Mergers in the Presence of Adverse Selection

Conor Ryan *

Abstract

In the presence of adverse selection, mergers can increase welfare through a reduction in inefficient sorting. I characterize the sorting externality internalized between merging firms in a tractable discrete choice model. Mergers benefit consumers when the firms are small, willingness to pay is moderately increasing in cost, and consumer costs are skewed. Applying the model to the non-group health insurance market, 13% of potential mergers would improve consumer surplus. In markets where the sorting distortion exceeds $5 per person, nearly one-third of mergers improve consumer surplus, highlighting the importance of considering adverse selection in merger evaluation.

1 Introduction

When can a merger between two firms improve social welfare? In some cases, a merger generates cost synergies through economies of scale or other production complementarities that allow the firms to earn both a greater profit and increase total output. Because identifying and screening harmful mergers is an ongoing policy concern, a growing empirical literature has sought to evaluate and quantify these types of cost efficiencies due to a merger. In this paper, I study a context in which mergers can improve consumer surplus that has received little empirical attention: adverse selection.

A merger in the presence of adverse selection may improve total welfare through a reduction in inefficient sorting. To understand this channel, consider that a firm has an incentive to offer a product that appeals to low-cost consumers and encourages high-cost consumers to purchase from a competitor. This distortion only exists when other firms are available to absorb high-cost

*Thank you to the paper's editor, Ying Fan, and to the anonymous referees throughout the life of the project in helping to guide the paper to its current form. Thank you to Tom Holmes, Amil Petrin, Paul Grieco, Naoki Aizawa, Pietro Tebaldi, and seminar participants at SITE, IIOC, SEA, ARIA, ASHEcon, and the FTC. Contact: Pennsylvania State University, Department of Economics, conor.ryan@psu.edu

consumers. In the event of a merger, there are fewer competitors that might attract the high-cost consumers, and this distortion decreases. Through this reduction in inefficient sorting, a merger has the potential to benefit consumers despite the greater markups charged by the merged firm. Whether the net welfare effect of a particular merger is positive or negative is an empirical question.

I specify an empirically tractable discrete choice model and apply the model to the non-group (individual) health insurance market in the United States. There are two main findings. First, there are potential mergers that lead to greater consumer surplus and social welfare. Mergers are more likely to generate meaningful welfare benefits in markets which suffer from large welfare costs due to inefficient sorting. Second, I show that a generalized pricing pressure measure that accounts for selection is an effective screen for harmful mergers.

I model the insurance market as strategic firms that compete in price with a fixed set of differentiated insurance products. Absent any adverse selection, a merger between any two products creates an incentive to raise prices. Prior to the merger, each product sets a price that balances earning more from infra-marginal consumers with the loss of marginal consumers. The merged firm now internalizes that some consumers lost due to a price increase in one product will be recaptured in the newly acquired product (and vice-versa). These recaptured consumers represent an incentive to raise prices.

The presence of adverse selection complicates this intuition because the cost of a product depends on the set of consumers that purchase it. Therefore, the price incentives that result from a merger depend on the cost of the recaptured consumers. If these consumers are low cost, the standard merger intuition remains. These recaptured, low-cost consumers are an additional benefit to the acquired product and represent another incentive to raise prices.

However, if these recaptured consumers have high costs, a price increase will increase the average cost of the acquired product. The merged firm internalizes the net of this harm due to greater costs and the benefit of the recaptured revenue. If recaptured consumers are especially high cost, the merger can create an incentive to lower the price of one or both of the merged products. I term this sorting externality as Selection Pricing Pressure (SPP).

To see how a merger could reduce prices, consider a stylized example with two symmetrically differentiated insurance plans—*A* and *B*. The two products are identical in their insurance coverage, but differentiated in non-risk related qualities such as brand names, physician locations, billing practices, etc. Consumers are heterogeneous in their costs to insure and choose between *A*, *B*, and the outside option of no insurance. In a duopoly, the two firms set a markup over the average cost of the marginal consumers for each product, which include both high- and low-cost consumers. Among these marginal consumers, those with high costs are more likely to be indifferent between the two products *A* and *B* and those with low costs are more likely to be indifferent between one of the products and the outside option. Because a monopolist only considers consumers that are marginal to the outside option, the marginal cost that matters to the monopolist is lower which could reduce its optimal price.

The net effect of the merger on consumer surplus and total welfare depends on both the effect of the cost externality between firms and the greater markup that the merged firm is able to charge. In simulations of a stylized model similar to the one described above, I show that mergers are most beneficial to consumers when the merging firms are initially smaller, the willingness to pay for insurance products is moderately increasing in consumer costs, and the distribution of consumer costs in the population is more skewed.

To determine the empirical relevance of this mechanism, I estimate the model using data on household health insurance choices in the non-group health insurance market made through a private marketplace in 2015 (Ryan et al. (2021)). The setting is an important and policy-relevant market to study questions of selection and competition. Adverse selection and its consequences are a first order concern that motivated many elements of the Affordable Care Act (ACA). The health reform law targets symptoms of selection in the non-group health insurance market that have been identified in the literature (Obama (2009), Cutler and Zeckhauser (2000), Van de ven and Ellis (2000), Gruber (2008)). Competition is also a focus of policy-makers. Local insurance markets vary widely in their market concentration. The largest firm has a market share of over 85% in five states and less than 33% in another five states. More importantly, managed competition in insurance markets is a common tool to provide health insurance in many market segments in the U.S. and around the world

While the data lack consumer-level detail on health expenditures, I combine several data sets on the distribution of costs among consumers and across firms to identify the key selection parameters governing the correlation between demand and cost. I supplement insurance choice data with moments on the consumer risk scores—output from the Health and Human Services Hierarchical Condition Categories (HHS-HCC) risk prediction model. In particular, I include the average HHS-HCC risk score across product categories and the relative risk score of beneficiaries across firms.[1] I combine these estimates with data on average firm costs and moments on the distribution of costs and risk in the Medical Expenditure Panel Survey to capture how medical risk is related to cost of providing insurance.

To evaluate when mergers may be beneficial to consumers and social welfare, I simulate every potential horizontal merger across all 107 local markets in the data. In the baseline policy scenario, 15% of 1186 merger-market combinations lead to greater social welfare and 13% lead to greater consumer surplus. Even among the largest mergers (in terms of pre-merger market share), about 1-in-20 lead to greater consumer surplus in the baseline policy scenario.

The markets where mergers are most likely to be beneficial are those where the welfare cost of inefficient sorting is greatest. For markets where the welfare cost of inefficient sorting exceeds \$10 per person per month (9% of markets), 43% of mergers lead to greater consumer surplus. And in markets where the cost exceeds \$5 per person per month (29% of markets), 28% of mergers lead to greater consumer surplus.

From the perspective of antitrust enforcement, it is useful to have a measure that can reliably predict when a merger will reduce consumer surplus. In standard markets without selection, a common index for consumer harm is an upward pricing pressure (UPP) measure net of cost efficiencies (Farrell and Shapiro (2010)). In the presence of adverse selection, the standard UPP

[1] The HHS-HCC risk prediction model is used to administer the risk adjustment transfer system in the non-group market. A similar risk adjustment system exists for Medicare (CMS-HCC), which has been used in a similar demand specification (Aizawa and Kim (2018), So (2019)).

measure should be compared to the additional SPP incentive. However, SPP is challenging to compute without a full model of inter-product selection. I show that antitrust agencies can still use the standard (and easier to compute) UPP with a higher threshold for harm. Moreover, the empirical distribution of SPP between products provides a reasonable upper-bound on the magnitude of negative SPP in the non-group health insurance market. The presence of adverse selection is similar to other sources of marginal cost efficiencies that result from a merger. Antitrust agencies can combine qualitative and quantitative evidence of selection with other evidence of potential efficiencies when determining the threshold for UPP that would lead to a harmful merger.

Relation to the Literature

This paper makes two main contributions. First, I provide intuition for inefficient sorting in an empirically tractable model of differentiated product markets with adverse selection. I build on a theoretical literature on contract design in markets with adverse selection that documents the ways in which private firms deviate from the socially optimal (e.g., Akerlof (1970), Rothschild and Stiglitz (1976), Veiga and Weyl (2016), Lester et al. (2019), Chade et al. (2022)) and an empirical literature measuring the effects of these deviations in health insurance markets (e.g., Einav et al. (2010), Lustig (2010), Handel et al. (2015), Layton (2017)). In addition to empirical tractability, this paper extends theoretical results of the literature to a setting where the product characteristics are fixed but firms compete by setting the prices of a menu of products (Chade et al. (2022)).

While U.S. health insurance markets are highly concentrated, there has been less focus in the literature on the effects of market power on adverse selection and policy design. Some recent theoretical work has shown that welfare in markets with adverse selection may be U-shaped in the degree of competition (Mahoney and Weyl (2017), Veiga and Weyl (2016), Lester et al. (2019)). And some empirical work investigates the effects of adverse selection in the presence of imperfect competition (Lustig (2010), Kong et al. (2023)). The potential for welfare benefits that arise from increased concentration highlights the importance of an empirically tractable model that can capture this possibility. This paper presents such a model and allows for flexibility in between-firm selection, the key determinant of whether a particular merger will improve welfare. This paper also builds on Geruso et al. (2018) and Saltzman (2021)—which evaluate the relationship between intensive and extensive margin selection—by introducing the relationship between these welfare costs and market power.

Second, I build on a literature that uses structural models of differentiated products to analyze the welfare impacts of policies addressing adverse selection and market concentration in health insurance markets. This draws from a large literature on estimating the demand for insurance (Gruber and Poterba (1994), Town and Liu (2003),Marquis et al. (2004),Handel and Kolstad (2015), Handel et al. (2019), Geruso (2017), DeLeire et al. (2017), Frean et al. (2017), Drake (2019), Ryan et al. (2021)). There is a growing literature on evaluating policies in regulated health insurance markets with a model of imperfect insurance competition (Miller et al. (2019), Jaffe and Shepard (2020), Shepard (2016), Tebaldi (2023), Ericson and Starc (2015), Starc (2014), Saltzman (2021)), and a related literature that studies health insurance firms' specific mechanisms and incentives to engage in risk selection (Cao and McGuire (2003), Brown et al. (2014), Newhouse et al. (2015), Newhouse et al. (2013), Aizawa and Kim (2018), Decarolis and Guglielmo (2017), Geruso et al. (2019)).

There is a substantial empirical literature on the effects of competition in insurance markets (Cutler and Reber (1998), Town (2001), Dafny et al. (2012)). Much of the recent work in this area is motivated by the two-sided nature of the market—insurance firms with market power may be able to raise markups but also lower costs through hospital bargaining (Capps et al. (2003), Gowrisankaran et al. (2015), Ho and Lee (2017), Chorniy et al. (2020)). Empirical work argue that concentration typically leads to higher prices (Dafny et al. (2015), Abraham et al. (2017), Chorniy et al. (2020)). This paper shows that the effects of market power may be uneven across different product offerings. In particular, the effect of concentration on the most comprehensive plan offerings may be small and possibly negative, even without accounting for potential cost efficiencies in payments to providers.

2 Model

2.1 Environment

Firms, indexed by $f \in F$, each sell a fixed set of differentiated insurance products, indexed by $j \in J^f \subset J$. Firms compete by setting prices, p_j. For the full vector of prices, I will write $\boldsymbol{p} = \{p_j\}_{j \in J}$.

There is a continuum of households, indexed by i. Households make a discrete choice among the set of insurance products, selecting the product that provides the greatest indirect utility. I will write the probability that household i chooses product j, given $\boldsymbol{p}$, as $S_{ij}(\boldsymbol{p})$. For ease of notation, I will drop the i subscript to denote the aggregate functions, e.g. $S_j = \int_i S_{ij} di$.

Each consumer i costs c_{ij} to insure with a product j. Selection arises from the relationship between demand, S_{ij}, and cost, c_{ij} (Akerlof (1970), Cardon and Hendel (2001), Azevedo and Gottlieb (2017)). The average cost of a particular product (AC_j) depends on how consumers are distributed across the products.

$$AC_j(\boldsymbol{p}) = \frac{1}{S_j(\boldsymbol{p})} \int_i S_{ij}(\boldsymbol{p}) c_{ij} di \tag{1}$$

Firms and Equilibrium

A firm, f, competing in a particular market has a profit function defined as,

$$\Pi^f(\boldsymbol{p}) = \sum_{j \in J^f} S_j(\boldsymbol{p})(p_j - AC_j(\boldsymbol{p})) \tag{2}$$

The equilibrium vector of prices $\boldsymbol{p}^*$ solves the Nash-Bertrand competitive equilibrium between the firms such that for every j,

$$p_j^* \in \arg\max_{p_j} \Pi^{f(j)}(\{p_j, p_{-j}\}).$$

2.2 Welfare Cost of Sorting

In this section, I propose a definition for the welfare cost of inefficient sorting. The key objective is to provide a useful summary of a distortion that can potentially be alleviated by a merger. To do so, I construct a constrained efficient benchmark that holds fixed two other distortions. First, I hold fixed the information friction that creates adverse selection (Akerlof (1970), Rothschild and Stiglitz (1976)), i.e. the allocation must be implemented using product-level prices rather than consumer-level prices that could take into account consumer costs. Second, product differentiation confers market power to the firms, and I hold fixed the total producer surplus that they earn. For detailed derivations of each of the conditions presented in this section, see Online Appendix Section B.3.

Let $SW(\cdot)$ be total utilitarian social welfare as a function of product-level prices. The welfare function is equal to the sum of consumer surplus and producer profits.[2] The prices that produce the greatest welfare are equal to the average cost of marginal consumers with respect to p_j across all products in the market, which I will write as MC_j^W.

$$\begin{aligned} SW(\boldsymbol{p}) &= \int_i CS_i(\boldsymbol{p})di + \sum_{k \in J} S_k(p_k - AC_k) \\ p_j^W &= \int_i \sum_k \frac{\partial S_{ik}}{\partial p_j} c_{ij} di \Big/ \int_i \frac{\partial S_{ij}}{\partial p_j} di = MC_j^W(\boldsymbol{p}^W) \end{aligned} \tag{3}$$

Consider the problem of a constrained social planner that chooses product-level prices to maximize consumer surplus subject to a promise of the same total profit $\overline{\Pi}^*$ earned by the insurance firms in equilibrium.

$$\begin{gathered} \max_{\{p_j\}_{j \in J}} \int_i CS_i(\boldsymbol{p})di \\ \text{such that } \sum_{k \in J} S_k(p_k - AC_k) \geq \overline{\Pi}^* \end{gathered} \tag{4}$$

[2] The results of this section do not depend on the specifics of a demand or consumer surplus specification, only that $\partial CS_i(\boldsymbol{p})/\partial p_j = -S_{ij}(\boldsymbol{p})$, which holds under much less restrictive assumptions on demand (Small and Rosen (1981)).

The constrained efficient price of product j, conditional on a given level of profit, is given by

$$p_j^{CE} + \frac{\lambda - 1}{\lambda} \frac{S_j}{\frac{\partial S_j}{\partial p_j}} = MC_j^W(\boldsymbol{p}^{CE}) \tag{5}$$

where λ is equivalent to a Pareto welfare weight on profit.

I define the welfare cost of inefficient sorting as the difference in welfare between the constrained efficient optimum and the competitive equilibrium, $SW(\boldsymbol{p}^{CE}) - SW(\boldsymbol{p}^*)$. This welfare cost is non-negative by construction.

The cost is closely linked to market structure. If there is only a single firm, the cost is zero. The monopolist internalizes all of the same between-product sorting as the planner. In the constrained planner problem, λ approaches infinity as $\overline{\Pi}^*$ approaches the monopolist's maximum level of profit, and Equation (5) converges to the monopolist's first order condition. This is consistent with the result from Veiga and Weyl (2016), who show that a monopolist has an optimal sorting incentive when choosing the quality of a single product offering.[3]

In a similar manner, the welfare cost of sorting can be reduced via a merger because the newly merger firm internalizes consumer sorting across a broader set of products. However, because the newly merged firm also has greater market power and greater profits, the balance of the effect of a merger on consumer surplus and total welfare is ambiguous. The next section studies this in detail.

2.3 Effect of a Merger

A merger between two firms will be modeled as a new entity that jointly maximizes the profit over the existing products offered by the two firms pre-merger.[4] Consider a potential merger between two single-product firms which own the products j and k. The optimal post-merger price for product j is given by the following first order condition:

[3] In Veiga and Weyl (2016), the monopolist sorts consumers optimally using endogenous product quality holding the quantity sold as fixed. In this context, the monopolist sorts consumers optimally using a menu of prices holding the industry profit fixed. Total quantity sold is related to equilibrium total profit, but its not held fixed in this comparison.

[4] I do not directly allow product-level entry or exit. Exit is possible in this model, as firms can set arbitrarily high prices, but the introduction of new products is not. I argue in Section 4.4.1 that this leads the positive consumer surplus results to be conservative.

$$p_j = \underbrace{-\frac{S_j}{\frac{\partial S_j}{\partial p_j}}\left(1-\frac{\partial AC_j}{\partial p_j}\right)+AC_j}_{\text{Pre-Merger First Order Condition}} + \text{GePP}_{jk} \tag{6}$$

$$\text{GePP}_{jk} = \frac{\frac{\partial S_k}{\partial p_j}}{-\frac{\partial S_j}{\partial p_j}}(p_k - AC_k) + \underbrace{\frac{S_k}{\frac{\partial S_j}{\partial p_j}}\frac{\partial AC_k}{\partial p_j}}_{\text{SPP}_{jk}} \tag{7}$$

The final term in Equation (6) is the Generalized Pricing Pressure (GePP) of the merger on product j. GePP is defined as the difference between the pre-merger and post-merger first order conditions for a particular product's price, both normalized to be quasi-linear in marginal cost (Jaffe and Weyl (2013)). For a complete derivation of these equations, see Online Appendix Section B.2.

In this setting, GePP captures two externalities that a price increase in product j exerts on product k, captured by the two terms in Equation (7). A price increase in product j exerts a positive externality on product k by diverting profitable consumers. As captured by the first term in Equation (7), a fraction of consumers lost due to the price increase ($\frac{\partial S_k}{\partial p_j}/\frac{\partial S_j}{\partial p_j}$) will switch to product k and generate profit, $p_k - AC_k$. A newly merged firm now internalizes this positive externality, increasing the benefit from raising prices. This is the standard "upward pricing pressure" that results from a merger (Farrell and Shapiro (2010)).

In the presence of selection, there is a sorting externality to consider: the effect of the diverted consumers on the rival product's average cost. An increase in the price of j diverts consumers to k, but those consumers may also change the average cost of k, as captured by the second term in Equation (7). If the externality is positive—i.e., lowers the cost of the rival product—it will lead to larger price effects from the merger than in the absence of selection. However, if the externality is negative—i.e., raises the cost of the rival product— it reduces the price effects of the merger. If the sorting externality is negative and large enough to outweigh the recapture of diverted consumers (first term), the merger creates an incentive to reduce price. For ease of reference and in keeping with the literature on merger evaluation, I will refer to this term as Selection Pricing Pressure (SPP).

2.3.1 When is Selection Pricing Pressure Negative?

SPP is negative whenever the price of product j raises cost of rival product k. Let MC_{kj} be the cost to product k of the consumers indifferent between products j and k. If MC_{kj} is greater than the average cost of product k, then $\frac{\partial AC_k}{\partial p_j} > 0$ and SPP$_{jk}$ is negative. Otherwise, $\frac{\partial AC_k}{\partial p_j} \leq 0$ and SPP$_{jk}$ is positive. With continuously differentiable demand functions, S_{ij}, this is a straightforward derivation from the definition of average cost.

$$\begin{aligned}\frac{\partial AC_k}{\partial p_j} &= \frac{\frac{\partial S_k}{\partial p_j}}{S_k}\left(MC_{kj} - AC_k\right) \\ MC_{kj} &= \frac{1}{\frac{\partial S_k}{\partial p_j}} \int_i \frac{\partial S_{ik}}{\partial p_j} c_{ik} di\end{aligned} \tag{8}$$

Because the sign of SPP depends on the specifics of the distributions of consumer demand and costs for differentiated products, it is challenging to make a general statement about when it is positive or negative. The following two examples show intuition for market conditions that lead to positive or negative sorting externalities between products.

An Example with Vertical Differentiation

Example 2.1. *Consider a market with two insurance products: a generous insurance option (H), a bare-bones insurance option (L), and an outside option. Each product is offered by a separate firm.*

Consumers costs are uniformly distribution, $c_i \sim U(0,1)$. Consumers are willing to pay βc_i for insurance plan L, and consumers are willing to pay $\beta \nu c_i$ for insurance plan H, where $\nu > 1$.

Consumers' indirect utility is given as follows:

$$u_{iH} = \beta \nu c_i - p_H$$
$$u_{iL} = \beta c_i - p_L$$
$$u_{i0} = 0$$

Consumers are more costly to insure in the H plan. The cost of consumer i is c_i in plan L and νc_i in plan H. Profit generated by each product is as follows:

$$\pi_H = \int S_{iH}(p_H - \nu c_i)di$$
$$\pi_L = \int S_{iL}(p_L - c_i)di$$

Proposition 2.2. *In any equilibrium of the market described of Example 2.1 in which both products have positive quantity sold, $AC_L < MC_{LH} < MC_{HL} < AC_H$. Therefore, $\frac{\partial AC_L}{\partial p_H} > 0$ and $\frac{\partial AC_H}{\partial p_L} < 0$. Selection pricing pressure (SPP_{HL}) from product L to H is positive, and SPP_{LH} is negative.*

Proof. See Online Appendix Section B.4. □

In this example, the consumers that are indifferent between products H and L are less costly than all other consumers that select H and more costly than all other consumers that select L. This model is similar to Azevedo and Gottlieb (2017), with consumer preferences differentiated only along their willingness to pay for insurance. As in Azevedo and Gottlieb (2017), the heterogeneity in those preferences is related to consumer cost variation, which has micro-foundation in models of insurance demand (Rothschild and Stiglitz (1976), Cardon and Hendel (2001)).

