The IO of Selection Markets[*]

Liran Einav, Amy Finkelstein, and Neale Mahoney

1. Introduction

The vast majority of theoretical and empirical research in IO focuses on markets where the identity of the buyer does not affect the cost – what we will call "conventional" markets. In conventional markets, sellers care about how consumer demand affects how many units they sell, but they do not care about which consumers buy their products. In other words, consumer demand (based on how much they are willing to pay) and producer cost are independent objects.

This chapter focuses on selection markets. By "selection markets" we refer to markets in which consumers vary not only in how much they are willing to pay for a product but also in how costly they are to the seller. In such settings, sellers not only care about the number of units they sell, but also about *who* the buyers are. That is, they care about selection.

Insurance markets and credit markets are leading examples of selection markets, and the applications we cover in this chapter are primarily drawn from these markets. Insurers sell insurance to consumers who vary in their risk, with higher-risk consumers more expensive to insure. As a result, competition among insurers is focused not only on selling more insurance policies, but also on identifying and attracting buyers who are less costly to cover. Through a variety of methods, insurers try to attract "good" consumers who are associated with lower expected costs ("cherry picking" or "cream skimming") and avoid "bad" ones whose expected costs are high ("lemon dropping"). Similarly, when lenders offer credit to potential borrowers, their goal is not only to maximize the volume of loans they originate, but also to try to make loans to borrowers who are likely to subsequently repay the loans, and to avoid lending to those borrowers who are likely to default.

Insurance and credit markets are ubiquitous, and the IO of these markets is important for many fields in economics. The IO of insurance markets is a central topic in health economics, and also relevant for topics in environmental and development economics. The IO of credit markets is important for questions in macroeconomics, corporate finance, household finance, development economics, as well as the study of entrepreneurship and innovation.

Yet, while insurance and credit markets are the prime examples of selection markets, and perhaps those where selection concerns are most clearly first order, selection concerns are also relevant for many other, more traditional markets. For example, a car dealership may find local consumers more profitable because they may use the dealership (post purchase) to service their cars. Restaurants may prefer certain groups of consumers because they tip more generously or eat faster, lowering the opportunity cost of occupying the dinner table. Labor markets are another important market where

[*] This manuscript is an invited chapter for the forthcoming Volume 4 of The Handbook of Industrial Organization by Elsevier. We thank the editors for including this chapter in the volume. We are grateful to Ignacio Cuesta, Katja Hofmann, and Chuan Yu for very helpful comments on earlier drafts.

selection concerns could be important. When setting benefits, such as family leave policies, firms may consider the type of potential employees who will find such benefits attractive. Education markets are another natural application, since school costs likely depend on which students the school enrolls. Following the focus of most of the recent work on selection markets, we will concentrate our discussion on insurance and credit markets. We will return to these other markets in the conclusion as potentially fruitful areas to apply some of the insights that have been developed.

We begin in Section 2 with a brief description of the history of work on selection markets prior to the last two decades. The early contributions include the seminal theoretical insights of Akerlof (1970) and Rothschild and Stiglitz (1976) that asymmetric information between buyers and sellers may lead to inefficient market outcomes and provide a rationale for welfare-improving government intervention. This work formed the basis of the 2001 Nobel Prize in economics. However, it was only several decades after this pioneering theory that researchers (though not many IO scholars yet) started to provide empirical content to the theoretical predictions. This initial work focused on testing whether and where selection existed in particular markets.

We focus primarily on the second wave of empirical research, after this initial "testing" literature. This is also when IO economists became more heavily involved in the study of selection markets. Over the last two decades, there has been a flurry of empirical work and progress on incorporating the theoretical insights from selection models into more complete equilibrium models of demand and supply in order to produce quantitative results. Our primary goal is to provide a common framework, terminology, and notation that can be used to understand many of these papers, and that hopefully can be usefully applied going forward.

To that end, we start in Section 3 by developing a generic framework for analyzing selection markets. This framework is closely related to the familiar empirical IO framework that has been used broadly, which allows us to illustrate the similarities but also to emphasize the key aspects along which selection markets differ. Section 4 describes two broad classes of empirical demand models that have been used to estimate demand and the links between demand and costs, and Section 5 analyzes equilibrium pricing, again emphasizing the aspects that make selection markets different.

Section 6 turns to questions of welfare, focusing on two topics that have received a reasonable amount of empirical attention. The first is the welfare effects of customized pricing, in which the firm varies the price the consumer faces based on some of their observable attributes. The second involves welfare analysis when consumers are behavioral and demand does not reveal preferences. Finally, Section 7 provides a brief conclusion, and speculates about some potential opportunities for applying the models and approaches covered here in future work.

Before turning to the substance of this chapter, we make a brief case for why selection markets are an excellent area of study for IO economists. First, selection markets are often characterized by high-quality data. For instance, most insurance companies maintain detailed data on consumers' claims and most lenders closely track delinquency and default. Moreover, because these costs are determined by downstream behavior, it is often possible to obtain data that links product choice with downstream utilization or repayment decisions. Second, there is a large body of consumer theory that can guide the analysis of product choice and utilization in selection markets. Insurance markets are perhaps the most natural setting to apply expected utility theory and credit markets are a natural setting for studying intertemporal preferences. Third, as emphasized by the original theoretical work, equilibria in selection markets may be inefficient. These and other concerns have led to considerable regulation and other policy activity in selection markets. For IO economists looking for a setting with rich data, theoretical underpinnings, and policy relevance, selection markets are a good place to work. Moreover, the increased importance of “big data,” artificial intelligence, and concerns about individual privacy are likely to further the need for quantitative modeling frameworks that can allow scholars and policymakers assess the various tradeoffs and challenges that these recent trends present.

2. Some (brief) intellectual history

The study of selection markets has evolved in three distinct waves. The first wave, starting in the 1970s, was primarily theoretical, and contained seminal contributions that highlighted the fact that informational asymmetries may lead to market inefficiencies, justifying potential efficiency-enhancing market interventions. Motivated by these theoretical insights, the second wave, which started several decades later, focused on empirical testing: whether adverse selection actually existed in real-world settings. The third wave has been more quantitative in nature and began to incorporate selection markets into more conventional empirical frameworks.

This chapter is focused on the last wave. In this section, we provide a very brief and informal overview of the two earlier waves. We focus on aspects that are most important in motivating and guiding the “third wave” empirical work and provide references to more comprehensive discussions of these topics for the interested reader.

2.1 Selection market theory

Some of the earliest work on selection markets dates back to Ken Arrow’s foundational paper “Uncertainty and the Welfare Economics of Medical Care” (Arrow 1963). One of the key insights that came out of this early work, and has influenced decades of subsequent work, is the idea that consumers are risk averse, while firms are diversified and therefore can be reasonably approximated as risk-neutral, expected profit maximizers. Under these assumptions, insurance markets can improve welfare by shifting risk from risk averse consumers to risk neutral insurers. This motivates the importance and interest in the efficient operation of these potentially welfare-improving markets.

Important follow-up work by Pauly (1968) made the key observation that full health insurance will not be optimal if demand for health care services rises as health insurance coverage reduces the out-of-pocket price of health care faced by the consumer. (Despite being not ideal or even appropriate, the literature often refers to this price elasticity of demand for health care as “moral hazard.”) That is, when consumers can mitigate their risk by taking non-contractible actions, full insurance provides no incentive for taking such actions, thus making it socially optimal to have consumers exposed to at least some risk. Taken together, efficient insurance contracts trade off risk exposure against moral hazard, and the socially optimal level of insurance coverage becomes an empirical question (Zeckhauser (1970) is an early quantitative application).

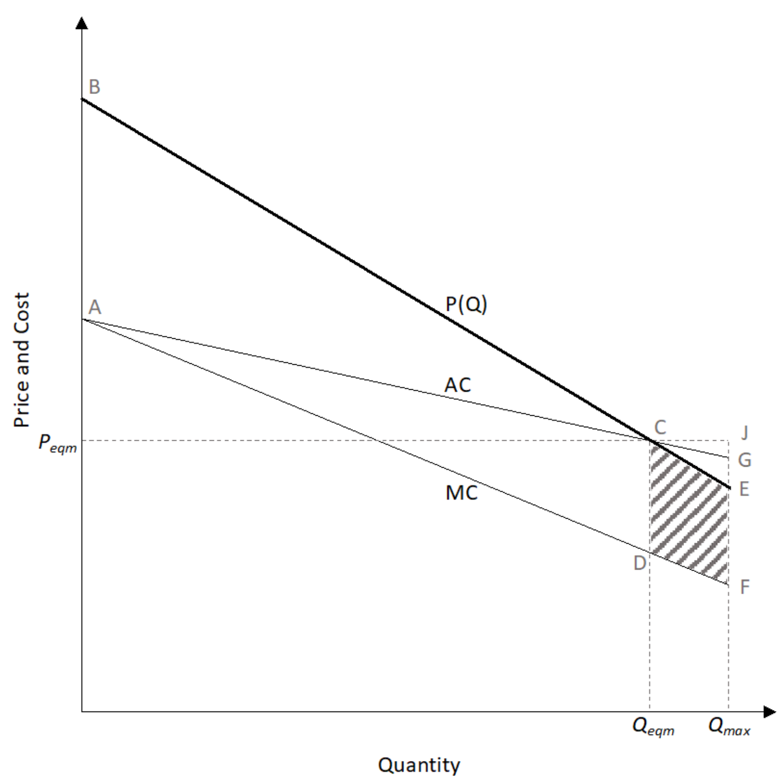
The two seminal contributions that are most important for the rest of the chapter are those of Akerlof (1970) and Rothschild and Stiglitz (1976, henceforth RS). They provide the theoretical frameworks for analyzing equilibrium in selection markets. Both frameworks assume that markets clear via Nash Equilibrium, and in principle can become special cases in a single unifying framework (Hendren 2014). The key distinction between the two frameworks, which we will emphasize later in the discussion of the “third wave” of empirical work, is the assumption which features of the product are endogenized in equilibrium.

There is a parallel here to empirical work in conventional markets, which predominantly relies on static Nash Equilibrium predictions, and makes little use of more subtle theoretical models that dominated the field of Industrial Organization in the 1980s, such as games of incomplete information or dynamic and repeated games. Similarly, in selection markets, the original contributions of Akerlof and RS are almost everything that one needs to know for understanding the subsequent empirical work. The vast follow-up theoretical literature has yet to have much applied impact. (For interested readers, we recommend Dionne et al. (2013) for a review of some of this subsequent theory.)

The key assumption in Akerlof is that products are fixed, and markets clear only through prices. In contrast, RS assume that changes in contract design are part of the equilibrium consideration. In both, the key insight is that firms’ costs from selling a product depend on the set of consumers who select the product. Crucially, consumer type is private information and is unobserved by the firm. This generates a selection market.

In the Akerlof setting, because firms cannot price on consumer type, they are restricted to pricing based on the costs of the average, not the marginal consumer. Figure 1 – which we lift from Einav and Finkelstein (2011) – illustrates the basic intuition of the Akerlof model in a stylized insurance market. (We discuss the analog in credit markets below.) Specifically, we assume that perfectly competitive, risk-neutral firms offer a single insurance contract that covers some probabilistic loss, and that risk-averse individuals differ in their (privately known) expected loss.

Figure 1: Adverse Selection



Consumers make a binary choice of whether or not to purchase this contract (so the market "quantity" of insurance is simply the fraction of insured individuals), and firms compete only over what price to charge for the contract (but not whether to adjust its coverage details). The resulting demand curve in Figure 1 is standard and reflects (one minus) the cumulative distribution function of individuals' willingness to pay for the coverage contract.

The distinguishing feature of selection markets is the cost curve, and, specifically, its link to demand. It is natural to assume that individuals' willingness to pay for insurance is increasing in their (privately known) expected costs. This is illustrated in Figure 1 by plotting the marginal cost (MC) curve as downward sloping, representing the well-known adverse selection property of insurance markets. The cost curve highlights the key distinction of insurance markets (or selection markets more generally) from conventional product markets: marginal cost is driven by demand and thus varies with the selection of consumers. That is, the shape of the cost curve is not driven by the nature of the increasing, decreasing, or constant returns to scale of the production technology. Rather, it is a function of demand-side customer selection. (Of course, selection markets may also exhibit advantageous selection, in which the marginal cost curve is upward sloping. This could be the case, for example, when demand for insurance is primarily driven by heterogeneous risk aversion, and the most risk-averse consumers are associated with the lowest risk; insurance brokers quip that long-term care insurance is purchased by the "healthy, wealthy, and anxious").

