. In this case, the bank has an incentive to raise x^* arbitrarily high so an interior x^* may not exist.

For an intuition, recall that u increases in both ω and a . Marginal consumers are either high- ω and low- a , or they are low- ω and high- a . Within the margin \mathcal{M} , a and ω have a negative relationship, as in Figure 2, despite being positively affiliated in the population overall.²² The most profitable marginal consumers are always those with high- a (and low- ω), because they pay larger overdraft fees xa . When $\lambda(\omega)$ is constant, high- a consumers are also the consumer most affected by changes in x and thus most responsive to them. Targeting these profitable high- a marginal consumers requires lowering the overdraft fee x , so sorting is adverse. However, if poor consumers are much less sophisticated than the wealthy, the most profitable marginal consumers will have high a , low ω (as before), and also low λ . Now, these profitable consumers are the least responsive marginal consumers to changes in x because of their low sophistication λ . In this case, raising x dissuades the less profitable low- a and high- λ marginal consumers, so sorting is advantageous. The dividing line between these cases is unit elasticity of sophistication with income, given our multiplicative specification for u .

What is a realistic value for this elasticity? In the context of attention to sales taxes, Rees-Jones and Taubinsky (2015) estimate it to be slightly above one third. Stango and Zinman (Forthcoming) consider a setting more similar to what we have in mind (overdraft fees), but do not precisely measure the income elasticity of sophistication. If we use their regression coefficients, we can calculate a partial elasticity of roughly two thirds, controlling for other characteristics such as education and race. Both of these cases suggest income elasticity of

²²This reversal is a stark illustration that the properties of \mathcal{M} depend on WTP u , not on the distribution $f(\theta)$.

sophistication is not great enough to eliminate all restraint on overdraft fees, but they do suggest heterogeneous inattention may significantly erode the incentive to restrain inefficient price discrimination.

3.3 Competition in insurance

Our third application illustrates how market power can improve welfare by blunting harmful sorting incentives. [Rothschild and Stiglitz \(1976, henceforth RS\)](#) and a large subsequent literature argues competing insurers have excessive incentives to offer cheap but low-quality insurance, because such insurance attracts low-risk consumers. RS also suggest frictions, such as multidimensional types or imperfect competition, might reduce the magnitude of harmful cream-skimming incentives. We confirm this intuition in a model of health insurance provision with explicit competition.²³

We extend our monopoly model to a simple competitive Hotelling environment, as in [Villas-Boas and Schmidt-Mohr \(1999\)](#) and [Bénabou and Tirole \(Forthcoming\)](#). Two horizontally differentiated (and otherwise symmetric) insurers are present, indexed by $i, j \in \{0, 1\}$. Each offers a single contract that absorbs a share $x \in [0, 1]$ of each consumer’s medical bills for a premium $p \geq 0$.²⁴

The individual’s expected medical bill is μ . The expected cost of an individual to the insurer is $c = \mu k(x)$. We assume $k(x) = x(1-x)^{-e}$, which implies mean expenditures of individuals have a constant “moral hazard” elasticity $e > 0$ with respect to the price paid by consumers, $1-x$.²⁵

Appendix [D.1](#) shows that further assuming CARA utility and Gaussian wealth shocks with mean μ implies WTP of type (μ, v) for coverage x for an insurer at “distance” d is

$$u = \mu h(x) + \psi(x)v - td.$$

The increasing convex function $h(x) = \frac{1-(1-x)^{1-e}}{1-e}$ is obtained in Appendix [D.1](#) by integrating consumer surplus. This function captures moral hazard because consumer value for mean care falls below the cost of providing it: $h(x) \leq k(x)$.²⁶ The function $\psi(x) = \frac{1}{2}(1 - (1-x)^2) \geq 0$ captures the increasing but concave willingness to pay for a reduction

²³[Einav, Finkelstein and Cullen \(2010\)](#) and [Azevedo and Gottlieb \(2015\)](#) have shown how multidimensional types can mitigate or even reverse the welfare effects of adverse selection under perfect competition. [Villas-Boas and Schmidt-Mohr \(1999\)](#) and [Bénabou and Tirole \(Forthcoming\)](#) have studied the benefits of market power in capital and labor markets with adverse selection.

²⁴In this application, higher x benefits consumers.

²⁵If the amount paid by consumers is y , so $k = xy^{-e}$, the elasticity is $-\frac{dk}{dy} \frac{y}{k} = -\frac{-exy^{-e-1}}{xy^{-e}} y = e$.

²⁶Without moral hazard ($e = 0$), then $h(x) = k(x) = x$.

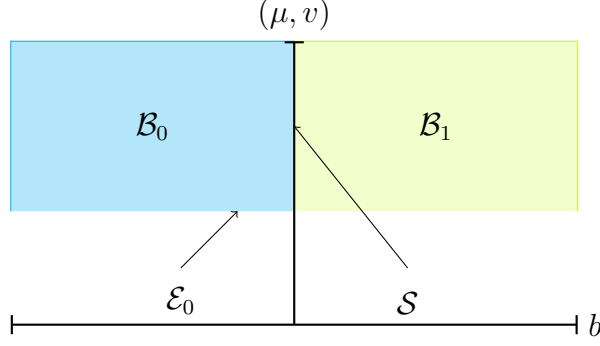


Figure 6: The set of types purchasing from each insurer at an LSE. The two dimensions of type (μ, v) are collapsed into a single axis for clarity.

in risk, which is maximized at full insurance ($x = 1$). The individual's type dimension v captures “insurance value” and scales how much each individual values reduction in risk.²⁷ The “travel cost” t captures market power in this Hotelling environment. A consumer incurs cost td to purchase from an insurer at “distance” d . Each consumer has a Hotelling location parameter $b \in [0, 1]$. The distances d_i to each insurer i are defined as $d_0 = b\mathbf{I}\{b > \frac{1}{2}\}$ and $d_1 = (1 - b)\mathbf{I}\{b \leq \frac{1}{2}\}$, where $\mathbf{I}\{\cdot\}$ is the indicator function. That is, consumers incur no cost if they purchase from their preferred insurer. Types are $\theta = (\mu, v, b) \in \mathbb{R}_+^2 \times [0, 1]$. We assume $b \sim \mathcal{U}[0, 1]$, so θ has PDF $f(\mu, v) > 0$. We assume (μ, v) are positively affiliated.

Following Section 2.3, i 's buyers are the set $\mathcal{B}_i \equiv \{u_i - p_i > \max\{u_j - p_j, 0\}\}$ and i 's exiters are the set $\mathcal{E}_i \equiv \{u_i - p_i = 0 > u_j - p_j\}$. The set of switchers is common to both plans, $\mathcal{S} = \{u_i - p_i = u_j - p_j > 0\}$. The measures of these sets are q_i, E_i , and S , respectively, with $M_i = E_i + S$.

Insurer i maximizes profit $\pi_i = \int_{\mathcal{B}_i} (p_i - c_i) f(\mu, v) d\mu dv db$, taking (x_j, q_j) as given. For simplicity, we focus on local deviations from a symmetric equilibrium in which both insurers choose the same (x^*, q^*) . We call this a “local symmetric equilibrium” (LSE).²⁸ At an LSE, the two symmetrically differentiated providers offer a single insurance contract. We check numerically for non-local deviations from LSE in Appendix D.6. Given our definition of distances d_i , at any LSE, consumers always purchase from their preferred insurer and therefore incur no travel costs. Given symmetry, we henceforth omit the subscripts i, j .

²⁷We assume moral hazard does not affect the variance of expenditures, and thus $\psi(x)v$.

²⁸Formally, an LSE is a pair (x^*, q^*) such that if firm i chooses $(x_i, q_i) = (x^*, q^*)$, an $\epsilon > 0$ exists such that (x^*, q^*) maximizes the profit of firm j in the region $(x, q) \in (x^* - \epsilon, x^* + \epsilon) \times (q^* - \epsilon, q^* + \epsilon)$. Then, if the FOC and SOC are satisfied at (x^*, q^*) , it is an LSE. This concept builds on a recent literature in the design of mechanisms for environments with multidimensional types, such as Erdil and Klemperer (2010) and Carroll (2012). Considering LSEs greatly increases the model's tractability without eliminating the effects of interest present in the literature following RS, such as cream-skimming incentives by firms.