If the price of product H goes up, consumers that will shift into product L are higher cost than the average L consumer, thereby increasing its average cost. The opposite is true for a price increase in product L. For products in a vertical relationship like this, the more generous products (H) that attract high-cost consumers exert negative sorting externalities on the less generous products (L), and the less generous products exert positive sorting externalities on the more generous products.

This example highlights an intuitive mechanism through which a merger improves consumer sorting: reducing the spread between the price of product L and H. When product H and L are operated by separate firms, product H is adversely selected relative to product L, driving up the relative price and driving down the relative quantity. If these products are operated by the same firm, it internalizes these sorting externalities between the products and as a result, sets a smaller and more efficient price difference between the products.

An Example with Symmetric Differentiation

Example 2.3. *Consider a market with two insurance products, A and B, and an outside option. Each product is offered by a separate firm. The two products are identical in their insurance coverage, but differentiated in non-risk related qualities such as brand names, where the covered physicians are located, billing practices, etc.*

Consumers costs are uniformly distribution, $c_i \sim U(0,1)$*. More costly consumers are more willing to pay for both insurance products. The differentiation between the products appears through idiosyncratic preferences,* $(\varepsilon_{iA}, \varepsilon_{iB})$*, which are orthogonal to costs and independently and identically distributed across consumers and products. Consumers select the option that provides the maximum indirect utility, given as follows:*

$$\begin{aligned} u_{iA} &= \beta c_i - p_A + \varepsilon_{iA} \\ u_{iB} &= \beta c_i - p_B + \varepsilon_{iB} \\ u_{i0} &= 0 \end{aligned}$$

The profit generated by each product is as follows:

$$\begin{aligned} \pi_A &= \int S_{iA}(p_A - c_i)di \\ \pi_B &= \int S_{iB}(p_B - c_i)di \end{aligned}$$

Proposition 2.4. *In any equilibrium of the market described of Example 2.3, $MC_{AB} = MC_{BA} > AC_A = AC_b$. Therefore, $\frac{\partial AC_B}{\partial p_A} = \frac{\partial AC_A}{\partial p_B} > 0$. SPP_{AB} and SPP_{BA} are identical and negative.*

Proof. Online Appendix Section B.5 contains proofs for uniform, type 1 extreme value, and normally distributed idiosyncratic preferences. □

This example is a stylized, symmetric version of many discrete choice demand models for insurance that are estimated in the empirical literature (Miller et al. (2019), Jaffe and Shepard (2020), Shepard (2016), Tebaldi (2023), Saltzman (2021)). In this example, SPP is *always* negative. High-cost consumers are disproportionately represented among the group of consumers indifferent between products A and B, relative to the average consumer purchasing either A or B. Among high-cost consumers, most of those that are on the margin of purchasing a product are indifferent between products A and B, rather than indifferent to the outside option. Among low-cost consumers, most of the marginal consumers are indifferent between one of the products and the outside option. Because more costly consumers are more likely to switch to the other product rather than the outside good, a price increase of either product A or product B increases the average cost of the rival product by shifting these higher-cost consumers. This occurs despite the fact that the average cost of all marginal consumers, including those indifferent to the outside option, is lower than the products' average cost.

For another perspective on the intuition underlying negative SPP, note that the two firms set an equilibrium markup over the average cost of the marginal consumers for each product, which include both high- and low-cost consumers. Among these marginal consumers, those with high costs are more likely to be indifferent between the two products A and B and those with low costs are more likely to be indifferent between one of the products and the outside option. If these two firms were to merge to a monopoly, the monopolist will only be concerned with consumers that are marginal to the outside option. Thus, the marginal cost that matters to the monopolist is lower.

The proof relies on the assumption that heterogeneity in demand occurs through the value of insurance rather than price sensitivity. While this is a common assumption in similar frameworks (Azevedo and Gottlieb (2017), Einav et al. (2013)), price sensitivity for health insurance products is often empirically found to be greater among demographics that are typically lower cost, i.e. younger consumers (Tebaldi (2023), Saltzman (2019)).[5] A sufficient negative correlation between price sensitivity and cost can create the opposite result: selection pricing pressure between symmetrically differentiated products that are always positive. In this case, the presence of adverse selection is amplifying the harm from mergers.

Combining the intuition of these two examples provides guidance for what we should expect in these more complicated environments. Approximately speaking, the most generous plans exert the most negative (or smallest positive) sorting externalities on all other products, and the least generous plans exert the weakest negative (or most positive) sorting externalities on all other products. In Section 4, I specify a model that can capture both positive and negative values of SPP between products, both through the importance of vertical differentiation and the potential correlation between price sensitivity and cost. The estimation procedure will identify the direction and magnitude of these externalities from the data.

The presence of negative SPP between merging products is necessary but not sufficient for a merger to improve consumer surplus, which depends on all of the price effects of the merger. In fact, the presence of negative SPP does not guarantee that a merger will reduce the welfare cost of sorting. To see this, consider a symmetric merger among two generous insurance products, similar to Example 2.3, but with a set of bare-bones insurance products also offered in the market by other firms. The selection pricing pressure between the two products is negative, for the same reason as outlined in Example 2.3. However, if the merger leads to higher prices for the merging products, the welfare cost of sorting will be greater after the merger.

As this counterexample shows, deriving clean theoretical predictions about the welfare effects of a merger is challenging as many features of the market—elasticities of demand, the distribution of consumer costs, pre-merger market structure—combine and interact in complex ways. In the next section, I explore this question using simulations.

[5] In the framework of Rothschild and Stiglitz (1976), price sensitivity is increasing in expected health costs. Higher expected medical expenses effectively reduce consumers' wealth and increase the marginal utility of income. But more flexible models of state dependent utility can generate either relationship (Cardon and Hendel (2001)).

2.3.2 When do Mergers Improve Consumer Surplus?

Many discrete-choice demand models in the empirical health insurance literature, including the one that is estimated in Section 4.1, combine the features of these two frameworks. Consumers choose between vertically differentiated products within menus that are offered by horizontally differentiated insurance firms. I simulate mergers in a simple version of this environment—two firms that each offer a menu of two insurance products—in order to illustrate the intuition for the relationship between market primitives and mergers that benefit consumers.

Because the potential for a merger to improve consumer surplus requires negative selection externalities, I focus on an example where this is always the case. I simulate a merger to monopoly, which is also guaranteed to reduce the welfare cost of sorting to zero. However, the intuition that connects model primitives to the potential for beneficial mergers still holds in more complex environments.[6]

An Example Combining Vertical and Horizontal Differentiation

Example 2.5. *Consider a duopoly to monopoly merger between two firms, A and B, that each offer a menu of two products, H and L. Consumer costs are bounded between 0 and 1, and distributed according to a beta distribution, $c \sim Beta(\sigma^c, 1)$. Consumers have idiosyncratic preferences over the products, which capture any heterogeneity in preferences among consumers that is orthogonal to cost.*

Consumers select the option that provides the maximum indirect utility, given as follows:

$$u_{ijH} = \delta_0 + \beta \nu c_i - p_{jH} + \varepsilon_{ijH}$$

$$u_{ijL} = \delta_0 + \beta c_i - p_{jL} + \varepsilon_{ijL}$$

$$u_{i0} = \varepsilon_{i0}$$

[6] For example, Figure 1 looks qualitatively similar in simulations of a merger from 4 single product firms to 2 single product firms offering the L products and one merged firm offering the H products. This merger frequently generates greater welfare costs of sorting, but the intuition for when merger benefit consumers remains the same.

Pre-merger, the firm $j \in \{A, B\}$ seeks to maximize the joint profit of products jH and jL. After the merger, a monopolist maximizes the joint profit of all four products. The profits generated by each of the products are given by

$$\pi_{jH} = \int S_{ijH}(p_{jH} - \nu c_i)di$$
$$\pi_{jL} = \int S_{ijL}(p_{jL} - c_i)di$$

The costs and preferences of the products are identical, with the exception of the term ν, and the menus offered by the two firms are identical with the exception of the ε terms. As such, I focus the simulations on symmetric equilibria with equal prices for the two types of products, (p_L^*, p_H^*).

In total, the model is specified by four parameters. The base utility term δ_0 determines the value of insurance relative to the outside good, and therefore shifts some degree of the firms' pre-merger market power. The preference term β governs the degree to which more costly consumers have a higher willingness to pay. The vertical term, $\nu \geq 1$ determines the degree of vertical differentiation between the two products in the menu.

The beta distribution term, σ^c, determines the skewness of the consumer cost distribution. In the results presented here, I measure skewness as the ratio of the distances of the 10^{th} and 90^{th} to the median:$\frac{P_{90}-P_{50}}{P_{50}-P_{10}}$. A value of one represents a perfectly symmetric distribution, which in this case is uniform. Value greater than one represent a shift of the distribution towards zero with a tail stretching towards one.

In Figure 1, I plot the consumer surplus effect of a merger, varying each one of these parameters while holding the others fixed. Panel (a) varies the pre-merger market share of the firms. Panel (b) varies the β term governing preferences. Panel (c) varies the degree of vertical differentiation ν. In each panel, I show the effect for four levels of skewness in the cost-distribution: 1, 1.5, 3, and 6.5. In panels where the terms are not varying, I fix the pre-merger market share to 20% (panels (b) and (c)), the β parameter to 0.5 (panels (a) and (c)), and the vertical differentiation parameter ν to 0.2 (panels (a) and (b)). For more details on the simulation, see Online Appendix Section B.6.

Mergers are better for consumers when the merging parties are initially small (Figure 1a). The benefit from new market power is not as great when the merging parties are both small, and the firms are more adversely selected when they serve a small section of the market in equilibrium. The large percent increases in consumer surplus are partially driven by smaller pre-merger consumer surplus when the size of the market is small.

Mergers are better for consumers when willingness to pay is significantly related to consumer costs but does not increase 1-for-1 (Figure 1b). When β is very low, adverse selection does not play a role in the equilibrium. When β is high, demand for insurance is strong enough that equilibrium prices are greater than most consumer costs. Since all consumers are profitable, this also reduces the importance of adverse selection.

Finally, mergers are more beneficial to consumers when the distribution of costs is more skewed, i.e. high-cost consumers are less common. This appears across a broad range of the other three parameters. When high-cost consumers are a minority among all consumers, they make up a disproportionately large fraction of consumers marginal between two products relative to their presence among all consumers. However, this relationship has a limit. If the distribution of consumer costs is too skewed, the equilibrium outcome is to price-out the lower-cost consumers, reducing the influence of selection overall. The benefit to consumers does not substantially vary with respect to the degree of vertical differentiation in the product menu (Figure 1c).

Each of these parameters—in particular skewness, willingness to pay, and the firm size—combine to create an equilibrium where the welfare cost of sorting is more or less important. When the cost is high, the potential efficiency gains from the merger are also high. And while the new monopolist will capture some of this as profit, consumers can still stand to benefit.

In Figure 2, I simulate 1000 mergers using randomly drawn values for $(\delta_0, \beta, \sigma^c, \nu)$. See Online Appendix Section B.6 for details on the Monte Carlo simulations.

In panel (a) of Figure 2, the relationship between the welfare cost of sorting and the benefit to consumers is clear. When welfare cost of sorting is low, market power effects dominate and consumers rarely benefit. When the welfare cost of sorting is high, consumers share directly in the efficiency gain from the merger. The stark, near-linear relationship for high values of the sorting cost is due to the fact that the welfare cost of sorting falls to zero in a monopoly. Greater welfare costs of sorting are also associated with smaller merging firms that cannot capture as much of the extra surplus. If there are other firms in the market, a merger between two small firms will not have such a predictable effect on sorting efficiency.

Panel (b) of Figure 2 shows mergers that reduce the price spread between products L and H have greater benefits effect on consumer surplus. As discussed in Example 2.1, a reduction in the price spread is one mechanism through which a merger can improve sorting efficiency. In a model with both vertical and horizontal differentiation, this correlation between the spread reduction and a consumer surplus benefit is not particularly strong.

While mergers frequently generate spread reductions, the price effects between L and H products are still positively correlated. Mergers that benefit consumers are generally those that lower prices overall and increase total quantity sold. A reduction in the price spread does not create a large enough welfare benefit to compensate consumers if prices are higher overall. This relationship also holds in more asymmetric merger scenarios, like the one studied in Example 2.1.

While this section demonstrates that a merger can increase consumer surplus under reasonable market conditions, it remains to be determined whether or not these conditions are empirically relevant. The following sections of this paper are devoted to establishing the extent to which mergers have the potential to increase consumer surplus in the setting of the non-group health insurance market. Ultimately, I will specify and estimate an empirical model that has the same components as the one studied here: multi-product firms setting the prices for a fixed menu of vertically differentiated products.

3 Non-group Market Data

The non-group insurance market is the only source of health insurance for any individuals or households that do not receive an offer for insurance through their employer or a government program. Consumers can purchase insurance by contacting an insurance firm directly, visiting the government-run marketplace, or shopping for insurance through a third-party marketplace. Not all plans are offered on all platforms, and insurance firms may elect to list some products on certain platforms and not on others. However, apart from insurers that do not list on the government marketplace at all, the kinds of plans listed by insurers typically have only small differences across platforms.[7]

Since the implementation of the ACA, all insurance products in this market must fit within one of five categories known as "metal" levels: Catastrophic, Bronze, Silver, Gold, and Platinum, listed in increasing level of generosity. Households (or individuals) may purchase products that are offered in their local rating area for a price that depends on the size and age composition of the household, the household income, and whether or not the members are smokers. Insurance prices are adjusted by an age-rating factor for each member of the household which, in 2015, increases from 0.635 for children under the age of 21 to 3 for a 64 year old. Some states add additional premium increases of up to 50% for household members that smoke.

Households that earn 100% of the federal poverty level (FPL) receive a subsidy that is sufficient for the household to buy the second-lowest price Silver plan in their rating area for roughly 2% of their household income. This subsidy declines non-linearly to 9.5% for households that earn 400% of FPL, and subsidies are zero for households that earn greater.[8] Households that earn less than 250% of FPL also receive additional subsidies to cover reduced cost-sharing.

To address selection between products, the ACA implemented "risk adjustment," a system of risk-based subsidies (taxes) that compensate firms for enrollees with higher (lower) than average expected costs. The government collects claims data throughout the year from every insurance firm in the market to assess the average risk at the plan level using the HHS-HCC risk prediction methodology. This method attributes to each individual a risk score based on age, sex, and a set of diagnosis codes that are organized into hierarchical condition categories. Plans that have lower

[7] Analysis of the Robert Wood Johnson Foundation HIX 2.0 data on plan offerings shows minimal differences between plan offerings on and off the exchange in premiums or deductibles.

[8] In recent years, California has extended subsidies to higher income households as well.

than average levels of risk are taxed and plans that have higher than average levels of risk receive subsidies. The formula that determines the taxes and subsidies is constructed to be budget neutral at the state-level: the total taxes across all firms within a state are mechanically equivalent to the total subsidies.

Risk-based subsidies are a common policy instrument to reduce adverse selection in health insurance markets (McGuire et al. (2011), Van de ven and Ellis (2000), Ellis and McGuire (2007)). The intention is to "eliminate the influence of risk selection on the premiums that plans charge," and see Section F for more detail on how risk adjustment works in a model of imperfect competition (Pope et al. (2014), Kautter et al. (2014a)).

3.1 Choice Data

The data on health insurance purchases come from a third-party private online marketplace. The private marketplace sells plans that are offered both on and off the ACA health insurance exchanges. In 2015, the private marketplace was authorized to sell subsidized health insurance plans in most states. I observe the choices of subsidized and unsubsidized consumers across 48 states. After dropping observations because of missing data or incomplete choice sets, the remaining data includes roughly 75,000 individual and family health insurance choices across 14 states and 107 rating areas.

The data contain information on the age of the consumer, the first three digits of the consumers' zip code, the household's income, the plan purchased by the consumer, and the subsidy received. A single observation in the data represents a household, but I observe only one member's age. I assume that this is the age of the head-of-household, i.e., the oldest member of the household. I assume that every household that contains more than one individual contains two adults of the same age, and any additional persons are children under the age of 21.[9]

[9]The choice data contains information on the premium paid for a subset of the observations. In combination with the base premium of the purchased product, the premium paid can be used to impute household composition. Using the median base premium in the selected firm and metal-level, I construct an imputed household age-rating measure. The correlation between this imputation and the more simple age-rating rule applied to the rest of the sample is 0.90. The results are robust to alternative assumptions about age rating.

The data from the private marketplace are a selected sample of all the consumers facing a particular firm. Using the same data set, Ryan et al. (2021) find that income is a primary determinant of driving selection into the private online marketplace. In order to create a sample of consumers that is representative of the consumer population facing firms in this market, I treat the choice data as a random sample conditional on subsidy eligibility and geographic market. Each observation from the choice data within a particular subsidy eligibility category and market is given an equal weight such that the weights sum to the size of the population as determined by the 2015 American Community Survey (ACS)[10]. The ACS also provides a sample of the uninsured population.

Online Appendix Section C.1 contains descriptive statistics of the data sample, as well as more detail on sample selection, missing data, and constructing the choice sets. For more detailed information on processing the ACS, see Online Appendix Section C.2.

3.2 Cost Data

Ideal data would link consumer health insurance purchases to consumer health information. Because these data are not available, I instead use more aggregated moments on consumer medical risk and cost in both the demand and cost estimations to identify the relationship between marginal cost and demand, the key feature of adverse selection.

The 2015 Medical Expenditure Panel Survey (MEPS) Medical Conditions File (MCF) contains self-reported diagnosis codes, which can be linked to information on household demographics, insurance coverage, and medical expenses in the Full Year Consolidated File. I apply the HHS-HCC risk prediction model coefficients, published by Center for Medicare and Medicaid Services (CMS), to the self-reported diagnoses to compute risk scores. For details on the processing of the MEPS data, see Online Appendix Section C.3.

To identify the relationship between risk scores and demand, I use aggregate moments on the risk distribution among market enrollees. I target 5 moments that CMS publishes in annual reports

[10]The weights do not significantly alter the price elasticity and risk preference estimates from demand estimation. They are important for how well the model predicts untargeted moments like aggregate insurance rates and the firm first-order conditions.

on the results of the risk adjustment transfer program: the national average risk score for enrolled beneficiaries and the average risk score of consumers in Bronze, Silver, Gold, and Platinum plans. CMS only began to publish risk scores by metal-level in 2017. In order to make it comparable to my data, I use the average of on- and off-exchange market segments, and scale the risk scores by the ratio of the 2015 national average risk score to the 2017 national average risk score.

I also target the risk distribution across firms using risk adjustment transfers. These transfers are related to the relative risk of the enrolled beneficiaries for each firm within a state (See Section F). In Online Appendix Section C.4, I detail how these moments are constructed from Medical Loss Ratio data submitted to the government.

I compliment these moments on the risk distribution with similar moments on costs. I match moments on the relative costs of individuals by age and risk, which come from MEPS (Online Appendix Section C.3). Additionally, I match moments on the average cost of insurance product categories and the average costs of each firm. Product category level data come from rate filings to state insurance regulators (Online Appendix Section C.5), and the average firm-level costs come from the Medical Loss Ratio data (Online Appendix Section C.4).

4 Empirical Model

In this section, I estimate an empirical model of supply and demand in the market for non-group health insurance. Using the estimated parameters, I simulate new equilibria for potential mergers between existing firms. I show that SPP is typically negative and that a non-negligible portion of potential mergers are beneficial to consumers.

4.1 Demand Specification

Households have characteristics $(a_i, y_i, Z_i, r_i^{HCC})$, where a is an average age-rating of all household members, y is household income, Z is a vector of demographic indicator variables that include three age buckets, whether or not the household includes only one person, and whether or not the household is subsidy eligible.[11] Households have an unobserved medical risk score, r^{HCC}.

Households make a discrete choice among the set of insurance products that are available in their market, m, or being uninsured. Products are characterized by their price p, observable characteristics, X, and unobservable characteristics, ξ. Households are heterogeneous in their preferences for the price, α_i, and insurance characteristics, β_i. Households have additive idiosyncratic preferences over products ε, which I assume are independently and identically distributed by type I extreme value. The indirect utility that household i receives from purchasing a product j is given by

$$u_{ijm} = \gamma' Z_i + \alpha_i (a_i p_{jm} - B(y_i)) + \beta_i' X_{jm} + \xi_{jm} + \varepsilon_{ijm}$$
$$u_{i0m} = \alpha_i M(y_i) + \varepsilon_{i0m}$$

where $B(y)$ is a function that maps income to subsidies. The indirect utility from being uninsured is u_{i0m}, where $M(y)$ maps income to the penalty for choosing not to buy health insurance. Observed characteristics X_{jm} include the actuarial rating of the plan and a firm fixed effect.