The efficient allocation is for everyone whose willingness to pay exceeds their own expected cost to be insured. In Figure 1 this means everyone. However, because the expected cost of each individual is private information, the firms must offer a single price for pools of observationally identical (but in fact heterogeneous) individuals. The competitive equilibrium is therefore given by the intersection of demand and average cost (point C), where the average cost curve is defined as the average expected costs of all those individuals who would buy insurance at a given price.

Figure 1 illustrates the fundamental inefficiency created by adverse selection. It arises because the efficient allocation is determined by the relationship between *marginal* cost and demand, while the equilibrium allocation is determined by the relationship between *average* cost and demand. Under adverse selection marginal costs are always less than average costs; as a result, the equilibrium price is too high and equilibrium quantity is too low, and welfare is reduced by the area CDEF. Indeed, it is possible to construct a cost curve where marginal costs are always (weakly) less than demand, while average costs are always (weakly) above it. In this extreme case, as in Akerlof's famous "lemons" example, the market would fully unravel even though it would be efficient for everyone to obtain insurance coverage. Given the potential for inefficiency, there is a natural rationale for government intervention in selection markets to insure individuals whom it is efficient to insure (willingness to pay is above own expected cost) but who are not insured in equilibrium (willingness to pay is below the average costs of all those who would enter the market at a given price).

The RS adverse selection model generates the same under-insurance result as the Akerlof model. But because they allow firms to also adjust the contract design (rather than only the contract price), RS show that the same adverse selection situation, which leads to equilibrium prices that are too high in the Akerlof framework, may manifest itself in equilibrium with socially suboptimal contract design that is aimed at generating self-selection. This is very much in the spirit of the literature on second-degree price discrimination, with the goal of screening low-risk types from high-risk types. Indeed, in their insurance setting, RS show that in equilibrium low-cost consumers are offered incomplete insurance coverage, which allows them to reveal their type without attracting higher-cost types.

The distinction between Akerlof and RS has strong parallels to empirical work in IO in more conventional markets. Much of the empirical literature in recent decades in empirical IO has assumed that product characteristics are taken as given, and markets clear through Nash Equilibrium in prices, as in Akerlof. But both insurance and credit contracts are financial contracts – with many of their product attributes financial in nature (e.g., deducible, co-insurance rate, the duration of a loan, and so on); in principle, these financial product attributes could be as endogenous as price. Nonetheless, as we will see below, the Akerlof framework remains quite appealing and relevant empirically. Indeed, the vast majority of empirical work on selection markets – just as in conventional markets – is focused on price competition and fixed contracts, thus essentially adopting the Akerlof framework.

One appealing feature of the Akerlof framework is that it is quite simple and robust, and thus can fairly easily accommodate the additional components that come up as we move from the realm of pure theory to the messier reality of data and empirical work. For example, with consumers that are heterogeneous along more than a single dimension, or products that vary horizontally and vertically, an Akerlof framework is still viable, and many of the equilibrium predictions and analyses remain, more or less, intact. In contrast, the RS framework is not as robust, and many of its equilibrium predictions and comparative statics become much more difficult to analyze. Indeed, it may well be that endogenous contracting, which is the focus of RS, rather than mispricing alone (as in Akerlof), is the most important dimension to consider in the context of selection markets, but the empirical intractability of the original RS framework has so far limited its impact on empirical work. Recent work by Azevedo and Gottlieb (2017) applies the Akerlof idea to a setting with a continuum of contracts. By adding a clever (mostly technical) assumption, their framework maintains the RS equilibrium predictions but makes it less fragile. It is possible that this extension will make the RS framework more tractable empirically in the future.

Empirical work in finance has relied primarily on the contributions of Stiglitz and Weiss (1981) and, to a lesser extent, Jaffee and Russell (1976) as theoretical antecedents. These are loosely the credit market analogs of Akerlof and RS, respectively, who focused on insurance markets. Stiglitz and Weiss build a model of lending markets with a single endogenous product characteristic – the interest rate (the analog to the insurance price). What makes their model well suited to finance applications is that adverse selection arises naturally from the borrower's default decision. Because the possibility of default truncates borrower's downside risk, holding upside risk fixed, a higher interest rate will screen for borrowers with riskier projects (i.e. those associated with more dispersed returns). Because lenders' expected profits are decreasing in riskiness, there can arise an equilibrium in which the demand for loans outstrips its supply, resulting in credit rationing. The finance analog to RS is Jaffee and Russell, who allow for endogenous interest rates *and* limits on loan amounts. Although the Jaffee and Russell model is more involved, allowing for both adverse selection and a repayment effect (that is, default risk increases with interest rates in contrast to RS, who abstract from any potential impact of insurance coverage on risk), the key insight is similar: there can be under-provision of lending in equilibrium and issues of equilibrium non-existence arise.

2.2. The empirical testing literature

Following the influential theoretical work of the 1970s, a large empirical testing literature emerged to test whether or not the adverse selection that was modeled in the theory actually existed in practice. Cohen and Siegelman (2010) and Chiappori (2013) provide reviews of this empirical testing literature in insurance markets.

This literature focused on generating and executing empirical tests of whether asymmetric information or adverse selection exists in specific markets. The initial work focused on testing equilibrium predictions of the theoretical model, with the key idea being the so-called "positive correlation test." Some version of this idea has been implemented earlier (e.g., Pueltz and Snow 1994), but the work of Chiappori and Salanie (2000) had the most enduring influence. Their basic idea is to compare the risk profile of consumers who (endogenously) choose more coverage against those who choose less coverage. The central prediction of the selection models – including those we discussed above as well as many others – is that higher-risk consumers will have more coverage in equilibrium.

The development of the positive correlation test represented a key breakthrough in taking the predictions of adverse selection models to the data. It was also relatively easy to implement, requiring "only" that the researcher observed potential customers' equilibrium contracts, their choices, and ex-post proxies for risk (i.e., accident or default rates). As a result, it generated a small cottage industry of papers testing for the presence of adverse selection in particular markets.

However, the positive correlation test faces three key challenges. One is that in many empirical settings prices are unobserved, or the researcher only observes the transacted prices and does not observe prices for products that aren't purchased. In conventional markets this may be sufficient, but in many selection markets prices are individually customized, so that observably higher-risk individuals are required to pay higher prices (for all products). This makes imputing prices challenging. If the researcher has access to the offered (customized) prices, this concern can be addressed by controlling for prices faced, so that the test for adverse selection occurs within sets of individuals who face the same menu of contracts. Absent data on the offered price menu, one can try instead to control flexibly for all the variables that enter the pricing formula. The main concern is that imperfect conditioning on prices may make the researcher conclude that high-risk individuals are no more likely than low-risk individuals to select greater insurance coverage, while in practice this is only because they faced higher relative prices. That is, one needs to control for supply when identifying demand. A familiar point.

A second challenge is to identify adverse selection separately from a direct impact of insurance coverage on risk (often described as "moral hazard"). Both would lead to a positive correlation. Adverse selection implies that higher-risk individuals are more likely to select greater coverage. Moral hazard generates reverse causality: greater coverage causes individuals to take more risky actions, also leading to a positive correlation between the amount of coverage and the ex-post risk occurrence. One way to address this concern is to focus on a setting where moral hazard is less likely (e.g., Finkelstein and Poterba (2002) who study annuities), but in most cases the positive correlation test is a joint test for adverse selection and moral hazard, and some research design or additional assumptions are needed to distinguish between the two.

A third challenge is how to interpret a null result; i.e., the inability to detect a positive correlation and to reject the null of symmetric information. As Finkelstein and McGarry (2006) document, one can find no correlation between coverage and risk occurrence even if there is in fact private information about risk type. This can arise when there are multiple dimensions of heterogeneity, which impact the correlation in different directions and offset each other. In other words, although the original theory assumed individuals differed only in their risk type, in practice, of course, individuals also differ in their preferences. Preference heterogeneity – which is the bread and butter of much of empirical IO – can complicate the predictions of the theory. For example, with heterogeneity in risk aversion, willingness to pay for insurance is not only increasing in risk but also in risk aversion. If risk aversion is negatively correlated with risk, lower-risk individuals can purchase more insurance in equilibrium, even when there is private information about risk type. Fang et al. (2008) provide another early example of this point, and Fang and Wu (2018) provide a careful analysis of the positive correlation test in the presence of multidimensional heterogeneity.

In some settings, one can overcome the first and second issues by using variation in prices that is exogenous to demand. Consider premiums in insurance markets. Under the assumption that premiums affect product choice, but do *not* affect costs conditional on product choice, the researcher can use variation in premiums to trace out the demand curve and the average cost curve in Figure 1. Under the assumption of a monotone marginal cost curve, a downward sloping average cost curve implies downward sloping marginal costs and therefore adverse selection, while an upward sloping average cost curve indicates upward sloping marginal costs and advantageous selection. The inability to reject the null of a flat average cost curve would imply that one cannot reject the null of symmetric information. This so-called "cost curve test" (Einav, Finkelstein, and Cullen 2010) has been applied in many settings, including health insurance (Fischer et al. 2018; Panhans 2019), flood insurance (Wagner 2020), unemployment insurance (Landais et al. 2020), worker's compensation insurance (Cabral et al. 2019), and consumer lending (Liberman et al. 2020).

The assumption that premiums affect contract choice but not costs conditional on product choice is what allows this method to separately identify selection from any direct effect of the coverage contract on costs ("moral hazard"). To see this, recall that the average cost curve is defined as the average costs of all those individuals who buy a specific insurance contract. Since the cost curve is defined over a sample of individuals who all have the same insurance contract, differences in the shape of the cost curve are not directly affected by moral hazard. [1] As we discuss later, this is a reasonable assumption in insurance markets. Premiums are sunk conditional on product choice, so as long as the income effects are ignorable, they should not affect costs conditional on product choice. In other settings, it may be harder to find a "price" that has these properties. For instance, in credit markets, contract terms like interest rates, credit limits, and down payment requirements likely affect both product choice and the downstream probability of default in most models.

Beyond testing. Suppose we are able to reject the null of symmetric information and find evidence of adverse selection in a specific insurance or credit market. What of it? The original theory was motivated by the prospect that adverse selection could impair the efficient operation of markets and

[1] Of course, it is possible that the moral hazard effect of the contract is greater for some individuals than others and that, anticipating this, individuals whose behavior is more responsive to the contract may be more likely to buy insurance. We would still view this as selection, however, in the sense that individuals are selecting the contract based on their anticipated behavioral response to it. We discuss this in more detail in the next section.

open up scope for potentially welfare-improving government intervention. The third wave of the literature – which is our focus in the rest of this chapter – takes up this gauntlet. It investigates the impact of adverse selection (once detected) on equilibrium prices and quantities, and on consumer welfare and social surplus. It also investigates the efficacy of policy instruments in these settings. All of this requires a quantitative economic framework.

3. Theoretical framework

3.1. Setting and notation

Throughout this chapter, we will try to stay close to the familiar static discrete choice setting, commonly used for empirical work in Industrial Organization. We will highlight the aspects that make selection markets different from standard, more traditional product markets.

We denote consumers by $i \in \mathcal{I}$ and products by $j \in \mathcal{J}$, where, as usual, $j = 0$ denotes the outside good. Each product j is defined by a vector of characteristics x_j and a vector of (possibly customized) prices $p_j = \{p_{ij}\}_{i\in\mathcal{I}}$. Unlike conventional markets, selection markets often condition prices on consumer characteristics.

The notation is deliberately abstract – so as to be generally applicable. But to fix ideas we offer two specific examples. In the insurance context, consider consumers choosing across a set of health insurance contracts *j*, with a vector of characteristics x_j that may include cost-sharing provisions (such as deductibles, co-pays, and out-of-pocket maximums), as well as non-financial features (such as the set of health care providers that are in-network, the contract length, and requirements for referrals or prior authorization before certain tests will be covered). Prices for these different insurance contracts may vary based on consumer characteristics, such as age, sex, income, or pre-existing health conditions. In the credit space, consider a choice among different types of mortgages, which could vary in duration, down-payment requirement, or the ability to refinance. Prices (or interest rates in this case) would typically vary based on the value of the house or the credit score of the potential borrower.