In this application, business-stealing externalities exist between firms. Following Section 2.3, each consumer in \mathcal{B}_i and \mathcal{E}_i contributes to welfare the WTP for i 's plan ($w = u_i$). However, consumers in \mathcal{S} contribute $w = u_i - (u_j - c) = 0$, by symmetry of the LSE. We can then apply Proposition 2 and equation 1 directly to obtain the inefficient cream-skimming distortion. At an LSE, the set of buyers is the same conditional on each b , as illustrated by Figure 6. In particular, the composition of \mathcal{S} is the same as that of \mathcal{B} , so $\mathbb{E}[\cdot | \mathcal{S}] = \mathbb{E}[\cdot | \mathcal{B}]$. Also by symmetry, at any LSE, the cost of the competitor's product is $c_0 = c$. Then, equation 1 yields

$$\underbrace{\frac{DCov[u', c_0 | \mathcal{S}]}{MS}}_{\text{cream-skimming distortion}} = \frac{DCov[u', c | \mathcal{B}]}{MS}.$$

Using this result, Propositions 1 and 2 imply the quality incentive distortion at any LSE is²⁹

$$\frac{1}{q} \left(\frac{\partial W}{\partial x} - \frac{\partial \pi}{\partial x} \right) = \frac{DCov[u', c | \mathcal{B}]}{MS} + (1 - D) (\mathbb{E}[u' | \mathcal{B}] - \mathbb{E}[u' | \mathcal{E}]).$$

Appendix D.3 shows that because u' , u , and c are each increasing in (μ, v) , affiliation of (μ, v) implies $Cov[u', c | \mathcal{B}] > 0$. Moreover, by Proposition 4, the second term is positive. Intuitively, reducing quality x sorts in favor of low-cost buyers (RS) and extracts rents from infra-marginal customers (Stiglitz, 1977). Together, these imply that competing insurers tend to under-provide quality, as formalized below.

Proposition 5. *Let the market quantity be fixed at a given q . At any LSE, for a fixed q , insurers under-provide quality: $x^* < x^{**}$.*

Like the results of Spence and Sheshinski, this conclusion is weak. It only compares optimal quality for a fixed quantity because the comparative statics interactions between x^* and q^* are complex. In a market where all individuals are mandated to have coverage, such as the United States under the Patient Protection and Affordable Care Act, Appendix D.2 shows $x^* < x^{**}$ always, even without the affiliation assumption. It also shows market power always brings x^* closer to x^{**} , so welfare always increases with market power.

However, we focus on the case of an incompletely covered market, where quantity is also endogenous. This setting is more realistic, but also more interesting because competition policy involves a real trade-off: market power may increase quality, but it may also decrease quantity, as emphasized by Mahoney and Weyl (2014) and Handel, Hendel and Whinston (2015, henceforth HHW).³⁰

²⁹This result uses $\mathbb{E}[u' | \mathcal{S}] = \mathbb{E}[u' | \mathcal{B}]$ and $\mathbb{E}[u' | \mathcal{M}] = D\mathbb{E}[u' | \mathcal{S}] + (1 - D)\mathbb{E}[u' | \mathcal{E}]$, where $D = \frac{S}{M}$.

³⁰Mahoney and Weyl (2014) emphasize that market power can directly increase price. Handel, Hendel and Whinston (2015) emphasize that if market power induces higher x^* , such an increase may likewise increase adverse selection and cost and thereby indirectly contribute to an increase in price.

In this setting, analytically determining the impact of market power on quality and welfare is cumbersome. We thus study these outcomes numerically in a calibrated model. We assume (μ, v) have a joint log-normal distribution with parameters matched to the means and variance-covariance matrix of the HHW data.³¹ HHW find μ and v correlate positively, which, given log-normality, implies affiliation.³² We do not analyze the perfectly competitive case here, because an LSE often fails to exist, for reasons similar to those in the analysis of RS. In particular, inspection of second-order conditions (SOCs) shows no LSE exists for small enough t given the HHW data.³³

Following the classic RAND experiment, we calibrate $e = 0.2$. We numerically compute the LSE FOCs and corresponding values of (x^*, q^*) for several values of the market power parameter t . We test whether each (x^*, q^*) actually constitutes an LSE by checking SOCs, and find these conditions hold for $t \geq 3450$.³⁴ We also report the implied LSE price p^* .

The welfare-maximizing contract has quantity $q^{**} \approx 1$ and quality $x^{**} \approx 0.90$ (price is $p^{**} \approx \$272$), and we refer to this benchmark as the “first-best.” The quality x^* , quantity q^* , welfare, and consumer surplus (CS) at each LSE are shown in Figure 7. To ease interpretation, we present these results as a function of the markup to average cost at each LSE, $R = \frac{p^* - \mathbb{E}[c|\mathcal{B}]}{\mathbb{E}[c|\mathcal{B}]}$, which increases with t over most of the relevant range. We consider LSE welfare $W(x, q) = 2q\mathbb{E}[u - c | \mathcal{B}]$ and graph it as a share of first-best welfare, $\frac{W^*}{W^{**}} = \frac{W(x^*(R), q^*(R))}{W(x^{**}, q^{**})}$. Similarly, we consider LSE consumer surplus $CS(x, q) = 2q\mathbb{E}[u - p | \mathcal{B}]$ and graph $\frac{CS^*}{CS^{**}} = \frac{CS(x^*(R), q^*(R))}{CS(x^{**}, q^{**})}$. Further computational details are provided in Appendix D.5.³⁵

For sufficient market power ($t \geq 3450$), we find a unique LSE whose properties we report. The x^* value that satisfies the firm FOCs continues to fall as $t \rightarrow 0$ beyond this

³¹Summary statistics in the HHW data are $\mathbb{E}[\mu] = 6.6 \times 10^3$, $\mathbb{E}[v] = 68 \times 10^3$, $\mathbb{V}[\mu] = 50 \times 10^6$, and $\text{Cov}[\mu, v] = 630 \times 10^6$. For additional details, see Appendix D.4.

³²In Appendix D.2, we show analytically that, in the case of a covered market, this positive correlation implies cream-skimming is more harmful in a multidimensional model in which consumers differ in (μ, v, b) than in a simpler unidimensional model where consumers differ only in μ (along with b). That is, heterogeneity in v exacerbates rather than mitigates cream-skimming, which is consistent with the findings of Azevedo and Gottlieb for the case of perfect competition.

³³Appendix D.2 shows that, in a covered market, the firm’s FOCs are satisfied at a unique x , for any $t > 0$, if $\beta \equiv \frac{\text{Cov}[\mu, v]}{\mathbb{V}[\mu]} < -1$. However, this point may often fail to actually be an LSE if the induced value of x^* is small enough that second-order conditions are violated. If SOCs are satisfied, in the limit as $t \rightarrow 0$, $x^* \rightarrow 1 + \frac{1}{\beta}$, which is the unique value of x where $\text{Cov}[u', c] = 0$. These results are highly consistent with those obtained by Azevedo and Gottlieb, who allow for general contracts and use a price-taking competitive equilibrium concept that always exists.

³⁴We compute the SOCs using a combination of analytical methods and numerical quadrature in Mathematica. The analytic portions are cumbersome so we do not report them here, but they appear in the code posted on our websites at www.andreveiga.com and www.glenweyl.com.

³⁵We round all figures to two significant digits.

