Sorting on Plan Design:
Theory and Evidence from the ACA

Chenyuan Liu*

Tsinghua University

Click Here for the Latest Version

Abstract

Health insurance plans often differ in coverage levels and the combinations of cost-sharing attributes to achieve that level. In this paper, I show that the proliferation of plan designs can result from distortion under asymmetric information. Though optimal risk protection requires concentrating coverage in large loss states (i.e., straight-deductible plans), low-risk types signal by sorting into plans with more coverage for smaller losses. Standardizing plans to vary only along a single dimension may exacerbate welfare loss from asymmetric information. Consistent with the model, I show that a large variation in plan designs exists in the ACA federal exchange and that straight-deductible plans attract individuals with significantly higher ex-post medical spending and ex-ante risk scores. I calibrate the potential welfare effects of standardizing plan designs in the ACA when asymmetric information and consumer confusion exist. *JEL Codes*: D82, G22, I13.

*Department of Economics, School of Economics and Management, Tsinghua University (e-mail: liuchy3@sem.tsinghua.edu.cn). I thank Naoki Aizawa, David Anderson, Thomas DeLeire, Ben Handel, Ty Leverty, David Molitor, John Mullahy, Casey Rothschild, Joan Schmit, Dan Sacks, Mark Shepard, Alan Sorensen, and Justin Sydnor for their generous comments. Financial support from the Geneva Association is gratefully acknowledged. All errors are my own.

There is a growing policy debate on why and when to offer choices in insurance markets. One area of policy attention is the design of health insurance plans' cost-sharing attributes. A typical health insurance plan often has a deductible, coinsurance rate, maximum out-of-pocket, etc. The combination of these cost-sharing attributes leads to vertically and horizontally differentiated contracts: plans differ in their coverage level (fraction of losses covered) and their cost-sharing designs to achieve that coverage level. Motivated by findings that consumer confusion is prevalent (Abaluck and Gruber 2011, 2023; Bhargava, Lowenstein, and Sydnor 2017), some health insurance markets choose to simplify plan choices into only vertically ranked choices. For example, in the Netherlands and Switzerland's health insurance markets, plans are vertically ranked by the deductible level.[1] In some state-based ACA exchanges (e.g., California), a single design is allowed per coverage level. Whether these regulations improve efficiency depends on how consumers evaluate and sort along different designs. If more complex cost-sharing attributes provide financial value to certain individuals, removing them may reduce social surplus.

In this paper, I develop a conceptual framework to show that the proliferation of plan designs can result from distortion under asymmetric information. The model setup follows Rothschild-Stiglitz (1978), where insurers use different cost-sharing rules to screen individuals with unknown risk types. Two key differences exist between my model and previous models: first, I allow individuals to have multiple loss states, a common feature in health insurance markets. Second, insurers offer plans with multiple cost-sharing attributes. My model puts no restriction on the cost-sharing designs: contracts vary by the covered losses in each state and may not be vertically ranked. I also assume plans have fixed, positive loading.

My model predicts that different risk types sort into different cost-sharing designs under asymmetric information. In a Rothschild-Stiglitz style separating equilibrium, the high-risk type sorts into their first best plan. Given positive loading, the first-best plan for the high-risk type is less than full insurance. Optimal risk protection requires concentrating

[1] In the Netherlands health insurance markets, plans have a single deductible and no coinsurance rates, making plans vertically ranked by the deductible level. In the Switzerland health insurance market, plans are vertically ranked by six deductible levels, followed by 10% coinsurance rates and a common cap of the out-of-pocket co-insurance amount.

coverage in larger loss states, so such a plan has a straight-deductible, a classic result from Arrow (1963). Under the straight-deductible plan, individuals pay full losses below the deductible and are fully insured once they reach the deductible level. The low-risk type distorts their coverage to avoid pooling with the other type. Asymmetric information creates a force that pushes lower-risk consumers to choose plan designs with more coverage for smaller losses (in the form of coinsurance and lower deductible) while forgoing coverage on larger losses (in the form of higher MOOP).[2] In summary, the equilibrium plan desired by the high-risk type has a straight-deductible design, while the plan desired by the low-risk type has a lower deductible, some coinsurance, and a higher MOOP. I demonstrate that this theoretical prediction holds both in an unregulated competitive separating equilibrium and regulated markets with perfect risk adjustment.

My model predicts that restricting plan designs to be vertically ranked may create large welfare losses. In unregulated competitive markets, plan design variation—specifically, the existence of plans with low deductibles and high MOOP—helps sustain a more efficient separating equilibrium. When consumers can sort along only one dimension of cost-sharing (i.e., deductibles), low-risk individuals sacrifice substantially more coverage to avoid pooling with higher-risk individuals. When there is perfect risk adjustment, restricting plans to be only straight-deductible plans also reduces the surplus of the low-risk type. However, because under risk adjustment, the marginal costs of insurance to the individual might differ from the social costs, the impacts on the overall social surplus are ambiguous.

In the second part of the paper, I examine the empirical relevance of sorting by plans launched in the Affordable Care Act (ACA) Federal Exchange (healthcare.gov), a market with risk-adjustment regulations. I combine publicly available data on the cost-sharing attributes, premiums, enrollment, and claims costs for plans launched between 2014 and 2017 in this market. The ACA Federal Exchange organizes plans into four "metal tiers" based on the level of coverage they provide for a benchmark average population: Bronze (60%), Silver (70%), Gold (80%), and Platinum (90%). Within these tiers, insurers have

[2] The sorting result relies on the insight that low-risk individuals signal themselves by accepting less coverage in the states they are less likely to experience. This insight is also documented by theoretical works studying other selection markets, including the English annuity markets (Rothschild 2007; Finkelstein, Poterba, and Rothschild, 2009) and bundled coverage for property and casualty insurance (Crocker and Snow, 2011).

significant latitude in designing the cost-sharing attributes of their plans in different combinations.

I use this empirical setting to examine two predictions from the model. First, in a market with heterogeneity in risk distributions and limited regulation in plan designs, there will be a proliferation of plans with different cost-sharing designs. Indeed, a large variation exists in plan designs in the ACA Exchange. For example, the within-county variation in the 2017 Silver deductible is more than $3,000 for half of the counties. Though previous models can also explain the variation in coverage levels, my model helps rationalize the fact that there are often multiple cost-sharing designs within and across coverage levels.

Second, variation in plan design creates room for sorting by risk type in the ACA market. My theoretical model predicts that plans with straight-deductible designs will be attractive to those with average to above-average risk but unattractive to lower-risk consumers. Using plan-level claims costs and insurer-level risk transfer information, I find that individuals enrolled in straight-deductible plans or similar designs have significantly higher ex-post medical expenditure, and insurers offering straight-deductible plans receive significantly larger risk transfers. Other confounding factors cannot fully explain these differences, including moral hazard, plans' provider network, health savings account (HSA) eligibility, and geographic variation in plan availability.