The preference over observed characteristics, β_i, depends on a household risk score, r^{HCC}, and the preference for the utility-value of money, α_i, depends on household demographics.

$$\alpha_i = \alpha_z' Z_i$$
$$\beta_i^k = \beta_0^k + \beta_r^k r_i^{HCC}$$

Risk scores are treated as an unobserved household characteristic, and are distributed according to a distribution that can depend on household demographics, Z_i.

$$r_i^{HCC} \sim G(Z_i)$$

Because of the distributional assumption on ε, the probability that an individual will purchase

[11] I use the demographics of the head-of-household as the representative demographics for the household.

a particular product, S_{ij}, is given by the standard multinominal logit formula.

4.2 Cost Specification

Since I do not observe micro-level data on consumer costs, I follow Tebaldi (2023) and specify a parametric function for costs that can be estimated through Method of Simulated Moments (MSM). As in Tebaldi (2023), I use a log-linear cost function.[12] I specify the costs as

$$\log(c_{ijm}) = \phi_f + \phi_{AV} AV_{jm} + \phi_{age} Age_i + \phi_r r_i^{HCC}$$

where ψ_f is a firm-state specific fixed effect, AV_{jm} is the actuarial value of the product, Age_i is the average age of the household, and r_i^{HCC} is the risk score of household.

The only mechanisms through which cost and preferences are correlated are through age and risk scores. If this assumption is violated and the remaining correlation is consistent with adverse selection, then the coefficient on actuarial value will be biased upward and attribute some portion of the selection differences to cost differences across products. In the context of this study, this attribution leads to conservative conclusions about the implications of adverse selection and underestimate the importance of SPP. For a more detailed discussion, see Online Appendix Section E.1.

In order to match the data, I also model the reinsurance program that was in place in 2015. For details on this implementation, see Online Appendix Section E.

4.3 Estimation Results

The estimation proceeds in three steps. First, I calibrate the distribution of consumer risk scores using the MEPS-MCF data and the HHS-HCC risk prediction model. For details on the specification, estimation, and fit, see Online Appendix Section D.1. Next, I follow Grieco et al. (2021) to combine a micro-data log-likelihood function with product-level moments on the average risk score

[12] A key difference between this specification and that in Tebaldi (2023) is that costs depend on consumer risk scores rather than directly on willingness to pay for insurance.

across different products and firms. I use a control function approach to address potential price endogeneity, following Tebaldi (2023). For details on the identification concerns and estimation methodology, see Online Appendix Section D.2.

Finally, I estimate the parameters of the cost function using Simulated Method of Moments to target moments on average firm costs and health care expenditures by age and risk score. This method does not require the assumption that firms are playing optimal strategies according to the specification of the model. For details on the moment construction, estimation, and model fit, see Online Appendix Section E.

Table 1 presents the results from the demand estimation. The specification used throughout the rest of the paper is specification 2. I estimate three specifications to demonstrate the sensitivity of parameter estimates to the level of fixed effect that controls for cross-product heterogeneity. The most detailed specification (specification 3) includes fixed effects for every firm-market-category, where category indicates whether the insurance plan is a high coverage plan that covers more than 80% actuarial value. It is a challenge to use this specification in counterfactual simulations, because not all firm-market-category combinations are chosen in the data. Instead, I use specification 2, which has very similar parameter estimates.

The median consumer willingness to pay for a 10% increase in the actuarial value of an insurance plan is \$134 per month. This actuarial increase is roughly equivalent to switching from a Bronze plan to a Silver plan (or Silver to Gold). The median price difference to consumers between Bronze and Silver plans is about \$52 per month. There is substantial variation in willingness to pay. The 10th percentile of willingness to pay is \$84.1 per month, and the 90th percentile is \$346 per month.

The average own-price elasticity of consumers is -3.96, and the semi-elasticity of purchasing any insurance at all is -0.03, i.e. a \$10 increase in monthly price of every insurance product will decrease insurance enrollment by 3%. These elasticities are similar to other estimates in the literature (Tebaldi (2023), Saltzman (2019)).

Consistent with other literature, I find that price sensitivities are greater among younger consumers which also have lower average medical risk. The flexibility in price sensitivity across demographics allows for the possibility that SPP between symmetrically differentiated products is positive, which occurs when lower cost consumers are also substantially more price sensitive. While there are no truly symmetric products in the data, these parameters will affect the estimated distribution of SPP. I describe the estimated distribution in more detail in Section 4.4.2.

Table 2 displays the results of the cost estimation. The table presents results for two demand specifications ((1) and (2)) used to simulate the moments targeted by the cost estimation. [13] The estimation implies a substantial amount of variation in consumer costs. The 10^{th} percentile of consumer costs is \$60.7 per month, the median consumer costs is \$161 per month, and the 90^{th} percentile of costs is \$455 per month. The skewness as measured in Section 2.3.2 is 2.9. However, costs increase very quickly in the tail. The 99^{th} percentile of consumers costs is \$1,623 per month.

The standard mechanism of adverse selection is present. The 50 percent most elastic consumers with respect to purchasing any insurance are 15% less costly than the least elastic consumers. The consumers that are infra-marginal in the insurance purchase decision are more expensive than the consumers that are more marginal to leaving the insurance market (Einav et al. (2010)).

As in the simulations presented in Section 2.3.2, willingness to pay is increasing in consumers costs, but at less than a one-for-one rate. A \$1 increase in consumer cost is associated with a \$0.38 increase in consumer willingness to pay. This is also in line with other empirical work on the willingness to pay for insurance (Finkelstein et al. (2019)).

There is substantial heterogeneity in preferences across firms that is correlated with cost. A one standard deviation increase in a consumers risk score leads to a 30% increase in cost and up to a \$37 dollar increase in the willingness to pay to switch from the lowest quality firm to the highest quality firm, as ranked by risk preferences (β_r). As shown in Section 2.3, close substitution among high-cost consumers is a key feature that leads to negative SPP between products.

[13] The specification 3 is not included. The average firm-level costs cannot be simulated in the same way because not all firm-market-category fixed effects present in the choice sets are chosen in the estimation data.

4.4 The Welfare Effects of Mergers

In this section, I simulate every potential horizontal merger between firms that compete in at least one local market and compute the effects on total welfare and consumer surplus. Because competitive equilibrium is not assumed in the estimation of demand and supply, I first re-solve the baseline equilibrium. Next, I solve the post-merger equilibrium for each potential merger. In the data, there are 243 potential bilateral, horizontal mergers between competing firms, each of which affect an average of 4.8 local markets.[14]

A merger between two firms is characterized as jointly maximizing the profit over a set of products that is fixed in both the pre-merger and post-merger equilibrium. In other words, the merger only affects the ownership of existing products, while all other qualities and characteristics are held fixed. The model does not allow product entry or exit due to a merger. I abstract from this mechanism for two reasons. First, a model with flexible product offerings is computationally difficult to solve and would require stronger assumptions on equilibrium selection. Second, such a model is conceptually challenging to characterize, especially in a setting where the degree of unobservable product differentiation may be strategically chosen for new products (Rysman and Ackerberg (2005)).[15]

In the following analysis, I make two assumptions about subsidies. First, I ignore any changes in government spending in the welfare computation. At the estimated parameters, the average consumer surplus generated from a dollar of additional government spending is less than a dollar, a result consistent with other work on government sponsored health insurance (Finkelstein et al. (2019)). To avoid comparing to a benchmark where the optimal outcome is zero government spending and very little insurance enrollment, I treat the government's subsidy policy as fixed and

[14] In markets with adverse selection, there may be multiple equilibria. To verify that the simulated effects of a merger are due to the changing market structure rather than equilibrium selection, I resolve for the pre-merger equilibrium from the post-merger prices. In every case, the solution returns to the original pre-merger equilibrium.

[15] The results of this section may be exaggerated if merged firms remove products from the market, and this leads to a reduction in consumer surplus. Due to regulation, insurance firms typically "merge" rather than remove products, preserving much of the important unobserved qualities such as the provider network. And when selection is important and gains price discrimination for the merged firm are large, removing products is less likely. In Medicare Part D, a market with many of the same firms, Chorniy et al. (2020) find that plan consolidation does not significantly increase after a merger.

assume the planner cares only about consumer and producer surplus.

Second, I assume that the price-linked subsidies are treated as vouchers by both the consumers and firms. Subsidies are tied to an order statistic of the equilibrium prices in each market: the second-lowest price silver plan. In merger simulations, I allow the equilibrium subsidy to adjust according to the rule, but firms do not behave strategically with respect to the subsidy. Because the price-linked subsidies introduce many new mechanisms through which mergers affect prices and welfare that are unrelated to adverse selection, I relegate discussion of these policies to Online Appendix Section G.

Finally, I include the ACA risk adjustment policy in the equilibrium simulations. For details on how I model this policy and how the policy relates to the distortions caused by adverse selection, see Online Appendix Section F.

4.4.1 Mergers Can Lead to Greater Consumer Surplus

Local markets for non-group insurance are quite concentrated. The median market has a Herfindahl–Hirschman Index (HHI) of 4,500 in the baseline equilibrium. As a result, the sorting problems due to adverse selection are significant but not exceptionally large. The inter-quartile range of the welfare cost of sorting across markets is \$1.7 to \$5.1 per person per month. For reference, the inter-quartile range of average consumer surplus is \$22.3 to \$86.6. In the median market, the welfare cost of sorting is about 5% of consumer surplus. The welfare costs of sorting are also mitigated by the risk adjustment transfer system implemented by the ACA. But even without risk adjustment transfers, the mean sorting cost would be only \$7.0 per person per month.

Despite the high levels of initial concentration and low welfare cost of sorting, many mergers are predicted improve *both* consumer and social welfare. The fraction of mergers that improve welfare is displayed in Table 3, broken down by the size of the merger (measured by the change in HHI predicted by pre-merger market shares) and the welfare cost of sorting in the pre-merger equilibrium. I follow the 2010 Merger Guidelines and classify the size of mergers into three categories: those unlikely to be of concern (change in HHI of less than 100), those that are potentially concerning (change in HHI of between 100 and 200), and those that are presumed to be harmful to consumers (change in HHI of greater than 200).[16]

There are three important facts to learn from Table 3. First, the majority of mergers reduce the welfare cost of sorting, and mergers are more likely to confer this benefit when they occur in markets with larger per-merger welfare costs of sorting. Across all potential mergers in the data, 71% lead to a reduction in the welfare cost of sorting. These mergers decrease the welfare cost of sorting by an average of \$0.32 per person-month, 7% of the pre-merger welfare cost of sorting. However, as discussed in Section 2.3, a merger can exacerbate the problems of adverse selection as well. In 29% of potential mergers, the welfare cost of sorting increases by an average of \$0.55 per-person per-month. And since these mergers typically occur in markets where the welfare cost of sorting is small, this amounts to a 28% increase over the pre-merger welfare cost.

Second, across all dimensions, *some* mergers are beneficial to consumers. Even in the current policy environment where policies are in place to address adverse selection and among large mergers likely to draw intense antitrust scrutiny, consumers are better off in 1 out of 20 mergers. While this is a small fraction, it demonstrates the importance of heterogeneity in merger effects. Just as heterogeneity in consumer substitution patterns can generate heterogeneous merger effects in apparently similar mergers, so too can heterogeneity in consumer selection patterns.

Third, the mergers that lead to greater consumer and social welfare are typically in markets with larger pre-merger welfare costs of sorting, as shown in the third panel of Table 3. Only in markets with a welfare cost of sorting per-person per-month of at least \$5 do an economically significant fraction of mergers benefit consumers. Where the welfare cost of sorting is already small, there is little room for a merger to benefit consumers, and the dominant effect is greater markups.

One mechanism through which a merger might reduce the welfare cost of sorting and potentially improve consumer surplus is by reducing the spread between generous insurance products (Gold and Platinum plans) and the less generous options (Bronze and Silver plans). Figure 3

[16]These thresholds are applied to markets that are already concentrated and are guides for scrutiny rather than hard rules.

demonstrates how this narrowing occurs in the price spread: typically through significant price increases in Bronze and Silver plans and lower or negative price changes in Gold and Silver plans.

However, while these differential price effects are important for increasing the level of insurance that people purchase, consumer surplus is more closely tied to the total quantity of insurance sold. Figure 4 demonstrates that the effect of the merger on the fraction of consumers that purchase any insurance is a driving predictor of whether consumers benefit from the merger. In contrast, the change in the average actuarial value of the purchased insurance product (conditional on purchasing any product) is relatively uncorrelated with the change in consumer surplus.

Figure 5 shows the distribution of merger effects on consumer surplus in each of the three categories of merger size designated in the Merger Guidelines, plotted against the welfare cost of sorting pre-merger. Each dot represents a merger-market.

Among the smallest mergers with a change in HHI of less than 100, it is rare for a merger to lead to significant consumer harm, and occasionally a merger leads to substantial consumer benefits. It is unsurprising that these merger-markets do not generate much consumer harm. However, the presence of large benefits to consumers means that markets with less overlap in market share should still be considered as a source of possible benefits from a merger.

Even among larger mergers that would be presumed harmful under current guidelines, there exist mergers that generate substantial benefits to consumers. And regardless of the change in HHI, mergers in markets with a pre-merger sorting cost greater than $10 per person per month are frequently beneficial to consumers with economically significant magnitudes.

4.4.2 Screening For Mergers in the Presence of Adverse Selection

As derived in Section 2.3, GePP captures the incentive of a merger in the presence of adverse selection, analogous to upward pricing pressure (UPP). Farrell and Shapiro (2010) argue that, while not a perfect predictor of actual price effects, UPP can accurately predict the direction of the effect of a merger on prices.[17] This logic extends relatively easily to the effect on consumer surplus

[17]The 2010 Merger Guidelines also adopt this view: "[t]he Agencies rely more on the value of diverted sales than on the level of the HHI for diagnosing unilateral price effects in markets with differentiated products." Typically, the

(Jaffe and Weyl (2013)). I follow Farrell and Shapiro (2010) and Miller et al. (2017) in normalizing the pricing pressure measure relative to pre-merger prices as a screening index, redefined below.

$$\text{GePP}_{jk} = \underbrace{\frac{-\frac{\partial S_k}{\partial p_j}}{\frac{\partial S_j}{\partial p_j}} \frac{(p_k - AC_k)}{p_j}}_{\text{UPP}_{jk}} + \underbrace{\frac{S_k}{\frac{\partial S_j}{\partial p_j}} \frac{\partial AC_k}{\partial p_j} \frac{1}{p_j}}_{\text{SPP}_{jk}} \tag{9}$$

I propose that SPP can be considered similar to efficiencies in the standard application of UPP. The SPP term enters linearly, similarly to marginal cost efficiencies that are frequently at issue in merger analysis.

In this section, I will compare the fully specified GePP to a naive UPP measure that is computed without taking selection into account, equal to the product of the average diversion ratio and the profit margin, as shown underlined in Equation (9). I consider only mergers that are *presumed* to be harmful due to a change in HHI of greater than 200.[18] While most of these mergers lead to significant harm to consumers, 7.2% of mergers lead to greater consumer surplus than pre-merger, even absent any efficiency gain. The goal is to screen for the harmful mergers in this group without investigating or blocking mergers that benefit consumers.

The full GePP measure is an imperfect prediction of the magnitude of the effects of a merger on consumer surplus, but it is an accurate prediction of the direction of the effect. In Online Appendix Figure A2, I plot the change in consumer surplus relative to the average GePP created by the merger. GePP is a conservative screen in the sense that no mergers with a negative average GePP are harmful. As GePP grows larger, the merger deserves more scrutiny.

SPP is challenging to compute, as it requires a model of consumer-level adverse selection. This may be an unrealistic expectation with the time and data available to antitrust authorities. However, qualitative and aggregate data may be available to determine the sign and potential mag-

UPP measure is compared to claims of potential cost-efficiencies, which I am abstracting from here.

[18] The 2010 Merger Guidelines state: "Mergers resulting in highly concentrated markets that involve an increase in the HHI of more than 200 points will be presumed to be likely to enhance market power."

nitude of SPP. In this setting, most product pairs have relatively small selection pricing pressure. The 5^{th} percentile is -8.3%, and the 10^{th} percentile is -4.5%. Therefore, if the naive UPP measure calculated from the merger is greater than 10%, it is unlikely that the merger would benefit consumers.[19]

Figure 6 plots the distribution of SPP_{jk} for all potential pairs of merging products. While this paper focuses on the potential for mergers to be beneficial to consumers, SPP is positive in many cases, leading to greater pressure to increase prices than may be presumed. Out of all the product-market-mergers studied in the counterfactual, 23.4 percent of products have greater incentives to raise price due to a merger than would be measured by the standard UPP measure. However, most of these harmful mergers are already deemed harmful by the standard UPP measure. For example, only 1% of products (across mergers and markets) with a UPP of less than 5 percent of the pre-merger price have a GePP measure of greater than that threshold.

Figure 7 shows the fraction of mergers that benefit consumers that would be "investigated" under a particular measure (GePP or naive UPP) and screening threshold. On the x-axis are screening thresholds which select for investigation only mergers with an average pricing pressure measure that exceeds that level. For example, the data points at 0.025 represent the percent of mergers that benefit consumers among those mergers with an average GePP or UPP value that exceeds a screening threshold of 0.025.

If the full GePP measure is used to screen mergers, it is unlikely that beneficial mergers will be investigated with any screening threshold that is greater than 0. UPP can also still be an effective screen. Using a threshold of 0.05, very few mergers that are beneficial to consumers will be investigated and potentially blocked. However, it is important to note that the threshold depends on the degree of adverse selection in a particular market. To demonstrate this, Figure 7 also contains estimates from a scenario in which there is no risk adjustment between firms. In this case, UPP is an even less accurate prediction of merger harm.

[19] The naive UPP measure and SPP are not strongly correlated.

5 Conclusion

This paper demonstrates that additional concentration due to a merger in markets with adverse selection has the potential to benefit consumers. In simulations of a stylized model, I show that mergers are most beneficial to consumers when the merging firms are initially smaller, the willingness to pay for insurance products is moderately increasing in consumer costs, and the distribution of consumer costs in the population is more skewed.

In the non-group insurance market, the potential for beneficial mergers is economically relevant. Even in the case of large mergers in concentrated markets, more than 1 out of 20 mergers lead to an improvement in consumer surplus. In markets where the welfare distortion due to sorting is greater than $5 per person, nearly 1 out of 3 mergers improve consumer surplus.

These results provide important insight for policy makers. From the perspective of antitrust enforcement, the degree of adverse selection in a market should be considered when evaluating a merger. The reduction in inefficient sorting due to additional concentration is a kind of merger efficiency that might lead some mergers to provide benefits for consumers.

This paper builds on a large theoretical and empirical literature and is itself only a small additional step towards understanding managed competition in an environment with adverse selection. I show that complex selection mechanisms can be studied rigorously even in settings where contract characteristics are fixed (Chade et al. (2022)). More work remains to fully understand the effect of mergers in this market. For instance, firm entry or exit in the context of mergers (Caradonna et al. (2023)), adverse selection, and their interaction is an important area for future research.

References

Abraham, Jean, Coleman Drake, Jeffrey S Mccullough, · Kosali Simon, B Jean, and Marie Abraham, "What drives insurer participation and premiums in the Federally-Facilitated Marketplace?," *International Journal of Health Economics and Management*, 2017, *17*, 395–412.

Aizawa, Naoki and You Suk Kim, "Advertising and Risk Selection in Health Insurance Markets," *American Economic Review*, mar 2018, *108* (3), 828–867.

Akerlof, George A., "The Market for "Lemons": Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 1970, *84* (3), 488.

Azevedo, Eduardo M. and Daniel Gottlieb, "Perfect Competition in Markets With Adverse Selection," *Econometrica*, jan 2017, *85* (1), 67–105.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, jul 1995, *63* (4), 841.

Brown, Jason, Mark Duggan, Ilyana Kuziemko, and William Woolston, "How does risk selection respond to risk adjustment? New evidence from the Medicare Advantage Program," *American Economic Review*, 2014, *104* (10), 3335–3364.

Cao, Zhun and Thomas G. McGuire, "Service-level selection by HMOs in Medicare," *Journal of Health Economics*, nov 2003, *22* (6), 915–931.

Capps, Cory, David Dranove, and Mark Satterthwaite, "Competition and market power in option demand markets," *RAND Journal of Economics*, 2003, *34* (4), 737–763.

Caradonna, Peter, Nathan Miller, and Gloria Sheu, "Mergers, Entry, and Consumer Welfare," 2023.

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

Chade, Hector, Victoria Marone, Amanda Starc, and Jeroen Swinkels, "Multidimensional Screening and Menu Design in Health Insurance Markets," *SSRN Electronic Journal*, 2022, (April).

Chorniy, Anna, Daniel Miller, and Tilan Tang, "Mergers in Medicare Part D: Assessing market power, cost efficiencies, and bargaining power," *International Journal of Industrial Organization*, jan 2020, *68*, 102548.

Cutler, D. M. and S. J. Reber, "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection," *The Quarterly Journal of Economics*, may 1998, *113* (2), 433–466.

Cutler, David M. and Richard J. Zeckhauser, "Chapter 11 The anatomy of health insurance," in "Handbook of Health Economics," Vol. 1, Elsevier, 2000, pp. 563–643.

Dafny, Leemore, Jonathan Gruber, and Christopher Ody, "More Insurers Lower Premiums: Evidence from Initial Pricing in the Health Insurance Marketplaces," *American Journal of Health Economics*, 2015, *1* (1), 53–81.

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

de ven, Wynand P.M.M. Van and Randall P. Ellis, *Risk adjustment in competitive health plan markets*, Vol. 1, Elsevier Science B.V., 2000.