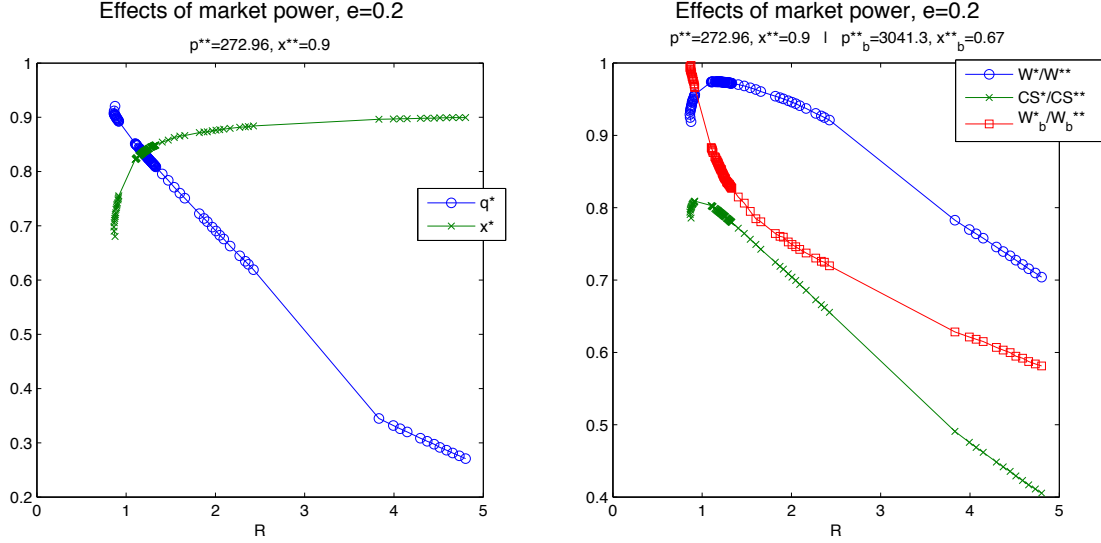


Figure 7: On the left panel, as a function of relative markup R , is the share of individuals who purchase insurance (q^*), insurance quality (x^*). On the right panel, as a function of R , is welfare as a share of its maximum ($\frac{W^*}{W^{**}}$), consumer surplus as a share first-best CS ($\frac{CS^*}{CS^{**}}$), and behavioral welfare as a share of its maximum ($\frac{W_b^*}{W_b^{**}}$). For details, see Appendix D.5.

point, but these values are no longer an LSE, because SOC's are violated.³⁶ Azevedo and Gottlieb study a concept of equilibrium that always exists in our environment even as $t \rightarrow 0$, and find insurance quality is severely downward distorted from the first-best under perfect competition. Thus, interpreting x^* as continuing to fall as t falls along the non-existence region seems natural. That is, when markets are very competitive, sorting incentives loom large as emphasized in Section 2.2, causing firms to offer low-quality insurance.

As R is allowed to increase, x^* increases and q^* decreases as expected. The LSE with the largest welfare has markup $R \approx 1.1$, quantity $q^* \approx 0.82$, and quality $x^* \approx 0.83$ ($t = 7500$, $p^* \approx 11000$). At this LSE, welfare is over 97% of its first-best level. The LSE with the largest CS has $R \approx 0.92$, $q^* \approx 0.91$, and $x^* = 0.76$ ($t = 4600$, $p^* \approx \$8100$), and CS is 81% of its first-best level. Thus, a large share of the first-best welfare can be achieved through a degree of market power higher than in present markets, according to the markups reported in Dafny, Duggan and Ramanarayanan (2012).

We then use the flexibility of our approach to consider an alternative welfare standard. We consider the same equilibrium pairs of (x^*, q^*) as above, because these are generated by observed behavior. However, we now incorporate Handel and Kolstad (2015) evidence that

³⁶This failure of existence is consistent with the non-existence of the pooling equilibria result under perfect competition derived by RS.

consumers are excessively risk averse over the small stakes involved in comparing different insurance plans. We compute welfare assuming consumers derive utility according to true insurance value $v_b = .054v$.³⁷ Behavioral welfare at an LSE is $W_b(x, q) = 2q\mathbb{E}[w_b - c \mid \mathcal{B}]$, where we still use $\mathcal{B} = \{u \geq p\}$, but $w_b = \mu h(x) + \psi(x)v_b$ reflects true insurance value v_b .

W_b is maximized at quantity $q_b^{**} \approx 0.98$ and quality $x_b^{**} \approx 0.67$ (price is $p_b^{**} \approx \$3000$). Figure 7 graphs behavioral welfare as a share of its maximum, $\frac{W_b^*}{W_b^{**}} = \frac{W_b(x^*(R), q^*(R))}{W_b(x_b^{**}, q_b^{**})}$. Optimal quality is lower than under-revealed preference ($x_b^{**} < x^{**}$), as is quantity ($q_b^{**} < q^{**}$), to mitigate insurance overconsumption. Because insurance quality increases with market power, the LSE that maximizes behavioral welfare features a lower level of market power ($R \approx 0.88$, $t = 3550$). This LSE induces quality $x^* \approx 0.68$ and quantity $q^{**} \approx 0.92$ (and price $p^* \approx \$6600$).³⁸ Because the optimal quantity is lower under the behavioral welfare standard, this level of market power induces behavioral welfare, which is 99.7% of its maximum. The level of market power implied by this more realistic calibration (priced 88% above cost) is not far above the levels observed in practice. In sum, some significant degree of market power is optimal and can achieve close to the first-best in an otherwise unregulated market, but the quantitatively optimal degree of market power naturally depends on the standard of welfare evaluation, which can be flexibly adjusted in our model.³⁹

4 Conclusion

This paper makes three contributions. First, we propose a modeling framework with a simplified contract space and multidimensional consumer heterogeneity that makes analyzing sorting incentives in selection markets tractable. Second, we derive from this model a characterization of sorting incentives as the ratio of two terms. The numerator is the covariance, among marginal consumers, between marginal willingness-to-pay for quality and cost (or welfare in the case of a social planner). The denominator is marginal consumer surplus, a common measure of market power. Finally, we use this characterization to quantify the

³⁷ $\frac{8.6 \times 10^{-5}}{1.6 \times 10^{-3}} = .054$ is the ratio of the average CRRA risk-aversion parameters estimated in the behavioral and neoclassical versions of their empirical model. [Handel and Kolstad \(2015\)](#) find significant heterogeneity in this behavioral component, although we abstract from this heterogeneity in our analysis. [Spinnewijn \(2015\)](#) shows these heterogeneous biases have important welfare effects as they lead to distortions in the sorting of individuals into purchase.

³⁸ This is the point at which behavioral welfare is highest among those satisfying both FOC and SOC, thus constituting an LSE. Behavioral welfare is actually slightly higher at the unique point that satisfies the FOC when $t = 3400$, but that point is not an LSE as SOC fails.

³⁹ These results should be taken with caution, however, because a variety of other regulatory interventions may help prevent cream-skimming in these markets and make market power a less attractive remedy. We discuss the implications of these results for competition policy in detail in a companion piece ([Mahoney, Veiga and Weyl, 2014](#)).

importance of sorting empirically in subprime lending, determine its direction analytically for a firm engaging in add-on pricing, and calibrate its implications for competition policy in insurance markets.

Perhaps the greatest weakness of our analysis is that firms are restricted to offering a single contract. This assumption is clearly counterfactual to many prominent selection markets: life insurers offer a variety of types of insurance (whole, term, universal) with a variety of financial terms, mortgage contracts come in a variety of flavors, and so on. Our analysis cannot give any insight into the range of products that are offered or adopted. This weakness is particularly acute in extremely competitive markets where the sort of symmetric pooling equilibrium we study in Section 3.3 often fails to exist but asymmetric equilibria often do. [Azevedo and Gottlieb \(2015\)](#) define and study the perfectly competitive equilibria of a model with multidimensional heterogeneity with a rich contract space. Although many of their substantive results (about the statistical properties of the type distribution that affect selection and sorting) are consistent with and complementary to ours, an exciting direction for future research is to develop a framework allowing for multidimensional types, rich contracting spaces and market power. Such a framework could, for example, be used to study more fully and formally our conjectures in Section 2.3 comparing privately and socially optimal product diversity.

Beyond the markets studied in the applications of Section 3, selection occurs in many other contexts, including employment relationships, matching markets with non-transferable utility, and platform markets.⁴⁰ Research on these markets has, in recent years, increasingly focused on the possibility of multidimensional types to which our analysis is adapted.

A number of theoretical issues remain unaddressed in our analysis. First, it would be interesting to study alternatives to our assumption, in the competitive setting of Section 3.3, that switching consumers are representative of the full population of consumers, as [Bonatti \(2011\)](#) does for the [Rochet and Stole \(2002\)](#) context. Second, it would be useful to generalize the conditions of Section 2.4, under which the sorting incentive and Spence distortion can be signed. Third, it would also be useful to expand our approach to a small number of contracts and to asymmetric equilibria in multi-firm markets. Finally, additional analysis of the interaction between sorting and selection distortions seems like a promising avenue for future research.