As in the standard setting, consumer i's willingness to pay for (or value from) product j is given by their indirect utility $v_{ij} = v(x_j, \zeta_i)$, which is a function of product *j*'s (observed or unobserved) characteristics x_j, consumer i's characteristics ζ_i, and any potential interaction between the two. In writing indirect utility in this form, we implicitly assume that it is quasi-linear; as a result, there are no income effects and the Marshallian and Hicksian demand curves coincide. While restrictive, this assumption is almost always made in the empirical IO analysis of both conventional and selection markets.[2] Unless otherwise noted, we impose it throughout this chapter.

In selection markets, the motivation for this assumption is both conceptual and practical. Conceptually, most of the empirical research on selection markets is focused on products that compose at most a modest share of the consumer's budget; in such settings, the income effects from counterfactual changes in prices can be plausibly ignored. From a practical perspective, the focus in the literature has

[2] It is equivalent to the standard specification of indirect utility in empirical IO, in which the indirect utility is additive and separable in price, and price enters linearly. To see this, consider a standard specification of indirect utility $u_{ij}(p_{ij}, x_j, \zeta_i) = f(x_j, \zeta_i) - \alpha_i p_{ij}$. This corresponds to a willingness-to-pay function of $v_{ij} = v(x_j, \{\zeta_i, \alpha_i\}) = f(x_j, \zeta_i)/\alpha_i$.

been to acquire data that have rich information on demand and cost from firms or administrative sources, and it is rare to have high-quality income data (and variation) to estimate income effects.

We define the surplus that consumer i receives from product j as $v_{ij} - p_{ij}$; this is their willingness to pay minus the (potentially customized) price. As in the standard context, we normalize $v_{ij} = 0$ and $p_{ij} = 0$ for $j = 0$ (for all i) and assume a discrete choice, in which consumer i chooses the product that maximizes her surplus. That is, consumer i chooses product j if and only if $j = \underset{k \in \mathcal{J}}{\operatorname{argmax}}\{v_{ik} - p_{ik}\}$. Integrating over consumers, demand for product j is then given by

$$\mathrm{D_j}\left(\{\mathrm{p_j}, \mathrm{x_j}\}_{\mathrm{j} \in \mathcal{J}}\right) = \int \mathrm{I}\left\{\mathrm{j} = \underset{\mathrm{k} \in \mathcal{J}}{\operatorname{argmax}}(v(x_k, \zeta_i) - p_{ik})\right\} \mathrm{d}\zeta_\mathrm{i}. \tag{1}$$

Similarly, product j's revenues are given by

$$\mathrm{R_j}\left(\{\mathrm{p_j}, \mathrm{x_j}\}_{\mathrm{j} \in \mathcal{J}}\right) = \int \mathrm{p_{ij}} \mathrm{I}\left\{\mathrm{j} = \underset{\mathrm{k} \in \mathcal{J}}{\operatorname{argmax}}(v(x_k, \zeta_i) - p_{ik})\right\} \mathrm{d}\zeta_\mathrm{i}. \tag{2}$$

In conventional markets, the quantity of demand is the only element from the demand side that enters the firm's costs. In contrast, the key distinguishing aspect of selection markets is that consumers are associated with subsequent events that feed back directly into firms' objective function. To reflect this aspect of selection markets, we denote by $c_{ij} = c(p_j, x_j, \zeta_i)$ the (expected) cost incurred by the seller from selling product j to consumer i. In conventional markets, these costs are simply the marginal cost of production, and do not depend on the identity of the consumer who purchases the product (although they may of course depend on the aggregate quantity purchased). Selection markets are defined as such precisely because costs also depend on consumer characteristics, thus making the seller care about customer selection; all else equal, the seller prefers selling her product to consumers who are associated with lower costs.

In order to focus on the aspects of firm costs that are the essence of selection markets, we assume (unless explicitly noted otherwise) that there are no economies or dis-economies of scale. Under this assumption, cost is simply the expected amount of claims in the case of insurance and the expected loss (relative to full repayment) associated with defaults in credit markets.

In the general case, product j's costs are given by

$$C_j\left(\{p_j, x_j\}_{j \in \mathcal{J}}\right) = \int c_{ij} I\left\{j = \underset{k \in \mathcal{J}}{argmax}(v(x_k, \zeta_i) - p_{ik})\right\} d\zeta_i. \tag{3}$$

The variation of cost $c_{ij} = c(p_j, x_j, \zeta_i)$ across individuals *i* who buy the same contract *j* is what distinguishes selection markets from conventional markets.

Within the set of individual characteristics that make up ζ_i, it can be sometimes useful to distinguish between the baseline risk components that characterize costs independently of contract characteristics

x_j, and components that characterize the "behavioral response" through which characteristics x_j affect costs.

In many markets, the baseline risk components of ζ_i are the primary source of cost variation. For example, in the context of an auto insurance market, where a product insures against traffic accidents and the cost captures the expected amount of claims the insurance company would pay out, consumers may vary in how long their commute is, how safe the roads in which they typically drive are, or their innate driving ability. In the context of consumer loans, individual default rates may vary across individuals due to the nature of their employment, how stable their life conditions are, or their ability to rely on their relatives when they are faced with financial distress.

Behavioral characteristics include those that are often referred to as "moral hazard." For instance, in the context of health insurance, these could be characteristics that affect consumers' responsiveness to a change in out-of-pocket costs due to a higher deductible such as their marginal rate of substitution between health and consumption. However, some behavioral characteristics may reflect constraints rather than preferences. For example, in response to a higher deductible, an individual may cut back on care because they are liquidity constrained. In most empirical settings, this distinction is difficult to define, and even more difficult to econometrically identify. This is presumably why the term "moral hazard" gets used rather loosely across contexts.

Whether it is important to separate out the components of ζ_i that reflect "moral hazard" depends on the application. When counterfactual changes in contract characteristics are of interest, isolating the moral hazard component can be important. However, for applications where the product characteristics are held fixed, recovering $c_{ij} = c(p_j, x_j, \zeta_i)$ may be sufficient, although understanding the potential channels through which ζ_i can affect costs can provide guidance on how to specify $c_{ij} = c(p_j, x_j, \zeta_i)$.

A fairly common restriction on the cost function is to assume that the price variable p_{ij} does not enter so that we can write $c_{ij} = c(x_j, \zeta_i)$. With this restriction, prices only affect profits through their impact on demand. This is a natural assumption for modeling premiums in an insurance market. While premiums affect product choice, it is reasonable to assume that they don't have a direct impact on costs conditional on plan choice.

In principle, this restriction can also be satisfied when applying the framework to credit markets if we define terms appropriately. Consider for example a firm that chooses the size of the loan L_{ij}, holding fixed the repayment schedule x_j. The size of the loan can be thought of as a negative price that affects product choice and generates upfront negative revenue but does not affect repayment conditional on the repayment schedule. Conditional on product choice, the firm receives negative costs from the realized stream of repayments $c(x_j, \zeta_i)$, which depends on the repayment schedule x_j and consumer characteristics ζ_i but does not depend on loan size L_{ij} directly. Thus, in principle, the lending problem can be written using the framework above with the sign on the revenue and cost terms reversed.

However, it is much more common in credit markets to view loan size as part of the contract terms x_j and view the interest rate as the "price variable" p_{ij} which clears the markets. Doing so implies that the above restriction no longer holds, as now the price (that is, the interest rate) does affect the repayment schedule, and subsequently the (negative) costs associated with a product purchase.

3.2. Definitions and terminology

To simplify the notation, let q_{ij} be an indicator that individual i chooses plan j:

$$q_{ij} = 1 \iff j = \underset{k \in \mathcal{J}}{\operatorname{argmax}}(v(x_k, \zeta_i) - p_{ik}) \, .$$

Formally, we define *selection markets* as markets in which there exists at least one product j for which c_{ij} is not constant across i:

Definition 1: A market is a selection market if and only if there exists a product j such that $Var(c_{ij}|j) > 0$.

We make three observations about this definition. First, it makes use of our assumption of no scale economies. If we allowed for economies or dis-economics of scale, then we would need to modify the definition to also hold fixed the quantity supplied, in order to eliminate variation in costs due to scale economies.

Second, the definition takes as given the set of products offered in the market. A change in the type of products in a given market could in principle make a conventional market become a selection market or vice versa. To see this, consider a market for home security. Decades ago, serving a household who lived in a remote area was much more costly as it required frequent patrolling around the (remote) house. With the advent of remote security cameras, the location of the house no longer materially affects costs, turning the market into a more conventional market. Similarly, a market with Internet Service Providers with slack capacities would be a conventional market, but as internet use increases and capacity at peak hours is binding, customers with high internet use become more costly, potentially turning it into a selection market.

Third, this definition – while theoretically clear – does not have much bite from an applied perspective. Taken literally, almost every product market is a selection market. A retail store may have lower costs from serving a customer who never returns any item relative to serving a customer who makes generous use of the return policy, and a souvenir seller in a tourist market may prefer to sell to a tourist who bargains faster down to the final (identical) transaction price. Yet, we probably wouldn't model these as selection markets because, while c_{ij} varies across consumers, it presumably doesn't vary enough to make the selection aspect of the market essential for applied analysis.

Whether the market is a conventional market or a selection market (as defined above) is a property of the market primitives. If a market is a selection market, the nature of selection becomes an object of interest, with familiar concepts to describe selection, such as adverse selection or advantageous selection. These concepts need to be applied carefully; in most cases they describe the market's equilibrium outcomes rather than market primitives. We define product j as being adversely selected in equilibrium by examining whether product j draws a riskier pool of customers relative to a randomly drawn pool from the population of customers in the market. That is,

Definition 2: Product j is adversely selected if and only if

$$E_{\zeta_i}(c_{ij}|j, q_{ij} = 1) > E_{\zeta_i}(c_{ij}|j).$$

The product would be advantageously selected if the inequality is reversed. As can be seen from Definition 2, whether product j is adversely selected is a statement about the market equilibrium, as it depends on the entire set of products and prices available in the market (which all enter the contract choice, as reflected by the conditioning on $q_{ij} = 1$).

To empirically test whether product *j* is adversely selected (based on the above definition) would require not only observing consumer costs (or proxies for them) but also observing costs of all *potential* consumers (including those who did not purchase any other product). To circumvent this challenge, the positive correlation test for adverse selection that we described in Section 2.2 often relies on comparing the risk profile – a component of ζ_i (or a proxy for it) – of consumers who select one contract to the risk profile of consumers who select a different contract, which provides less coverage. Under the null of no adverse selection, both contracts would be associated with the same risk profile of consumers, but under adverse (advantageous) selection, the higher-coverage contract would be selected by consumers associated with higher (lower) risk. We previously discussed some of the empirical challenges with implementing and interpreting this test.

Sometimes, the concept of adverse selection is invoked and used in more nuanced ways. For example, scholars may describe a specific product j as more adversely selected than a different product h. In the context of the general framework described above, this statement would require researchers to pick a specific cost structure as the basis for comparison. (At the extreme, j and h could be completely different projects with very different cost structures, rendering a comparison of insurer costs c_{ij} across products meaningless.) However, the choice of the costs structure could be consequential. In fact, one could find that both products in a two-product market are adversely selected if, for example, $E_{\zeta_i}(c_{ij} | j, q_{ij} = 1) > E_{\zeta_i}(c_{ij} | j, q_{ih} = 1)$ and $E_{\zeta_i}(c_{ih} | h, q_{ih} = 1) > E_{\zeta_i}(c_{ih} | h, q_{ij} = 1)$, which is a plausible situation when the products are horizontally differentiated.

However, in many applications consumers' characteristics can be projected on a single, vertical dimension (e.g., a risk score), so that this comparison is more meaningful. In other words, if ζ_i can be summarized by a scalar risk θ_i and $c_{ij} = c(x_j, \theta_i)$ is monotone in θ_i for *every* j, then one can in principle compare $E_{\theta_i}(\theta_i | q_{ij} = 1)$ to $E_{\theta_i}(\theta_i | q_{ih} = 1)$.[3] For example, in government-run health insurance marketplaces, such as Medicare Advantage or the health insurance exchanges implemented by the Affordable Care Act, consumers are characterized by a risk score that is designed to predict health care utilization. To the extent that these risk scores are predictive, we can use the average risk score of consumers across plans to describe the selection in the market.