Decarolis, Francesco and Andrea Guglielmo, "Insurers' response to selection risk: Evidence from Medicare enrollment reforms," *Journal of Health Economics*, dec 2017, *56*, 383–396.

DeLeire, Thomas, Andre Chappel, Kenneth Finegold, and Emily Gee, "Do individuals respond to cost-sharing subsidies in their selections of marketplace health insurance plans?," *Journal of Health Economics*, dec 2017, *56*, 71–86.

Drake, Coleman, "What Are Consumers Willing to Pay for a Broad Network Health Plan?: Evidence from Covered California," *Journal of Health Economics*, 2019, *65*, 63–77.

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

_ , _ , Stephen P Ryan, Paul Schrimpf, and Mark R Cullen, "Selection on Moral Hazard in Health Insurance," *American Economic Review*, feb 2013, *103* (1), 178–219.

Ellis, Randall P. and Thomas G. McGuire, "Predictability and predictiveness in health care spending," *Journal of Health Economics*, jan 2007, *26* (1), 25–48.

Ericson, Keith M Marzilli and Amanda Starc, "Pricing Regulation and Imperfect Competition on the Massachusetts Health Insurance Exchange," *Review of Economics and Statistics*, 2015, *97* (3), 667–682.

Farrell, Joseph and Carl Shapiro, "Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition," *The B.E. Journal of Theoretical Economics*, 2010, *10.*

Finkelstein, Amy, Nathaniel Hendren, and Mark Shepard, "Subsidizing Health Insurance for Low-Income Adults: Evidence from Massachusetts," *American Economic Review*, apr 2019, *109* (4), 1530–1567.

Frean, Molly, Jonathan Gruber, and Benjamin D. Sommers, "Premium subsidies, the mandate, and Medicaid expansion: Coverage effects of the Affordable Care Act," *Journal of Health Economics*, may 2017, *53*, 72–86.

GAO, "CHILDREN'S HEALTH INSURANCE: Opportunities Exist for Improved Access to Affordable," *Government Accountability Office*, 2012, pp. 1–45.

Geruso, Michael, "Demand heterogeneity in insurance markets: Implications for equity and efficiency," *Quantitative Economics*, 2017, *8*, 929–975.

— , **Timothy J Layton, Grace Mccormack, and Mark Shepard**, "Trade-offs between Extensive and Intensive Margin Selection in Competitive Insurance Markets," 2018.

— , **Timothy Layton, and Daniel Prinz**, "Screening in contract design: Evidence from the ACA health insurance exchanges," *American Economic Journal: Microeconomics*, may 2019, *11* (2), 64–107.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town, "Mergers when prices are negotiated: Evidence from the hospital industry," *American Economic Review*, jan 2015, *105* (1), 172–203.

Grieco, Paul L E, Charles Murry, Joris Pinkse, and Stephan Sagl, "Efficient Estimation of Random Coefficients Demand Models using Product and Consumer Datasets," 2021.

Gruber, Jonathan, "Covering the uninsured in the United States," sep 2008.

— **and James Poterba**, "Tax Incentives and the Decision to Purchase Health Insurance: Evidence from the Self-Employed," *The Quarterly Journal of Economics*, aug 1994, *109* (3), 701–733.

Handel, Ben, Igal Hendel, and Michael D. Whinston, "Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk," *Econometrica*, 2015, *83* (4), 1261–1313.

Handel, Benjamin R and Jonathan T Kolstad, "Health Insurance for Humans: Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, 2015, *105* (8), 2449–2500.

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

Ho, Kate and Robin S Lee, "INSURER COMPETITION IN HEALTH CARE MARKETS," *Econometrica*, 2017, *85* (2), 379–417.

Jaffe, Sonia and E. Glen Weyl, "The First-Order approach to merger analysis," *American Economic Journal: Microeconomics*, nov 2013, *5* (4), 188–218.

— **and Mark Shepard**, “Price-Linked Subsidies and Imperfect Competition in Health Insurance,” *American Economic Journal: Economic Policy*, aug 2020, *12* (3), 279–311.

Lustig, Josh , “Measuring Welfare Losses from Adverse Selection and Imperfect Competition in Privatized Medicare,” 2010.

Kautter, John, Gregory C. Pope, and Patricia Keenan, “Affordable Care Act Risk Adjustment: Overview, Context, and Challenges,” *Medicare & Medicaid Research Review*, 2014, *4* (3), 1–11.

— **, Gregory Pope, Sara Freeman, Lindsey Patterson, Michael Cohen, and Patricia Keenan**, “The HHS-HCC Risk Adjustment Model for Individual and Small Group Markets under the Affordable Care Act,” *Medicare & Medicaid Research Review*, 2014, *4* (3), E1–E46.

Kong, Edward, Timothy J Layton, and Mark Shepard, “Adverse Selection and (un) Natural Monopoly in Insurance Markets,” 2023.

Lavetti, Kurt, Thomas DeLeire, and Nicolas Ziebarth, “How Do Low-Income Enrollees in the Affordable Care Act Marketplaces Respond to Cost-Sharing?,” *National Bureau of Economic Research*, 2019.

Layton, Timothy J., “Imperfect risk adjustment, risk preferences, and sorting in competitive health insurance markets,” *Journal of Health Economics*, dec 2017, *56*, 259–280.

Lester, Benjamin, Ali Shourideh, Venky Venkateswaran, and Ariel Zetlin-Jones, “Screening and Adverse Selection in Frictional Markets,” *https://doi.org/10.1086/700730*, jan 2019, *127* (1), 338–377.

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

Marquis, M. Susan, Melinda Beeuwkes Buntin, José J. Escarce, Kanika Kapur, and Jill M. Yegian, “Subsidies and the demand for individual health insurance in California,” *Health Services Research*, oct 2004, *39* (5), 1547–1570.

McGuire, Thomas G., Joseph P. Newhouse, and Anna D. Sinaiko, "An Economic History of Medicare Part C," *Milbank Quarterly*, jun 2011, *89* (2), 289–332.

_ , _ , Sharon-Lise Normand, Julie Shi, and Samuel Zuvekas, "Assessing incentives for service-level selection in private health insurance exchanges," *Journal of Health Economics*, may 2014, *35*, 47–63.

Miller, Keaton, Amil Petrin, Robert Town, and Michael Chernew, "Optimal Managed Competition Subsidies," 2019.

Miller, Nathan H., Marc Remer, Conor Ryan, and Gloria Sheu, "Upward pricing pressure as a predictor of merger price effects," *International Journal of Industrial Organization*, 2017, *52*, 216–247.

Newhouse, Joseph P, J Michael Mcwilliams, Mary Price, Jie Huang, Bruce Fireman, and John Hsu, "Do Medicare Advantage Plans Select Enrollees in Higher Margin Clinical Categories?," *Journal of Health Economics*, 2013, *32* (6).

Newhouse, Joseph P., Mary Price, J. Michael Mcwilliams, John Hsu, and Thomas G. McGuire, "How much favorable selection is left in medicare advantage?," *American Journal of Health Economics*, nov 2015, *1* (1).

Obama, Barak, "Obama's Health Care Speech to Congress," sep 2009.

Petrin, Amil, "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, jul 2002, *110* (4), 705–729.

Planalp, Colin, Julie Sonier, and Brett Fried, "State-Level Trends in Employer-Sponsored Health Insurance: A State-by-State Analysis," *Robert Wood Johnson Foundation*, 2015, pp. 1–83.

Pope, Gregory C, Henry Bachofer, Andrew Pearlman, John Kautter, Elizabeth Hunter, Daniel Miller, and Patricia Keenan, "Risk Transfer Formula for Individual and Small Group

Markets Under the Affordable Care Act," *Medicare & Medicaid Research Review*, 2014, *4* (3), 1–23.

Rothschild, Michael and Joseph Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 1976, *90* (4), 629.

Ryan, Conor, Roger Feldman, and Stephen Parente, "Estimating the Demand for Individual Health Insurance: Evidence from a Private Online Marketplace," 2021.

Rysman, Marc and Daniel A. Ackerberg, "Unobserved Product Differentiation in Discrete Choice Models: Estimating Price Elasticities and Welfare Effects," *RAND Journal of Economics*, 2005, *36* (4), 771–788.

Saltzman, Evan, "Demand for Health Insurance: Evidence from the California and Washington ACA Marketplaces," *Journal of Health Economics*, 2019, *63*, 197–222.

_ , "Managing adverse selection: underinsurance versus underenrollment," *The RAND Journal of Economics*, jun 2021, *52* (2), 359–381.

Shepard, Mark, "Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange," 2016.

Small, Kenneth A. and Harvey S. Rosen, "Applied Welfare Economics with Discrete Choice Models," *Econometrica*, 1981, *49* (1), 105–130.

So, Jaemin, "Adverse Selection, Product Variety, and Welfare," *Workiing Paper*, 2019.

Starc, Amanda, "Insurer Pricing and Consumer Welfare: Evidence from Medigap," 2014.

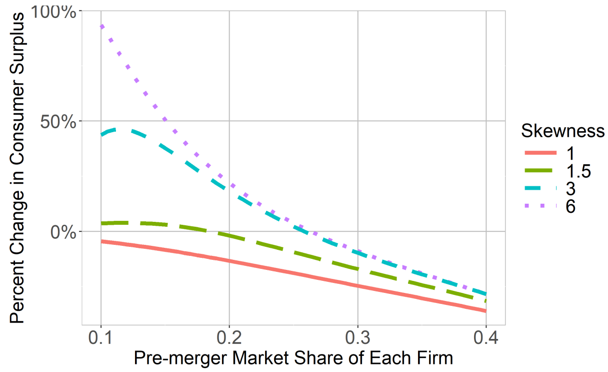
Tebaldi, Pietro, "Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA," *Review of Economic Studies*, 2023, *Forthcomin.*

Town, Robert, "The welfare impact of HMO mergers," *Journal of Health Economics*, nov 2001, *20* (6), 967–990.

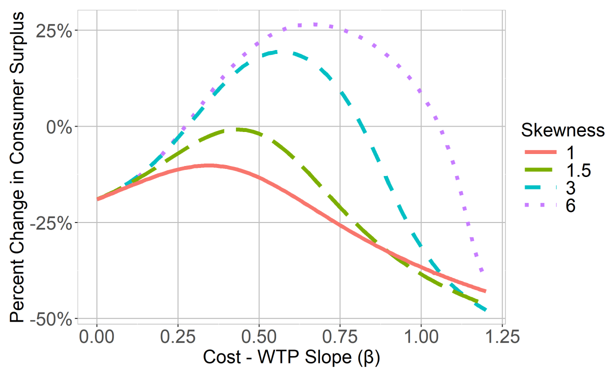
_ **and Su Liu**, "The Welfare Impact of Medicare HMOs," *The RAND Journal of Economics*, 2003, *34* (4), 719.

Veiga, André and E. Glen Weyl, "Product design in selection markets," *Quarterly Journal of Economics*, may 2016, *131* (2), 1007–1056.

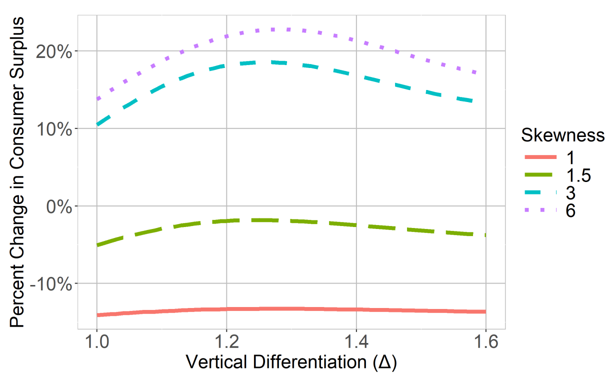
Waldfogel, Joel, "Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets," *The RAND Journal of Economics*, 2003, *34* (3), 557.



(a) Better for consumers when firms are smaller



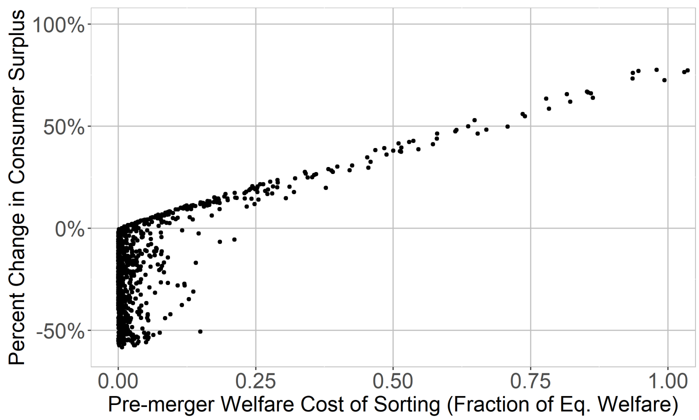
(b) Better for consumers when cost-WTP slope is moderate



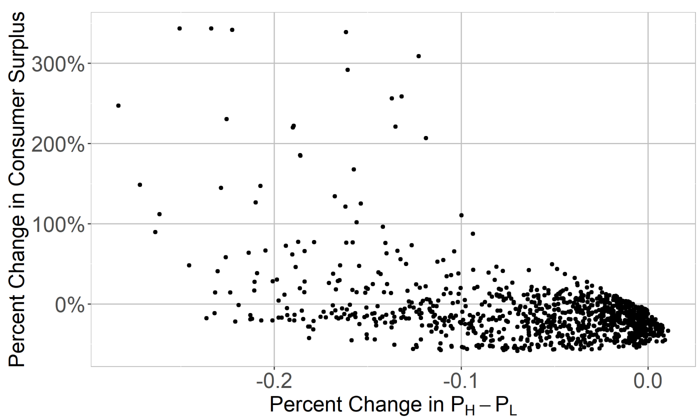
(c) Limited relationship to degree of vertical differentiation

Figure 1: When Do Mergers Improve Consumer Surplus: Intuition from Parameters

Note: These plots show the consumer surplus effect of a merger in the symmetric merger example presented in Section 2.3.2. The y-axis shows the percent change in consumer surplus. Each panel varies a single parameter of the model for multiple levels of skewness. Panel (a) varies β, the relationship between cost and willingness to pay. Panel (b) varies the initial size of the firms, pre-merger. Panel (c) measures the degress of vertical differentiation between the two products offered by each firm. Skewness is defined as the ratio of distances of the 90th and 10th percentiles to the median: $\frac{P_{90}-P_{50}}{P_{50}-P_{10}}$.



(a) Mergers Benefit Consumers when Cost of Sorting is High



(b) Weak Correlation Between Price Spread and Consumer Benefit

Figure 2: When Do Mergers Improve Consumer Surplus: Intuition from Monte Carlo Simulations

Note: These plots show the consumer surplus effect of a merger in the symmetric merger example presented in Section 2.3.2. Every point represents a merger simulated with randomly drawn parameters. The y-axis shows the percent change in consumer surplus. Panel (a) plots the merger effects relative to the pre-merger welfare cost of sorting, measured as a percentage of pre-merger total welfare. Panel (b) plots the merger effects relative to the percent change in the price difference between the generous insurance products (H) and the less generous insurance products (L).

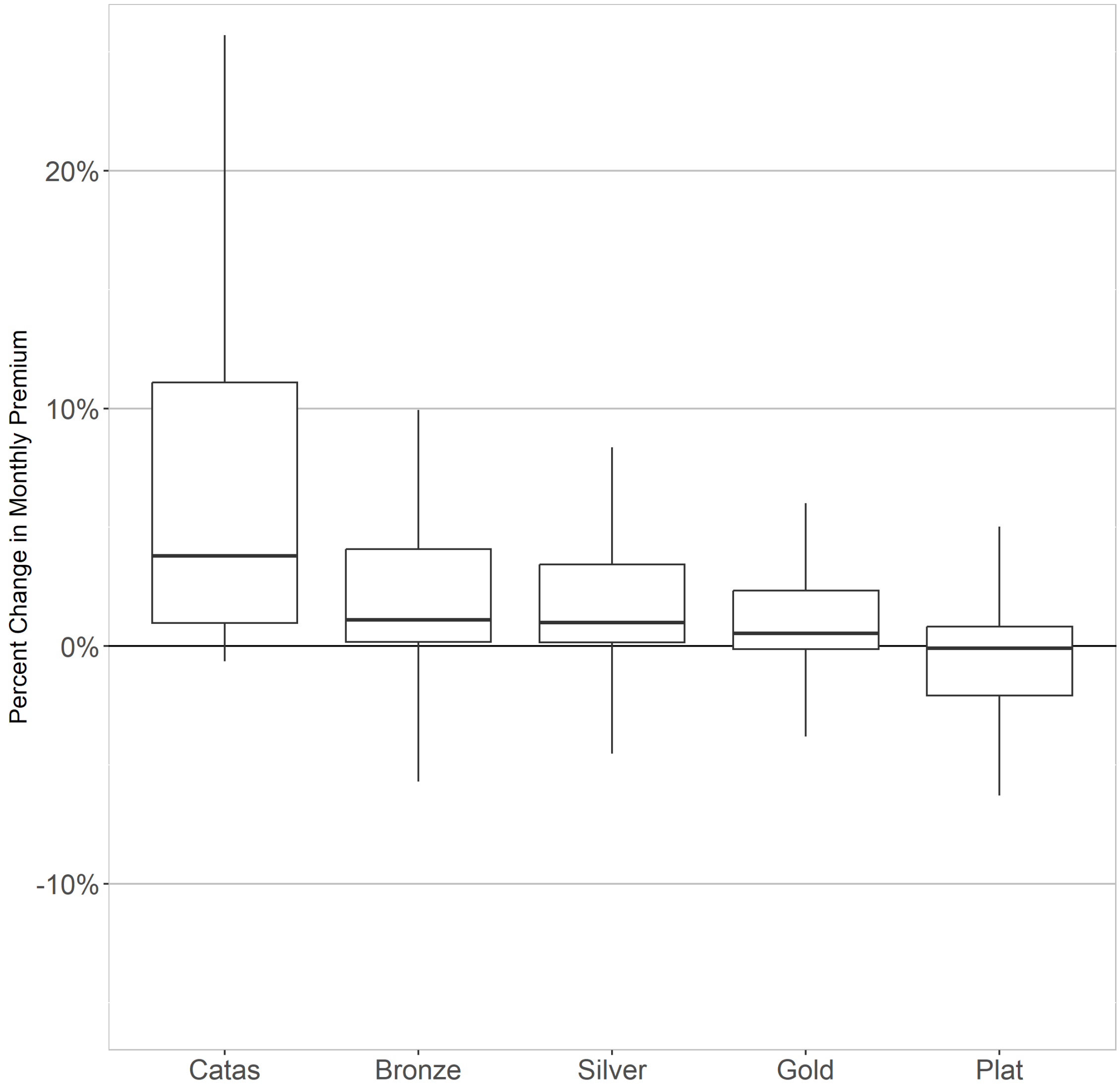


Figure 3: Mergers Decrease Price Spread Between Most and Least Generous Plans

Note: One mechanism through which mergers reduce inefficient sorting is by reducing the spread plan of different generosity. This figure shows the distribution of price effects across all merger-market-products in the simulation. The plan categories are ranked from least to most generous: catastrophic (catas), bronze, silver, gold, and platinum (plat). The dark black line represents the median effect, the box contains the inter-quartile range, and the lines extend to the 10^{th} and 90^{th} percentiles.

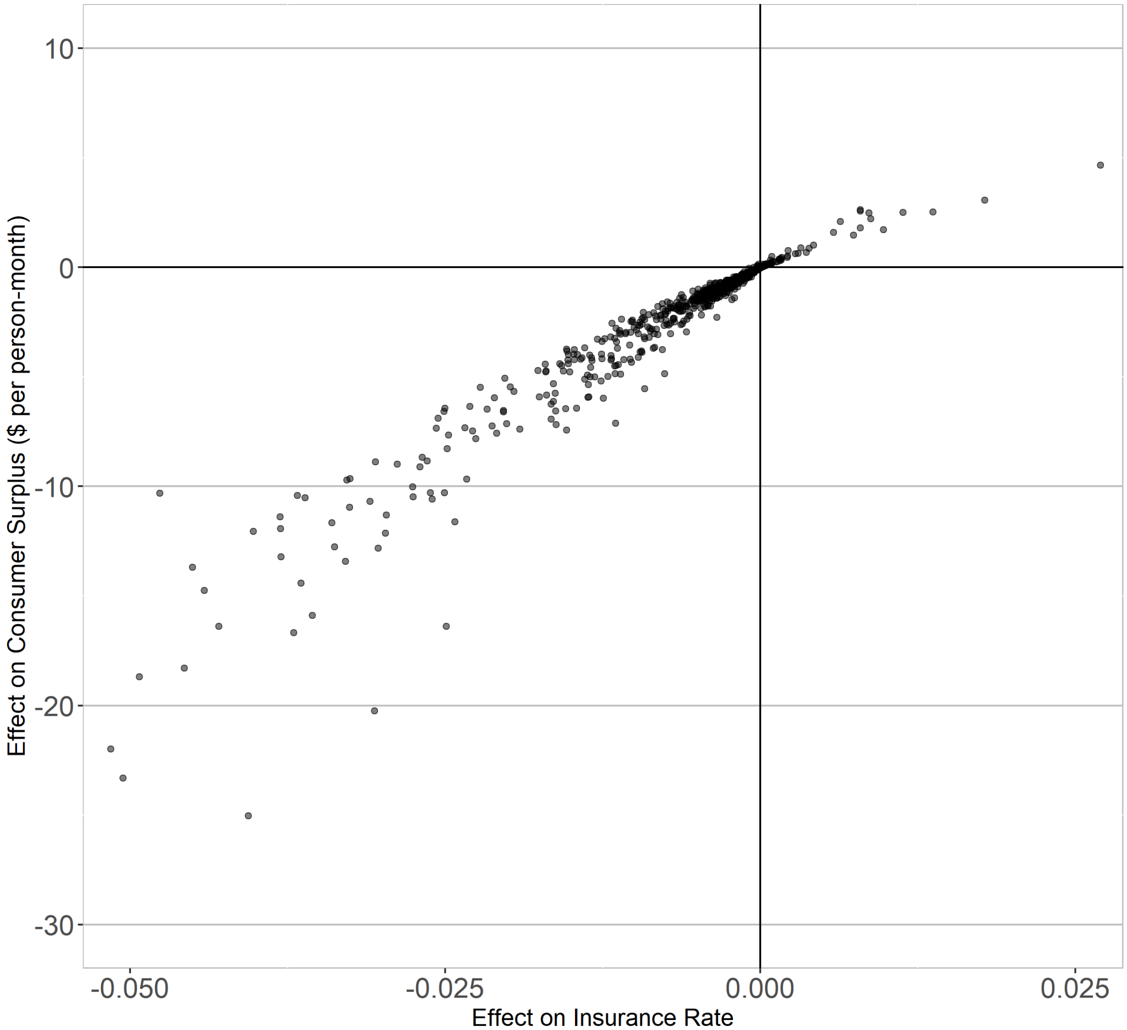


Figure 4: Effect on Insurance Rate Determines Effect on Consumer Surplus

Note: Welfare gains to consumers are tightly linked to the change in the insurance rate. This figure shows the change in consumer surplus relative to the change in the fraction of consumers that buy any insurance.