NUFFIELD COLLEGE, UNIVERSITY OF OXFORD

MICROSOFT RESEARCH NEW ENGLAND AND UNIVERSITY OF CHICAGO

Appendix

⁴⁰An earlier version of this paper ([Veiga and Weyl, 2012](#)) contains several analyses of such settings.

A Monopoly Sorting

A.1 First-order conditions

Proof of Proposition 1 and Proposition 2. First, we use the implicit function theorem (IFT) on the equation that defines the margin, $u(x, \zeta, \tilde{\tau}(x, p, \zeta)) - p = 0$, to obtain

$$\frac{\partial \tilde{\tau}}{\partial p} = \frac{1}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}}, \quad \frac{\partial \tilde{\tau}}{\partial x} = -\frac{u'(x, \zeta, \tilde{\tau})}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}}.$$

Second, we derive $\frac{\partial P(x, q)}{\partial x}$. To do so, we differentiate $Q = \int_{\zeta} \int_{\tilde{\tau}(x, p, \zeta)}^{\bar{\tau}} f(\theta) d\theta$ with respect to x , borrowing $\frac{\partial Q}{\partial p} = -M$ from our analysis in the text:

$$\frac{\partial Q}{\partial x} = \int_{\zeta} \frac{u'(x, \zeta, \tilde{\tau}) f(\zeta, \tilde{\tau})}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}} d\zeta = M \mathbb{E}[u' \mid \mathcal{M}].$$

Applying the IFT to the equation that defines inverse demand, $Q(x, P(x, q)) = q$, yields

$$\frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial p} \frac{\partial P(x, q)}{\partial x} = 0 \Leftrightarrow \frac{\partial P}{\partial x} = -\frac{M \mathbb{E}[u' \mid \mathcal{M}]}{-M} = \mathbb{E}[u' \mid \mathcal{M}].$$

Third, we tackle profit and welfare maximization simultaneously by considering

$$Z(x, q) = \int_{\zeta} \int_{\tilde{\tau}(x, P(x, q), \zeta)}^{\bar{\tau}} z(x, q, \theta) f(\theta) d\theta.$$

Simple differentiation yields (dropping arguments where possible)

$$\begin{aligned} \frac{\partial Z}{\partial x} &= \int_{\zeta} \int_{\tilde{\tau}}^{\bar{\tau}} \frac{\partial z(x, q, \theta)}{\partial x} f(\theta) d\theta + \int_{\zeta} -\left(\frac{\partial \tilde{\tau}}{\partial x} + \frac{\partial \tilde{\tau}}{\partial p} \frac{\partial P}{\partial x}\right) z(x, \zeta, \tilde{\tau}) f(\zeta, \tilde{\tau}) d\zeta \\ &= q \mathbb{E}\left[\frac{\partial z}{\partial x} \mid \mathcal{B}\right] + \int_{\zeta} \left(\frac{u'}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}} - \frac{\mathbb{E}[u' \mid \mathcal{M}]}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}}\right) z f d\zeta \\ &= q \mathbb{E}\left[\frac{\partial z}{\partial x} \mid \mathcal{B}\right] + M (\mathbb{E}[u' z \mid \mathcal{M}] - \mathbb{E}[u' \mid \mathcal{M}] \mathbb{E}[z \mid \mathcal{M}]) = q \mathbb{E}\left[\frac{\partial z}{\partial x} \mid \mathcal{B}\right] + M \text{Cov}[u', z \mid \mathcal{M}]. \end{aligned}$$

Profit maximization considers $z = P(x, q) - c(x, \theta)$, so given P is invariant to θ ,

$$\frac{\partial \pi}{\partial x} = q \mathbb{E}[u' \mid \mathcal{M}] - q \mathbb{E}[c' \mid \mathcal{B}] - M \text{Cov}[u', c \mid \mathcal{M}].$$

Welfare maximization considers $z = w(x, \theta) - c(x, \theta)$, so

$$\frac{\partial W}{\partial x} = q \mathbb{E}[w' - c' \mid \mathcal{B}] + M \text{Cov}[u', w - c \mid \mathcal{M}].$$

□

First Order Conditions with respect to q . By the IFT $\frac{\partial P}{\partial q} = -\frac{1}{M}$. Then,

$$\frac{\partial Z}{\partial q} = \int_{\zeta} \int_{\tilde{\tau}} \frac{\partial z(x, q, \theta)}{\partial q} f(\theta) d\theta + \int_{\zeta} \left(-\frac{\partial \tilde{\tau}}{\partial p} \frac{\partial P}{\partial q} \right) z f d\zeta = q \mathbb{E} \left[\frac{\partial z}{\partial q} \mid \mathcal{B} \right] + \frac{1}{M} M \mathbb{E} [z(x, \zeta, \tilde{\tau}) \mid \mathcal{M}],$$

where $\frac{\partial \tilde{\tau}}{\partial p} = \frac{1}{\frac{\partial u(x, \zeta, \tilde{\tau})}{\partial \tau}}$ was derived in the previous proof. Profit maximization considers $z = P(x, q) - c(x, \theta)$ so the FOC becomes $\frac{\partial \pi}{\partial q} = -\frac{q}{M} + (p - \mathbb{E}[c \mid \mathcal{M}]) = 0$. Welfare maximization considers $z = w(x, \theta) - c(x, \theta)$, so the FOC becomes $\frac{\partial W}{\partial q} = \mathbb{E}[w - c \mid \mathcal{M}] = 0$. \square

A.2 Signing the sorting and Spence terms

Proof of Proposition 3. By assumption, $\text{Cov}[u', c \mid X] = \text{Cov}[\hat{u}(x, g), \hat{c}(x, g) \mid X]$ for any subset of consumers X . If $\hat{c}(x, g)$ and $\hat{u}(x, g)$ are co-monotonic in g , from Schmidt (2003), the covariance of two co-monotone functions of a single variable is positive. If $\hat{c}(x, g)$ and $\hat{u}(x, g)$ are anti-monotone, their covariance is negative. \square

Proof of Corollary 1. The set marginals is defined by $u(x, \zeta, \tilde{\tau}(p, x, \zeta)) = p$. Thus, conditional on the margin, marginal WTP is $u'(x, \zeta, \tau) = u'(x, \zeta, \tilde{\tau}(p, x, \zeta)) \equiv \hat{u}(x, p, \zeta)$. Similarly, $c(x, \zeta, \tau) = c(x, \zeta, \tilde{\tau}(p, x, \zeta)) \equiv \hat{c}(x, p, \zeta)$. Then,

$$\frac{\partial \hat{u}}{\partial \zeta} = \frac{\partial u'}{\partial \zeta} + \frac{\partial u'}{\partial \tau} \frac{\partial \tilde{\tau}}{\partial \zeta} = \frac{\partial u'}{\partial \zeta} + \frac{\partial u'}{\partial \tau} \left(-\frac{\frac{\partial u}{\partial \zeta}}{\frac{\partial u}{\partial \tau}} \right),$$

where $\frac{\partial \tilde{\tau}}{\partial \zeta} = -\frac{\frac{\partial u}{\partial \zeta}}{\frac{\partial u}{\partial \tau}}$ by applying the IFT to $u(x, \zeta, \tilde{\tau}(p, x, \zeta)) = p$. Because $\frac{\partial u}{\partial \zeta} > 0$, $\frac{\partial \hat{u}}{\partial \zeta}$ has the same sign as $\frac{\partial u'}{\partial \zeta} \frac{1}{\frac{\partial u}{\partial \zeta}} - \frac{\partial u'}{\partial \tau} \frac{1}{\frac{\partial u}{\partial \tau}}$, which is the first term in the expression given by Proposition 3. The same result follows for \hat{c} , yielding the second term in the Proposition. \square

Proof of Proposition 4. Because $u(x, \theta)$ and $u'(x, \theta)$ are increasing functions of affiliated random variables, they are affiliated by the results of Milgrom and Weber (1982). Thus $\mathbb{E}[u' \mid u = p] \leq \mathbb{E}[u' \mid u > p]$. \square

B Credit Markets

For the left panel of Figure 4, we simulate 10^6 consumers, each with an initial car value and initial liquidity, following the distribution estimated by EJJL. For each consumer, we simulate a stream of stochastic income shocks, also following the model in EJJL. For each consumer, we compute whether she purchases the car in three scenarios: (1) when $(x, p) = (\$1000, \$11000)$, (2) when $(x, p) = (\$1010, \$11000)$, and (3) when $(x, p) = (\$1010, \$11500)$. For instance, those types who purchase the car in the baseline but do not purchase after the increase in p are “ p -marginals.” The graphs plot only a subset (1%) of the simulated consumers, for clarity.