Another common use of the concepts of adverse (or advantageous) selection is with respect to a contract characteristic. Here, scholars describe the nature of selection by asking how $E_{\zeta_i}(c_{ij} | q_{ij} = 1)$ varies with small changes in p_{ij} or in x_j. Refer back, for example, to Figure 1 where we graphed the market equilibrium in which consumers face the binary choice whether to buy or not a particular insurance product. That insurance product was described as adversely selected because the expected costs of those who purchased the insurance at a given price were increasing as the price increased. In other words, the cost curve was downward sloping. Naturally, as Geruso et al. (2021) highlight, selection can be adverse along one dimension but advantageous on another, so any description of selection needs to carefully define the relevant benchmark. For instance, mortgages might be advantageously selected

[3] Yet, the interpretation of such a statement may still be tricky because such comparison may be sensitive to non-linear transformations of θ_i.

with respect to higher down-payment requirements but adversely selected with respect to longer loan durations.

3.3 Road map

In the remaining sections, we will use this framework to try to synthesize a number of applications from the quantitative empirical literature on selection markets. Unlike traditional markets, in which the demand model provides a way to obtain choice probabilities and consumer welfare, in selection markets the demand model is tightly connected to firms' costs. As a result, in addition to generating choice probabilities and consumer welfare, the demand model also generates estimates of firms' expected costs from individual buyers. Estimation of demand therefore essentially requires a joint estimation of the demand structure $v_{ij} = v(x_j, \zeta_i)$ and the cost structure $c_{ij} = c(p_j, x_j, \zeta_i)$. This is precisely what the tests of adverse selection described in Section 2.2 were designed to capture. The models of demand (and cost) that we discuss next impose additional structure on this relationship, which allows the researcher to make quantitative, rather than just qualitative, statements about adverse selection.

We will not discuss identification of the models that we describe below. The specific identification assumptions naturally vary on a case-by-case basis. Broadly speaking, with data on costs, identification of demand models in the context of selection markets is conceptually similar to identification of demand models in more conventional markets, which are covered in much more detail in a separate chapter in this Handbook (Berry and Haile, 2021).

4. Empirical models of demand in selection markets

We follow the organization of Einav, Finkelstein, and Levin (2010) and classify demand models into two broad categories. One follows the more traditional empirical IO demand modeling approach, and directly models willingness to pay $v_{ij} = v(x_j, \zeta_i)$ (or equivalently, indirect utility), without engaging in the "deeper" utility primitives from which willingness to pay is derived. A second category of empirical demand models derives the indirect utility from deeper primitives.

For most IO-related applications, the main object that the demand model needs to "deliver" is the joint distribution of willingness to pay and cost, $F(v_{ij}, c_{ij})$. Thus, there is little a-priori reason to prefer a model of willingness to pay or a model of deeper primitives, and the choice should primarily depend on the nature of the products and the richness and granularity of the post-purchase data. Models of deeper primitives are clearly essential when the object of interest is those deeper primitives themselves (e.g., estimates of risk aversion). They may also be useful when economic theory can guide the choice of functional form of the indirect utility function.

Models of willingness to pay. As in modeling demand for cars or breakfast cereals, the researcher can model demand in a selection market by characterizing each product j by a set of product characteristics x_j, making assumptions about the scope and nature of consumer heterogeneity ζ_i, and specifying a convenient functional form through which willingness to pay varies with product and consumer attributes.

The simplest, and perhaps most trivial, such model is offered by Einav, Finkelstein, and Cullen (2010). In their setting there is only a single product (a health insurance contract) in addition to the outside option. Therefore, both individual willingness-to-pay v_i and individual cost c_i are scalars. Given the simplicity of their setting, their demand system boils down to the two-dimensional joint distribution of v_i and c_i, and they estimate the conditional distribution $F_c(c_{ij}|v_{ij})$, which is sufficient for their purposes.

Some recent applications of this approach in health insurance include Fischer et al. (2018), who conduct a randomized experiment in which they randomly vary the price of health insurance offered to different people in rural Pakistan; Panhans (2019), who uses discontinuities in premiums for subsidized health insurance for low-income individuals at regulatory borders in Colorado; and Finkelstein et al. (2019), who use discontinuities in premiums at specific income levels in Massachusetts. In the finance literature, Liberman et al. (2020) use a modified Einav, Finkelstein, and Cullen (2010) model to study the effects of a policy in Chile that deleted default information from consumer credit reports.

A similar approach to Einav, Finkelstein, and Cullen (2010) is taken by Bundorf, Levin, and Mahoney (2012) in a much richer setting, but also in the context of health insurance. In their setting, consumers face four possible choices of health insurance that are both vertically and horizontally differentiated. A completely flexible demand system implies that each consumer is characterized by an eight-dimensional type (a willingness to pay for each of the products, and a cost associated with each product). Adapting their notation to fit the notation we use in this chapter, the willingness to pay of individual i for enrolling in plan j is given by

$$v_{ij} = v(x_j, \zeta_i) = x_j\alpha_x + z_i\alpha_{z,j} + f(r_i + \varepsilon_i; \alpha_{r,j}) + \omega_{ij},$$

and costs for individual i at plan j are given

$$c_{ij} = c(x_j, \zeta_i) = \beta_0 + \beta_j(r_i + \varepsilon_i) + \epsilon_{ij},$$

where the α's and β's are parameters to be estimated. In this specification, x_j are plan characteristics (e.g., coinsurance, deductible, brand), z_i are observed consumer characteristics (e.g., age, sex), r_i is an observed risk score of each consumer (a risk score is an actuarial prediction of health care spending given demographics and prior medical diagnoses), ε_i is an unobserved risk shifter of each consumer, and ω_{ij} and ϵ_{ij} are iid consumer-specific terms that shift, respectively, their valuation and cost associated with each product j. That is, their model of individual characteristics $\zeta_i = \{z_i, r_i, \varepsilon_i, \{\omega_{ij}\}_{j\in\mathcal{J}}, \{\epsilon_{ij}\}_{j\in\mathcal{J}}\}$ is quite rich.

A key assumption in this parameterization is that the common term $r_i + \varepsilon_i$ is the only component that generates correlation between costs and valuations. Another observation about this specification is that it takes a somewhat mixed approach to demand, combining characteristic-space and product-space approaches, with product characteristics entering the willingness to pay function both through their characteristics x_j and via the product-specific coefficients on consumer characteristics (i.e., they allow $\alpha_{z,j}$ and $\alpha_{r,j}$ to depend on j). In the finance literature, Crawford et al. (2018) use a Bundorf et al. type framework to study the small business lending market in Italy.

Models of deeper primitives. An alternative to specifying the willingness to pay function directly is to derive it from deeper primitives based on product utilization. In many conventional product markets this approach is not tractable for two related reasons. First, precisely because post-purchase consumer behavior does not directly affect firms' profits, it is rare to have data on post-purchase behavior (e.g., whether milk expires before it is consumed). Moreover, even if such additional information were available, economic theory does not provide much guidance for why consumers may like certain products over others. For example, consumers may prefer to buy electric cars because they value saving on gas, but also because they may enjoy the less noisy ride or may derive intrinsic value from conserving energy.

In contrast, in many selection markets – such as insurance and credit markets – core economic theory motivates why consumers would demand insurance or credit. Willingness to pay can be clearly linked to deeper primitives of the utility function, such as risk aversion, discount rate, or other aspects of the environment (e.g., the probability of an adverse event). An advantage of building the demand model from these deeper primitives, which subsequently add up to a willingness to pay function, is that counterfactual analysis of demand for out-of-sample products is guided by consumer theory (e.g., assumptions about the functional form of the utility function, such as CRRA or CARA) rather than statistical assumptions as in the willingness to pay approach. (By the same token, to the extent the researcher does not have confidence in the assumptions about the utility function, this may also be viewed as a downside of the approach.)

As mentioned earlier, an additional motivation for this approach is that researchers may also be intrinsically interested in the primitives of the model. One such example is Cohen and Einav (2007). They attempt to estimate the distribution of risk aversion using coverage choices made by customers of an Israeli automobile insurance company. Each individual i chooses between a high-deductible contract with a price and (per claim) deductible of $p_{i,HD}$ and $x_{i,HD}$, respectively, and a low deductible contract, $\{p_{i,LD}, x_{i,LD}\}$. Cohen and Einav assume that claims arrive according to a Poisson process that is not affected by the choice of deductible (i.e., there is no effect on driving behavior or claims behavior from a change in coverage). Combining this with an assumption of CARA utility over wealth,[4] they write the expected utility from a contract $\{p_{ij}, x_{ij}\}$ as:

$$Eu_{ij} = (1 - \varepsilon_i)u_i(a_i - p_{ij}) + \varepsilon_i u_i(a_i - p_{ij} - x_{ij}),$$

where ε_i is the individual's Poisson risk rate, a_i is their wealth, and $u_i(w) = -\exp(-\psi_i w)$, with ψ_i denoting the coefficient of absolute risk aversion. With CARA preferences, the consumer's wealth does not affect their insurance choices, so the relevant consumer characteristics are given by $\zeta_i = \{\psi_i, \varepsilon_i\}$.

The main object of empirical interest in Cohen and Einav (2007) is the joint distribution of risk aversion and risk rate $G(\psi, \varepsilon)$. However, with a little algebra, this set up allows them to write the willingness to pay of consumer i for upgrading their coverage contract from a high deductible $x_{i,HD}$ to a lower deductible $x_{i,LD} < x_{i,HD}$ as:

$$v_{ij} = {\psi_i}^{-1} ln\left[(1 - \varepsilon_i) + \varepsilon_i \exp(\psi_i x_{i,HD})\right] - {\psi_i}^{-1} ln\left[(1 - \varepsilon_i) + \varepsilon_i \exp(\psi_i x_{i,LD})\right]$$

and the incremental cost to the insurance company from such an upgrade as

$$c_{ij} = \varepsilon_i(x_{i,HD} - x_{i,LD}).$$

Thus, this model of deeper primitives simply translates to a particular functional form that links consumer types, defined by $\zeta_i = \{\psi_i, \varepsilon_i\}$, to willingness to pay and cost. A very similar model is applied by Barseghyan et al. (2011) to study deductible choice for auto and home insurance contracts in the US.

In lending markets, it is natural to build up demand from a consumer's intertemporal optimization problem. One example of this is Bachas' (2019) model of demand to refinance student loans. Bachas writes down an intertemporal model in which, in each period, borrowers receive income and have to

[4] We note that the baseline model of Cohen and Einav (2007) is of quadratic utility, which carries certain computational advantages. But in order to be consistent with the price separability we use throughout this chapter, we illustrate the same ideas in the context of CARA utility.

make a repayment decision on their student loan. They have CRRA utility over period consumption, parameterized by intertemporal elasticity of substitution (IES) γ, and future periods are discounted by ρ. The lender offers borrowers a schedule of risk-based loan maturities and interest rates. The borrower then chooses the loan that maximizes their expected discounted sum of period utility, given their preference parameters. Bachas parameterizes the IES as a function of consumer observables and estimates it using variation over time in the contract schedule, calibrating the discount factor and income process. Two other examples are Einav, Finkelstein, and Schrimpf (2010) and Illanes and Padi (2021), who use a similar approach to model demand for annuities.

Two-period models. In addition to the two main classes of demand models, there is a commonly used set of demand models in selection markets that is based on a simplified two-period model and that can be placed somewhere between the willingness-to-pay models and the "deeper primitives" models described above. In these models, individuals make two sequential decisions. In the first period they face uncertainty about future realizations and choose a contract. In the second period, uncertainty is realized, and they then choose their optimal behavior, taking as given the contract they chose in the first period. This follows the spirit of the "deeper primitives" approach in that there is a model of behavior under each possible contract, which is then mapped to the willingness to pay for each contract, which drives demand. However, instead of fully describing the behavior under each contract, these models summarize period-two behavior under each contract with a simple model (e.g., with health care utilization depending on the amount of cost sharing). These types of two-period models in selection markets share many common features with similar models of demand for durable goods in more traditional markets, where product purchase and product utilization are modeled jointly (Dubin and McFadden 1984; Davis 2008).

This two-period modeling approach is particularly useful when the contract choice in period 1 may affect behavior in period 2 (which, as mentioned earlier, is often referred to as "moral hazard"). As a result, it is commonly used in the context of demand for health insurance. Cardon and Hendel (2001) were the first to introduce such a model. Einav et al. (2013) and Handel (2013) use the Cardon and Hendel two-period modeling approach – albeit with richer data and a richer model of heterogeneity – to study, respectively, the heterogeneity in moral hazard across individuals and how inertia in plan choice interacts with adverse selection. This two-period modeling approach is also used in models of credit markets, such as Einav, Jenkins, and Levin (2012) and Cuesta and Sepulveda (2020).