The theory and empirical analysis highlight how asymmetric information in risk types can explain the variation in plan designs. However, moral hazard is another rationale for the existence of non-straight-deductible plans. Theoretical research has shown that moral hazard can affect the optimal plan design, changing either the deductible level or the form of coverage (Zeckhauser 1970). Although models with moral hazard can help explain why plan designs are complex, they offer no ready explanation for the simultaneous existence of different plan designs. Empirically, my results using risk scores illustrate that the expenditure differences within an ACA coverage tier are mainly driven by selection and cannot be explained by moral hazard alone. Interesting dynamics might be at play when incorporating moral hazard responses into my model. Those considerations are outside the scope of this paper but could be a valuable direction for future research.

In the last part of the paper, I calibrate the likely impacts of simplifying plans' cost-sharing into vertically-ranked options in the ACA Federal Exchange. Specifically, I compare the market outcome under two menus: The actual 2017 plans offered in the ACA Exchange and a hypothetical choice set replacing all options with a straight-deductible plan of the same premium. I assume consumers have different risk distributions (calibrated using the Truven MarketScan data) and allow a fraction of "behavioral types" who randomly pick plans available in the choice set instead of choosing the plan maximizing their expected utility. The numeric exercise highlights the following trade-off: Restricting plans to be straight-deductibles reduces the chance that the behavioral high-risk types choose the wrong plan; however, that also removes valuable options for the low-risk types and might hurt them. The aggregate impact depends on the fraction of these different types.

I estimate that when there are no behavioral types, the overall efficiency of the ACA would be only slightly higher ($10 per person per year) with regulated plan designs. The increase in the higher-risk types' surplus because of the availability of straight-deductible plans is largely offset by the decrease in the lower-risk types' surplus. However, when there are more behavioral types, the benefits to the higher-risk types dominate. This is because plans with high out-of-pocket limits create the possibility of a costly mistake for higher-risk consumers, who are disproportionately adversely affected by such plans. I show that the efficiency benefits of regulating plan design in the ACA Exchange are significantly higher if a moderate share of consumers makes plan-choice mistakes.

This paper contributes to a growing body of literature studying optimal menu design in selection markets. The existing literature examines why and when to offer vertical choices instead of a single mandated option, highlighting economic forces including adverse selection, moral hazard, and consumer confusion (Ericson and Sydnor, 2017; Marone and Sabety, 2022; Chade et al., 2022; Ho and Lee, forthcoming). All these works model plans' financial attributes as vertically ranked options.[3] I supplement these works by highlighting an under-appreciated consequence of adverse selection: when heterogeneity exists in the likelihood of incurring larger and smaller losses, simplifying plan designs into vertically-ranked options puts restrictions on the market equilibrium and may reduce welfare. The

[3] Modeling contracts' cost-sharing designs as vertical choices is sufficient when individual faces a binary loss, because plans can only differ by the fraction of losses covered for this single loss state. Such a setting is natural for some selection markets, e.g. unemployment insurance, while is an abstraction for other markets, e.g. health insurance market.

finding provides a new angle in evaluating the plan standardization policies in health insurance markets.

More broadly, the paper contributes to the literature studying sorting by quality under asymmetric information (Veiga and Weyl, 2016). Existing literature documents, both theoretically and empirically, that adverse selection and contract distortion happen along dimensions like coverage generosity for different medical services and providers (Frank, Glazer, and McGuire 2000; Ellis and McGuire 2007; Shepard 2016), drug formulary (Lavetti and Simon 2014; Carey 2016; Geruso, Layton, and Prinz 2019), and overall plan generosity along a single dimension (Decarolis and Guglielmo, 2017). My paper highlights that a similar adverse selection force applies to the distortion in cost-sharing designs: the need for low-risk types to avoid pooling with the high-risk types creates demand for multiple plan designs. My model also explains sorting patterns in empirical findings in other markets.[4]

Finally, my conceptual framework connects two long-standing theoretical literature: optimal risk protection with multiple loss states (Arrow 1963) and plan distortion under asymmetric information (Rothschild and Stiglitz 1976). My model extends Arrow (1963) by allowing for asymmetric information in loss distributions and extends Rothschild and Stiglitz (1976) by allowing for multiple loss states. The insight that asymmetric information distorts plans' multi-dimensional cost-sharing attributes illustrates a new mechanism shaping the complex cost-sharing attributes in insurance markets. Previous studies find that moral hazard (Pauly 1968; Zeckhauser 1970), nonlinear loading factors or risk-averse insurers (Raviv 1979), background risk (Doherty and Schlesinger 1983), and liquidity constraints (Ericson and Sydnor 2018) can lead people to select into various types of plan designs. My model shows that adverse selection may also create a proliferation of plan designs.

The rest of the paper is organized as follows: In Section 2, I lay out the conceptual framework and derive the conditions leading to design distortion. In Section 3, I examine

[4] Decarolis and Guglielmo (2017) documented that 5-star Medicare Part C plans increase MOOPs and decrease deductibles in the face of the pressure of worsening risk pools. This paper's conceptual framework predicts that the incentives to attract low-risk types can drive this movement towards non-straight-deductible plans.

the issue empirically using the ACA Federal Exchange data. In Section 4, I discuss the implications for regulating plan designs. The final section concludes.

2 Conceptual Framework of Optimal Plan Design

In this section, I present a stylized model of insurance markets where individuals have hidden information about their loss distributions. A key difference of my model from previous works (Rothschild and Stiglitz, 1976) is that I assume the losses are not binary and can take on multiple values. I show that under asymmetric information (community rating), individuals with different loss distributions sort into different cost-sharing designs. I then illustrate the welfare implications of policies removing such cost-sharing complexities and restricting to a single design. All proofs are in Appendix A.

2.1 Model Setup

The market consists of two risk types, L and H with equal population size. Each type i faces uncertainty in their medical expenditure x_s in state s. The realization of $s \in S$ is uncertain, with state s obtaining with probability f_s^i for individual i.

I consider a general state-dependent insurance plan that captures the wide range of potentially complex plan designs consumers could desire. Specifically, an insurance plan is defined as a function mapping loss states to non-negative real number: $\boldsymbol{l}: s \rightarrow R_0^+$, where $l_s \equiv l(s)$ is the value of the function evaluated at s, and l_s represents the insurer payment in state s. l_s satisfies the condition $0 \leq l_s \leq x_s$.

DEFINITION 1 *(Straight-Deductible Plan): A straight-deductible plan with a deductible of* d *is defined as:*

$$l(x_s) = \begin{cases} 0, if\ x_s \leq d, \\ x_s - d, if\ x_s > d. \end{cases}$$

Under such plans, individuals pay full losses out-of-pocket below the deductible level and get full insurance once the losses reach the deductible level. Full insurance is a straight-deductible plan with zero deductible. All other plans are non-straight-deductible plans.

The financial outcome (consumption) after insurance in each loss state is $w_i - x_s + l_s - p(\boldsymbol{l})$, where w_i is the non-stochastic initial wealth level and $p(\boldsymbol{l})$ represents the premium of plan $\boldsymbol{l}$. I assume individual i has a concave utility function u_i over the financial outcome of each loss state: $u_i' > 0, u_i'' < 0$. Individuals are offered a menu of contracts C and choose the plan maximizing their expected utility:

$$\max_{\boldsymbol{l} \in C} \sum_s u_i(w_i - x_s + l_s - p(\boldsymbol{l})) f_s^i. \tag{1}$$

2.2 Model Predictions

We now turn to predictions from the model. I consider three cases: symmetric information, asymmetric information with no risk adjustment, and asymmetric information with perfect risk adjustment.