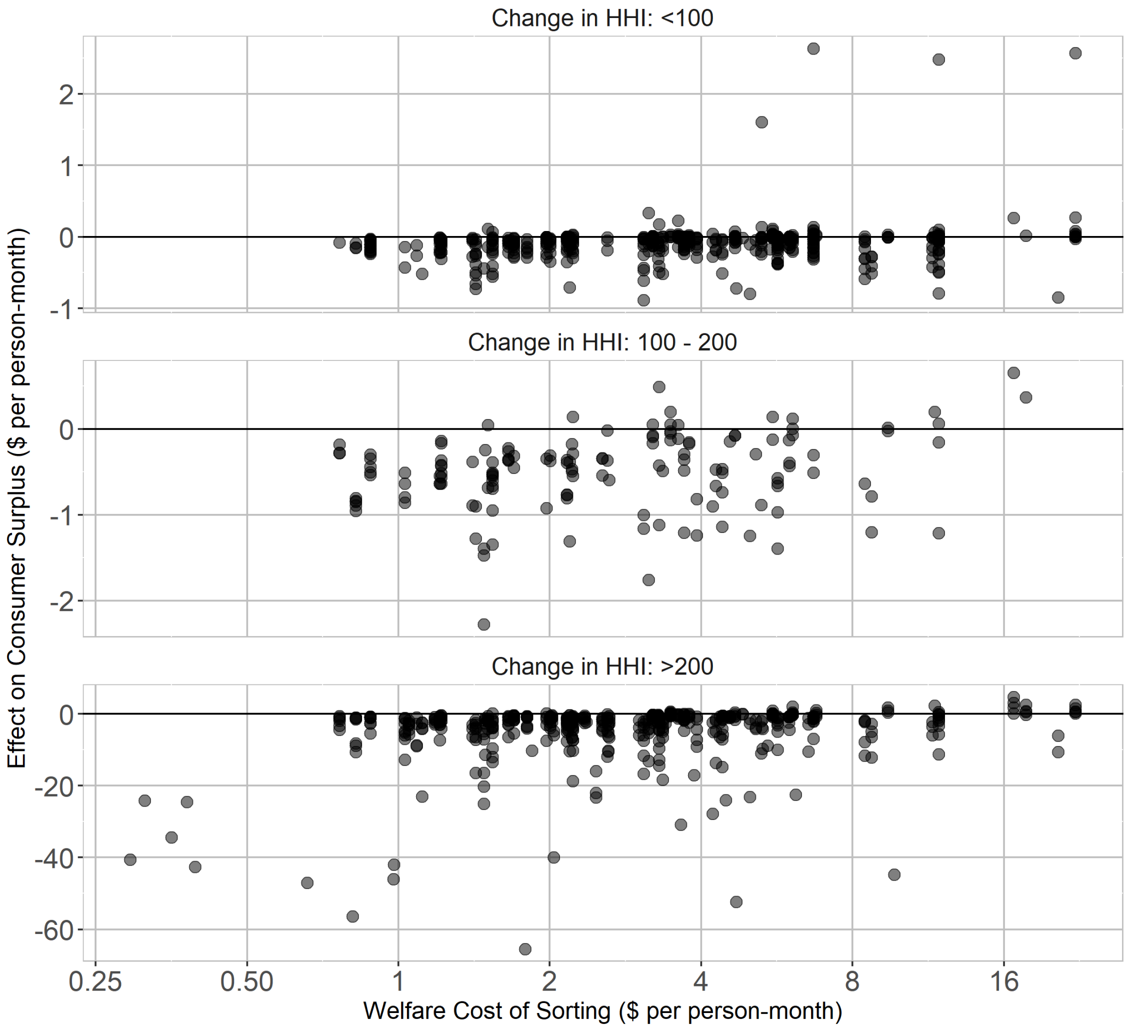


Figure 5: Mergers Benefit Consumers in Markets with Large Welfare Costs of Sorting

Note: Markets where the welfare cost of sorting is larger are more likely to have mergers that benefit consumers, and greater welfare costs of sorting lead to greater benefits. This figure shows the effect of each merger on consumer surplus by the welfare cost of sorting, separated by the magnitude of the change in HHI created by the merger. Each dot represents a single merger-market. Sorting cost is displayed on a log scale. Both the welfare cost of sorting and the effect on consumer surplus are measured in dollars per person per month.

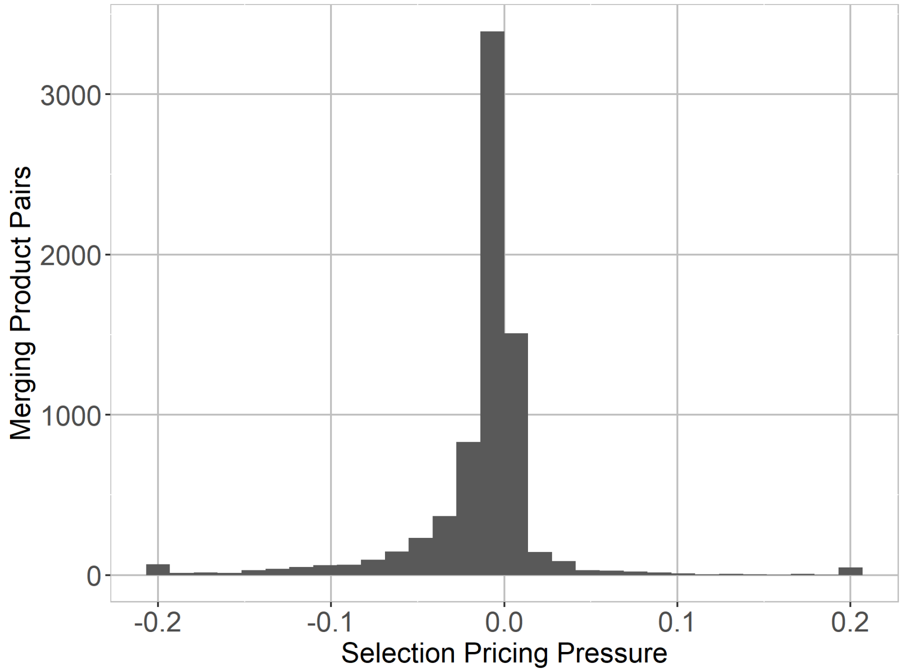


Figure 6: Selection Pricing Pressure is Typically Negative

Note: Typically, accounting for adverse selection in the pricing incentive of a merger leads to less upward pricing pressure. This plot shows the distribution of Selection Pricing Pressure across all product-market-mergers in the counterfactual simulation.

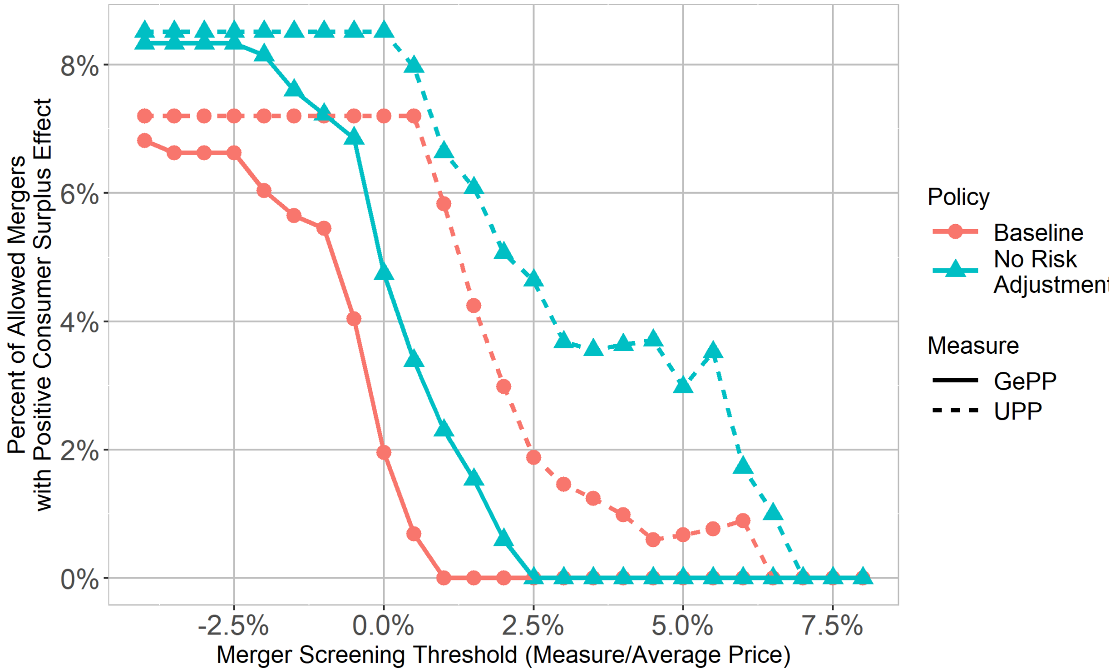


Figure 7: Traditional Screening Methods for Merger Harm Can Still Apply

Note: In the presence of adverse selection, a merger screen for investigation based on UPP may mis-predict the direction of the effect of a merger. Each dot represents the percent of mergers which exceed a threshold of each pricing measure (GePP and UPP) and lead to greater consumer surplus, displayed for the baseline and a no-risk-adjustment policy scenario.

Table 1: Demand Estimation Results

	(1)	(2)	(3)
Premium	-2.03	-1.41	-1.32
	(0.02)	(0.02)	(0.01)
Age 31 - 40	0.36	0.36	0.29
	(0.02)	(0.02)	(0.02)
Age 41 - 50	0.70	0.54	0.43
	(0.02)	(0.02)	(0.02)
Age 51 - 64	1.24	0.83	0.70
	(0.02)	(0.02)	(0.01)
Family	0.01	0.05	0.04
	(0.01)	(0.01)	(0.01)
Subsidized	0.29	0.32	0.29
	(0.01)	(0.01)	(0.01)
Acutarial Value (AV)	7.21	11.03	11.70
	(0.05)	(0.08)	(0.08)
Risk Preference			
AV	0.55	0.50	0.54
	(0.00)	(0.00)	(0.00)
Firm - Risk Interaction	Y	Y	Y
Fixed Effects			
Age, Fam., Inc.	Y	Y	Y
Firm	Y		
Firm-Market		Y	
Firm-Category		Y	
Firm-Mkt-Cat.			Y

Note: The top row of price coefficients corresponds to the estimate for households that do not fall into any of the listed subgroups (single, high income, 18 to 30 year olds). The price coefficients for other households are obtained by adding the relevant demographic adjustments to the top line. Premiums are in thousands of dollars per year.

Table 2: Cost Estimation Results

	(1)	(2)
Age	0.41	0.42
	(0.01)	(0.01)
Risk Score	0.11	0.11
	(0.00)	(0.00)
Actuarial Value	4.15	4.31
	(0.02)	(0.02)
State-Firm	Y	Y

Note: This table displays the estimates of the marginal cost function. Standard errors are computed using the GMM formula, accounting for demand estimation error in the simulation.

Table 3: Many Mergers are Predicted to Improve Consumer Surplus and Social Welfare

	Number of Mergers	Fraction $\Delta SW > 0$	Fraction $\Delta CS > 0$	Fraction Δ Sort Cost < 0
Total	1186	0.15	0.13	0.71
Δ HHI				
<100	557	0.22	0.19	0.84
100 - 200	143	0.13	0.10	0.73
200 - 1000	306	0.10	0.09	0.64
>1000	180	0.05	0.05	0.39
Sorting Cost				
<$5	847	0.09	0.07	0.67
$5 - $7.5	200	0.23	0.22	0.75
$7.5 - $10	36	0.28	0.22	0.81
>$10	103	0.52	0.43	0.90

Note: Many mergers lead to improvements in consumer surplus and social welfare. This table displays the fraction of mergers with positive welfare effects, both in terms of consumer surplus and total social welfare, and the fraction of mergers that reduce the welfare cost of sorting. The top line displays the average across all mergers, and the following two panels breakout the results by the size of the merger and the pre-merger welfare cost of sorting. The change in HHI is computed using pre-merger market shares to reflect pre-merger size of merging firms, and the sorting cost is measured in dollars per consumer per month.

Mergers in the Presence of Adverse Selection: Online Appendix

A Appendix Tables and Figures

Table A1: Control Function Estimation

	(1) All Products	(2) Products Purchased in Choice Data
Fraction of Consumers Under 35	-249	-272
	(94.5)	(120)
Firm	✓	✓
Metal Level	✓	✓
State	✓	✓

Note: This table displays the first stage of the control function estimation for the demand estimation procedure presented in Section 4.1. The dependent variable is the annual base premium. The first column shows estimates using all products in the markets included in the sample. The second column shows estimates using only products purchased in the choice data used in estimation. The control function is constructed from the residual of specification (1).

Table A2: Demand Estimation Results

	(1)	(2)	(3)
Premium	-2.07	-1.35	-1.32
	(0.02)	(0.02)	(0.01)
Age 31 - 40	0.30	0.28	0.29
	(0.02)	(0.02)	(0.02)
Age 41 - 50	0.62	0.44	0.43
	(0.02)	(0.02)	(0.02)
Age 51 - 64	1.20	0.71	0.70
	(0.01)	(0.01)	(0.01)
Family	0.01	0.06	0.04
	(0.01)	(0.01)	(0.01)
Subsidized	0.34	0.29	0.29
	(0.01)	(0.01)	(0.01)
Actuarial Value (AV)	7.03	11.98	11.70
	(0.05)	(0.08)	(0.08)
Risk Preference			
AV	0.59	0.55	0.54
	(0.00)	(0.00)	(0.00)
Firm - Risk Interaction	Y	Y	Y
Fixed Effects			
Age, Fam., Inc.	Y	Y	Y
Firm	Y		
Firm-Market		Y	
Firm-Category		Y	
Firm-Mkt-Cat.			Y

Note: This table displays the demand estimation results without using the control function as an additional source of identification. The specifications match those estimated in the three GMM columns of Table 1. The estimates are qualitatively similar.

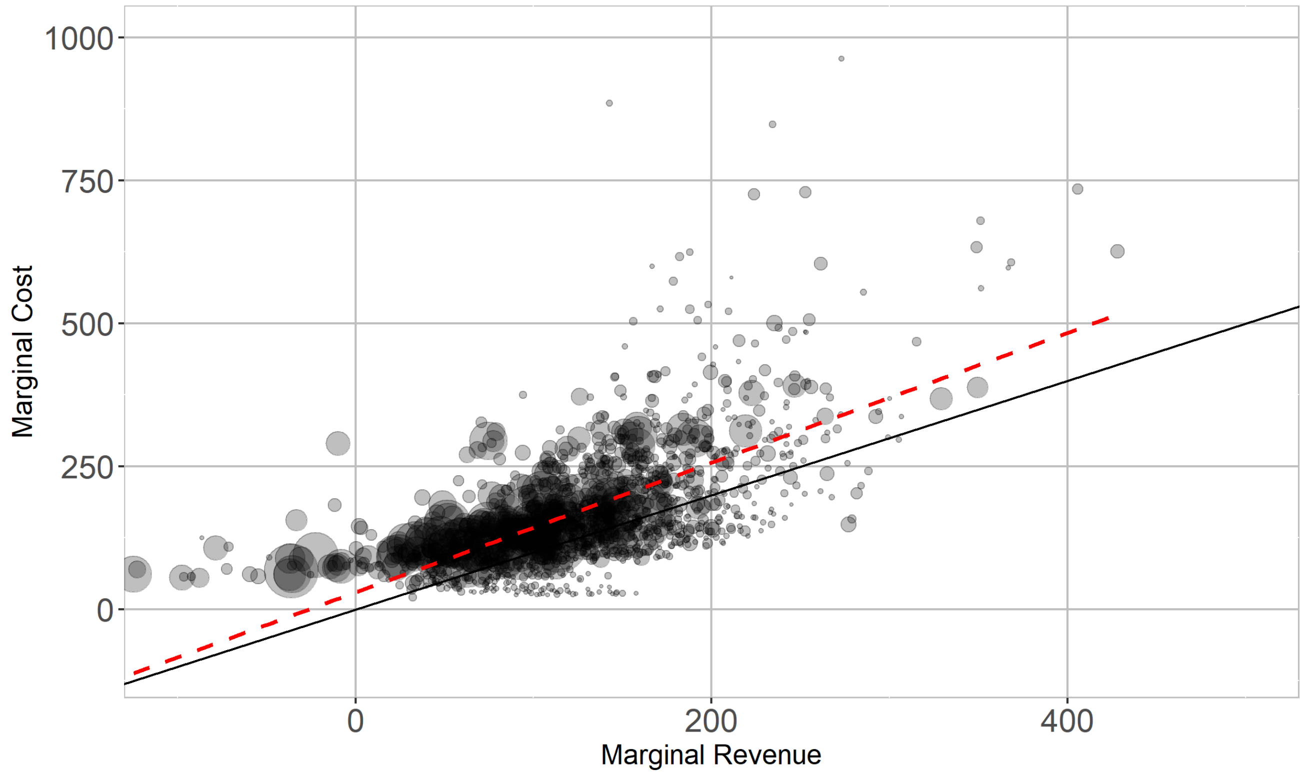


Figure A1: Marginal Revenue vs Marginal Cost in Baseline Model

Note: The product-level marginal cost and marginal revenue predicted by the estimated model are roughly equal on average. Each dot represents a product in a market. The size of the dots is proportional to the quantity sold. The model does relatively well with products that are close to the mean marginal revenue and costs but struggles to fit the outliers.

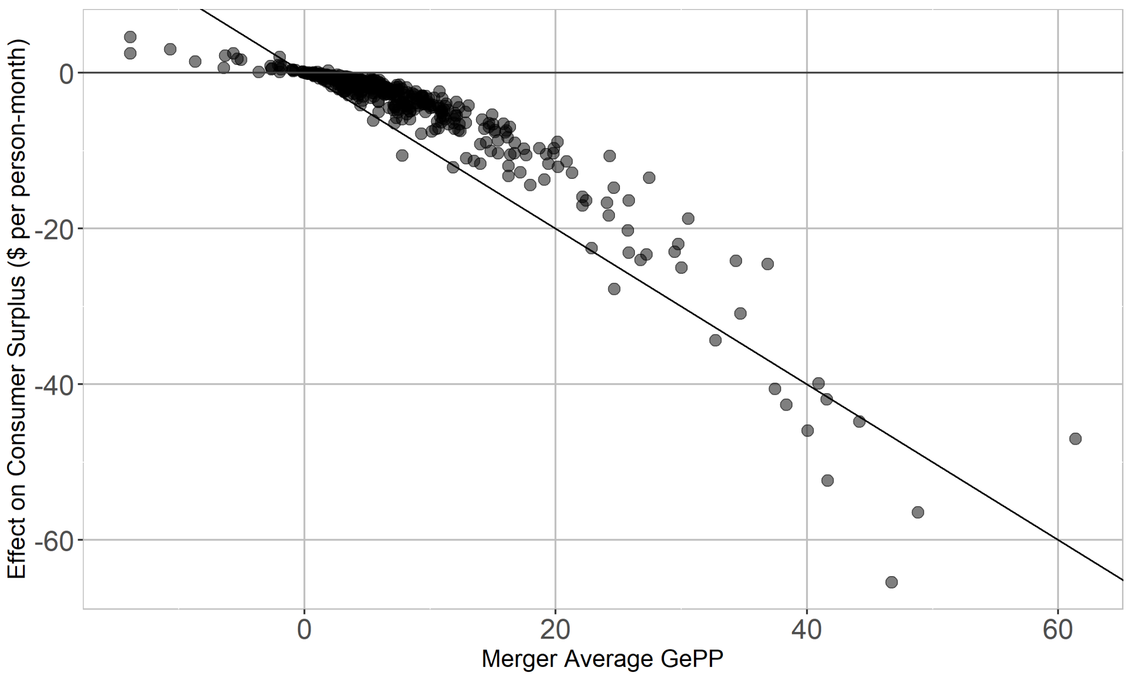


Figure A2: GePP Predicts Direction of Consumer Surplus Effect

Note: Average GePP forms a good prediction for the direction of the effect on consumer surplus. This figure compares the effect of a merger in a particular market relative to the average GePP across all products of the merging parties. Each dot represents a single merger-market. Sorting cost is displayed on a log scale. Both the welfare cost of sorting and the effect on consumer surplus are measured in dollars per person per month.The dark line represents the 45-degree line.