5. Pricing and equilibrium with selection

We now discuss models of supply in selection markets. We focus first on the Akerlof-style setting in which products are fixed. Pricing is therefore the only supply-side decision. As we will see, the impact of selection interacts quite closely with the nature of competition. We start by analyzing a case of perfect competition. Here, in the absence of selection, the market allocation is efficient. This makes it a natural benchmark from which to analyze the impact of introducing selection. Perhaps for this reason, some version of perfect competition (or non-strategic pricing) was the primary focus of most of the initial analyses of equilibrium in selection markets.

We then consider three separate extensions to this baseline framework. First, we introduce market power. The assumption of perfect competition – while appealing for the reasons just mentioned – contrasts with many (most?) empirical IO applications. It may also not be a good approximation of many real-world selection markets. Second, we depart from the Akerlof setting and allow for non-price contract elements (as well as the price) to respond endogenously on the supply-side. Finally, we briefly explore models in which prices are customized to consumers. While such third-degree price discrimination is also studied in conventional IO markets, it is particularly central (and somewhat distinct) in selection markets where customers vary not only in their willingness to pay, but also in their costs.

5.1 Perfect competition

We start with a single, homogenous product and multiple identical firms, $J \geq 2$, competing in prices. As in the above willingness to pay framework, with a single product, consumers are defined by two values: their willingness to pay for the product v_i and the cost they would impose on the firm from buying the product c_i. We make the natural assumption that (with identical firms), if consumers purchase the product, they do so from the lowest priced firm, and if there is more than one firm offering this lowest price, customers are split equally and randomly across them.

This leads to a standard homogenous-product Bertrand-Nash Equilibrium, in which profits are zero and prices are set equal to average costs, $p = E(c_i | v_i > p)$.[5] Figure 1 above illustrated this situation. Because average cost declines with price (the high-cost buyers are those who are willing to pay the most for the product), marginal buyers are more attractive to sell to relative to infra-marginal buyers. This leads to an equilibrium price that is too high and quantity that is too small, in the sense that under full information (and thus no selection) there would be additional surplus-producing trades.

While we have focused thus far on a single contract (and a binary choice by the consumer of whether or not to buy the contract at a given price), in many contexts the "outside option" is not "no contract" but an alternative contract (typically one that is free, or highly subsidized). For example, in many applications to employer-provided health insurance, employees are choosing between different contracts (Einav, Finkelstein, and Cullen 2010; Bundorf et al. 2012; Handel 2013); of course, choosing no contract is also an option, but given the large tax subsidy for employer-provided health insurance, the vast majority of employees purchase some coverage.

In this situation, the above framework remains the same, except that some renormalizations are needed. To see this, consider a case in which the offered product is high quality and consumers could obtain a lower-quality product at a lower price. For ease of exposition (but not for necessity), assume that the two products are vertically (but not horizontally) differentiated, denoted by H and L. Consumers are now defined by four objects: $\zeta_i = \{v_{iH}, v_{iL}, c_{iH}, c_{iL}\}$.

[5] Two "technical" notes may be in order. First, we assume throughout that both the demand relationship and the average cost relationship are continuous, which guarantees the existence of an equilibrium. We note that there could be multiple prices at which prices could equal average cost, but only the lowest of these prices constitute a (unique) Nash Equilibrium. Second, we assume (for now) a single price (for a given product in a given market); the analysis remains the same for customized prices as long as (and this is critical) all firms have the same information set on consumers and partition the market in the same way (or alternatively face the same regulatory constraints on the pricing structure). That is, for example, if consumer gender is observed and all sellers offer gender-specific prices, one can still analyze the market as if there is a single price, except that markets will be defined and will equilibrate separately for males and females. In Section 5.4 we consider the firm's decision to customize prices.

Einav, Finkelstein, and Cullen (2010) (EFC) illustrate a special case of this multi-contract framework. They assume that product L is the default coverage and that firms offering product H are only exposed to the *incremental* cost associated with each buyer. We can define incremental price ($\Delta p = p_H - p_L$), incremental willingness to pay ($\Delta v_i = v_{iH} - v_{iL}$), and incremental cost ($\Delta c_i = c_{iH} - c_{iL}$) and then proceed to analyze the market using these incremental objects in the single contract framework described above. One natural application would be for elderly individuals in the United States who are given public health insurance by default (Traditional Medicare), but have the option to buy additional private coverage (supplementary Medigap policies).

A more common situation, however, may be when the two vertical products, H and L, are provided by separate firms (single-product firms each offering one of the products). In this situation, a competitive equilibrium implies that each seller needs to break even, and selection feeds into both the price of product *L* and the price difference Δp (rather than only the latter as in EFC). This is the setting considered in Handel, Hendel, and Whinston (2015) (HHW). An equilibrium now requires two zero-profit conditions: $p_H = E(c_{iH}|v_{iH} > p_H, \Delta v_i > \Delta p)$ and $p_L = E(c_{iL}|v_{iL} > p_L, \Delta v_i < \Delta p)$.[6]

As described in Weyl and Veiga (2017), these two different market clearing assumptions can lead to very different results. The EFC framework considers the market for incremental coverage, which clears at a Δp which satisfies $\Delta p = E(c_{iH} - c_{iL}|\Delta v_i > \Delta p) = E(c_{iH}|\Delta v_i > \Delta p) - E(c_{iL}|\Delta v_i > \Delta p)$. In contrast, in the HHW framework, the market clears by the conditions above, which imply (assuming that nobody selects the outside option, so $v_{iH} > p_H \ and \ v_{iL} > p_L$ are trivially satisfied) that $\Delta p = p_H - p_L = E(c_{iH}|\Delta v_i > \Delta p) - E(c_{iL}|\Delta v_i < \Delta p)$.

Comparing these two equilibrium conditions, one can see that an adverse selection effect would be much greater in the HHW framework in terms of the equilibrium price difference Δp between the *H* and *L* contracts. By adverse selection we mean (as before) a positive correlation between willingness to pay and cost, which naturally implies that $E(c_{iL}|\Delta v_i > \Delta p) > E(c_{iL}|\Delta v_i < \Delta p)$; in other words, the expected costs under the *L* contract are higher for those who purchase the *H* contract than for those who purchase the *L* contract. By “greater” adverse selection under the HHW framework than the EFC framework, we mean that the difference in expected costs, $E(c_{iH}|\Delta v_i > \Delta p) - E(c_{iL}|\Delta v_i < \Delta p)$, is larger.

Why is adverse selection larger (in this sense) under the HHW framework than the EFC framework? Consider the thought experiment in which the quality difference between the two products shrinks to zero. In the market for incremental coverage imagined by the EFC framework, the incremental price will approach zero as the quality difference between the two products shrinks to zero. However, in the HHW framework in which the market clears separately for each of the two contracts, as the quality difference between the two products shrinks to zero, the higher-quality product (which attracts the more expensive consumers) would unravel and only the low-quality product would remain in the market.

---

[6] Existence of a Nash Equilibrium in such a situation is not guaranteed. HHW analyze this situation in great detail and propose using the Riley Equilibrium definition (Riley 1979) instead. A Riley Equilibrium is the Nash Equilibrium when a Nash Equilibrium exists, but the Riley Equilibrium always exists (even when Nash Equilibrium does not).

5.2 Imperfect competition

When imperfect competition is added to selection markets, the result is a classic illustration of the theory of the second best (Lipsey and Lancaster 1956): In the presence of multiple market failures, reducing or eliminating one market failure does not necessarily produce a more efficient allocation.

To illustrate some of the key ideas, we consider a monopolist pricing decision within the framework introduced in Section 3 and a single (non-customized) price. As usual, the monopolist's pricing decision is the standard solution to its profit maximization problem:

$$p^* = \underset{p}{\operatorname{argmax}}\, D(p)p - C(p),$$

where $D(p)$ is demand and $C(p)$ is total cost. To see how selection changes the firm's pricing decision, the firm's first-order condition can be written as

$$p^* = AC(p) + \frac{D(p)}{-D'(p)}\big(1 - AC'(p)\big),$$

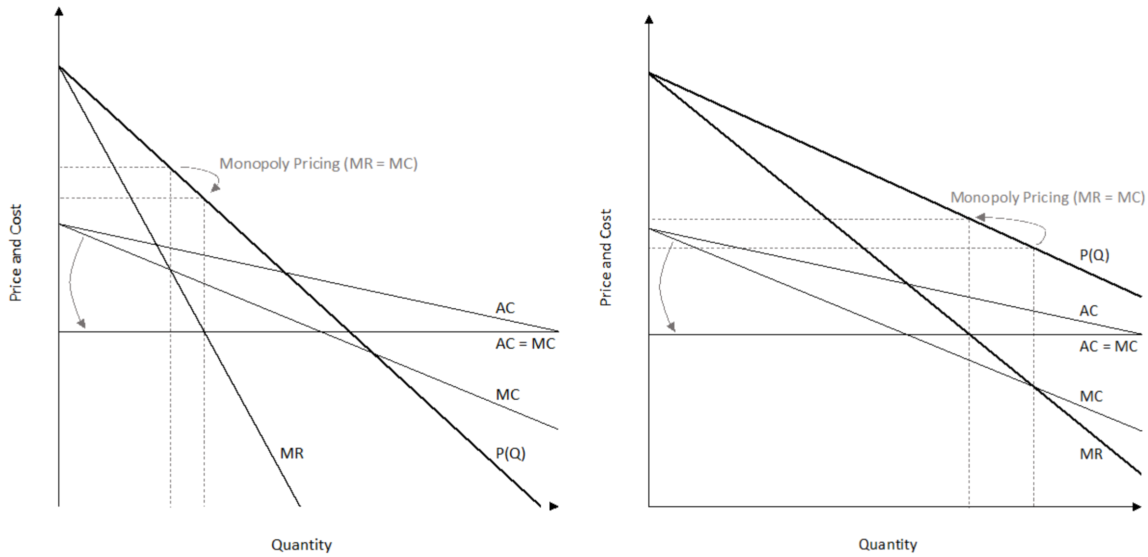
where $AC(p) = \frac{C(p)}{D(p)}$ is average costs. When there is no selection, average costs are constant (recall we have assumed no economies or diseconomies of scale) and the first-order condition reduces to the standard costs plus markup expression: $p^* = AC(p) + \frac{D(p)}{-D'(p)}$.

Because selection implies a non-zero slope of average costs ($AC'(p) \neq 0$), it is tempting to think that changes in the degree of selection modify the firm's first-order condition by re-scaling the markup term $\frac{D(p)}{D'(p)}$ by $\big(1 - AC'(p)\big)$. For instance, one might think that adverse selection, which is characterized by average costs that increase with price ($AC'(p) > 0$), reduces markups and equilibrium prices.

However, a change in the degree of selection typically involves not only a change in the slope of average costs at a given price, $AC'(p)$, but also a change in the level of average costs at that price, $AC(p)$. This complicates the analysis. To deal with this issue, Mahoney and Weyl (2017) consider changes in the degree of selection, holding fixed average costs for the population of potential consumers. A decrease in the degree of selection in their framework therefore corresponds to a mean-preserving decrease in the spread of population costs; equivalently, it can be thought of as a rotation of average costs around the point $AC(p = 0)$ where there is full takeup of the product. With this formulation, the impact of a decrease in the degree of adverse selection depends on the equilibrium quantity at baseline. Figure 2 – taken from Mahoney and Weyl (2017) – depicts the monopolist pricing problem, which is to set the price that equates marginal revenue and marginal cost. When equilibrium quantity is low, the marginal consumer is more expensive than the population average and a decrease in the spread lowers the cost of the marginal consumer, lowering prices (Panel A). When equilibrium quantity is high, the marginal consumer is less costly than average and a decrease in the spread raises average costs, pushing up prices (Panel B). The comparative statics with advantageous selection are similarly subtle.

Figure 2: Decrease in Adverse Selection with Monopolist

(A) Low Baseline Quantity (B) High Baseline Quantity



One way to avoid this complexity is to model changes in selection as a rotation of costs around the equilibrium price (Starc 2014). In this case, average cost $AC(p)$ does not change, and the impact on equilibrium prices is solely pinned down by changes in $AC'(p)$. In such a situation, increasing adverse selection makes $AC(p)$ steeper (and thus $AC'(p)$ becomes more negative), and therefore unambiguously reduces prices and markups, a straightforward application of the theory of the second best. The analysis of this type of rotation is clean, but it has two drawbacks. First, a rotation of the average cost curve around the equilibrium price would have no effect on equilibrium in a perfectly competitive market, since $AC'(p)$ drops out of the first-order conditions. This counterfactual thus conflicts with the intuition that selection affects equilibrium outcomes in competitive markets. Second, it is difficult to motivate why the rotation of the average costs curve would occur exactly at the equilibrium price.