Case 1 - Symmetric Information/Risk-Based Pricing

For this single-risk-type case, I drop subscript i for simplicity of exposition. Assume perfectly competitive insurers set premiums as a linear function of the expected covered expenditure:

$$p(\boldsymbol{l}) = \theta \sum_s f_s l_s + c. \tag{2}$$

where $\theta \geq 1$ is a proportional loading factor, and $c \geq 0$ is a fixed loading factor. Suppose further that all possible insurance contracts are available and priced this way.

PROPOSITION 1. *Under risk-based pricing, for any fixed loading factors, the contract maximizing expected utility is a straight deductible plan.*

The result is a direct application of Arrow (1963) and Gollier and Schlesinger (1996). When there is no loading, the optimal contract will be full insurance. When there is positive loading, the expected-utility-maximizing contract has some cost-sharing. Proposition 1 states that such a contract has a straight-deductible design: optimal risk protection requires that coverage is concentrated on larger losses.

Case 2 - Asymmetric Information/Community Rating

Now consider the case where there are two risk types (L and H) in the market, and insurers cannot distinguish L from H ex-ante, or they cannot charge different premiums for the same plan because of community rating regulations. The plan premiums are a mechanical function of the expected covered losses given who sort into that plan, plus loading.[5]

[5] The assumption rules out equilibrium concepts with cross-subsidization among plans (as in Spence 1978).

A key component of the model is how L and H are defined. It is not the purpose of the model to fully characterize that equilibrium. Instead, I consider a potential separating equilibrium similar to Rothschild and Stiglitz (1976), where one risk type (H) gets the first-best contract under symmetric information, and the other type (L) distorts their coverage to prevent the higher-risk type from pooling with them. H and L types are defined such that the incentive compatibility constraint is constrained for H and slack for L. In equilibrium, H sorts into the first-best plan, which, according to Proposition 1, is a straight-deductible plan. L chooses the incentive-compatible plans that maximize the expected utility. I further assume that both types face multiple loss states, and there exist at least two non-zero, positive coverage loss states, s and t, where $x_s \neq x_t$ and $f_s^L / f_s^H \neq f_t^L / f_t^H$.

PROPOSITION 2. *Among all incentive-compatible plans for H, the one that maximizes the expected utility of L has a non-straight-deductible design.*

The intuition can be illustrated starting from the first-best plan for L, which has a straight-deductible design. Such a plan will not be incentive compatible, however, because it is priced based on the loss distribution of L, and makes H deviate. Therefore, L needs to change their coverage to prevent pooling with H. They could achieve this by either reducing coverage for larger loss states or reducing coverage for small loss states. Doing the former would make the plan less attractive to H since larger losses are more likely to happen for H. Sacrificing coverage for large losses and transferring to coverage for small losses is less problematic for L, though, since most of their losses are likely to be small.

Case 3 - Asymmetric Information/Community Rating with Perfect Risk Adjustment

In many markets, regulators impose risk adjustment regulations to flatten premium differences among plans and to remove screening incentives for insurers. I consider a market with perfect risk adjustment where the premium reflects the market average risk and is a linear function of the expected costs obtained if both risk types enroll in the plan.[6] This setting approximates the regulatory environment in many US health insurance markets, including Medicare Advantage, Medicare Part D, and the ACA Exchange.

[6] This definition is a special case of Einav, Finkelstein and Tebaldi (2018), which defines risk adjustment as a transfer r_i to the insurer if individual i enrolls in the plan. My setting is equivalent as setting r_i as the difference between the cost of insuring that type, $\theta \sum_s f_s^i l_s$, and the market average cost. Geruso et al. (2023) also uses the same formula to define perfect risk adjustment.

Under perfect risk adjustment, the premium is:

$$p(\boldsymbol{l}) = \frac{\theta}{2}\left(\sum_s f_s^L l_s + \sum_s f_s^H l_s\right) + c. \tag{3}$$

To obtain the sorting result, I assume that H and L have loss distributions with monotone likelihood ratio property in losses: for any two loss states s and t where $x_t > x_s$, it is also true that $\frac{f_s^L}{f_s^H} > \frac{f_t^L}{f_t^H}$. Further, assume that there exist at least two non-zero loss states for both types. I also assume that the feasible plans imply non-decreasing out-of-pocket spending when the loss increases: $x_s - l_s \leq x_t - l_t, \forall x_s < x_t$. This is a common feature for health insurance plans because the losses are cumulative within a year.

PROPOSITION 3. Under perfect risk adjustment, H sorts into a straight-deductible plan; L sorts into a non-straight-deductible plan.

Under perfect risk adjustment, the premiums are effectively "shared" between the two types. The marginal cost of reducing out-of-pocket spending depends on the spending of both types. Ideally, both types want to have the premium covering more of their own spending than the spending of the other type. The utility-maximizing plans for each type thus direct more coverage into states where that type is relatively more likely to experience. Since H is more likely to experience larger losses, they sort into straight-deductible plans, which offer full coverage for large losses. The opposite is true for L.

The proposition can be extended to a scenario where both risk types are choosing from plans with the same premium:

COROLLARY 1. Under perfect risk adjustment and among all plans have the same premium, H sorts into a straight-deductible plan; L sorts into a non-straight-deductible plan.

In summary, the complexity of plan designs can be motivated by multiple loss states and asymmetric information (community rating). Under perfect information, all types desire straight-deductible plans. Under asymmetric information, L has an incentive to deviate to non-straight-deductible designs.

2.3 Implications for Plan Standardization Regulation

The sorting result presented in 2.2 has implications for evaluating the welfare impacts of the plan standardization policy. To illustrate, I first define the welfare notion as follows.

The consumer surplus of individual i choosing plan $\boldsymbol{l}$, cs_{il}, is defined as the certainty equivalent of choosing plan $\boldsymbol{l}$ relative to no loss:

$$\sum_s u_i(w_i - x_s + l_s - p(\boldsymbol{l}))f_s^i = u_i(w_i + cs_{il}). \tag{4}$$

The social surplus of individual i choosing plan $\boldsymbol{l}$, ss_{il} is defined as:

$$\sum_s u_i\big(w_i - x_s + l_s - \tau_i(\boldsymbol{l})\big)f_s^i = u_i(w_i + ss_{il}), \tag{5}$$

where $\tau_i(\boldsymbol{l}) = \theta \sum_s f_s l_s + c$, is the social cost of offering plan $\boldsymbol{l}$ to individual i. Note that when there is no risk adjustment, $\tau_i(\boldsymbol{l}) = p(\boldsymbol{l})$, and cs_{il} is the same as ss_{il}. Under risk adjustment, this relation is, in general, not true. The overall social surplus is defined as the sum of ss_{il} for all individuals in the market.