For the right panel of Figure 4, we simulate 5×10^6 consumers. We determine the set of all marginal consumers (x - and p -marginals together) following the procedure above. For each consumer, we determine whether they default on the car payments, given their type

and (random) draw of income shocks. We then segment marginal consumers into several subgroups, based on their initial car value. We consider car value in intervals of \$250 within [\$0, \$6000]. Within each interval, we compute the default rate. We then plot a subset (1%) of all marginal types, according to the default probability of their subgroup. The distribution of default probabilities is not monotonic because of the noise in the stochastic income shocks simulated for each consumer.

To produce Figure 5, we simulate 10^6 consumers. We then consider x in equally spaced intervals of \$250 within [\$0, \$2000], and consider p in equally spaced intervals of \$300 within [\$5000, \$17000]. We compute the set of buyers at each combination of (x, p) . We identify x - and p -marginals again by considering a \$10 increase in x and a \$500 increase in p . We then compute, at each point, the default rates among all buyers, x -marginals and p -marginals.

C Add-On Pricing

We assumed $\lambda'(\omega) < \frac{\alpha}{x}$ to guarantee u increases in ω . This condition is endogenous but quite weak. Suppose, for example, that individuals are perfectly sophisticated once they have reached an income of \$100k (if they were not perfectly sophisticated at this income, the results would be biased in favor of our assumption) and that sophistication never increases at more than twice the average rate between an income of 0 and \$100k ($\lambda' < 2\mathbb{E}[\lambda']$). Then a sufficient condition for our assumption to hold is that x is no more than half of the total value of the activity to an individual earning \$100k ($x \leq \frac{\alpha x}{2}\lambda(100k)$), which strikes us as a very weak assumption. Primitive conditions can also be given that ensure this condition holds, but they are too strong for our purposes. For example, if $\eta_\lambda \leq 1$, then for any individual purchasing given that price is strictly positive, $\omega > \lambda x$ so that $\frac{\alpha}{x} > \frac{\lambda}{w} = \frac{\lambda'}{\eta_\lambda} \geq \lambda'$.

Proposition 6. *$Cov[u', c \mid \mathcal{M}]$ has the same sign as $1 - \eta$, where $\eta \equiv \omega \frac{\lambda'(\omega)}{\lambda(\omega)}$, if $1 - \eta$ is signed globally.*

Proof. By the reasoning in the text, we must show $1 - \eta$ has the same sign as

$$x \left(\frac{\lambda \gamma}{\alpha \omega - \lambda x} - \frac{\lambda' a}{a(\alpha - \lambda' x)} \right) \left(\frac{1}{\alpha \omega - \lambda x} - \frac{0}{a(\alpha - \lambda' x)} \right) \propto \frac{\lambda}{\alpha \omega - \lambda x} - \frac{\lambda' a}{a(\alpha - \lambda' x)} \propto$$

$$\lambda a(\alpha - \lambda' x) - \lambda' a(\alpha \omega - \lambda x) = \alpha \lambda (1 - \eta).$$

Here, \propto denotes preservation of sign, and our claim relies on the fact that $\frac{1}{\alpha \omega - \lambda x}$ and the hypothesis that $\lambda'(\omega) < \frac{\alpha}{x}$.⁴¹ This completes the proof as $\alpha, \lambda > 0$. \square

⁴¹The result can be proven without this hypothesis, but we omit that proof for brevity (it is available upon request).

D Competition in Insurance

D.1 Deriving WTP and cost

We derive the expressions of WTP and cost from a more primitive model. Constant absolute risk aversion (CARA) consumers maximize the expected value (over realizations of a shock) of wealth equivalents $-e^{-aw}$, where w is final wealth equivalent and $a > 0$ is absolute risk aversion.

Individuals obtain wealth-equivalent value ϕ from mean health care expenditures of m , where $\phi = \mu^{\frac{1}{e}} \frac{1}{e-1} e m^{\frac{e-1}{e}}$, where μ is an individual-specific expenditure-shifter, and e is a common constant moral hazard elasticity. If an individual has insurance with actuarial rate x , the cost of health expenditures m is $(1-x)m$ dollars. Optimal mean expenditures are thus given by the first-order condition

$$\mu^{\frac{1}{e}} m^{-\frac{1}{e}} = 1 - x \Rightarrow m^* = \frac{\mu}{(1-x)^e}.$$

The expected cost of covering an individual of type μ with actuarial rate x is thus $c(x, \mu, v) = \frac{x\mu}{(1-x)^e}$. Moral hazard also implies net utility achieved under an actuarial rate of x of

$$h(\mu, x) = -\frac{\mu(1-x)^{1-e}}{(1-e)}.$$

In addition to this mean utility, individuals face an unpredictable Gaussian variance in their health expenditures with individual-specific variance σ^2 . By the standard CARA-normal formula, the mean wealth-equivalent loss for this variance is $\frac{a\sigma^2}{2}$ without insurance, or $(1-x)^2 \frac{a\sigma^2}{2}$ if the individual is insured at actuarial rate x . Thus, the wealth equivalent of no insurance ($x = 0$) is $-\frac{\mu}{1-e} - \frac{a\sigma^2}{2}$, and the net WTP for an actuarial rate of x relative to no insurance is

$$u = \frac{\mu}{1-e} [1 - (1-x)^{1-e}] + \frac{a\sigma^2}{2} [1 - (1-x)^2].$$

D.2 Analytic results with a covered market

This section derives stronger results than those presented in the main text, under the assumption that the market is covered; that is, consumers may only choose which of the firms to purchase from and may not choose to forgo purchasing.

Proposition 7. *In a covered market, any LSE x^* solves*

$$\mathbb{E}[u' - c'] = \frac{1}{t} \text{Cov}[u', c]. \quad (2)$$

For any $t > 0$, exactly one solution $x^ \in (0, 1)$ to this equation exists.*

Proof. At an LSE, by symmetry, $q = \frac{1}{2}$, and because the market is covered, the set of switching consumers are all those with $b = \frac{1}{2}$ because the products are identical. This set

of consumers has the same $\theta = (\mu, v)$ distribution as all consumers, and thus the relevant expectations are all unconditional. Furthermore, $M = S = -\frac{\partial Q}{\partial p} = \frac{1}{2t}$.

The first part of the proposition follows directly from Proposition 1. Note these results hold regardless of the specification of the utility and cost functions.

We now plug in $u' = \frac{\mu}{(1-x)^e} + v(1-x)$ and $c' = \frac{\mu}{(1-x)^e} + e\frac{x\mu}{(1-x)^{e+1}}$. Thus, equation 2 yields

$$(1-x)\mathbb{E}[v] - \frac{e\mathbb{E}[\mu]}{(1-x)^{e+1}}x = \frac{1}{t}\text{Cov}\left[\frac{\mu}{(1-x)^e} + v(1-x), \frac{x\mu}{(1-x)^e}\right] = \frac{x\mathbb{V}[\mu] + x(1-x)\text{Cov}[\mu, v]}{t(1-x)^e} \Rightarrow$$

$$\frac{(1-x)^e}{x}\mathbb{E}[v] = \frac{1}{t}\frac{\mathbb{V}[\mu]}{1-x} + \frac{e\mathbb{E}[\mu]}{(1-x)^2} + \frac{1}{t}\text{Cov}[\mu, v].$$

The left-hand side of this equation is strictly decreasing in x and approaches ∞ as $x \rightarrow 0$, whereas the RHS is strictly increasing in x and approaches ∞ as $x \rightarrow 1$. Thus, precisely one solution to the equation must exist by the mean value theorem given that all terms are continuous. \square

In what follows, we use a natural measure of the statistical relationship between μ and v repeatedly, namely, the ordinary least squares regression coefficient of v onto μ , $\beta \equiv \frac{\text{Cov}[\mu, v]}{\mathbb{V}[\mu]}$.