Einav, Finkelstein, and Tebaldi (2019) provide a (rare) example where a rotation at the equilibrium price is a plausible thought experiment. They analyze alternative, budget-neutral policy interventions in an insurance market. Specifically, they analyze premium subsidies for consumers and risk adjustment for private insurers. Under risk adjustment, the government pays the private insurer a fixed amount per customer, with that amount varying ("adjusted") based on the customer's observable characteristics ("risk"). Because they compare optimal risk adjustment under a fixed government budget against a budget-neutral alternative of a uniform subsidy and no risk adjustment, the fixed budget assumption makes the equilibrium a natural rotation point.

Mahoney and Weyl (2017) show that these "theory of the second best" intuitions extend beyond the case of monopoly that we have considered thus far to symmetric forms of imperfect competition, such as symmetric Cournot and symmetric differentiated-products Bertrand competition. In particular, increases in the degree of market power can have differential impacts on social surplus depending on the sign and magnitude of selection. Similarly, the welfare effects of changes in the degree of selection are subtle and depend on the direction of selection, the prevailing quantity in the market, and the degree of market power. As a result, predicting equilibrium effects in settings with market power and selection is ultimately an empirical question. Recent empirical work that emphasizes the interaction of market power and selection include Cabral et al. (2018) on privately provided Medicare health insurance in the U.S.; Nelson (2020) and Cuesta and Sepulveda (2019) on consumer lending in the U.S. and Chile, respectively; and Crawford et al. (2018) on small business lending in Italy.

Making progress with richer models of oligopoly than the symmetric oligopoly models studied by Mahoney and Weyl (2017) is challenging. In selection markets with oligopolistic competition, not only does a firm's price affect its own cost by attracting specific types of consumers, but it also affects their rivals' cost at the same time. For this reason, solving for equilibrium prices in oligopolistic selection markets is more nuanced, and some of the familiar results regarding existence and uniqueness (e.g., Caplin and Nalebuff 1991) do not necessarily hold without additional assumptions. More work is needed to generate tractable restrictions on selection that permit us to characterize equilibria in such environments.

5.3 *Endogenous contract design*

Our discussion of the supply side has thus far concentrated on pricing, taking other product characteristics as fixed. This focus reflects the nature of much of the empirical work on selection markets and is similar to empirical IO work on more traditional markets. There, the focus is on pricing, while the set of products and their attributes are taken as given (and sometimes – somewhat uncomfortably – are even assumed exogenous to demand in order to justify the oft-used so-called "BLP instruments").

Endogenizing product attributes is an important direction for work in conventional IO markets (recent examples include Sweeting (2013) and Fan and Yang (2020)). It may be even more important when studying selection markets. Here, the products – such as insurance and credit – are often financial products, and many of the product attributes are price-like – e.g., deductible and copayment rates in insurance markets, or minimum down payment requirements and minimum loan amounts in consumer credit markets. These product characteristics in principle can be changed as frequently and as easily as the price itself, unlike say, the design of a car.[7] Another motivation for endogenizing product attributes in empirical analysis of selection markets is that it has played a central role in the theoretical analysis of these markets. As we discussed in Section 2, although the seminal contribution of Akerlof (1970) focused on fixed contracts, the influential work of Rothschild and Stiglitz (1976) – as well as the subsequent theoretical literature it spawned – focused on endogenous contract design in analyzing equilibrium in selection markets.

Conceptually, endogenizing contract design elements is straightforward. Consider, for example, a case in which markets clear through a single contract design element, instead of through price. Denote by ψ this endogenous contract feature and assume that all other product features (including the price) are taken as given, so that each product j is defined by a vector (ψ_j, p_j, x_j).

[7] Recent examples of endogenous product design in the context of health insurance include Carey (2017), Decarolis and Guglielmo (2017), Lavetti and Simon (2018), and Geruso et al. (2019).

There is no distinction between a setting in which price is the endogenous supply-side market clearing variable and a setting with a non-price endogenous variable if the price is allowed to enter costs. However, if we assume, as is typical in insurance applications, that prices do not have an impact on costs (i.e., $c_{ij} = c(x_j, \zeta_i)$) while we allow for non-price elements to matter (that is, $c_{ij} = c(\psi_j, x_j, \zeta_i)$), then there are important distinctions.

To see this, consider the monopolist's objective, $\max_{\psi} D(\psi)p - C(\psi)$, and observe that the first-order condition – instead of the usual form equating marginal revenue to marginal cost, $D'(\psi)p = C'(\psi)$ – can be written as

$$D'(\psi)E(p - c_i(\psi)|v_i(\psi) = p) - D(\psi)E(c'_i(\psi)|v_i(\psi) > p) = 0.$$ [8]

The first term is the familiar marginal impact of ψ on demand multiplied by the profit of the marginal customer. What distinguishes this term from the analog in a non-selection market is that marginal profits are determined by the marginal consumer (defined as having $v_i(\psi) = p$), which could be very different from the profits from the average customer who purchases the product. The second term captures the effect of the change in ψ on all infra-marginal consumers. A similar cost effect would show up in non-selection markets, but in selection markets this impact is likely heterogeneous across consumers, so again it is important to take selection into account.

Depending upon the endogenous contract feature, either (or both) of these terms could be important. For example, Einav, Jenkins, and Levin (2012) analyze the optimality of two auto-loan design components and show that one (minimum required down payment) primarily operates through its impact on selection (the first term), while the other (which can be thought of as the interest rate) primarily operates through its impact on repayment rate by infra-marginal consumers (the second term).

Endogenizing contract design becomes more difficult when there are multiple dimensions of the product that are allowed to be endogenous, such as price and one or more non-price characteristics.[9] At this point, just like in any traditional market, one has to take a stance as to whether prices equilibrate simultaneously with other product features, or whether firms are playing a two-stage game in which they first select product design and then choose prices to clear markets.

Finally, yet another feature that distinguishes an endogenous contract design element from price is in generating scope for direct externalities. We have already discussed how selection markets generate indirect externalities from one competitor to another through consumer selection (see Section 5.2). With an endogenous contract design, such externalities may have a more direct (and perhaps more first-order) impact on competitors. For example, in the context of credit markets, when consumers take loans from multiple lenders, the credit terms of one lender may directly impact the default probability of loans taken from the other lender, thus altering equilibrium interest rates in the credit market (Bizer and DeMarzo 1992). A similar situation may arise in insurance markets. For example, Cabral and Mahoney

---

[8] To see this, rewrite the objective as $\max_{\psi} \int_{\{i: v_i(\psi) > p\}} (p - c_i(\psi))\, di$ and apply the Leibniz's rule for differentiation under the integral sign.

[9] Moreover, as Veiga and Weyl (2016) point out, this can be even more complicated when consumers are heterogeneous on multiple dimensions, so that the nature of selection could differ depending on which contract element is being changed.

(2018) show that taking up supplemental Medicare coverage (often referred to as Medigap) increases demand for healthcare services, thus generating negative externalities on the original Medicare insurer.

5.4 Rejections and customized pricing

In conventional IO markets, empirical analysis often assumes that any consumer is able to purchase any product j if they are willing to pay the price p_j. Empirical analyses typically take advantage of variation in prices across products or within a product across distinct markets separated in time or geography. Because costs in traditional markets do not depend on the consumer who purchases the product, consumers are rarely denied the option to purchase a product, and customized pricing is not a central feature in many markets. This is, of course, a simplification; for example, price negotiation is a very common feature of the US car market, and there is entire literature that studies third-degree price discrimination. However, it is often a reasonable abstraction.

In selection markets, the assumption that everyone faces the same price usually does not hold. Credit card interest rates vary by more than a factor of three across individuals with different credit scores (Agarwal et al. 2015), and, absent regulation, the price of private health insurance would likely vary by more than a factor of five across age groups (Orsini and Tebaldi 2017). One reason that customized pricing is more central in selection markets is that potential customers differ not only in their willingness to pay (as in traditional markets) but also in their costs. Cost-based pricing raises a host of conceptually interesting and policy-relevant questions, as well as tradeoffs for economic analysis.

The most extreme, but not uncommon, case is where the market does not exist for some type of consumers, which can be thought of as an infinite price faced by those consumers. The foundational theoretical work by Akerlof (1970) emphasized the possibility that selection can lead to market nonexistence for everyone, even if there is surplus to be had for every possible consumer in the market. Graphically, this can be illustrated by moving the curves in Figure 1 so that the average cost curve is everywhere above the demand curve but the marginal cost curve is everywhere below the demand curve (see Einav and Finkelstein 2011, Figure 2.B). Empirically, Hendren (2017) argues that individuals' private information about potential future job loss leads to the complete nonexistence of the market for private unemployment insurance.

In many settings, this so-called "unraveling" happens for a specific set of consumers, while leaving the market in existence for others. This generates the common – and ostensibly puzzling – phenomenon of insurance rejections, in which companies refuse to sell insurance to applicants who have certain observable, often high-risk, characteristics (despite the absence of any regulatory restriction against charging them a higher price). For example, people who have had a stroke may find it impossible to purchase long-term care insurance, and people with pre-existing health conditions were often unable to purchase private health insurance until regulation was put in place that outlawed this practice (Hendren 2013). In credit markets, it is common for lenders to only provide loans to people with credit scores above a certain level.

In settings where consumers have to apply for a product and can be rejected, the modeling framework needs to be extended to account for the fact that consumers' choice sets are endogenously determined by the application decisions of consumers and the rejection decisions of firms. Cuesta and Sepulveda (2019) augment the standard framework to account for applications and rejections in their study of consumer loans in Chile. Adapting their model to our notation, let $A \subseteq J$ indicate the set of products that the consumer chooses to apply to and let $A' \subseteq A$ be the set of these applications that are approved (not rejected) by the lender. Consumers choose the set of loans to apply to (A) to maximize utility

$$u_i(A) = \Pr(A' \neq \emptyset)\, E[\max_{j \in A'} \{v(x_j, \zeta_i) - p_{ij}\} |\, A' \neq \emptyset] + \Pr(A' = \emptyset)\, v_{i0} - \kappa_i(A) + \epsilon_i$$

where the first term is the probability that the set of approved loans is non-empty multiplied by expected value of loan approval, the second term is the value from rejection, the third term $\kappa_i(A)$ is the application cost, and ϵ_i is a shock to the utility of applying for loans relative to the outside option. The outside option of not applying is the same as the value of being rejected, v_{i0}. Cuesta and Sepulveda (2019) combine this equation with product pricing and cost equations that are similar to those presented in Section 4. They allow for selection by permitting correlation between the shock to the application equation ϵ_i and the shock in the cost equation.

At first blush, cost-based pricing in selection markets may seem more palatable than pricing variation that is based solely on willingness to pay in traditional markets, which is sometimes deemed unfair and has been known to create regulatory or consumer backlash. Because variation in marginal costs can lead to market unraveling and the elimination of all of the potential surplus from the market, there are potential benefits of allowing firms to reject some applications or charge them cost-based prices. Indeed, there may even be a role for government to aid in the collection and distribution of the relevant data needed to implement cost-based prices. The development of credit registries to record consumer borrowing and repayment behavior is a salient example. De Janvry et al. (2010) show how the availability of credit histories has affected the credit market in Guatemala.

However, customization of prices can also have negative consequences. Indeed, it is possible that allowing firms to price on additional consumer characteristics can *increase* the welfare costs of adverse selection if it does not eliminate private information entirely. Intuitively, pricing on an additional binary characteristic segments the market from one (pooled) adversely selected market into two, and there is no a-priori reason why the sum of the welfare losses from adverse selection in each market will be less than the welfare loss in the pooled market. Again, this is an application of the theory of the second best. One recent empirical application has considered the consequences of introducing finer risk-based prices in the Medicare Advantage program on selection in this market (McWilliams et al. 2012; Brown et al. 2014). In credit markets, Liberman et al. (2020) have studied the effects of a coarsening of pricing due to information deletion.

Another concern with cost-based pricing is that the type of information that facilitates cost-based pricing can also allow firms to take advantage of market power. For example, Nelson (2020) shows that the information credit card issuers acquire about borrowers via their use of the card not only allows them to better tailor their prices to predicted default risk, which can reduce market unraveling, but also allows them to raise prices to extract rents from inelastic borrowers.