B Derivations for Section 2

B.1 The Average Cost Function

The cost to a particular product j of enrolling a household i is given by c_{ij}. The average cost of an insurance plan is the share-weighted average cost of all consumers that select that plan.

$$AC_j(\boldsymbol{p}) = \frac{1}{S_j(\boldsymbol{p})} \int_i S_{ij}(\boldsymbol{p}) c_{ij} di$$

The derivative of the average cost of product j with respect to the price of product k depends on the demand derivatives of the consumers of product j.

$$\frac{\partial AC_j}{\partial p_k} = \frac{1}{S_j}\int_i \frac{\partial S_{ij}}{\partial p_k} c_{ij} di - \frac{\frac{\partial S_j}{\partial p_k}}{S_j} \frac{\int_i S_{ij}(\boldsymbol{p}) c_{ij} di}{S_j}$$
$$\frac{\partial AC_j}{\partial p_k} = \frac{\frac{\partial S_j}{\partial p_k}}{S_j}\left(\frac{1}{\frac{\partial S_j}{\partial p_k}}\int_i \frac{\partial S_{ij}}{\partial p_k} c_{ij} di - AC_j\right)$$

B.2 Generalized Pricing Pressure

The optimal price for single-product firm, j, is derived below.

$$\Pi_j = S_j(\boldsymbol{p})(p_j - AC_j(\boldsymbol{p}))$$
$$0 = \frac{\partial S_j}{\partial p_j}(p_j - AC_j) + S_j(1 - \frac{\partial AC_j}{\partial p_j})$$
$$p_j = AC_j + \frac{S_j}{\frac{\partial S_j}{\partial p_j}}(1 - \frac{\partial AC_j}{\partial p_j}) \tag{10}$$

Consider now a multi-product firm with two products j and k that are offered to the same set of consumers. This is equivalent to the merged entity in the example given in Section 2.3. The optimal price for product j in this multi-product firm is derived below.

$$\Pi_j = S_j(\boldsymbol{p})(p_j - AC_j(\boldsymbol{p})) + S_k(\boldsymbol{p})(p_k - AC_k(\boldsymbol{p}))$$
$$0 = \frac{\partial S_j}{\partial p_j}(p_j - AC_j) + S_j(1 - \frac{\partial AC_j}{\partial p_j}) + \frac{\partial S_k}{\partial p_j}(p_k - AC_k) - S_k\frac{\partial AC_k}{\partial p_j}$$
$$p_j = AC_j + \frac{S_j}{\frac{\partial S_j}{\partial p_j}}(1 - \frac{\partial AC_j}{\partial p_j}) - \frac{\frac{\partial S_k}{\partial p_j}}{\frac{\partial S_j}{\partial p_j}}(p_k - AC_k) + \frac{S_k}{\frac{\partial S_j}{\partial p_j}}\frac{\partial AC_k}{\partial p_j} \tag{11}$$

GePP is defined as the difference between the pre-merger and post-merger first order conditions for a particular product's price, both normalized to be quasi-linear in marginal cost (Jaffe and Weyl (2013)). In the example given in Section 2.3, this is given by the price defined in Equations 11 less the price defined in Equation (10).

$$GePP_{jk}(\boldsymbol{p}) = -\frac{\frac{\partial S_k(\boldsymbol{p})}{\partial p_j}}{\frac{\partial S_j(\boldsymbol{p})}{\partial p_j}}(p_k - AC_k(\boldsymbol{p})) + \frac{S_k(\boldsymbol{p})}{\frac{\partial S_j(\boldsymbol{p})}{\partial p_j}}\frac{\partial AC_k(\boldsymbol{p})}{\partial p_j}$$

Importantly, GePP is a function of prices. In practice, it is typically evaluated at pre-merger prices.

B.3 Socially Optimal and Constrained Optimal Prices

The social welfare function, $SW(\cdot)$, is given by the sum of consumer surplus and producer profits.

$$SW(\boldsymbol{p}) = \int_i CS_i(\boldsymbol{p})di + \sum_{k\in J} S_k(p_k - AC_k)$$

The socially optimal price for a particular product j is derived below, using the result that $\frac{\partial CS_i}{\partial p_j} = -S_j$.

$$\begin{aligned}
0 &= \int_i \frac{\partial CS_i}{\partial p_j}di + \frac{\partial S_j}{\partial p_j}(p_j - AC_j) + S_j(1 - \frac{\partial AC_j}{\partial p_j}) + \sum_{k\neq j}\frac{\partial S_k}{\partial p_j}(p_k - AC_k) - S_k\frac{\partial AC_j}{\partial p_j} \\
0 &= \frac{\partial S_j}{\partial p_j}(p_j - AC_j) - S_j\frac{\partial AC_j}{\partial p_j} + \sum_{k\neq j}\frac{\partial S_k}{\partial p_j}(p_k - AC_k) - S_k\frac{\partial AC_j}{\partial p_j} \\
p_j^W &= AC_j + \frac{S_j}{\frac{\partial S_j}{\partial p_j}}\frac{\partial AC_j}{\partial p_j} + \left(\sum_{k\neq j}\frac{S_k}{\frac{\partial S_j}{\partial p_j}}\frac{\partial AC_j}{\partial p_j} - \frac{\frac{\partial S_k}{\partial p_j}}{\frac{\partial S_j}{\partial p_j}}(p_k - AC_k)\right) = \int_i \sum_k \frac{\partial S_{ik}}{\partial p_j}c_{ij}di \Big/ \int_i \frac{\partial S_{ij}}{\partial p_j}di
\end{aligned}$$

The problem of a constrained social planner that chooses product-level prices subject to a promise of total profit $\bar{\Pi}$ to the insurance industry is given below.

$$\begin{aligned}
&\max_{\{p_j\}_{j\in J}} \int_i CS_i(\boldsymbol{p})di \\
&\text{such that } \sum_{k\in J} S_k(p_k - AC_k) \geq \overline{\Pi}
\end{aligned}$$

The constrained optimal price for product j is derived below, where λ is the Lagrange multiplier on the profit constraint.

$$\begin{aligned}
\mathscr{L} &= \int_i CS_i(\boldsymbol{p})di + \lambda(\sum_{k\in J} S_k(p_k - AC_k) - \overline{\Pi}) \\
0 &= \int_i \frac{\partial CS_i}{\partial p_j}di + \lambda\left(\frac{\partial S_j}{\partial p_j}(p_j - AC_j) + S_j(1 - \frac{\partial AC_j}{\partial p_j}) + \sum_{k\neq j}\frac{\partial S_k}{\partial p_j}(p_k - AC_k) - S_k\frac{\partial AC_j}{\partial p_j}\right) \\
0 &= S_j(1-\frac{1}{\lambda}) + \frac{\partial S_j}{\partial p_j}(p_j - AC_j) - S_j\frac{\partial AC_j}{\partial p_j} + \sum_{k\neq j}\frac{\partial S_k}{\partial p_j}(p_k - AC_k) - S_k\frac{\partial AC_j}{\partial p_j} \\
p_j^{CE} &= -\frac{\lambda-1}{\lambda}\frac{S_j}{\frac{\partial S_j}{\partial p_j}} + AC_j + \frac{S_j}{\frac{\partial S_j}{\partial p_j}}\frac{\partial AC_j}{\partial p_j} + \left(\sum_{k\neq j}\frac{S_k}{\frac{\partial S_j}{\partial p_j}}\frac{\partial AC_j}{\partial p_j} - \frac{\frac{\partial S_k}{\partial p_j}}{\frac{\partial S_j}{\partial p_j}}(p_k - AC_k)\right)
\end{aligned}$$

B.4 Proof of Proposition 2.2

In Example 2.1, it is possible that preferences and costs are such that only one product is purchased in equilibrium, and it is no longer a meaningful example for a potential merger. In any equilibrium in which both products are purchased, it must be that $p_H^* > p_L^*$. If this is not the case, the demand for product L is 0.

The marginal consumers between the two products are those that are indifferent between product H and product L. Specifically, the marginal consumers are those with $c_i = \hat{c}$, where

$$\hat{c} = \frac{p_H^* - p_L^*}{(\nu-1)\beta}$$

Thus that marginal cost, $MC_{LH} = \hat{c}$, and the marginal cost $MC_{HL} = \nu\hat{c}$. Since $\nu > 1$, $MC_{HL} > MC_{LH}$.

All consumers with $c_i > \hat{c}$ prefer product H and all consumers with $c_i < \hat{c}$ prefer product L. Because we have assumed that the mass of consumers that purchase each product is non-zero, there exist other consumers that purchase both H and L with $c_i \neq \hat{c}$. Therefore, the average value of c_i for consumers that purchase H must be strictly greater than $\hat{c}$, and $AC_H > MC_{HL}$. Similarly, the average value of c_i for consumers that purchase L must be strictly less than $\hat{c}$, and $AC_L < MC_{LH}$.

Thus, it must be that $AC_L < MC_{LH} < MC_{HL} < AC_H$. Following Equation (8), this implies that

$\frac{\partial AC_L}{\partial p_H} > 0$ and $\frac{\partial AC_H}{\partial p_L} < 0$.

B.5 Proof of Proposition 2.4

In any symmetric equilibrium, $p_A^* = p_B^* = p^*$. The quantities, average costs, marginal costs, and other features of the equilibrium will be identical across the two products. As such, I will drop the product subscripts in the following proof.

Given the equilibrium prices, let $s(c)$ be the share of consumers with $c_i = c$ that purchase either product A or B. Let $m(c)$ be the share of consumers with $c_i = c$ that both purchase either A or B and are indifferent between the two options. The average cost of both products is given by

$$AC = \frac{1}{S}\int_i s(c_i)c_i di$$
$$S = \int_i s(c_i)di$$

The average cost of consumers on the margin between products A and B is given by

$$MC_{AB} = \frac{1}{M}\int_i m(c_i)c_i di$$
$$M = \int_i m(c_i)di$$

Let $\tilde{s}$ and $\tilde{m}$ be normalized weighting functions that determine the composition of consumers that make up the average cost and marginal cost.

$$\tilde{s}(c) = \frac{s(c)}{S}$$
$$\tilde{m}(c) = \frac{m(c)}{M}$$

Following Equation (8), if $MC_{AB} > AC$ then $\frac{\partial AC_A}{\partial p_B} > 0$ and $\frac{\partial AC_B}{\partial p_A} > 0$.

In order to show that $MC_{AB} > AC$, it is sufficient to demonstrate that $\tilde{s}(c)$ and $\tilde{m}(c)$ satisfy a single crossing property such that there exists some point in the support $\hat{c}$ where for all $c < \hat{c}$, $\tilde{m}(c) < \tilde{s}(c)$ and for all $c > \hat{c}$ $\tilde{m}(c) > \tilde{s}(c)$. If this is the case, then the weighting function $\tilde{m}(c)$ first-order stochastic dominates the weights given by $\tilde{s}(c)$, and by the properties of first-order stochastic dominance, $MC_{AB} > AC$.

The proof relies on demonstrating that $s(c)$ is "more concave" than $m(c)$. To see why this might be true, consider the distribution of idiosyncratic preferences F. Because preferences are i.i.d., the quantity demanded is equal to the share of consumers for whom both idiosyncratic preference terms fall below a threshold value. The share of marginal consumers are those for whom this threshold condition is true and the two draws of ε_A and ε_B are equal. See the expressions below for an unbounded distribution of ε.

$$s(c_i) = 1 - F(p^* - \beta c_i)^2$$
$$s(c_i) = 1 - \left(\int_{-\infty}^{p^* - \beta c_i} f(\varepsilon) d\varepsilon \right)^2$$
$$m(c_i) = 1 - \frac{1}{\Gamma} \int_{-\infty}^{p^* - \beta c_i} (f(\varepsilon))^2 d\varepsilon$$

where $\Gamma = \int_{-\infty}^{\infty} (f(\varepsilon))^2 d\varepsilon$. Very roughly speaking, the marginal function $m(c_i)$ is the expected value of the concave function $f(x) = -x^2$, while the share function $s(c)$ follows the concave function. The technical details of rigorously establishing the single crossing condition for an arbitrary distribution are non-trivial. Here, I will show that this condition holds for three commonly used distributions for idiosyncratic preferences: uniform, type I extreme value, and normal.

In each of these proofs, I will assume that c has infinite support, which allows me to exploit limiting properties of the distributions (or bounds in the case of uniform distribution.) However, because $\tilde{s}(c)$ and $\tilde{m}(c)$ are strictly increasing functions that always integrate to 1, the crucial single crossing property is maintained for any bounded range of c.

The example also assumes that the distribution of c is uniform. Establishing the single cross condition of $s(c)$ and $m(c)$ does not depend on the distribution of c and is maintained for any continuous distribution. However, some distributions of c may preclude the existence of an equilibrium.

Uniform Distribution

Suppose that the distribution of idiosyncratic preferences is uniform, i.e. $\varepsilon \sim U(l,u)$. In this case,

$$s(c) = 1 - \left(\frac{p^* - \beta c - l}{u-l}\right)^2$$
$$m(c) = F(p^* - \beta c) = \frac{p^* - \beta c - l}{u-l}$$

Because consumers are uniformly distributed, the probability that $\varepsilon_{iA} = \varepsilon_{iB}$ is independent of the value of ε_{iA}. Therefore, the distribution of $F(\varepsilon_{iA}|\varepsilon_{iA} = \varepsilon_{iB}) = F(\varepsilon_{iA})$. Thus the fraction of consumers that are marginal is simply given by the c.d.f. of the preference shocks.

Because both $\tilde{s}(c)$ and $\tilde{m}(c)$ must both integrate to one and are not everywhere equal, these two lines must intersect at least once. Because $\tilde{s}(c)$ is strictly concave for values $c \in (\frac{p^*-h}{\beta}, \frac{p^*-l}{\beta})$, and $\tilde{m}(c)$ is linear, they can intersect only one time.

For every c, $s(c) \geq m(c)$, with a strict inequality whenever $s(c) \in (0,1)$. Therefore, $S > M$. And there exists some large enough c value such that $s(c) = m(c) = 1$. Thus, at this large value of c, it must be that $\tilde{s}(c) < \tilde{m}(c)$. Therefore, $\tilde{m}(c)$ must be strictly greater than $\tilde{s}(c)$ after this crossing point and strictly lesser before it.

Type I Extreme Value

Suppose that the distribution of idiosyncratic preferences is such that $\varepsilon_{iA} = \upsilon_{iA} - \upsilon_{i0}$ and $\varepsilon_{iB} = \upsilon_{iB} - \upsilon_{i0}$, where υ_{iA}, υ_{iB}, and υ_{i0} are all distributed type I extreme value with variance σ^{ε}. In this case,

$$
\begin{aligned}
s(c) &= \frac{e^{\frac{\beta c - p^*}{\sigma^{\varepsilon}}}}{1 + 2e^{\frac{\beta c - p^*}{\sigma^{\varepsilon}}}} \\
\tilde{s}(c) &= \frac{s(c)}{\int s(c_i)di} \\
m(c) &= \frac{1}{\sigma^{\varepsilon}} s(c)^2 \\
\tilde{m}(c) &= \frac{\frac{1}{\sigma^{\varepsilon}} s(c)^2}{\frac{1}{\sigma^{\varepsilon}} \int s(c_i)^2 di} \\
\tilde{m}(c) &= \frac{s(c)^2}{\int s(c_i)^2 di}
\end{aligned}
$$

These functions satisfy three conditions that are sufficient to establish the single crossing property. First, $\lim_{c\to-\infty} s(c) = \lim_{c\to-\infty} m(c) = 0$ and $\lim_{c\to\infty} s(c) = \lim_{c\to\infty} m(c) = 1$. Second, $s(c) > m(c)$ for all c.

Third, the derivatives of $s(c)$ and $m(c)$ satisfy a single crossing property such that there exists a $\bar{c}$ where for all $c < \bar{c}$, $s'(c) > m'(c)$ and for all $c > \bar{c}$, $s'(c) < m'(c)$. Note that $m'(c) = 2s(c)s'(c)$. Because s is a strictly increasing function, this single crossing point is the value $\bar{c}$ where $s(\bar{c}) = \frac{1}{2}$.

Because both $\tilde{m}(c)$ and $\tilde{s}(c)$ must integrate to 1 and are not everywhere equal, these two functions must intersect at least once. Because both s and m are strictly increasing with single crossing derivatives, the intersections must be countable. In other words, the functions only intersect at points rather than overlapping for an interval.

Because $s(c) > m(c)$ for all c, $\int s(c_i)di > \int s(c_i)^2 di$. Because the denominator for $\tilde{s}(c)$ is strictly larger and $\lim_{c\to\infty} s(c) = \lim_{c\to\infty} m(c) = 1$, there is some final crossing point $\hat{c}$ such that for all $c > \hat{c}$, $\tilde{s}(c) < \tilde{m}(c)$.

Suppose for a contradiction that there is another crossing point at some value $c' < \hat{c}$. It must be that s is less than m for $c < c'$ greater than m for $c > c'$ for values of c in the neighborhood of c'. Therefore, $\frac{\partial \tilde{s}(c')}{\partial c} > \frac{\partial \tilde{m}(c')}{\partial c}$. Because $\int s(c_i)di > \int s(c_i)^2 di$, it must also be that $\frac{\partial s(c')}{\partial c} > \frac{\partial m(c')}{\partial c}$. Because of the single crossing condition for these derivatives, it must be that for all values of $c < c'$, $\frac{\partial \tilde{s}(c')}{\partial c} > \frac{\partial \tilde{m}(c')}{\partial c}$. Thus, for values of $c < c'$, $\tilde{s}(c') < \tilde{m}(c')$, and because of the derivative inequality, these two functions are diverging as c goes to negative infinity. However, $\lim_{c\to-\infty} s(c) = \lim_{c\to-\infty} m(c) = 0$ and thus $\lim_{c\to-\infty} \tilde{s}(c) = \lim_{c\to-\infty} \tilde{m}(c) = 0$. This is a contradiction. Therefore, we know that these functions must cross only once at $\hat{c}$. Moreover, $\tilde{m}(c)$ must be strictly greater than $\tilde{s}(c)$ after this crossing point and strictly lesser before it.

Normal Distribution

Suppose that the distribution of idiosyncratic preferences are normal, i.e. $\varepsilon \sim N(\mu, \sigma^{\varepsilon})$. In this case,

$$s(c) = 1 - \Phi\left(\frac{p^* - \beta c}{\sigma^{\varepsilon}} - \mu\right)$$
$$m(c) = 1 - \Phi\left(\sqrt{2}\left(\frac{p^* - \beta c}{\sigma^{\varepsilon}} - \mu\right)\right)$$

That the marginal consumers follow a distribution scaled by the $\sqrt{2}$ can be seen from the square of the normal p.d.f.

$$
\begin{aligned}
\phi(x)^2 &= \left(\frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}} \right)^2 \\
&= \frac{e^{-x^2}}{2\pi} \\
&= \frac{e^{\frac{-(\sqrt{2}x)^2}{2}}}{2\pi} \\
&= \frac{1}{\sqrt{2\pi}} \phi\left(\sqrt{2}x\right)
\end{aligned}
$$

The extra $\frac{1}{\sqrt{2\pi}}$ will drop out when normalizing to the weighting function $\tilde{m}$, so we can drop it here.

The functions $s(c)$ and $m(c)$ satisfy the same three conditions as those generated by the type I extreme value preferences. First, the limiting values of both s and m are 0 as c approaches negative infinity and 1 as c approaches positive infinity. Second, $s(c) > m(c)$ for all c. This second condition is not generally true for an arbitrary scaling factor, but holds for $\sqrt{2}$.

Third, the derivatives of $s(c)$ and $m(c)$ satisfy a single crossing property such that there exists a $\bar{c}$ where for all $c < \hat{c}$, $s'(c) > m'(c)$ and for all $c > \hat{c}$, $s'(c) < m'(c)$. The derivatives are given by:

$$
\begin{aligned}
s'(c) &= \frac{2\beta}{\sigma^\varepsilon} 2\Phi\left(\frac{p^* - \beta c}{\sigma^\varepsilon} - \mu\right) \phi\left(\frac{p^* - \beta c}{\sigma^\varepsilon} - \mu\right) \\
m'(c) &= \frac{\sqrt{2}\beta}{\sigma^\varepsilon} \phi\left(\sqrt{2}\left(\frac{p^* - \beta c}{\sigma^\varepsilon} - \mu\right)\right)
\end{aligned}
$$

These are known parametric functions that meet this single crossing condition. For intuition, it is easy to verify that $2\Phi(x)\phi(x)$ and the standard normal p.d.f., $\phi(x)$, satisfy this condition by a similar argument to that used in the type I extreme value proof above. Consider all potential scaling factors d for $d\phi(dx)$. For small values of d, the p.d.f. is very disperse—large in the tails and small at the mean—and it intersects $2\Phi(x)\phi(x)$ twice. For large values of d, the p.d.f. is very concentrated—vanishing in the tails and large at the mean—and it also intersects $2\Phi(x)\phi(x)$

twice, but from the opposite directions. For values of d sufficiently close to 1—including $\sqrt{2}$—the two curves satisfy the single crossing property. Note that β, σ^{ε}, and μ, are all scaling parameters that apply equally and will not alter this relationship.