Proposition 8. *Suppose $\beta > -1$ (μ and v are not too negatively correlated). Then sorting is adverse at any LSE, and insurance is lower at the LSE than is socially optimal ($x^* < x^{**}$). Moreover, $\frac{dx^*}{dt} > 0$.*

Proof. For any $x \in (0, 1)$, sorting $\frac{x\mathbb{V}[\mu] + x(1-x)\text{Cov}[\mu, v]}{t(1-x)^e} > 0$ because it has the same sign as $\mathbb{V}[\mu] + (1-x)\text{Cov}[\mu, v]$ and thus of $1 + (1-x)\beta > 0$ as $\beta > -1$. Thus, sorting is adverse at any LSE.

Following the logic of Proposition 7, but applied to the competitive welfare-maximizing condition derived in Section 2.3, yields $\mathbb{E}[u' - c'] = 0$, which again has a unique solution for the same reasons above.

By the standard comparative static argument (Milgrom and Shannon, 1994), adding a strictly negative term to this equation must cause x^* to strictly fall, so that the welfare-maximizing x^* level is above the LSE level. Finally, by the same logic, because raising t strictly reduces the magnitude of this strictly negative term, x^* must rise in t . \square

Although we focus on imperfect competition, we briefly discuss the limit when $t \rightarrow 0$ in order to compare the results of our model with the perfectly competitive models of RS and Azevedo and Gottlieb (2015, henceforth AG). The low transport-cost limit of imperfect competition is frequently interpreted as perfect competition; for example, see Rochet and Stole (2002) and Rochet and Tirole (2003). For consistency with RS, we assume away moral hazard for the purpose of these results (i.e., $e = 0$).

Proposition 9. *Suppose $e = 0$. Let $\gamma \equiv \frac{1}{4}\frac{\mathbb{E}[v]^2}{\mathbb{V}[\mu]} > 0$, $\bar{\beta} \equiv \gamma - 1$ and $\underline{\beta} < -1$ be the unique solutions to $2\gamma = -(\beta^3 + \beta^2)$. In the limit as $t \rightarrow 0$, we obtain the following:*

First, if $\beta < \underline{\beta}$, the unique LSE has $x^ = 1 + \frac{1}{\beta}$.*

Second, if $\underline{\beta} < \beta < \bar{\beta}$, no LSE exists.

Third, if $\beta > \bar{\beta}$, the unique LSE has $x^* = 0$.

Given this result is very tangential to our analysis, we omit its proof here; we included it in the working version of this paper. We instead discuss only the relationship between this result and the work of AG. First, note that over a substantial range of β (in our calibration, down to $\underline{\beta} = -6$ and up to $\bar{\beta} = 22$), no LSE exists. This conclusion is consistent with the non-existence results of RS, given that in their model, effectively $\beta = 0$, which in our model always produces non-existence. Second, note that equilibrium insurance, when a positive insurance LSE exists, increases in the degree of negative correlation ($-\beta$) between μ and v . This result is consistent with AG's finding that the stronger the negative relationship between μ and v , the greater the average insurance quality in equilibrium. Finally, also consistent with their results, we find that when the correlation between μ and v is strongly positive, all insurance is driven out of the market. Thus, in the case in which they are comparable (in the perfectly competitive limit), our results closely parallel results, allowing for arbitrary contracts (RS and AG).

D.3 Equilibrium insurance is insufficient

Proof of Proposition 5. By the comparative static argument used to establish Proposition 8 above, for a fixed q , it suffices to show that both terms identified in the text in the divergence between the social and private first-order conditions are positive. The Spence distortion can be signed by the assumption of affiliation and Proposition 4.

For sorting, recall the distribution of $\theta = (\mu, v)$ in \mathcal{S} is the same as in \mathcal{B} . Let

$$u'_+(x, p, \theta) \equiv \begin{cases} u'(x, \theta) & u(x, \theta) > p \\ 0 & u(x, \theta) \leq p \end{cases},$$

and similarly for $c_+(x, p, \theta)$. Note that because $u', c > 0$ and u, u', c are monotone increasing in all components of θ , u'_+, c_+ are monotone increasing functions of both components of θ and thus are affiliated by the argument in the proof of Proposition 4. Furthermore, by definition, $\text{Cov}[u', c \mid \mathcal{B}] = \text{Cov}[u'_+, c_+] > 0$, because affiliated functions have positive covariance. \square

In a working version of the paper, we also showed that, regardless of the statistical relationship between μ and v , $\text{Cov}[u', c \mid \mathcal{E}] > 0$. The reason is that the insurance motive $\psi(x)$ is concave, whereas risk-transfer motive $k(x)$ is linear (with no moral hazard) or convex (with moral hazard, though we did not consider this case in the previous draft). We do not include the proof of this result here, because it is not relevant to any results we state in the text, but we note it here as an application of Corollary 1.

D.4 Simulating the HHW market

HHW use proprietary-claims data from a firm that uses cross-subsidies to achieve a variety of objectives other than single-plan profit maximization, such as employee retention and productivity. Those data are not from a competitive market.⁴² To the first two significant

⁴²Assuming a widely used proprietary risk-estimation package represents the information set of individuals, the authors are able to recover the joint distribution of μ and v for the entire population from the joint

digits, summary statistics in the HHW data are $\mathbb{E}[\mu] = 6.6 \times 10^3$, $\mathbb{E}[v] = 68 \times 10^3$, $\mathbb{V}[\mu] = 50 \times 10^6$, and $\text{Cov}[\mu, v] = 630 \times 10^6$. We do not know $\mathbb{V}[v]$, but the variation in a seems to be quite small and only weakly correlated with that in σ^2 , so we set $\mathbb{V}[v] = \mathbb{E}[a]^2 \mathbb{V}[\sigma^2] = 9.8 \times 10^9$.

We assume (μ, v) has a joint log-normal distribution:

$$\begin{pmatrix} \log(\mu) \\ \log(v) \end{pmatrix} \sim \mathcal{N} \begin{pmatrix} m_\mu & V_\mu & C \\ m_v & C & V_v \end{pmatrix}.$$

Then $\mathbb{E}[\mu] = e^{m_\mu + \frac{1}{2}V_\mu}$, $\mathbb{E}[v] = e^{m_v + \frac{1}{2}V_v}$, $\mathbb{V}[\mu] = \mathbb{E}[\mu]^2 (e^{V_\mu} - 1)$, $\mathbb{V}[v] = \mathbb{E}[v]^2 (e^{V_v} - 1)$, and $\text{Cov}[\mu, v] = \mathbb{E}[v]\mathbb{E}[\mu] (e^C - 1)$. This system of equations of five equations with five unknowns can be solved analytically and uniquely (calculations omitted here) to yield $m_\mu = 8.4$, $m_v = 11$, $V_\mu = .76$, $V_v = 1.2$, and $C = .63$.

D.5 Computing LSE

We use the joint density $f(\mu, v) = f_{\mu|v}(\mu | v)f_v(v)$, where $\log(v) \sim \mathcal{N}(m_v, V_v)$ and $\log(\mu) | v \sim \mathcal{N}(8.4 + 0.53 \times (\log(v) - 11), 0.43)$.

HHW do not consider moral hazard in their model. We adjust the distributions they obtain so that there is no moral hazard at their typical coverage level, $\hat{x} = 0.86$. Without moral hazard cost is $x\mu$, so we shift the mean of μ by α , so that $\hat{\mu} = \alpha\mu$ solves

$$c_i = \hat{x}\hat{\mu}(1 - \hat{x})^{-e} = \mu\hat{x} \Rightarrow \alpha = (1 - \hat{x})^e,$$

where $e = 0.2$ is the moral hazard elasticity. This formula corresponds to adding $\log(\alpha)$ to the mean of $\log(\mu) | v$ described above.

We then use MatLab to compute the required one- and two-dimensional integrals using its built-in numerical integration routines. We provide a few examples below. Because \mathcal{S} has the same composition as \mathcal{B} , we can compute expectations conditional on \mathcal{S} as if they referred to the entire population of buyers (e.g., $\mathbb{E}[u' | \mathcal{S}] = \mathbb{E}[u' | \mathcal{B}]$). We define $\tilde{\mu} = \frac{p - \psi(x)v}{h(x)}$ as the threshold value of μ conditional on v that makes a consumer indifferent about purchasing. Then, regarding the sets \mathcal{B} or \mathcal{S} , we can compute

$$Q = \int_0^\infty \int_0^\infty 1\{\mu \geq \tilde{\mu}\} f_{\mu|v}(\mu | v)f_v(v) d\mu dv$$

$$\mathbb{E}[c(x, \mu) | \mathcal{S}] = \mathbb{E}[c(x, \mu) | \mathcal{B}] = \frac{1}{Q} \int_0^\infty \int_0^\infty 1\{\mu \geq \tilde{\mu}\} c(x, \mu) f_{\mu|v}(\mu | v)f_v(v) d\mu dv.$$

Moreover, note $S = \frac{Q}{2t}$.