Customized pricing can also make it conceptually more difficult to analyze the pricing game, as it opens the door for many other, less traditional aspects of competition. That is, while in conventional markets the key pricing decision centers around the price level, in selection markets sellers also need to choose how finely to partition consumers into groups. For example, in the early 2000s, auto insurance companies realized that consumers' credit histories are predictive of their auto insurance claims risk and started incorporating this information into their price offers (Carter 2005). Relatedly, Einav, Jenkins, and

Levin (2012) estimate the extent to which a lender who can partition potential borrowers into risk groups (and price loan accordingly) benefits relative to a lender who cannot.

A broader concern with cost-based pricing is that if the information used for cost-based pricing or the way it can be utilized is proprietary, it may create a strategic barrier to entry or to competition across firms. To the extent that market leaders are able to acquire additional information to help set prices for their customers, a new entrant without such stock of customer information might find itself at an additional disadvantage in the market. For example, Einav, Jenkins, and Levin (2012) illustrate – in the context of a subprime auto lender – that such information advantages could raise barriers to entry by 25-30%.

Indeed, this aspect of competition in pricing strategies rather than in prices per se is a feature of selection markets and has not been explored much in the academic literature. The advent of big data and artificial intelligence is likely to make such pricing strategies become an increasingly important aspect of competition in selection markets, and its interaction with regulatory policies (e.g., policies to address concerns about discrimination and consumer privacy) make it an important topic for future work.

Finally, cost-based pricing can also raise important concerns about fairness and equity. In the US, these concerns are particularly salient when there are high rejection rates or higher prices for historically underprivileged groups. For example, there is concern that artificial intelligence in consumer lending will have a disparate impact by race and other protected classes (Klein 2020). Fairness and equity concerns also play a role in regulation of insurance products. For example, because of concerns about high health insurance prices for the nearly elderly, the Affordable Care Act restricted the amount that health insurance premiums could vary across age groups.

6. Welfare

An important result of the original theoretical work on selection markets was that private information may impair the efficient operation of markets, providing scope for welfare-improving government policy. Motivated by this result, a key focus of the empirical work on selection markets has been to develop frameworks for estimating the welfare cost of selection and the welfare consequences of the corresponding policy responses, such as mandates, subsidies, and risk adjustment. Einav, Finkelstein, and Levin (2010) provide a more detailed discussion of this empirical welfare analysis.

In this section, we briefly discuss two specific welfare questions that have been the subject of more recent attention. The first follows the work just discussed on customized pricing to consider its welfare implications. The second explores approaches to empirical welfare analysis when one cannot (or does not want to) use the demand curve to reveal consumer preferences.

6.1 Customized pricing and reclassification risk

The customized, cost-based pricing discussed in the last section creates a tradeoff between static and dynamic concepts of welfare. In lending markets, for instance, more information on past defaults can reduce the welfare cost of adverse selection, but it can also expose individuals to long-run consequences of bad luck or bad choices. To the extent that we want to insure people against long-run consequences, there is a natural tradeoff between the selection-reducing benefits of cost-based pricing and the insurance benefits of providing people with a “fresh start.” In accordance with this tradeoff, credit bureaus in many countries delete information after a specified amount of time (typically 7 years in the US). To help quantify these tradeoffs, empirical researchers have studied the impacts of this type of information deletion policies in a number of settings (Liberman et al. 2020; Dobbie et al. 2021).

A related downside of cost-based pricing is that it creates dynamic risk for a person over time. For instance, in US health insurance markets, contracts are typically one year in length. With full cost-based pricing, if an individual’s health condition changes, they will face a different premium in subsequent years. In the extreme, if health conditions are persistent, this type of “reclassification risk” can undo much of the financial protection that health insurance would ideally offer. One way to address this conflict is to make the coverage period longer (Hendel and Lizzeri 2003) or provide health-based severance payments (Cochrane 1995). When this is not feasible (e.g., because incomplete contracts over future provider networks create hold-up issues), there is a clear efficiency tradeoff between adverse selection and reclassification risk.

This tradeoff is the focus of the study by Handel, Hendel, and Whinston (2015). Using individual-level panel data on health insurance purchases and claims from a large employer, they estimate demand for annual health insurance coverage (applying a two-period model of the type described in Section 4) and the serial correlation in health insurance claims across years. They compute welfare under two alternative (feasible) contract designs and compare it to welfare under a long-term coverage contract, which is infeasible in this setting. In one case, cost-based pricing is not allowed, thus leading to adverse selection. In the other, cost-based pricing is admissible, which resolves the adverse selection concern but leads to changing premiums over time and thus to reclassification risk. They then compare the welfare loss associated with this reclassification risk with the welfare loss associated with adverse selection.

6.2 Demand vs. welfare

In many IO applications, both in conventional markets and selection markets, a key objective is to assess the impact of market interventions on consumer welfare. Having a complete model of individual utility (as in Section 3) makes the welfare exercise standard and natural. That is, the framework of Section 3 delivers estimates of primitives, which can be used to construct $v_{ij} = v(x_j, \zeta_i)$ and to examine how it varies in response to changes in market equilibrium and product allocation.

Yet, in many markets evidence of consumer mistakes and related calls for paternalistic regulation raise concerns about reliance on revealed preferences. While this is not unique to selection markets, it is an issue that has received considerable attention in these markets. This may be because mistakes are more prevalent in selection markets, perhaps because the products are more complex and it is harder for the consumer to learn their preferences from experience. If a consumer buys an expensive sweater and does not wear it, they may learn about their tastes and cut their clothing budget going forward. But if an individual buys a homeowner’s insurance policy with an extremely high deductible (as one of us has done), and two years in a row has leaks and floods that blow through the deductible, it’s not obvious that the high deductible policy is a mistake (no matter what their spouse may think).

Another reason for the emphasis on behavioral economics in selection markets may be purely practical. Because many selection markets involve financial products (e.g., insurance and credit) with rich data on post-purchase behavior, it may be easier for the researcher to try to establish “mistakes” empirically in selection markets than in traditional non-selection markets (e.g., choice of cars or flat screen TVs). For example, the choice of a dominated health insurance plan – one that costs the consumer more in every state of the world than another but is otherwise identical – seems like a clear mistake (Handel 2013; Bhargava et al. 2017), while documenting a dominated car choice is more challenging. One could imagine similar behavioral mistakes in more conventional markets – e.g., someone buying too much milk and having to throw it away, purchasing a large flat screen TV and never turning it on – but documenting these types of mistakes is much more difficult.

From a positive (rather than normative) perspective, this concern is unimportant. Counterfactual effects on price and allocations are the same regardless of whether the utility function representation in Section 3 represents "true" utility or instead is just used to describe demand. However, any consumer welfare exercise needs to take a strong stance on what the actual welfare function is and if it is different from the one revealed by demand.

Spinnewijn (2017) and Handel, Kolstad, and Spinnewijn (2019) use the Einav, Finkelstein, and Cullen (2010) framework to illustrate this point. In addition to the cost and demand curves shown in Figure 1, they draw a third curve – the "value curve" – that captures true utility. Knowledge of the value curve is now essential to make any consumer welfare statement. Estimating this value curve in actual applications, however, remains an important empirical challenge.

One application that illustrates this point is the work by Cuesta and Supeldova (2019) described earlier. Their paper analyzes the impact of interest rate caps on competition in the consumer credit market in Chile. From a positive perspective, the estimated effect of interest rate caps on competition and lender profitability is true regardless of utility interpretation. But analysis of the welfare impact of these changes requires a normative stance on the issue of potentially "excessive" credit demand by the poor, a topic that has attracted much academic and policy debate (Karlan and Zinman 2010; Zinman 2014).

Even in the absence of these "behavioral" concerns, insurance markets present another challenge for using demand to reveal the relevant construct of utility: the issue of when demand is measured. In conventional markets, we tend to think of preferences (and willingness to pay) as stable. However, as Hendren (2020) points out, in many insurance and credit settings, individual risk (or information about that risk) evolves over time, and therefore willingness to pay for insurance (or credit) evolves as well. The welfare gain from insurance can be measured at the time that the insurance is purchased (which is what demand for insurance may reveal), but this may be quite different from the welfare gain as evaluated from an earlier time point, prior to some (or any) realization of risk. For instance, when an individual with chronic health conditions purchases health insurance, some uncertainty has already been resolved and the current ("revealed") value of insurance may be lower than the ex-ante value before the onset of those chronic conditions.

7. Looking ahead

It has now been a half-century since the publications of Akerlof (1970) and Rothschild and Stiglitz (1976) launched the study of selection markets in economics. The first wave of research in the 1970s was mainly theoretical. It was followed, several decades later, by empirical research focused on testing whether or not selection actually existed in specific, real-world markets. In this chapter, we describe a third wave, which empirically models equilibrium demand and supply in these selection markets, allowing us to quantify the welfare cost of selection and the potential benefits of remedial policies.

We have tried to offer a simple empirical framework to synthesize this “third wave” of studies. The framework closely follows the traditional empirical IO framework, which we hope makes it accessible and intuitive to scholars familiar with empirical IO models in more traditional markets. Our aim is to help scholars better understand this latest wave of empirical work on selection markets, and to provide a foundation for those looking to continue the empirical analysis of selection markets. In this last section, we therefore briefly indulge in suggesting a few areas of challenge and opportunity for future work.

A core tension in the empirical literature on selection is that it typically focuses on markets where we can observe individual choices. However, as we noted in Section 2, the original Akerlof theory emphasized that selection can cause markets to unravel completely, in which case there are no choices to observe. The natural concern is that the markets that we study empirically are – by virtue of their very existence – those markets where selection is not large enough to cause market unraveling. To overcome this issue, researchers have taken to supplying the product oneself so as to estimate demand and costs (Fischer et al. 2018), directly eliciting private information about risk and modeling its potential implications for market equilibrium (Hendren 2013, 2017), calibrating utility models (Hosseini 2015), and using behavioral responses to shocks to infer the consumer surplus from insurance (Landais and Spinnewijn 2020). Such explorations seem an important area for future work.

Another opportunity lies in applying the approaches we have discussed to other markets. The examples we used in this chapter focused on insurance and credit markets – markets where selection concerns are undoubtedly first order. However, as mentioned in the Introduction, the central feature of selection markets – that customers vary in their cost so that firms care about *which* customers they attract – applies more broadly. Labor markets and education markets are two natural and important settings where selection concerns may well be quantitatively important. Another setting in which selection may become important is subscription markets (e.g., software as a service). In such markets, sellers try to associate potential and existing customers with their “lifetime value” and invest greater resources in attracting (or retaining) those with the highest lifetime profits. Although our framework emphasized selection on costs, the framework we provide in this chapter should be useful for analyzing these isomorphic situations, where selection is on future value.

Another promising area for study lies at the intersection of big data, machine learning, and artificial intelligence. The use of machine learning to engage in risk-based pricing in insurance and credit markets provides a new setting to engage on old questions of static vs. dynamic efficiency and in the equity and fairness of allowing price variation. The prospect of imperfect competition among firms with different amounts of data and different pricing algorithms has and will continue to raise important questions for competition policy. We hope that the frameworks described in this chapter may prove useful to scholars interested in pursuing such topics.

References

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel. 2015. “Regulating Consumer Financial Products: Evidence from the Card Act” *Quarterly Journal of Economics* 130 (1): 111-64.

Akerlof, George. 1970. “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism.” *Quarterly Journal of Economics* 84: 488–500.

Arrow, Kenneth J. 1963. “Uncertainty and The Welfare Economics of Medical Care.” *American Economic Review* 53 (5), 941–973.

Azevedo, Eduardo M., and Daniel Gottlieb. 2017. “Perfect Competition in Markets with Adverse Selection.” *Econometrica* 85: 67-105.

Bachas, Natalie. 2019. “The Impact of Risk-Based Pricing in the Student Loan Market: Evidence from Borrower Repayment Decisions.” Mimeo.

Barseghyan, Levon, Jeffrey Prince, and Joshua C. Teitelbaum. 2011. “Are Risk Preferences Stable across Contexts? Evidence from Insurance Data.” *American Economic Review* 101 (2): 591–631.

Berry, Steve, and Phil Haile. 2021. “Demand.” In *The Handbook of Industrial Organization*, Vol. 4.

Bhargava, Saurabh, George Loewenstein, Justin Sydnor. 2017. “Choose to Lose: Health Plan choices from a Menu with Dominated Options.” *Quarterly Journal of Economics* 132 (3): 1319-72.

Bizer, David S., and Peter M. DeMarzo. 1992. “Sequential Banking.” *Journal of Political Economy* 100 (1): 41-61.