The proof then proceeds in the same manner as the previous proof for the type I extreme value distribution.

B.6 Simulation Parameters

The illustrative model outlined in Section 2.3.2 can be governed four parameters: $(\delta_0, \beta, \nu, \sigma^c)$. For the parameter δ_0, I calibrate it to target an initial market share $\bar{S} = S_{AL} + S_{AH} = S_{BL} + S_{BH}$. For the parameter, σ^c, I calibrate it to match a value of a skewness measure of the cost distribution, skew $= \frac{P_{90}-P_{50}}{P_{50}-P_{10}}$.

Table A3 displays the range of targeted values for these parameters. The final column of the table also shows the fixed value of the parameter in all panels of Figure 1 where that parameter is held constant.

Table A3: Data Description

Parameter	Minimum Value	Maximum Value	Fixed Value
$\bar{S}$	0.05	0.4	0.2
β	0	1.2	0.5
ν	0	0.6	0.2
skew	1	6	-

Note: This table displays the range of parameters used to simulate the mergers studied in Section 2.3.2

In order to generate Figure 1, I hold fixed two of the three parameters (δ_0, β, ν) fixed and simulate a merger for every value of the third parameter along a fine grid between the minimum and maximum value shown in Table A3. In each of the three panels, I repeat this for four values of skewness: 1, 1.5, 3, and 6.

In order to generate Figure 2, I simulate 1000 mergers by drawing parameters in a Monte Carlo simulation. I each simulation, I draw each of the four parameters from independent uniform distributions bounded between the minima and maxima listed in Table A3.

In each simulation, I solve for the symmetric equilibrium for both pre-merger and post-merger. Because of selection, some parameter values could potentially generate asymmetric equilibria, but I ignore this possibility. A symmetric equilibrium with positive market shares exists for all parameter values considered. I use 100 Gaussian quadrature points to integrate over the distribution of c.

In these simulations, I must also set a fifth parameter, σ^{ε}, which governs the variance of the idiosyncratic preferences. Because the distribution of c affects demand and is fixed, this is a non-trivial normalization. I set $\sigma^{\varepsilon} = 0.125$. In the Monte Carlo simulations (explained below), this value of σ^{ε} produces reasonable average own-price elasticities that range from -3.9 to -9.2, with a mean value of -5.6. Smaller values of σ^{ε} lead to greater elasticities and make it more likely that mergers will benefit consumers. Greater elasticities lead to lower equilibrium prices and more potentially unprofitable consumers. Greater values of σ^{ε} lead to lower elasticities and have the opposite effect.

C Data Processing

C.1 Processing the Choice Data

The choice data contain only the ultimate choices made by the consumers, not the scope of available options. In order to construct choice sets, I use the HIX 2.0 data set compiled by the Robert Wood Johnson Foundation. This data set provides detailed cost-sharing and premium information on plans offered in the non-group market in 2015. The data set is nearly a complete depiction of the market for the entire United States, but there are some markets in which some cost-sharing information is missing, or insurance firms are absent altogether.

I restrict the analysis to markets in which I observe characteristics of the entire choice set and can be reasonably confident that the private marketplace presents nearly the complete choice set of health insurers. Using state-level market shares from the Medical Loss Ratio reporting data, I throw out any markets in which I do not observe any purchases from insurance firms that have more than 5% market share in the state. In this way, I hope to ensure that my sample of choices is not segmented to only a portion of the market.

In Table A4, I summarize the data sample used in estimation and compare it to other data on the non-group insurance market: the ACS and data reported by the Office of the Assistant Secretary for Planning and Evaluation (ASPE) at the U.S. Department of Health and Human Services. The ACS survey design offers the broadest depiction of the market across all market segments. ASPE publishes detailed descriptive statistics on purchases made through the federally-facilitated HealthCare.gov. Relative to the ACS, enrollment through HealthCare.gov is weighted heavily towards low-income, subsidy-eligible consumers. As a result, the plan type market shares reported by ASPE are weighted heavily towards Silver plans that have extra cost-sharing benefits at low incomes. While the private marketplace is tilted towards higher-income and younger households, the ACS weighting moves the demographic distributions and market shares closer to those in the other data sources. Ryan et al. (2021) investigate these relationships in more detail and show that the market shares, conditional on income and geography, are quite close to those reported by ASPE.

Choice Sets

A household's choice set depends on the age composition of its members and the household income. Since I observe only one age of the household, I use a simple rule to impute the age composition: any household with more than one individual contains two adults of the same age and additional persons are under the age of 21. For a subsample where I can infer the age composition based on their charged premium, this simple rule has a correlation with the inferred age composition of 0.9. The income information also contains some missing values. For subsidized consumers, income can be imputed from the observed subsidy value and the household size. I use this imputed income for subsidized consumers with missing income information. However, doing so is not possible for the consumers that do not receive a subsidy. I assume that those in the data without a reported subsidy amount have an income greater than the subsidy qualification threshold.

The choice set in each market is large. The typical market has about 150 plans to choose

Table A4: Data Description

	ACS	ASPE	Private Marketplace Un-weighted	Private Marketplace Weighted
		Age Distribution		
Under 18	0.0%	9.0%	0.0%	0.0%
18 to 25	7.6%	11.3%	11.1%	11.4%
26 to 34	17.2%	17.5%	30.8%	29.1%
35 to 44	22.2%	16.8%	21.4%	20.1%
45 to 54	25.3%	20.9%	19.9%	20.5%
55 to 64	27.7%	23.3%	16.8%	19.0%
		Income Distribution		
Under 250% FPL	32.1%	76.1%	30.8%	43.0%
250% to 400% FPL	24.5%	15.4%	9.1%	13.4%
Over 400% FPL	43.4%	8.5%	60.1%	43.6%
		Metal Level Market Shares		
Catastrophic		1.1%	5.0%	3.6%
Bronze		24.2%	39.2%	36.0%
Silver		66.4%	41.8%	48.8%
Gold		6.6%	11.1%	9.4%
Platinum		1.7%	2.9%	2.2%

Note: The table compares the weighted and unweighted distribution of consumers in the estimation data sample relative to other data sources on the non-group market. The age distributions reported are for the head-of-household with the exception of ASPE, which is the individual-level distribution.

from, and these plans do not necessarily overlap with other markets. Because I observe only a sample of choices, there are many plans that I do not observe being chosen. The lack of observed choices does not necessarily imply that these plans have a zero market share and may be due to the fact that the number of options is large relative to the observed number of choices. The median number of choices per market is 300.

To simplify this problem, I aggregate to the level of firm-metal offerings in a particular market. For example, all Bronze plans offered by a single insurance firm are considered a single product. While firms typically offer more than one plan in a given metal level, the median number of plan offerings per metal level is three, and the 75^{th} percentile is five. Wherever there is more than one plan per category, I aggregate by using the median premium within the category. The only other product attributes I use in estimation are common to all plans in each category.

C.2 American Community Survey

Data on the size and demographic distributions of both the uninsured and insured populations in each market come from 2015 American Community Survey (ACS). The population of individuals who might consider purchasing non-group health insurance is any legal US resident that is not eligible for Medicaid, Medicare, and is not enrolled in health insurance through their employer. An individual that is not enrolled in employer sponsored insurance but has an offer that they chose not to accept is assumed to be in the non-group market. These consumers may be ineligible for subsidies but can often obtain waivers to get the same treatment as those without an employer offer. This population is small (Planalp et al. (2015)), and I treat them identically to the rest of the non-group market.

In order to address under-reporting of Medicaid enrollment, any parent that receives public assistance, any child of a parent that receives public assistance or is enrolled in Medicaid, any spouse of an adult that receives public assistance or is enrolled in Medicaid or any childless or unemployed adult that receives Supplemental Security Income payments are assumed to be enrolled in Medicaid. Besides Medicaid and CHIP enrollment, an individual is considered eligible for either program if his or her household income falls within state-specific eligibility levels. If an individual is determined to be eligible for Medicaid through these means but reports to be enrolled in private coverage, either non-group coverage or through an employer, they are assumed to be enrolled in Medicaid. This accounts for those that confuse Medicaid managed care programs with private coverage, and Medicaid employer insurance assistance.

This paper follows the Government Accountability Office methods (GAO (2012)) to construct health insurance households. This method first divides households as identified in the survey data into tax filers and tax dependents, linking tax dependents to particular tax filers. A tax filing household, characterized by the single filer or joint filers and their dependents, is generally considered to be a health insurance purchasing unit. In some cases, certain members of a tax household will have insurance coverage through another source, e.g. an employer or federal program. In this case, the health insurance household is the subset of the tax household that must purchase insurance on the non-group market.

C.3 Medical Expenditure Panel Survey

The Medical Expenditure Panel Survey (MEPS) is a nationally representative household survey on demographics, insurance status, and health care utilization and expenditures. MEPS provides moments on the distribution of risk scores in the insured population and the relative costs of households by the age and risk score of the head of household and the risk. All moments are constructed using all surveyed households with a head of household under the age of 65.

The 2015 Medical Conditions File (MCF) of MEPS contains self-reported diagnosis codes. The publicly available data only list 3-digit diagnosis codes, rather than the full 5-digit codes. I follow McGuire et al. (2014) and assign the smallest 5-digit code for the purpose of constructing the condition categories. For example, I treat a 3-digit code of '571' as '571.00'. This implies that many conditions in the hierarchical risk prediction framework are censored. However McGuire et al. (2014) find that moving from 5-digit codes to 3-digit codes does not have a large effect on the predictive implications for risk scores.

I link the MCF to the Full Year Consolidated File to identify the age and sex of the individual, and then apply the 2015 HHS-HCC risk prediction methodology (Kautter et al. (2014b)). The risk coefficients are published by CMS and publicly available.

C.4 Medical Loss Ratio Data

CMS makes publicly available the state-level financial details of insurance firms in the non-group market for the purpose of regulating the MLR.[1] This information includes the number of member-months covered by the insurance firm in the state and total costs.

This paper uses two pieces of information from the Medical Loss Ratio filings: average cost and average risk adjustment transfers.

[1] Insurance firms in this market are restricted in how much premium revenue they may collect, relative to an adjusted measure of medical costs. In 2015, this constraint is not often binding. Excess revenue is returned to consumers via a rebate.

Firms are defined by operating groups at the state level. Some firms submit several medical loss ratio filings under for different subsidiaries in a given state. I group these filings together.

Average cost is defined as total non-group insurance claims divided by total non-group member months, current as of the first quarter of 2016. This computation includes claims and member months that may not be a part of the non-group market as it is characterized in this analysis. For instance, grandfathered insurance plans that are no longer sold to new consumers are included. These are likely to be a small portion of the overall market.

To compute the average risk adjustment payment, some adjustment to the qualifying member months is required. Unlike medical claims, grandfathered plans (and other similar non-ACA compliant plans) are not included in the risk adjustment system. Dividing the total risk adjustment transfer by the total member months will bias the average transfer towards zero.

The interim risk adjustment report published by CMS includes the total member months for every state. And the MLR filings separately list the risk-corridor eligible member months, which are a subset of the risk adjustment eligible member months. I define "potentially non-compliant" member months as the difference between risk-corridor eligible member months and total member months. I scale the potentially non-compliant member months of all firms in each state proportionally so that total member months is equal to the value published by CMS, with two exceptions. First, firms that opted not to participate in the ACA exchange in that state have zero risk-corridor eligible member months. I do not reduce the member months of these firms, as I cannot isolate the potentially non-compliant months. Second, if the risk-corridor eligible member months exceed the total member months published by CMS, I assume that the risk-corridor eligible member months are exactly equal to the risk adjustment eligible member months.

Computing Firm-level Risk

This paper firm-level risk transfers to infer the equilibrium distribution of risk across firms. With a bit of simplification, the ACA risk transfer formula at the firm level can be written as

$$T_f = \left[\frac{\bar{R}_f}{\sum_{f'} S_{f'}\bar{R}_{f'}} - \frac{\bar{A}_f}{\sum_{f'} S_{f'}\bar{A}_{f'}}\right]\bar{P}_s$$

where $\bar{R}_f$ is the firm level of average risk and $\bar{A}_f$ is the firm level average age rating, where the average is computed across all the firms products and weighted by members, a geographic adjustment, and a metal-level adjustment. S_f is the firm's state-level inside market share, and $\bar{P}_s$ is the average total premium charged in the state.

Every element of this formula is data available in the Interim Risk Adjustment Report on the 2015 plan year, except for the plan-level market shares, the plan-level average age rating, and the plan-level average risk. As a simplification, I assume that the average age rating is constant across all firms, and that the weighting parameters in the risk component are negligible. In reality, variation in the average age rating is not very large, and incorporating this variation in the moment matching dramatically increases the computational burden.

I compute the implied firm-level average risk as

$$\bar{R}_f = \left(\frac{T_f}{\bar{P}_s} + 1\right)\bar{R}$$

where the risk transfer T_f is the average firm-level risk adjustment transfer from MLR data, $\bar{P}_s$ is the average state level premium reported in the interim risk adjustment report, and $\bar{R}$ is the national average risk score reported in the interim risk adjustment report.[2] In the estimation, I target the difference between $\bar{R}_f$ and an adjusted average risk score for the state that accounts for grandfathered insurance products not sold through the marketplace.

[2]The formula implies that the state average risk score should go in place of the national average. However, I do not allow the risk distribution among consumers to vary by geography (other than through composition). I use the national risk score to abstract from these geographical differences.

C.5 Rate Filing Data

The Center for Medicare and Medicaid Services (CMS) tabulates the Premium Rate Filings that insurance firms must submit to state insurance regulators if they intend to increase the premiums for products they will continue to offer. In these filings, insurance firms include information on the cost and revenue experience of the insurance product in the prior year and projections for the following year.

The rate filing data are divided into two files—a firm-level worksheet and a plan-level worksheet—and contain information on the prior year experience of the plan and the projected experience of the plan in the coming year. I use projected firm-level average cost and the average ratio of experienced costs across metal levels for all firms. Using projected average costs for the firms leads to the best fit for the first order conditions, which are not imposed in estimation. This may be because it more accurately represents firms' expectations when setting their costs. While the decision to use projected or experienced costs does affect the marginal cost estimation, it does not qualitatively impact the results.

To construct moments on the ratio of average cost across metal level categories, I use the prior year experience submitted in the 2016 rate filings data. To recover the average cost after reinsurance, I subtract the experienced total allowable claims that are not the issuer's obligation and the experienced risk adjustment payments from the total allowable claims.

The ratio of average cost across each metal level category is computed as the weighted average of every within firm ratio. I compute the average cost across all plans within each metal level category in each firm, and then compute the weighted average of the ratios across each firm. Each step is weighted using the reported experienced member months. The model moments are constructed in the same manner.

To estimate firm average costs, this paper takes advantage of the firm's projected costs for the 2015 plan year. I use the projected firm level average cost from the 2015 plan year firm-level rate filing data. I compute post-reinsurance projected costs by subtracting projected reinsurance payments from "projected incurred claims, before ACA Reinsurance and Risk Adjustment."

Some firms do not appear in the risk filing data. For these firms, I compute the projected average cost for those firms by adjusting the experienced average cost reported in the Medical Loss Ratio filings by the average ratio of projected to experienced claims. In 2015, the average ratio of project to experienced claims for firms in my sample is 71.5%.

D Demand Estimation

The demand model closely follows that of Tebaldi (2023), with the difference that cost heterogeneity is modeled through this calibrated random coefficient on risk score rather than directly associated to the willingness to pay for insurance. This specification allows me to incorporate more data on the risk distribution of consumers to match heterogeneity in risk preferences across firms—via the firm fixed effect—in addition to the amount of insurance. In this section, I detail the calibration of the risk score distribution and the demand estimation methodology.

D.1 Risk Score Distribution

The risk scores in the demand model correspond to the output of the Health and Human Services Hierarchical Condition Categories risk adjustment model (HHS-HCC) used in the non-group market for the purpose of administering risk adjustment transfers. The HHS-HCC risk adjustment model is designed to predict expected plan spending on an individual based on demographics and health condition diagnoses. It is the result of a linear regression of relative plan spending on a set of age-sex categories and a set of hierarchical condition categories derived from diagnosis codes.

$$\frac{\text{Plan Spending}_{it}}{\text{Avg. Plan Spending}_t} = \gamma_0 + \sum_g \gamma_{tg}^{age,sex} Age_{ig} Male_{ig} + \sum_{g'} \gamma_{tg'}^{HCC} HCC_{ig'} + \eta_{it}$$

The prediction regressions are performed separately for different types of plans t, where t represents the metal category of the plan. The resulting risk score for an individual is a normalized predicted relative-spending value. Because all regressors take a value of either 1 or 0, the risk score is equal to the sum of all coefficients that apply to a particular individual.

$$r_{it} = \underbrace{\sum_g \gamma_{tg}^{age,sex} Age_g Male_g}_{r_{it}^{dem}} + \underbrace{\sum_{g'} \gamma_{tg'}^{HCC} HCC_{g'}}_{r_{it}^{HCC}}$$

Unless specifically noted, r_i^{HCC} will refer to the Silver plan HCC risk-score component and represent standard a measure of health status across all product types.

Parametric Distribution

The distribution of risk scores, $\hat{G}$, is estimated from the 2015 Medical Conditions File (MCF) of the Medical Expenditure Panel Survey. The MCF contains self-reported diagnosis codes and can be linked to demographic information in the Population Characteristics file. The publicly available data only list three-digit diagnosis codes, rather than the full five-digit codes. I follow McGuire et al. (2014) and assign the smallest five-digit code for the purpose of constructing the condition categories and matching the HHS-HCC risk coefficient.[3] See Appendix Section C.3 for detail on processing the data.

In the data, a majority of individuals have no relevant diagnoses, i.e., $r_i^{HCC} = 0$.In order to match this feature of the data, the distribution combines a discrete probability that an individual has a non-zero risk score and a continuous distribution of positive risk scores. With some probability $\delta(Z_i)$, the household has a non-zero risk score drawn from a log-normal distribution, i.e., $r_i^{HCC} \sim$ Lognormal$(\mu(Z_i), \sigma^r)$. With probability $1 - \delta(Z_i)$, $r_i^{HCC} = 0$. I allow the probability of having any relevant diagnoses and the mean of the log-normal distribution to vary by two age categories for the head of household and two income categories—above and below 45 years old, and above and below 400 percent of the federal poverty level.

Table A5 displays the moments of the risk score distributions for each metal level in the data.

[3]For example, I treat a three-digit code of '301' as '301.00'. McGuire et al. (2014) find that moving from five-digit codes to three-digit codes does not have a large effect on the predictive implications for risk score estimation. In this case, there is measurement error as the model used was originally estimated on 5-digit codes.

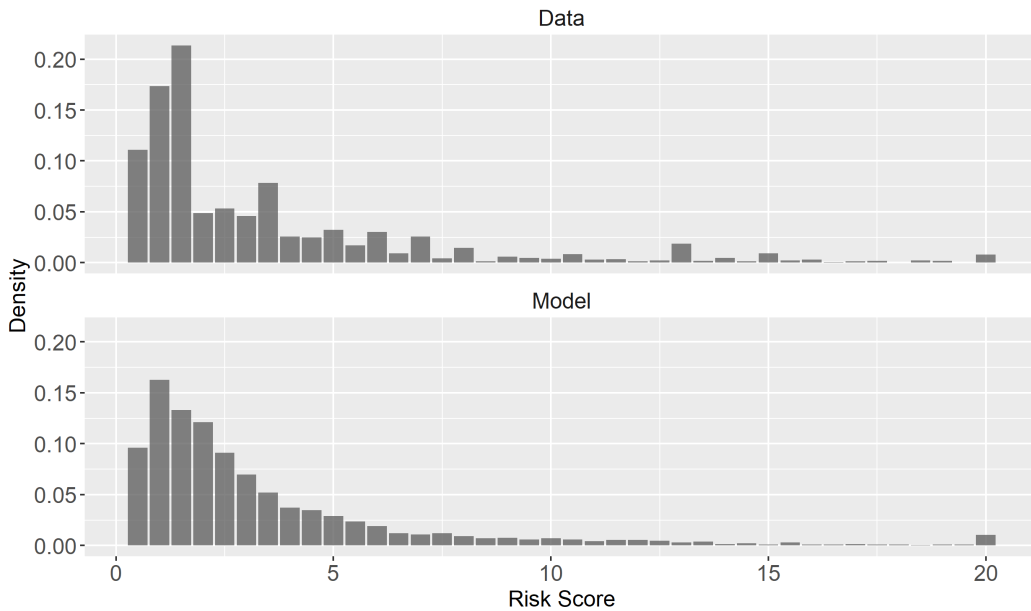


Figure A3: Risk Score Distribution Model Fit

Note: The data distribution comes from applying the HHS-HCC risk prediction methodology to the distribution of self-reported diagnoses in the 2015 Medical Conditions File of the MEPS. The model distribution comes from predicting the distribution of risk scores in the same MEPS sample. In both cases, the distribution of positive Silver metal-level risk scores are displayed.

Figure A3 compares the risk distribution in the MCF with the simulated risk distribution in the estimation sample.