First, consider a plan standardization regulation that restricts all plans to vary along a single dimension and are vertically ranked. A natural policy is to restrict plans to straight-deductible plans (as in the Netherlands health insurance markets). Under asymmetric information, *H* sorts into a straight-deductible plan, so they are unaffected. However, any straight-deductible plan *L* chooses under the design regulation makes them strictly worse off. When there is no risk adjustment, social surplus coincides with consumer surplus, implying a decrease in social surplus. Under perfect risk adjustment, however, exactly how the social surplus will change is ambiguous: perfect risk adjustment imposes an externality in the pricing because the marginal costs of the extra cost-sharing are shared by the other risk type. Restricting to straight-deductible plans may or may not reduce such externality, so the social surplus may or may not improve.

The social and consumer losses from such a regulation can be sizable. Consider the following numeric example. Two risk types are constructed using the 2013 Truven MarketScan data, where *L* has a mean spending of \$1,843 (SD=\$7,414), and *H* has a mean spending of \$7,537 (SD = \$22,444). I assume both types have CARA utility function and a risk aversion level of 0.0004. I then calculate the equilibrium plans under 1) risk-based pricing and 2) community rating with no risk adjustment. The details of the calculation are in Appendix B.

Table 1 shows the numeric example. Under community rating, L sorts into a coinsurance plan with a 23% self-paid coinsurance rate, which causes a \$561 welfare loss relative to the first-best plan. It happens to have the same fraction of losses covered as the first-best plan, so the welfare losses of community rating purely come from the design distortion. Note that the equilibrium plans under no regulation are not vertically ranked: a constant coinsurance plan with 23% coinsurance rates offers more coverage for smaller losses than the straight-deductible plan with a deductible of $1,820.

However, if L is forced to choose a straight-deductible plan, they sort into one with a $13,154 deductible, which offers much less coverage and the surplus reduction is more than doubled.

Table 1. Numeric Example: Community Rating and Design Regulation

		Risk Type	Plan	% losses covered	Surplus
Risk-Based Pricing		H	Straight-deductible, deductible = $1,820	82%	/
		L	Straight-deductible, deductible = $933	77%	0
Community Rating and No Risk Adjustment	No Regulation	L	Constant coinsurance, coinsurance rate = 23%	77%	-$561
	Straight-Deductible Only	L	Straight-deductible, deductible = $13,154	23%	-$1,256

Note: Risk types constructed from Truven MarketScan data. "Risk-based pricing" refers to the scenario where each plan is priced based on the risk type choosing it, and the premium for the same plan can vary for different risk types. "Community rating" refers to the scenario in which insurers cannot vary premiums for the same plan for different risk types, and the premium is a linear function of the expected spending of the risk type choosing the plan. "Straight-deductible only" refers to the scenario in which only straight-deductible plans are available. Surplus refers to either consumer or social surplus, as they are the same in this case. I rescale them so the value is the difference from the first-best plan.

Second, consider a market with perfect risk adjustment. In the market, a coverage level is defined as the fraction of losses covered for the average population. Plans with the same coverage level thus have the same premiums under perfect risk adjustment. The market regulates that there can be at most one plan design within a coverage (premium) level, as in the case of the California ACA Exchange. Given two risk types, the regulation is optimal only when the specified design coincides with the socially optimal design for each type. Generally, it is not obvious whether such a policy will improve or reduce welfare.

Moreover, such a regulation may reduce consumer surplus under certain conditions. Suppose L is more risk averse than H, such that under perfect risk adjustment, the plan desired by both types has the same premium. According to Proposition 3, however, the two types desire different plan designs. Table 2 shows a numeric example where, under perfect risk adjustment and no plan regulation, both types prefer plans with around $6,200 premium under perfect risk adjustment, while only H chooses a straight-deductible design. When this is the case, restricting to a single design per coverage level will at least make one of the risk types worse off than if all designs are allowed.

Table 2. Numeric Example: Community Rating and Perfect Risk Adjustment

	Risk Type	Risk Aversion Level	Plan	Premium
Community Rating and Perfect Adjustment	H	0.00005	Straight-deductible, deductible = $1,272	$6,193
	L	0.002	Deductible = $ 700, 10% coinsurance after deductible, MOOP = $3,100.	$6,152

Note: Risk types constructed from Truven MarketScan data.

In summary, restricting plans to be vertically ranked often reduces certain risk types' surplus. Besides, it may also reduce the overall social surplus. These regulations are often motivated by consumer confusion concerns, and the rationale is that restricting to simpler design makes it easier for consumers to choose. My conceptual framework suggests that a complete evaluation of such policies depends on the tradeoff of consumer confusion and welfare loss from a simplified menu. An open question is to what extent the sorting force matters in reality, which I now discuss in Section 3.

3 Empirical Analysis in the ACA Market

There are two key predictions from the conceptual framework. First, in insurance markets with community rating, different risk types prefer different plan designs. Second, high-risk types prefer plans concentrating coverage in larger losses (i.e., straight-deductible plans). In comparison, the low-risk types prefer plans with more coverage for smaller losses (i.e., non-straight-deductible designs.) In this section, I show the empirical relevance of the theory by illustrating that these predictions are consistent with the plan offering and sorting pattern observed in the ACA Federal Exchange. The ACA Federal Exchange is particularly suitable for studying plan design variation because the market allows insurers considerable freedom to offer different plan designs.

3.1 Institutional Background

The Affordable Care Act Exchange (the Exchange henceforth) was launched in 2014. Private insurers can offer comprehensive health insurance plans, and the federal government provides subsidies for certain low-income consumers who purchased plans. The Exchange regulates the actuarial value (AV) of plans, defined as the fraction of losses covered for the average population, but leaves insurers with latitude to offer a range of different plan designs. The Exchange has regulations on the market-average AV: Plans can only have a population-average AV of around 60%, 70%, 80%, and 90% and are labeled as Bronze, Silver, Gold, and Platinum plans, respectively. The Exchange also requires plans to have an upper limit on out-of-pocket spending ($7,150 in 2017). Some state Exchanges further regulate the plan designs.[7] Insurers are otherwise free to offer any cost-sharing attributes. Each state can either join the Federal Exchange or establish its own state exchange. I focus on plans launched via healthcare.gov, including the federally administered Individual Exchange and state exchanges operated via healthcare.gov. On this platform, insurers can offer any design satisfying the AV regulation and the MOOP limit.[8] The list of states in the sample is in Appendix Table C1.

ACA regulations also limit insurers' ability and incentive to do risk screening. The regulators calculate risk scores for enrollees and transfer money from insurers with a lower-cost risk pool to insurers with a higher-cost risk pool to equalize plan costs across insurers. Further, there is a single risk pool pricing regulation: the premiums of plans offered by the same insurer will be set based on the overall risk pool of that insurer, not the risk of individuals enrolled in each plan. Third, community rating limits insurers' ability to set premiums based on individual characteristics. Premiums can only vary by family composition, tobacco use status, and (partially) by age group.

---

[7] Regarding cost-sharing flexibility, insurers in Connecticut, the District of Columbia, Massachusetts, New York, Oregon, and Vermont must offer standardized options. They can offer a limited number of non-standardized options within a metal tier. California requires all insurers to offer only standardized plans (one per tier).

[8] Plans launched in states using healthcare.gov are still subject to each state's insurance regulation. For example, the essential health benefits that a plan must cover may differ across states.