Brown, Jason, Mark Duggan, Ilyana Kuziemko, and William Woolston. 2014. “How does Risk Selection Respond to Risk Adjustment? New Evidence from the Medicare Advantage Program.” *American Economic Review* 104 (10): 3335-64.

Bundorf, M. Kate, Jonathan Levin, and Neale Mahoney. 2012. “Pricing and Welfare in Health Plan Choice.” *American Economic Review* 102 (7): 3214–48.

Cabral, Marika, Michael Geruso, and Neale Mahoney. 2018. “Do Larger Health Insurance Subsidies Benefit Patients or Producers? Evidence from Medicare Advantage.” *American Economic Review* 108 (8): 2048-87.

Cabral, Marika, Can Cui, and Michael Dworsky. 2019. “What is the Rationale for an Insurance Coverage Mandate? Evidence from Workers' Compensation Insurance.” Mimeo.

Cabral, Marika, and Neale Mahoney. 2018. “Externalities and Taxation of Supplemental Insurance: A Study of Medicare and Medigap.” *American Economic Journal: Applied Economics* 11 (2): 33-73.

Caplin, Andrew, and Barry Nalebuff. 1991. “Aggregation and Imperfect Competition: On the Existence of Equilibrium.” *Econometrica* 59 (1): 25-59.

Cardon, James H., and Igal Hendel. 2001. “Asymmetric Information in Health Insurance: Evidence from The National Medical Expenditure Survey.” *RAND Journal of Economics* 32 (3): 408–27.

Carey, Colleen. 2017. “Technological Change and Risk Adjustment: Benefit Design Incentives in Medicare Part D.” *American Economic Journal: Economic Policy* 9(1): 38-73.

Carter, Adrienne. 2005. “Telling The Risky From The Reliable.” *BusinessWeek*, August 1, 2005.

Chiappori, Pierre-Andre. 2013. “Econometric Models of Insurance under Asymmetric Information.” In Dionne George (eds.), *Handbook of Insurance*, Springer: 365-393.

Chiappori, Pierre-Andre, and Bernard Salanie. 2000. “Testing for Asymmetric Information in Insurance Markets.” *Journal of Political Economy* 108 (1): 56-78.

Cochrane, John. 1995. “Time-Consistent Health Insurance.” *Journal of Political Economy* 103: 445-473.

Cohen, Alma, and Liran Einav. 2007. “Estimating Risk Preferences from Deductible Choice.” *American Economic Review* 97 (3): 745–88.

Cohen, Alma, and Peter Siegelman. 2010. “Testing for Adverse Selection in Insurance Markets.” *Journal of Risk and Insurance* 77 (1): 39-84.

Crawford, Gregory S., Nicola Pavanini, and Fabiano Schivardi. 2018. “Asymmetric Information and Imperfect Competition in Lending Markets.” *American Economic Review* 108 (7):1659–1701.

Cuesta, Jose Ignacio, and Alberto Sepulveda. 2019. Price Regulation in Credit Markets: A Trade-off between Consumer Protection and Credit Access.” Mimeo.

Davis, Lucas. 2008. “Durable Goods and Residential Demand for Energy and Water: Evidence from a Field Trial.” *RAND Journal of Economics* 39 (2): 530-46.

De Janvry, Alain, Craig McIntosh, and Elisabeth Sadoulet. 2010. “The Supply- and Demand-Side Impacts of Credit Market Information.” *Journal of Development Economics* 93 (2): 173-188.

Decarolis, Francesco, and Andrea Guglielmo. 2017. “Insurers’ Response to Selection Risk: Evidence from Medicare Enrollment Reforms.” *Journal of Health Economics* 56: 383-396.

Dionne, Georges, Neil Doherty, and Nathalie Fombaron. 2013. “Adverse Selection in Insurance Markets.” In Dionne George (eds.), *Handbook of Insurance*, Springer: 185-243.

Dobbie, Will, Paul Goldsmith-Pinkham, Neale Mahoney, and Jae Song. 2021. “Bad Credit, No Problem? Credit and Labor Market Consequences of Bad Credit Reports.” *Journal of Finance*. Forthcoming.

Dubin, Jeffrey A., and Daniel L. McFadden. 1984. “An Econometric Analysis of Residential Electric Appliance Holdings and Consumption.” *Econometrica* 52 (2): 345–62.

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 Evidence in Pictures.” *Journal of Economic Perspectives* 25 (1): 115–138.

Einav, Liran, Amy Finkelstein, and Jonathan Levin. 2010. “Beyond Testing: Empirical Models of Insurance Markets.” *Annual Review of Economics* 2 (1): 311–36.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2010. “Optimal Mandates and The Welfare Cost of Asymmetric Information: Evidence from The U.K. Annuity Market.” *Econometrica* 78 (3), 1031-1092.

Einav, Liran, Amy Finkelstein, Steven P. Ryan, Paul Schrimpf, and Mark R. Cullen. 2013. “Selection on Moral Hazard in Health Insurance.” *American Economic Review* 103 (1), 178–219.

Einav, Liran, Amy Finkelstein, and Pietro Tebaldi. 2019. “Market Design in Regulated Health Insurance Markets: Risk Adjustment vs. Subsidies.” Mimeo.

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

Fan, Ying, and Chenyu Yang. 2020. “Competition, Product Proliferation and Welfare: A Study of the U.S. Smartphone Market.” *American Economic Journal: Microeconomics* 12 (2): 99-134.

Fang, Hanming, Michael Keane, and Dan Silverman. 2008. “Sources of Advantageous Selection: Evidence from the Medigap Insurance Market.” *Journal of Political Economy* 116 (2): 303–50.

Fang, Hanming, and Zenan Wu. 2018. “Multidimensional private information, market structure, and insurance markets.” *RAND Journal of Economics* 49(3): 751-787.

Finkelstein, Amy, and Kathleen McGarry. 2006. “Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market.” *American Economic Review* 96 (4): 938–58.

Finkelstein, Amy, and James Poterba. 2002. “Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market.” *Journal of Political Economy* 112 (1): 183-208.

Finkelstein, Amy, Nathaniel Hendren, and Mark Shepard. 2019. “Subsidizing Health Insurance for Low-Income Adults: Evidence from Massachusetts.” *American Economic Review* 109 (4): 1530–67.

Fischer, Torben , Markus Frolich, and Andreas Landmann. 2018. “Adverse Selection in Low-Income Health Insurance Markets: Evidence from a RCT in Pakistan.” IZA Discussion Paper No. 11751.

Geruso, Michael, Timothy J. Layton, Grace McCormack, and Mark Shepard. 2021. “The Two Margin Problem in Insurance Markets.” *Review of Economics and Statistics*, forthcoming.

Geruso, Michael, Timothy J. Layton, and Daniel Prinz. 2019. “Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges,” *American Economic Journal: Economic Policy* 11(2): 64-107.

Handel, Benjamin R.. 2013. “Adverse Selection and Switching Costs in Health Insurance Markets: When Nudging Hurts.” *American Economic Review* 103: 2643–82.

Handel, Benjamin R., Igal Hendel, and Michael D. Whinston. 2015. “Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk.” *Econometrica* 83: 1261-1313.

Handel, Benjamin R., Jonathan T. Kolstad, and Johannes Spinnewijn. 2019. “Information Frictions and Adverse Selection: Policy Interventions in Health Insurance Markets.” *Review of Economics and Statistics* 101 (2): 326-40.

Hendel, Igal, and Alessandro Lizzeri. 2003. “The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance.” *Quarterly Journal of Economics* 118: 299–328.

Hendren, Nathaniel. 2013. “Private Information and Insurance Rejections.” *Econometrica* 81 (5): 1713-1762.

Hendren, Nathaniel. 2014. “Unraveling versus Unraveling: A Memo on Competitive Equilibriums and Trade in Insurance Markets.” *Geneva Risk and Insurance Review* 39 (2): 176-183.

Hendren, Nathaniel. 2017. “Knowledge of Future Job Loss and Implications for Unemployment Insurance.” *American Economic Review* 107 (7): 1778-1823.

Hendren, Nathaniel. 2020. “Measuring Ex-Ante Welfare in Insurance Markets.” *Review of Economic Studies*, Forthcoming.

Hosseini, Roozbeh. 2015. "Adverse Selection in the Annuity Market and the Role for Social Security." *Journal of Political Economy* 123 (4): 941-84.

Illanes, Gaston, and Manisha Padi. 2021. “Retirement Policy and Annuity Market Equilibria: Evidence from Chile.” NBER Working Paper No. 26285.

Jaffee, Dwight M., and Thomas Russell. 1976. “Imperfect Information, Uncertainty, and Credit Rationing.” *Quarterly Journal of Economics* 90(4): 651–66.

Karlan, Dean, and Jonathan Zinman. 2010. “Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts.” Review of Financial Studies 23 (1): 433-64.

Klein, Aaron. 2020. “Reducing bias in AI-based financial services.” Brookings Report. Available at https://www.brookings.edu/research/reducing-bias-in-ai-based-financial-services/

Landais, Camille, Arash Nekoei, Peter Nilsson, David Seim and Johannes Spinnewijn. 2020. “Risk-Based Selection in Unemployment Insurance: Evidence and Implications.” Mimeo.

Landais, Camille, and Johannes Spinnewijn. 2020. The Value of Unemployment Insurance.” Mimeo.

Lavetti, Kurt, and Kosali Simon. 2018. “Strategic Formulary Design in Medicare Part D Plans.” *American Economic Journal: Economic Policy* 10(3): 154-192.

Liberman, Andres, Chris Neilson, Luis Opazo, and Seth Zimmerman. 2020. “The Equilibrium Effects of Information Deletion: Evidence from Consumer Credit Markets.” Mimeo.

Lipsey, Richard, and Kelvin Lancaster. 1956. “The General Theory of Second Best.” *Review of Economic Studies* 24 (1): 11-32.

Mahoney, Neale, and E. Glen Weyl. 2017. “Imperfect Competition in Selection Markets.” *Review of Economics and Statistics* 99 (4): 637-651.

McWilliams, J. Michael, John Hsu, and Joseph P. Newhouse. 2012. “New Risk-Adjustment System was Associated with Reduced Favorable Selection in Medicare Advantage.” *Health Affairs* 31 (12): 2630–40.

Nelson, Scott. 2020. “Private Information and Price Regulation in the US Credit Card Market.” Mimeo.

Orsini, Joe, and Pietro Tebaldi. 2017. “Regulated Age-Based Pricing under the Affordable Care Act.” Mimeo.

Panhans, Matthew. 2019. “Adverse Selection in ACA Exchange Markets: Evidence from Colorado.” *American Economic Journal: Applied Economics* 11 (2): 1–36.

Pauly, Mark V. 1968. “The Economics of Moral Hazard: Comment.” *American Economic Review* 58 (3): 531–37.

Pueltz, Robert, and Arthur Snow. 1994. “Evidence on Adverse Selection: Equilibrium Signaling and Cross-Subsidization in the Insurance Market.” *Journal of Political Economy* 102 (2): 236–57.

Riley, John G. 1979. “Informational Equilibrium.” *Econometrica* 47: 331–59.

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: 629–49.

Spinnewijn, Johannes. 2017. “Heterogeneity, Demand for Insurance, and Adverse Selection.” *American Economic Journal: Economic Policy* 9 (1): 308–43.

Starc, Amanda. 2014. “Insurer Pricing and Consumer Welfare: Evidence from Medigap.” *RAND Journal of Economics* 45: 198-220.

Stiglitz, Joseph E., and Andrew Weiss. 1981. “Credit Rationing in Markets with Imperfect Information.” *American Economic Review* 71 (3): 393–410.

Sweeting, Andrew. 2013. “Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry.” *Econometrica* 81 (5): 1763-1803.

Veiga, Andre, and E. Glen Weyl. 2016. “Product Design in Selection Markets.” *Quarterly Journal of Economics* 131 (2): 1007-56.

Wagner, Katherine. 2020. “Adaptation and Adverse Selection in Markets for Natural Disaster Insurance. Mimeo.

Weyl, E. Glen, and Andre Veiga. 2017. “Pricing Institutions and the Welfare Cost of Adverse Selection.” *American Economic Journal: Microeconomics* 9 (2): 139-48.

Zeckhauser, Richard. "Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives." Journal of Economic theory 2.1 (1970): 10-26.

Zinman, Jonathan. 2014. “Consumer Credit: Too Much or Too Little (or Just Right)?" *Journal of Legal Studies* 43 (S2): 209-37.