Table A5: Parametric Distribution of Risk Scores

Age	Income		Bronze		Silver		Gold		Platinum	
	(% of FPL)	$\delta(Z_i)$	$\mu(Z_i)$	σ^r	$\mu(Z_i)$	σ^r	$\mu(Z_i)$	σ^r	$\mu(Z_i)$	σ^r
$\leq$45	$\leq$400%	0.15	2.86	19.7	2.99	19.5	3.11	19.8	3.31	20.8
	$>$400%	0.13	3.02	19.7	3.22	19.5	3.22	19.8	3.40	20.8
$>$45	$\leq$400%	0.31	3.49	19.7	3.73	19.5	3.73	19.8	3.97	20.8
	$>$400%	0.24	3.25	19.7	3.46	19.5	3.46	19.8	4.67	20.8

Note: This table displays three aspects of the distribution of HHS-HCC risk scores in the 2015 Medical Conditions File of the MEPS. The first column displays the portion of risk scores that are positive for four categories divided by age and income. The next columns display the mean and variance for each metal-level specific risk score. The mean depends on these same demographic groups, and the variance is calculated across the whole population.

D.2 Estimation and Identification

The demand model two primary identification concerns. First, a plan premium's price may be correlated with the unobserved quality ξ_{jm}, leading to biased estimates of α_i. In this environment, the premium regulations provide a source of variation in price, which is exogenous to variation in unobserved quality. The age-adjustment on premium, a_i, increases monotonically and non-linearly with age, and strictly increases with every age after 25. Income-based subsidies are available to households that earn below 400 percent of the federal poverty level. These subsidies decline continuously within the subsidy-eligible range. Moreover, the choice-set changes discretely at 250 percent, 200 percent, and 150 percent of the federal poverty level. At each of these income thresholds, the Silver plan becomes significantly more generous with no discontinuous change in the price (Lavetti et al. (2019)).

However, it is possible that preferences are also changing with age and income in a way that is hard to disentangle with the regulatory identification. I follow Tebaldi (2023) by augmenting this source of identification with variation in premiums due to local demographics (Waldfogel (2003)). The intuition is that areas with higher proportions of young consumers will have lower average costs and lower premiums. Conditional on the age of a consumer, the age of her neighbors does not affect her demand for insurance.

I implement the control function approach from Tebaldi (2023). In the first stage, I regress the base premium set for each product p_{jm} on the share of consumers in that market that are under the age of 35, and fixed effects for the firm, market, and metal-level of the product.

$$p_{jm} = \text{Share Under } 35_m + \gamma_{f(j)} + \gamma_m + \gamma_{metal} + \zeta_{jm}$$

The results of the first stage are presented in Appendix Table A1. I use the residual, $\widehat{\zeta_{jm}}$, in demand estimation as a control function. Specifically, I allow demand to depend on a 3rd order polynomial of the residual, with each term of the polynomial interacted with age group, as well as firm-market fixed effects. Appendix Table A2 includes demand estimates without the control function, which are qualitatively similar.

The second concern is the identification of the risk coefficients, $(\gamma_r, \{\beta_r^k\})$. These parameters are incorporated into the estimation equations in the same manner as variance parameters for distributions of unobserved consumer preferences (e.g. Berry et al. (1995)). However, because I have data on the distribution of risk scores in the market and moments on the average risk scores of individuals that choose certain products, I am able to incorporate these product-level moments to ensure that the model captures the appropriate risk-related substitution patterns and improve identification (Petrin (2002)).

The demand model targets eighty nine moments on the distribution of consumer risk scores: the average risk score of all insured consumers; the average risk score of enrollees in the Bronze, Silver, Gold, and Platinum plan categories; and the average risk score of each firm relative to the average risk score in the state for each firm-state combination in the data. Let l index the moments, let n index the $N = 500$ draws from the unobserved distribution of risk scores, and let $I(j)$ be the set of consumers that have product j in their choice set. For each group of products, J_l, I compute the moments as

$$M_l = \frac{\sum_{j\in J_l}\sum_{i\in I(j)}\sum_{n=1}^{N} w_i S_{inj} r_{inj}}{\sum_{j\in J_l}\sum_{i\in I(j)}\sum_{n=1}^{500} w_i S_{inj}} - R_m^{data}$$

where r_{inj} is a product-specific risk score draw to match the definition of the moments in the data and w_i is the weight consumer i (see Section 3.1 for more details on weighting).

To estimate the demand model, I follow Grieco et al. (2021) to combine a micro-data log-likelihood function with product-level GMM moments. The parameters maximize the sum of the log-likelihood of observed choices less the weighted moment objective value.

$$\hat{\theta} \in \text{argmax}_\theta \sum_i \sum_j Y_{ij} \log\left(\frac{1}{N}\sum_n S_{inj}\right) - M'WM \tag{12}$$

The estimation proceeds in two steps. First, I use the identity matrix as the weighting matrix,

W. Second, I set the diagonal of the weighting matrix equal to the inverse of the moment variances evaluated at the parameters estimated in the first stage. Because the moments do not apply to all consumers in the data, I cannot directly compute the moment variances. Instead, I follow Petrin (2002) by computing the variance of a separate set of moments that can be used to construct the intended moments for estimation. In this case, the predicted choice probabilities, S_{ij}, and the average predicted risk score for each product, $\frac{1}{N}\sum_n S_{inj} r_{inj}$ are sufficient. The variance of the targeted moments can then be computed using the delta method.

This estimation procedure is analogous to a GMM estimation that uses the first order conditions of the likelihood function as moments (Grieco et al. (2021)). Using the likelihood function in place of an additional set of moments allows the estimation to maintain the desirable convergence and identification properties of maximum likelihood estimation. However, to compute standard errors, I exploit the analogous GMM framework and compute the typical GMM standard errors where the weighting matrix is a block diagonal matrix with the Hessian of the likelihood function in one block and the moment weighting matrix W in the other.

E Cost Estimation

The method of simulated moments estimation procedure targets four sets of moments which each identify four sets of parameters. The age and risk parameters are identified using moments from the Medical Expenditure Panel Survey (Appendix Section C.3). For clear identification of costs by age separate from risk score, the estimation targets age moments among adults that have a risk score of zero. The moments are computed as the ratio of average covered expenditures within five-year age brackets for adults between 25 and 64 years old to the average covered expenditures of adults between 20 and 24 years old. The cost parameter on the risk score is identified using the ratio of average covered expenditures among adults with a positive risk score to those with a risk score of zero. This helps to separate sorting-related costs from firm-specific or product-specific costs. The parameter on actuarial value is identified using the ratio of experienced cost of each metal level to Bronze plans from the 2016 rate-filing data.

Conditional on these three cost parameters (ϕ_{AV}, ϕ_{Age}, ϕ_r), the firm-specific cost parameter, ϕ_f, is set to exactly match the projected average cost in the 2015 rate-filing data. See Appendix Section C.5 for more detail on the data.

When simulating moments that match data from the insurance firm rate filings, I use the reinsurance adjusted cost, c_{ijm}^{rein}. The moments from the Medical Expenditure Panel Survey are computed using total covered expenses across all insured individuals. Thus, I use the predicted cost c_{ijm} to compute these moments.

Cost is estimated using two-stage MSM to obtain the efficient weighting matrix. The estimated demand parameters are used to simulate the distribution of consumer age and risk scores throughout products in each market, using ACS data as the population of possible consumers (see Appendix Section C.2).

E.1 Identifying Assumption

The identifying assumption is that any unobserved cost variation is orthogonal to the idiosyncratic demand shocks. Given the specification of the demand function, this implies that the only mechanisms through which cost and preferences are correlated are through age and risk scores. If this assumption is violated and the remaining correlation is consistent with adverse selection, then the coefficient on actuarial value will be biased upward. For illustration, suppose I estimate $\hat{\phi}$ to solve for a single product and single observable type,

$$\frac{E[S_{ij}c_{ij}]}{S_j} - AC^{data} = 0$$

$$E[S_{ij}c_{ij}] = S_j AC^{data}.$$

This is equivalent to

$$S_j E[c_{ij}] - \text{cov}(S_i, c_{ij}) = S_j AC^{data}.$$

I assume that, conditional on age and risk score, this covariance term is 0. If there is an endogeneity problem consistent with adverse selection, this covariance term would be positive and increasing in plan generosity, leading to an upward bias in the estimated coefficient on adverse selection.

An alternative specification could treat expected total medical spending as a household characteristic. Then, I could allow preferences to vary with this characteristic instead of risk scores. Doing so has the advantage of circumventing this particular exogeneity assumption, but the principle concern that residual costs unobservable to the econometrician are correlated with demand errors would remain.

E.2 Reinsurance

In 2015, the ACA implemented a transitional reinsurance program, which mitigates a portion of the liability to insurance firms of very-high-cost enrollees. This policy was important in limiting the amount of realized adverse selection facing insurance firms and is included in cost estimation in order to match the post-reinsurance average firm costs. The federal government covered 45% of an insurance firm's annual liabilities for a particular individual that exceeded an attachment point, $\underline{c} = \$45,000$, and up to a cap, $\bar{c} = \$250,000$. For an individual with a cost c_{ijm}, the insurance firm is liable for the cost c_{ijm}^{rein} under the reinsurance policy.

$$
\begin{aligned}
c_{ijm}^{cov} &= \min\left(\max(c_{ijm} - \underline{c}, 0), \bar{c} - \underline{c}\right) \\
c_{ijm}^{exc} &= \max(c_{ijm} - \bar{c}, 0) \\
c_{ijm}^{rein} &= \min(c_{ijm}, \underline{c}) + 0.45 c_{ijm}^{cov} + c_{ijm}^{exc}
\end{aligned}
$$

E.3 Matching Firm Moments

Let $\bar{C}_f^{obs}$ be the observed projected firm-level average cost. The firm-specific cost parameters, $\tilde{\psi}(\phi)$, can be set such that these moments are matched exactly. Without incorporating reinsurance,

$\tilde{\psi}(\phi)$ can be computed analytically.

$$\bar{C}_f^{obs} = e^{\psi_f} \frac{1}{\sum_{j \in J^f} S_j} \sum_{j \in J^f} \int_i S_{ij} e^{\phi_1 AV_{jm} + \phi_2 Age_i + \phi_3 r_i^{HCC}} dF(i)$$

$$\tilde{\psi}_f(\phi) = \log \left(\frac{1}{\sum_{j \in J^f} S_j} \sum_{j \in J^f} \int_i S_{ij} e^{\phi_1 AV_{jm} + \phi_2 Age_i + \phi_3 r_i^{HCC}} dF(i) \right) - \log(\bar{C}_f^{obs})$$

When incorporating reinsurance, the parameters ψ can no longer be separated from ϕ because they interact in determining how much reinsurance an individual receives. Instead, $\tilde{\psi}$ can be found by iteration.

$$\tilde{\psi}_f^{n+1} = \tilde{\psi}_f^n + \left[\log \left(\frac{1}{\sum_{j \in J^f} S_j} \sum_{j \in J^f} \int_i S_{ij} c_{ijm}^{rein}(\psi_f, \phi) dF(i) \right) - \log(\bar{C}_f^{obs}) \right]$$

Without any reinsurance, this iteration method gives the analytic result at $n = 1$ given any feasible starting point, ψ^0. The reinsurance payments are not particularly sensitive to ψ which affects average payments and have less effect on the tails targeted by reinsurance. As a result, $\tilde{\psi}$ can be precisely computed with a small number of iterations.

E.4 Method of Simulated Moments

I will write the moments as $d(\phi)$ to represent the remaining moments on the cost ratios by metal level, age, and risk, incorporating the predicted parameters of $\tilde{\psi}(\phi)$. $\hat{\phi}$ is estimated by minimizing, for a weighting matrix W,

$$\hat{\phi} = \text{argmin}_{\phi} \; d(\phi)' W d(\phi)$$

The minimum of the function is found using the Neldermead method. I estimate $\hat{\phi}$ in two stages. In the first stage, I use the identity weighting matrix and obtain estimates of the variance of the moments, V. In the second stage, I use $W = V^{-1}$. Similar to the demand estimation, the moments do not necessarily apply to every observation of the data. I use the same procedure from

Petrin (2002) to compute the variance of the moments.

E.5 Model Fit

Table A6: Cost Estimation Fit of Cost-Ratio Moments

	Data	Model Fit	
		GMM-1	GMM-2
Age ($r^{HCC} = 0$)			
18 - 24	1.0	-	-
25 - 29	1.34	1.32	1.30
30 - 34	1.44	1.53	1.56
35 - 39	2.08	2.38	2.35
40 - 44	2.98	2.08	2.07
45 - 49	1.74	2.56	2.57
50 - 54	3.49	2.74	2.84
55 - 59	2.98	3.75	3.78
60 - 64	3.57	3.75	3.82
Risk Score			
$r^{HCC} = 0$	1.0	-	-
$r^{HCC} > 0$	3.57	3.26	3.29
Metal Level			
Bronze	1.0	-	-
Silver	2.28	1.67	1.77
Gold	3.80	3.39	3.41
Platinum	4.28	7.39	7.47

Note: This table displays the targeted and estimated cost ratios that are used to identify the marginal cost estimation. In each category—age, risk score, and metal level—the ratios are defined relative to the first row. The first row of each category is equal to one by construction. The two columns of estimated moments represent the two demand estimation specifications used to simulate the moments. Marginal costs are not estimation for the final specification, GMM-3, since this specification cannot be used in counterfactual analyses.

Table A6 presents the targeted and estimated moments used in the cost estimation. The age and risk moments are matched more closely than the metal-level ratio moments. In particular, the cost specification leads to overestimates of the cost of covering individuals with Platinum coverage. The combination of ordered risk preferences, age preferences, and log-linear costs in actuarial value lead to the implication that the difference in average costs among expensive and generous plans (Gold and Platinum) is much greater than the difference in average cost among the less comprehensive options (Silver and Bronze).

In estimating the parameters of demand and marginal cost, I do not use the assumption that firms are optimally setting Nash-Bertrand prices to maximize profit. This approach allows the demand and cost parameters to be identified from data, and shielded from potential model misspecification. In general, the demand model implies greater markups than the cost estimation. The median and mean implied markup from the demand estimation is 43 and 48 percent, respectively. The mean and median implied markups from the cost estimation are 34 percent and 28 percent, respectively. Appendix Figure A1 plots the marginal revenue and marginal cost implied by estimated parameters under the baseline policy regime, which includes risk adjustment and reinsurance.

The fact that the demand model and Bertrand-Nash competition imply greater markups could be an indication that state insurance agencies are successful in negotiating lower markups on behalf of consumers. This mechanisms are outside of the scope of this paper. To the extent that regulators can effectively discipline markups, the results that follow will underestimate the positive effects of consolidation.

F A Model of Risk Adjustment in the Affordable Care Act

The ACA includes a risk adjustment transfer policy specifically intended to mitigate between-firm adverse selection. The government administers a transfer between firms that is equal to the difference between the firm's own average cost and the implied average cost of the firm if it were to insure the same risk balance as the market as a whole (Pope et al. (2014)).[4] (For more details on the policy specifics, see Section 3.)

[4]The implemented policy has to approximate this transfer using a risk-scoring system, but I will assume for theoretical simplicity that the regulator has full information about consumer risk.

$$T_j(\boldsymbol{p}) = \underbrace{\frac{E[\sum_k S_{ik} c_{ik}]}{E[\sum_k S_{ik}]}}_{\text{Pooled Cost}} - \underbrace{\frac{E[S_{ik} c_{ij}]}{E[S_{ij}]}}_{\text{Average Cost}}$$

In the presence of risk adjustment transfers, the firm then faces a new average cost, $AC_j^T(\boldsymbol{p}) = AC_j(\boldsymbol{p}) - T_j(\boldsymbol{p})$. The equilibrium price can be written as

$$p_j^* + \frac{S_j}{S_j'} = \Psi_j \frac{E\left[(\sum_k \frac{\partial S_{ik}}{\partial p_j}) c_{ij}\right]}{\sum_k \frac{\partial S_j}{\partial p_j}} + \left(1 - \Psi_j\right) \frac{E[\sum_k S_{ik} c_{ik}]}{\sum_k S_k} \tag{13}$$

where,

$$\Psi_j = \frac{S_j}{\sum_k S_k} \frac{\sum_k \frac{\partial S_k}{\partial p_j}}{S_j'}$$

There are two important features of equilibrium under risk adjustment. First, the transfers adjust the private incentive of the firm according to how the marginal cost of its products deviates from the market-wide average cost. The policy-induced incentive is not the optimal sorting incentive in Equation (3) that penalizes or reward firms based on the profitability of their marginal consumers. Therefore, it does not eliminate the welfare cost of sorting.

Second, this particular policy converges to the firm's own private incentive as the market share of a particular product increases or if one firm merges with others in the market. The policy follows the importance of the sorting distortion by fading out with market concentration.

G Effects of a Merger under Price-Linked Subsidies

Price-linked subsidies have two important effects in the context of mergers and market power. First, it allows firms a greater ability to exploit their market power. Greater prices are partially covered by the government rather than consumers, which reduces consumers' effective elasticity and leads to greater markups (Jaffe and Shepard (2020)).

Second, when two firms merge, the price effect is greater not only due to the reduced elasticity of consumers, but also due to the increased probability that the merged firm will control the linked product that governs the subsidy. The merged firm now internalizes more of the subsidy policy, and as a result, it is as if the merged firm faces *less elastic* consumers than pre-merger, even at the same prices and among the same consumers.

Without considering selection, the first-order effects of a merger when firms internalize the price-linked nature of subsidies are that government spending increases substantially, firms capture some of this increased spending as an increase in profits, and subsidized consumers are protected against—and in some cases can benefit from—higher prices. The key group that is harmed due to a merger are higher-income, un-subsidized consumers, which make up a smaller portion of the market.

These effects are important in the context of this paper, because they greatly reduce the harm to consumers from increased markups. Consumers may benefit from mergers through less inefficient sorting without bearing the full burden of greater markups.

In this section, I repeat the main results of the paper, allowing for the subsidies to adjust with prices and for firms to internalize their probability of controlling the price-linked product, i.e. the silver plan with the second lowest price. I follow the methodology of Jaffe and Shepard (2020) and require the equilibrium to be an ex-post best-response. The firm that controls the second-lowest-price silver plan sets the optimal price conditional on the knowledge that the plan is linked to the market-wide subsidy level.

In order to smooth the computation of equilibrium, I assume that firms have an expectation over the probability that the silver plan they offer in each market is the benchmark silver plan. Let p^{2lps} represent the second lowest-price silver plan. All silver plans in the market are assigned a probability that the plan is the benchmark plan,π_j, given by

$$\pi_j = \frac{e^{-\chi|p_j - p^{2slps}|}}{\sum_k e^{\chi|p_k - p^{2slps}|}}. \tag{14}$$

The parameter χ governs the certainty with which firms' know if they offer the benchmark premium. In the limiting case of a very large χ, this probability distribution collapses to certainty. In the results in this section, I set $\chi = 0.1$, which corresponds roughly to a firm knowing with 99% probability that its plan is the benchmark silver plan if the absolute price difference of the next closest silver plan is at least \$40. At the observed prices, the benchmark plans in 53 out of 107 markets are assigned probabilities greater than 70%, and in 88 markets the probabilities exceed 50%. With a greater certainty parameter, the equilibrium is more difficult to solve but it does not substantially alter the results of this section.

The price-linked model performs similarly in rationalizing the observed equilibrium. As shown in Figure A1, average marginal revenue and average marginal costs are similar in the baseline model used in the body of the paper. The price-linked model has similar findings, but the marginal revenue variation matches slightly less of the marginal cost variation—33% versus 37%.

Table A7: Many Mergers are Predicted to Improve Consumer Surplus and Social Welfare

	Number of Mergers	Fraction $\Delta CS > 0$	Fraction $\Delta SW > 0$ (Govt Spending Excl.)	Fraction $\Delta SW > 0$ (Govt Spending Incl.)
			Baseline	
Total	1186	0.53	0.83	0.07
Δ HHI				
$<$100	533	0.74	0.91	0.08
100 - 200	144	0.53	0.87	0.05
200 - 1000	319	0.36	0.77	0.07
$>$1000	190	0.21	0.66	0.03

Note: When accounting for price-linked subsidies, mergers are generally beneficial to consumers. This table displays the fraction of mergers with positive welfare effects. The top line displays the average across all mergers, and the following rows breakout the results by the size of the merger. The change in HHI is computed using pre-merger market shares to reflect pre-merger size of merging firms.

Table A7 displays the main results in the price-linked model. In both policy environments, more than half of all mergers are beneficial to consumer, and more than 1 out of 5 of the largest mergers are still beneficial to consumers.

I present two total welfare measures: one that excludes government spending (as in the body of the paper) and another that includes spending. Because the primary costs of a merger in this environment are borne by the government, the vast majority of mergers increase the combined surplus accrued by consumers and firms. In the case that a dollar of government spending is associated with extra resource costs due to distortionary taxes, it may be the case that no mergers produce any welfare benefit. Similarly, if a dollar of government spending in this market is valued at less than a dollar due to preferences for redistribution, the fraction of mergers that produce a welfare benefit may be somewhere in between these two estimates.[5]

In these results, it is hard to disentangle the mechanisms created by adverse selection from those caused by the price-linked subsidies. Even in the absence of any adverse selection, mergers that lead to greater subsidies can be beneficial for both consumers—many prices fall in absolute terms due to greater subsidies—and for firms—due to greater profit at costs borne primarily by the government. These dynamics are important considerations for this market but outside of the primary scope of this paper.

[5]As mentioned in Section 4.4, the total surplus generated by an additional dollar of government spending is less than a dollar.