We define $v_{\max}(p, x) = \frac{p}{\psi(x)}$ as the point where the function $\tilde{\mu}$ crosses the v -axis (i.e., where $\tilde{\mu} = 0$). Then, regarding the set \mathcal{E} , we can compute

distribution of claims and risk estimates, based on the package and plan choice by individuals. They follow an approach analogous to that of [Cohen and Einav \(2007\)](#), using individuals' choices among available plans coupled with an assumption that a is distributed normally conditional on μ, σ^2 , and other covariates to estimate the joint distribution of μ, σ^2 , and a .

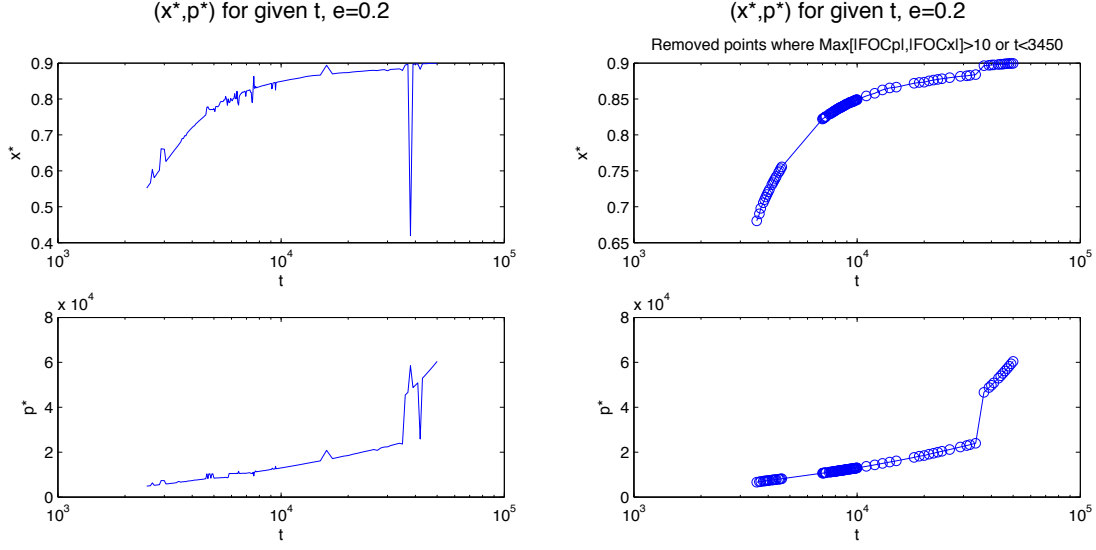


Figure 8: On the left, all computed values of (x^*, p^*) ; on the right is a graph where we removed those values for which $\xi > 10$ and those for which the SOC does not hold ($t < 3450$).

$$E = \frac{1}{2x} \int_0^\infty 1 \{v \leq v_{max}\} f_{\mu|v}(\tilde{\mu} | v) f_v(v) dv$$

$$\mathbb{E}[u'(x, \mu, v) | \mathcal{E}] = \frac{1}{E} \int_0^\infty 1 \{v \leq v_{max}\} u'(x, \tilde{\mu}, v) f_{\mu|v}(\tilde{\mu} | v) f_v(v) dv.$$

We then considered the function $\xi(x, p) = \text{Max}[|FOC_p|, |FOC_x|]$, where FOC_p and FOC_x are the RHS of the FOCs in Propositions 1 and 2. We used MatLab's optimization procedure to minimize $\xi(x, p)$ for different values of t . For $t \in [2500, 10000]$, we considered t at intervals of 50 (because this interval is the range in which the welfare and CS maxima occur). For $t \in [15000, 50000]$, we considered t in intervals of 1,000 to confirm the broad trends in (x^*, q^*) .

The minimization procedure occasionally failed to converge, so we do not consider those values of t for which $\xi > 10$, although most observations have $\xi < 0.01$. This failure of convergence affects only a small minority of the computed LSE. As a robustness check, we also computed the values of (x, p) that minimize ξ from a fine grid of values of (x, p) , for a few values of t . This procedure is computationally much more demanding and delivers results that are reassuringly similar but coarser (these results are not shown but are available upon request). To give a sense of this procedure, Figure 8 shows a graph of all the values of (x, p) computed for each t , next to a graph where we removed those values for which $\xi > 10$ and those for which the SOC does not hold ($t < 3450$).

We use Mathematica to obtain an analytic expression for the SOC, and determine that, in the HHW, the SOC holds for $t \geq 3450$. This expression is cumbersome and therefore not presented here but is available upon request.

We computed relative markups, welfare, and consumer surplus at each LSE, again by numerical integration. Table 1 below presents the characterization of several of the LSE we

t	x^*	p^*	q^*	R	W^*/W^{**}	CS^*/CS^{**}	W_b^*/W_b^{**}
2500*	.55	4800	.95	.85	.82	.71	.96
3500	.67	6900	.90	.94	.91	.78	1
4100	.73	7500	.90	.90	.94	.80	.98
4800	.77	10000	.85	1.29	.95	.78	.94
7000	.82	11000	.85	1.1	.97	.80	.88
10000	.85	13000	.81	1.3	.97	.78	.83
25000	.88	21000	.68	2.12	.94	.70	.74
40000	.90	50000	.32	4.06	.76	.47	.62

Table 1: LSE characteristics for different levels of market power t (used in Figure 7) rounded to two significant digits. Starred value of t are not included in the figure because they do not constitute equilibria given that second-order conditions fail.

computed (the full table is available upon request).

D.6 Non-local deviations

We simulate a market with 10^4 consumers, each with a type (μ, v) following the lognormal distribution described above. We augment types by the location parameter b , using 100 equally spaced points drawn from $\mathcal{U}[0, 1]$, for a total of 10^6 consumers. We then compute profit at non-local deviations from the LSE values in Table 1. We assume firm 0 plays the LSE that maximizes welfare, and then consider profit at a number of deviations by firm 1. We compute profit at 100×100 equally spaced points in the grid $(p, x) \in [0, 30000] \times [0, 1]$. The best deviation affords an increase in profit of only 0.2% relative to profit at the LSE. Moreover, the contour graph of Figure 9 show these deviations occur for values of (p, x) that are extremely close to the computed LSE (p^*, x^*) , so we attribute this small disparity to numerical rounding errors.

References

- Akerlof, George A. 1970. “The Market for “Lemons”: Quality Uncertainty and the Market Mechanism.” *Quarterly Journal of Economics*, 84(3): 488–500.
- Araujo, Aloisio, and Humberto Moreira. 2010. “Adverse Selection Problems without the Spence-Mirrlees Condition.” *Journal of Economic Theory*, 145(3): 1113–1141.
- Azevedo, Eduardo M., and Daniel Gottlieb. 2015. “Perfect Competition in Markets with Adverse Selection.” <http://goo.gl/k5HSSB>.
- Bénabou, Roland, and Jean Tirole. Forthcoming. “Bonus Culture: Competitive Pay, Screening, and Multitasking.” *Journal of Political Economy*.
- Bonatti, Alessandro. 2011. “Brand-Specific Tastes for Quality.” *International Journal of Industrial Organization*, 29(5): 562–575.