3.2 Data and Sample

I use the Health Insurance Exchange Public Use Files from 2014 to 2017. This dataset is a publicly available dataset of the universe of plans launched through healthcare.gov. I define a plan based on the plan ID administered by CMS, which is a unique combination of state, insurer, cost-sharing attributes, provider network, drug formulary, and covered benefits. For each plan, I observe its financial attributes (deductibles, coinsurance rates, copays, MOOPs, etc.), premium (which varies at the plan-rating area level), and enrollment numbers in that plan (at the plan-state level). I focus on the 2014-2017 year for the main analysis, but the results are similar for other years.

I study risk sorting using the Uniform Rate Review Data from 2016 to 2019.[9] The data include average premium and claim cost information at the plan level for 50% of plans and insurer-level claim costs and risk transfer information for 75% of the insurers.[10] For the rest of the insurers, I match almost all of them in the Medical Loss Ratio filings, another insurer-level dataset reporting premium and claim costs, but not risk transfers. I use the 75% insurers as the baseline because all variables of interest are available, and I use the Medical Loss Ratio filings as robustness checks. Appendix Table C2 summarizes the data sources used in the empirical analysis.

I focus on Bronze, Silver, Gold, and Platinum plans. Catastrophic plans are dropped from the analysis because they have no officially reported AV and are unavailable to most consumers. Each Silver plan has three cost-sharing reduction variations available to the low-income population. These plans have the same premium as the standard Silver plan and a higher AV. In the plan design analysis, I use the cost-sharing characteristics of the standard Silver plan. In studying the sorting pattern, I label the straight-deductible design based on the standard Silver plan because the straight-deductible design is consistent across the standard plans in almost all cases, and the cost-sharing variations and the claim costs are reported for all variations.

I study the cost-sharing features of a plan's first-tier in-network coverage for essential health benefits. The utilization rate of the first-tier in-network coverage is 94.59% on

[9] The reports have a two-year lag, so the 2016 -2019 reports match the 2014-2017 plan information.
[10] The plan level information is incomplete because only plans with more than 10% premium increase are required to report, while the insurer level information is required for all insurers unless they exit the market.

average for the sample plans, and 99.47% of the total premium is contributed to cover the essential health benefits on average. I exclude preventive care because all plans must cover it with no cost-sharing. The resulting benefits in Appendix Table C3 are consistent with the list of the AV calculator, a tool created by CMS to compute the AV of each plan.[11]

A straight-deductible plan is identified as one under which 1) all benefits are subject to the general deductible, 2) there is no coverage before hitting the deductible, and 3) there is no cost-sharing after the deductible. Screenshots of an example straight-deductible plan and a non-straight-deductible plan on the ACA Exchange are in Appendix Figure C1.

3.3 Analysis of Plan Design Variations in the ACA Market

The market is populated with both straight-deductible and non-straight deductible plans. Table 3 shows the market share of straight-deductible plans over time. Take the year 2016 as an example. There are around 4,000 unique plans offered in this market. Among them, 13% are straight-deductible plans. In total, 9.7 million consumers purchased a plan in this market, and about 7.6% selected a straight-deductible plan.

Table 3. Market Share of Straight-Deductible Plans

year	% plans that are straight-deductible	Total number of plans	Enrollment share in straight-deductible plans	Total number of consumers (mm)
2014	10.48%	2,871	5.40%	5.57
2015	9.58%	4,573	6.97%	9.22
2016	12.99%	3,966	7.63%	9.71
2017	11.14%	3,106	4.52%	9.00

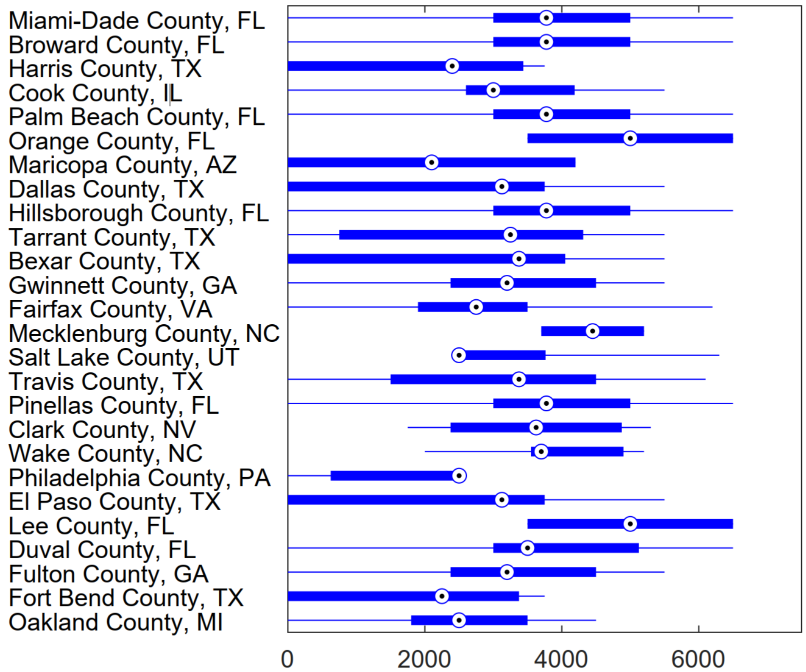
Note: The sample includes the universe of plans launched via healthcare.gov. The enrollment data of Silver plans represent four cost-sharing variations: The standard Silver plans and three cost-sharing reduction plans (which are only available to lower-income households). I classify straight-deductible for these plans based on the standard plan.

Consumers also face substantial variation in plan designs within a metal tier. Figure 1 shows the 2017 Standard Silver plans' deductible distribution for counties with the top 25 enrollment size via Healthcare.gov. In all these counties, consumers face over $2,500 differences in the Silver deductibles. The large variation in plan design faced by a particular consumer is prevalent for many other counties and metal tiers. For example, an average

[11] Accessed from https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/

consumer faces Gold plans with MOOPs ranging by more than $2,000. Appendix Figure C2 shows the distribution of deductible and MOOP across all counties.

Figure 1. Distribution of the 2017 Silver Deductible for the 25 Largest Counties



Note: Data from the 2017 CMS Health Insurance Exchange Public Use Files. Counties are ranked by the enrollments via Healthcare.gov and included counties have enrollment numbers larger than 50,000. Silver plans are standard Silver plans. The deductible refers to tier-one, in-network coverage for an individual, cumulative over a year. The circle in the center of the bar indicates the median, the lower and upper bounds of the bar indicate the 25th and 75th percentile and the lower and upper of the whiskers indicate the minimum and the max.