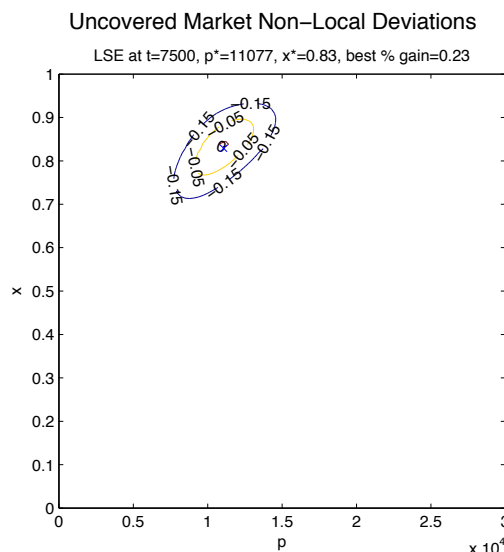


Figure 9: Contour graph of profit at non-local deviations from LSE in the uncovered HWW market, from the LSE that maximizes welfare. Contour lines correspond to points where deviation profit is equal to and 5% and 10% below the LSE profit, and equal to LSE profit (contour lines for deviation profit 5% and 10% above LSE profit are also included but cannot be seen). The “x” identifies the LSE (p^*, x^*) . The subtitle of the figure shows the profit at the best deviation from LSE as a percentage of the LSE profit (0.23%).

Carroll, Gabriel. 2012. “When are local incentive constraints sufficient?” *Econometrica*, 80(2): 661–686.

Cohen, Alma, and Liran Einav. 2007. “Estimating Risk Preferences from Deductible Choice.” *American Economic Review*, 97(3): 745–788.

Dafny, Leemore, Mark Duggan, and Subramaniam Ramanarayanan. 2012. “Paying a Premium on Your Premium? Consolidation in the US Health Insurance Industry.” *American Economic Review*, 102(2): 1161–1185.

de Meza, David, and David C. Webb. 1987. “Too Much Investment: A Problem of Asymmetric Information.” *Quarterly Journal of Economics*, 102(2): 281–292.

Diamond, Peter A. 1973. “Consumption Externalities and Imperfect Corrective Pricing.” *Bell Journal of Economics*, 4(2): 526–538.

Einav, Liran, Amy Finkelstein, and Mark R. Cullen. 2010. “Estimating Welfare in Insurance Markets Using Variation in Prices.” *Quarterly Journal of Economics*, 125(3): 877–921.

Einav, Liran, and Amy Finkelstein. 2011. “Selection in Insurance Markets: Theory and Empirics in Pictures.” *Journal of Economic Perspectives*, 25(1): 115–138.

Einav, Liran, Mark Jenkins, and Jonathan Levin. 2012. “Contract Pricing in Consumer Credit Markets.” *Econometrica*, 80(4): 1387–1432.

- Ellison, Glenn.** 2005. "A Model of Add-On Pricing." *Quarterly Journal of Economics*, 120(2): 585–637.
- Erdil, Aytek, and Paul Klemperer.** 2010. "A New Payment Rule for Core-Selecting Package Auctions." *Journal of the European Economic Association*, 8(2-3): 537–547.
- Flanders, Harley.** 1973. "Differentiation Under the Integral Sign." *American Mathematical Monthly*, 80(6): 615–627.
- Gabaix, Xavier, and David Laibson.** 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, 121(2): 505–540.
- Green, Jerry, and Eytan Sheshinski.** 1976. "Direct versus Indirect Remedies for Externalities." *Journal of Political Economy*, 84(1): 797–808.
- Grubb, Michael.** 2009. "Selling to Overconfident Consumers." *American Economic Review*, 99(5): 1770–1807.
- Handel, Ben, Igal Hendel, and Michael D. Whinston.** 2015. "Equilibria in Health Exchanges: Adverse Selection vs. Re-Classification Risk." *Econometrica*, 83(4): 1261–1313.
- Handel, Benjamin R., and Jonathan T. Kolstad.** 2015. "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare." *American Economic Review*, 105(8): 2449–2500.
- Lester, Benjamin, Ali Shourideh, Venky Venkateswaran, and Ariel Zetlin-Jones.** 2015. "Screening and Adverse Selection in Frictional Markets." https://dl.dropboxusercontent.com/u/8443871/LSVZ_draft2.pdf.
- Mahoney, Neale, and E. Glen Weyl.** 2014. "Imperfect Competition in Selection Markets." http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2372661.
- Mahoney, Neale, André Veiga, and E. Glen Weyl.** 2014. "Competition Policy in Selection Markets." *CPI Antitrust Chronicle*, 10(1): Article 2.
- Marshall, Alfred.** 1890. *Principles of Economics*. Macmillan and Co.
- Mian, Atif, and Amir Sufi.** 2014. *House of Debt*. Chicago: University of Chicago Press.
- Milgrom, Paul, and Chris Shannon.** 1994. "Monotone Comparative Statics." *Econometrica*, 62(1): 157–180.
- Milgrom, Paul R.** 1981. "Good news and bad news: Representation theorems and applications." *The Bell Journal of Economics*, 380–391.
- Milgrom, Paul R., and Robert J. Weber.** 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica*, 50(5): 1089–1122.

- Mirrlees, James A.** 1971. “An Exploration in the Theory of Optimum Income Taxation.” *Review of Economic Studies*, 38(2): 175–208.
- Mussa, Michael, and Sherwin Rosen.** 1978. “Monopoly and Product Quality.” *Journal of Economic Theory*, 18(2): 301–317.
- Myerson, Roger B.** 1981. “Optimal Auction Design.” *Mathematics of Operations Research*, 6(1): 58.
- Pigou, Arthur C.** 1920. *The Economics of Welfare*. . 4 ed., Macmillan and Co., Limited.
- Rees-Jones, Alex, and Dmitry Taubinsky.** 2015. “On Rationality and Heterogeneity of Attention to Non-Salient Taxes: Theory and Experimental Evidence.” <https://goo.gl/ssW26z>.
- Rochet, Jean-Charles, and Jean Tirole.** 2003. “Platform Competition in Two-Sided Markets.” *Journal of the European Economic Association*, 1(4): 990–1029.
- Rochet, Jean-Charles, and Lars A. Stole.** 2002. “Nonlinear Pricing with Random Participation.” *Review of Economic Studies*, 69(1): 277–311.
- Rothschild, Michael, and Joseph E. Stiglitz.** 1976. “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information.” *Quarterly Journal of Economics*, 90(4): 629–649.
- Schmidt, Klaus D.** 2003. “On the Covariance of Monotone Functions of a Random Variable.” <http://www.math.tu-dresden.de/sto/schmidt/dsvm/dsvm2003-4.pdf>.
- Shaked, Avner, and John Sutton.** 1982. “Relaxing Price Competition Through Product Differentiation.” *The Review of Economic Studies*, 49(1): 3–13.
- Shapiro, Carl.** 1996. “Mergers with Differentiated Products.” *Antitrust*, 10: 23–30.
- Sheshinski, Eytan.** 1976. “Price, Quality and Quantity Regulation in Monopoly Situations.” *Economica*, 43(170): 127–137.
- Spence, A. Michael.** 1973. “Job Market Signaling.” *Quarterly Journal of Economics*, 87(3): 355.
- Spence, A. Michael.** 1975. “Monopoly, Quality, and Regulation.” *Bell Journal of Economics*, 6(2): 417–429.
- Spinnewijn, Johannes.** 2015. “Heterogeneity, Demand for Insurance and Adverse Selection.” http://personal.lse.ac.uk/spinnewi/perceptions_welfare.pdf.
- Stango, Victor, and Jonathan Zinman.** Forthcoming. “Limited and Varying Consumer Attention Evidence from Shocks to the Salience of Bank Overdraft Fees.” *Review of Financial Studies*.

- Stiglitz, Joseph E.** 1977. “Monopoly, Non-Linear Pricing and Imperfect Information: The Insurance Market.” *Review of Economic Studies*, 44(3): 407–430.
- Uryasev, Stanislav.** 1995. “Derivatives of probability functions and some applications.” *Annals of Operations Research*, 56(1): 287–311.
- Veiga, André, and E. Glen Weyl.** 2012. “Multidimensional Production Design.” http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1935912.
- Villas-Boas, J Miguel, and Udo Schmidt-Mohr.** 1999. “Oligopoly with Asymmetric Information: Differentiation in Credit Markets.” *The RAND Journal of Economics*, 375–396.
- Weyl, E. Glen.** 2010. “A Price Theory of Multi-Sided Platforms.” *American Economic Review*, 100(4): 1642–1672.
- Weyl, E. Glen, and Jean Tirole.** 2012. “Market Power Screens Willingness-to-Pay.” *Quarterly Journal of Economics*, 127(4): 1971–2003.