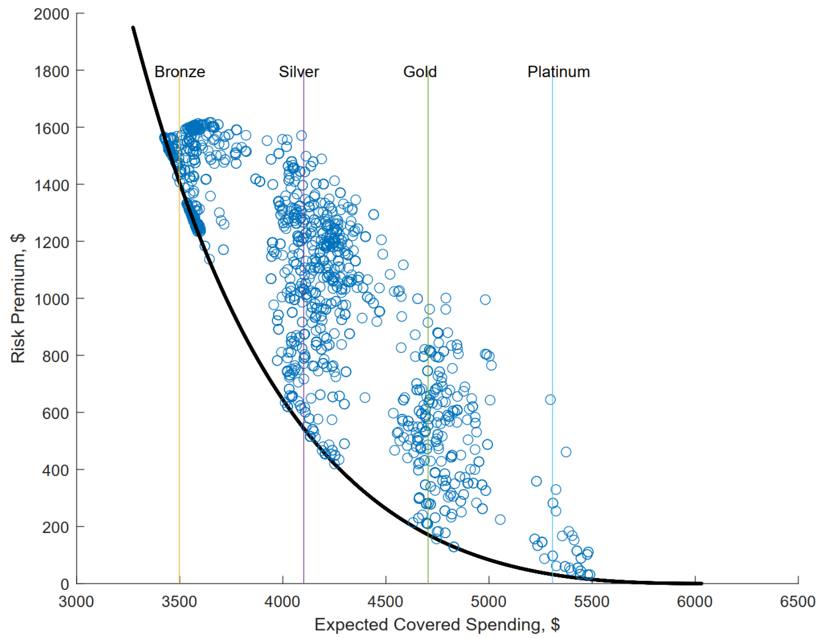
The plan design variation implies significant variations in plans' financial values to consumers within a coverage tier. To quantify, I evaluate each plan's financial value for the average ACA individual with a CARA utility function with a risk-averse coefficient of 0.0004 (Handel, 2013). I first apply the cost-sharing rules of all plans to this representative individual's risk distribution and calculate the stochastic out-of-pocket spending, a, for each plan. I then calculate the risk premium R, using the following formula:

$$E[u(w-a)] = u\,(w - E(a) - R), \quad (6)$$

where w represents the wealth level, and $u(\cdot)$ is the utility function. The risk premium represents the sure amount the individual needs to receive to be indifferent between enrolling in that plan and a full-insurance plan, when both are priced at their fair AV. It represents the risk protection of different designs (the smaller, the higher the value).

Straight-deductible plans have the lowest R, holding fixed $E(a)$ (Gollier and Schlesinger 1996). The calculation details are in Appendix D.

Figure 2. Risk Premium and Expected Covered Spending for 2017 Plans



Note: The sample includes plans launched in the individual market via healthcare.gov. A plan is a unique combination of insurer, covered benefits, cost-sharing designs, drug formulary, provider network, and state. Plans launched in multiple rating areas or counties are only counted once. Cost-sharing reduction plans and Catastrophic plans are excluded. The black solid line shows the lowest possible risk premium conditional on the expected spending level (achieved by straight-deductible plans) and does not represent actual plans. A dot might represent multiple plans if they have the same cost-sharing feature. The vertical lines show the targeted AV for each metal tier. Not all plans align with the vertical lines perfectly because the regulator allows for a two percent error margin and because of measurement error in my calculation.

Figure 2 shows the risk premium and the expected covered spending for all plans in the four metal tiers in 2017. The four clusters represent the four metal tiers. A substantial difference in risk premium exists for a range of AV levels. For example, among plans in the Silver tier, which have an AV of around 70%, the smallest risk premium relative to full insurance is around $500 and is achieved by the straight-deductible plan (black line in Figure 2). In contrast, the largest risk premium for Silver plans is nearly $1,000 larger, originating from plans that have lower deductibles and MOOP closer to the maximum allowed by the regulation.

3.4 Evidence of Sorting by Health into Different Plan Designs

The existence of the plan design variation may create room for selection. The theoretical analyses in Section 2 suggest that straight-deductible plans are more attractive to the higher-risk types. An ideal test for the sorting pattern requires observing the full distribution of individuals enrolled in different plans. Unfortunately, I don't have that information for all plans available in the ACA Exchange. Instead, I focus on testing the first moment of the loss distribution. I perform two sets of analyses: first, I compare the plan-level average total claim costs per member per month between straight-deductible and other designs. Second, I examine the differences in insurer-level risk transfers among those offering straight-deductible plans and the rest. The measure represents ex-ante risk scores of plans.

3.4.1 Plan-Level Analysis

A comparison in unconditional means of the total medical expenditure illustrates a strong correlation between average medical spending and plan designs, consistent with the theoretical predictions on sorting. Figure 3 shows the average monthly total medical expenditure for straight-deductible plans and the other designs across the metal tiers. Straight-deductible plans have significantly higher medical expenditures than other plans. The difference is more than $400 per month for Silver and Gold plans.[12]

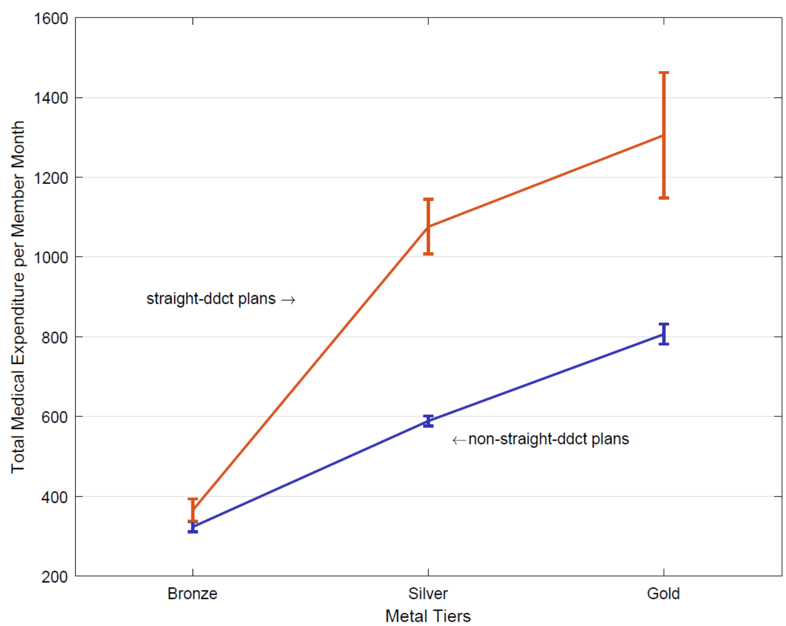
The correlation between plan design and expenditure might be driven by other confounding factors, which I address using the regression model. First, given that only plans with an excessive premium increase are subject to report the claim information, the mean difference of the reported plan may not be representative. To address the concern, I leverage the fact that insurers are subject to the single risk pool requirement and will spread out unexpected medical expenditures of a particular plan among all plans offered, making all plans subject to reporting. In the plan-level regression, I include insurer-year fixed effects so that the differences in claims costs between different plan designs are identified based on the within-insurer-year variation.

Second, the correlation might be driven by other plan characteristics. I first examine whether other plan attributes are correlated with straight-deductible design. Appendix Table C4 presents a balanced test of straight-deductible and other designs. I find that straight-deductible plans are not correlated with other plan attributes consumers might sort on, including plan types, whether having a national network, new or existing plan, offered

[12] The difference in the Bronze tier is smaller because the OOP-limit regulation limited the room for design difference. The market has few Platinum plans, so they are not shown in the graph.

to rural counties, etc. The only significant difference is that straight-deductible plans are more likely to have a health savings account (HSA), because these accounts require a high-deductible, and straight-deductible plans have high deductibles in a metal tier. Given that individuals with greater health needs may prefer HSA, the correlation between HSA-eligibility and straight-deductible may bias the difference away from zero. In the baseline regressions, I add HSA eligibility, dummies for plan type, and the service area fixed effects as control variables to address the concern.

Figure 3. Average Total Expenditure per Member Month by Plan Design



Notes: The graph shows the mean and 95% confidence interval of the total medical expenditure of plans launched through healthcare.gov in 2014-2017. Only plans with premium changes of more than 10% are reported in the Uniform Rate Review data. Such plans account for about 50% of the universe of plans launched.

Finally, and most importantly, the differences in the ex-post expenditure may reflect ex-post moral hazard instead of selection. Moral hazard is likely a primary concern for the differences in the ex-post expenditure *across* metal tiers. To address the issue, I control for metal tier fixed effects and actuarial values.[13] It is unclear whether straight-deductible designs imply more moral hazard than other designs *within* a metal tier. Straight-deductible plans have no coverage for small losses and may deter large expenditures because of that.

[13] The sorting into designs within a metal tier could either be driven by the fact that there is a negative correlation between risk aversion and risk levels, as illustrated in the numeric example in Table 2, or the sorting pattern conditional on coverage level, as stated in Corollary 1.

Table 4 column (1) shows the results with total medical expenditure per member per month as the dependent variable. On average, individuals enrolled in straight-deductible plans have significantly higher medical expenditures ($119 higher per month and $1,428 annually) relative to the mean spending of $555 per month. Table 4, columns (2) and (3) show that premiums are similar for straight-deductible and other plan designs in the same metal tier. The little difference in premiums suggests that the single risk pool requirement is well enforced and blunts the pass-through of these selection differences to consumers.

Table 4. Plan-Level Sorting Pattern

	(1)	(2)	(3)
	monthly total expenditure	monthly premium	
		collected	charged
straight-deductible	118.52	2.11	0.72
	(21.70)	(3.66)	(1.59)
N	7,842	7,842	72,829
R^2	0.55	0.85	0.73
y-mean	554.74	397.04	265.6
y-sd	382.28	120.56	93.05
Controls	AV, metal tier, network type, HSA-eligibility, insurer FE, year FE		
Fixed Effects	service area FE		rating area FE

Note: Straight-deductible is a dummy variable indicating whether the plan has a straight-deductible design. The AV of a plan is the fraction of losses covered for the average population, which varies no more than four percentage points within a metal tier. Columns (1) and (2) include plans between 2014 and 2017 with a premium increase of more than 10%. I dropped those reporting non-positive total expenditure or premium and plans with the top and bottom one percent of either value to avoid impact from extreme values. Each observation is a plan-state-year. The dependent variable in (1) is the average total medical expenditure per member month. The dependent variable in (2) is the average collected premium per member month. Column (3) includes all plans between 2014 and 2017. The dependent variable is the per-month premium of the single coverage for a 21-year-old non-tobacco user. Since premium varies by rating area, each observation is a plan-rating area-year. Standard errors are clustered at the insurer level and shown in parentheses.

The empirical findings hold when using other measures of plan designs. The conceptual framework shows that sorting into straight-deductible plans represents high-risk individuals' preference for designs concentrating coverage in larger losses. In reality, many other non-straight-deductible designs may provide similar coverage. I create three continuous measures of plan designs to capture the similarity of plan designs to straight-deductible plans. First, I calculate each plan's fraction of losses covered for the first $2,000 total medical expenditure (evaluated for the individual with market-average risk). Within a metal tier, the smaller the value, the more coverage is concentrated in larger losses. Second, I use the relative risk premium, calculated as risk premium minus the risk premium of the straight-deductible plan with the same AV for the average population. This measure is zero for straight-deductible designs and is larger when the design differs more from a straight-deductible design. Third, I calculate the ratio of deductible over MOOP. Straight-deductible plans will have a ratio of one, while a smaller value indicates the plan offers more coverage for smaller losses. Table 5 shows that enrollees in plans with more coverage in larger losses, smaller risk premiums, and larger deductible to MOOP ratio have significantly higher total medical expenditure, consistent with the baseline results.

Table 5. Plans' Total Expenditure and Different Design Measures

	(1)	(2)	(3)
	Dependent Variable: total medical expenditure per member month		
% losses covered for first $2,000	-5.49 (0.79)		
Risk premium, $100		-24.02 (3.73)	
Deductible to MOOP ratio			99.98 (20.03)
N	7,842	7,842	7,842
R^2	0.56	0.55	0.55
y-mean		554.74	
y-sd		382.28	
Controls	AV, metal tier, network type, HSA-eligibility, insurer FE, year FE, service area FE		

Note: The sample includes plans between 2014 and 2017 with a premium increase of more than 10%. I dropped those reporting non-positive total expenditure or premium and plans with the top and bottom one percent of either value to avoid impact from extreme values. Each observation is a plan-state-year. "% losses covered for first $2,000" measures each plan's fraction of losses covered for the first $2,000 total medical expenditure evaluated for the individual with market-average risk. "Risk premium" measures the difference in the risk premium of choosing the plan relative to the straight-deductible plan with the same actuarial value. The dependent is the average total medical expenditure per member per month. Standard errors are clustered at the insurer level and shown in parentheses.

The results are robust to controlling for various other plan attributes. In Appendix Figure C3, I use total medical expenditure as the dependent variable, add different plan characteristics one at a time, and plot the coefficient of straight-deductible plan. The figure shows that the estimates are stable when different controls are added. I further estimate Table 4 separately for plans with an HSA and those without. Appendix Table C5 shows that the baseline sorting pattern holds for both samples, though the magnitude is smaller among plans with no HSA.

3.4.2 Insurer-Level Analysis

There are two concerns with the plan-level analysis. First, despite controlling for metal tier fixed effects, the ex-post medical expenditure may reflect moral hazard rather than selection within a metal tier. Second, the plan-level analysis is based on half of the plans with large premium increases, not the whole sample. I address both issues using insurer-level analysis. I collect insurer-level total medical expenditures and, most importantly, the risk transfer payment information from the Uniform Rate Review filings.[14] Risk transfers are calculated based on the average risk scores of enrollees and reflect the ex-ante medical expenditure risk rather than moral hazard responses (Polyakova, 2016). All insurers are subject to the risk-transfer reporting, producing a more representative sample.

The key independent variable is whether an insurer offers any straight-deductible plan. I confirm that such insurers are similar to others along many observed dimensions. Appendix Table C6 shows that insurers offering at least one straight-deductible plan are similar to other insurers in terms of offering HMO plans, offering plans with a national network, operating in rural areas, and total enrollment size. The only difference is that they are more likely to offer HSA-eligible plans. I add enrollment share in HSA-eligible plan, state, and year fixed effects for the insurer-level analysis to control for the potential impacts of HSA-eligibility.

Table 6 shows the comparison at the insurer level. As in the plan-level analysis, insurers offering straight-deductible plans experience significantly higher total medical expenditure per member month (column 1) and pay more claims (column 2) than other insurers. Moreover, they also receive higher risk transfer payments ($41 per member per month, column 3.) The estimated risk transfer differences account for more than 75% of the estimated differences in insurers' liability. The premium differences are much smaller and

[14] Plan-level risk transfers are also estimated by insurers for a subset of plans. However, many insurers use plan premiums to allocate insurer-level risk transfers to plans. According to the single risk pool regulation, different designs are required to have similar premiums within a metal tier, making the allocated plan-level risk transfers inappropriate to capture selection within a metal tier.

indifferent from zero, suggesting that risk transfer regulations are well enforced to blunt the pass-through of these selection differences to consumers.

Table 6. Insurer-Level Sorting Pattern

	(1) Total Expenditure	(2) Insurer Liability	(3) Risk Transfers	(4) Average Premium
Offer Straight-Deductible Plan	71.92 (27.55)	53.84 (22.23)	40.58 (13.51)	14.34 (13.84)
N	617	617	617	617
R^2	0.262	0.239	0.144	0.604
Dep. Var. Mean	474.7	357.1	-6.201	381.1
Dep. Var. SD	124.1	102.5	66.03	97.42

Note: Each observation is an insurer year. "Offer Straight-Deductible Plan" is a dummy variable indicating whether an insurer offers any straight-deductible plan. In all columns, the dependent variables are measured using the per member per month value. The dependent variable in (1) is the average total medical expenditure of enrollees in a plan, including consumer cost-sharing and insurer payments. The dependent variable in (2) is the average medical expenditure paid by insurers. The dependent variable in (3) is the average risk transfer an insurer receives. The dependent variable of (4) is the average premium. All models include year fixed effects, state fixed effects, and the fraction of enrollees in plans with a health savings account. The regressions are weighted by the enrollment at each insurer-year. Standard errors are clustered at the insurer level.

Table 7. Insurers' Risk Transfers and Different Plan Design Measures

	(1)	(2)	(3)
	Dependent Variable: Risk Transfers		
Avg. % losses covered for first $2,000	-1.24 (0.60)		
Avg. risk premium, $100		-16.96 (3.75)	
Avg deductible to MOOP ratio			102.21 (34.48)
N	617	617	617
R^2	0.56	0.26	0.24
y-mean		-6.20	
y-sd		66.03	
Controls	HSA-eligibility enrollment share, year FE, state FE		

Note: Each observation is an insurer year. "Avg. % losses covered for first $2,000" is the insurer-year-level average of the plans' fraction of losses covered for the first $2,000 total medical expenditure evaluated for the individual with market-average risk. The unit is one percentage point. "Avg. risk premium" is the insurer-year-level average of the risk premium of plans offered by insurers. "Avg. deductible to MOOP ratio" is the insurer-year-level average of the deductible over MOOP of plans offered by insurers. The dependent variable is the average risk transfers per member per month. Standard errors are clustered at the insurer level and shown in parentheses.

The results are similar using continuous plan design measures. I calculate the insurer-level average of the three continuous plan design measures and use them as the independent variables. Table 7 shows that insurers with more plans covering larger losses, lower risk premiums, and larger deductible to MOOP ratios receive significantly larger risk transfers.

The results are robust when controlling for different sets of insurer characteristics and imputing missing observations from the Medical Loss Ratio files. I present the robustness checks in Appendix Table C7.

As a final note, the sorting pattern may be driven by specific (rational) choice heuristics. For example, the straight-deductible plans have the lowest MOOP within a metal tier. If high-risk individuals care only about the worst-case risk and choose based on MOOP, they sort into straight-deductible plans. My conceptual framework provides one rationale for such heuristics.

4 Calibrating Impacts of Plan Design Regulations for the ACA Federal Exchange

The plan offering and sorting pattern in the ACA Exchange suggests that limiting plan design variations might have economically meaningful impacts on consumer welfare. In this section, I calibrate the likely impacts of limiting plan designs in the ACA Federal Exchange. Specifically, I compare the market outcome under two menus: The actual 2017 plans offered in the ACA Exchange and a hypothetical choice set replacing all options with a straight-deductible plan of the same premium. This new choice set has the same number of options and premiums as the existing one. The only difference is that all plans have a straight-deductible design.

The exercise highlights the tradeoff between two factors. First, consumers are often confused about the plan design and fail to sort into suitable plans for them (Abaluck and Gruber 2011, 2019; Bhargava et al. 2017). The consequence of choosing the wrong plan is especially large for the higher-risk types. Given that their desired plans have a straight-deductible design, limiting only to these plans may help mitigate the consequence of sorting into the wrong plan for them. Second, as discussed in Section 2.3, limiting to straight-deductible plans will reduce the consumer welfare of the lower-risk types. The overall surplus then depends on which force dominates.

4.1 Setup

Demand side. Consumers are modeled as expected utility maximizers choosing plans based on the perceived utility:

$$v_{ij} = \overbrace{\underbrace{\int u\big(w - OOP_j(x) - p_j\big)dF_i(x)}_{welfare-relevant\ utility} + \beta\epsilon_{ij}}^{decision\ utility}. \tag{7}$$

The deterministic part, $\int u\big(-OOP_j - p_j\big)dF_i$, is a function of the wealth level, w, out-of-pocket spending, OOP_j, and net premium after subsidy, p_j. It determines the welfare-relevant value of each plan j for individual i. The second component of the choice utility is an error component, ϵ_{ij}. It affects the choice of each consumer but is not relevant to welfare. The error term creates the potential for consumer confusion. Consumers in the ACA Exchange often face a large choice set, typically around 20 options in each county, making confusion a likely concern. The larger the scaling parameter β, the more randomness there will be in plan choice. $\beta = 0$ represents the case where all consumers choose optimally.

Supply side. On the supply side, I assume a perfectly competitive market with perfect risk adjustment. In such a market, raw plan premiums are a mechanic function of the expected covered spending if *all* types choose the plan plus a loading factor:

$$p_j = \theta \sum_i \tau_{ij} w_i, \tag{8}$$

where τ_{ij} is the expected covered spending of individual type i under plan j, w_i is the population weights of each type, θ is the loading factor. Even though other literature finds that the risk adjustment is not perfect along dimensions like drug formulary (Geruso, Layton, and Prinz, 2020), the findings in section 3.3 suggest that single risk pool requirement and risk adjustment regulations are successful in flattening the premium level of different plan designs. The perfect risk adjustment assumption is a good characterization of the sorting pattern of this model. I also assume all insurers incur the same loading factor, representing the necessary transaction costs of providing insurance, which is motivated by the Medical Loss Ratio regulation. Under these assumptions, insurers are passive about which plans to offer.

Model calibration. I calibrate the model's key components to the observed data in the ACA market. First, I model each market as a county because the choice set varies at the county level in the ACA. I create 100 risk types using the k-means clustering method from Truven MarketScan data (see details in Appendix B.) I mean-shifted these distributions such that the overall medical expenditure level is benchmarked to the 2017 ACA Federal Exchange average. Since I do not have information about the risk distributions at each county, I assume that all counties have identical risk distributions. I create five sub-types for each risk type: consumers facing non-CSR plans and no premium subsidy, non-CSR plans, and premium subsidy, and both (3 CSR plan types). The county-level average premium subsidy amount (among eligible people) and relative weights of each type are collected from the 2017 Open Enrollment Period County-Level Public Use Files.[15] The CSR-eligible consumers can choose from CSR plans rather than the Standard Silver plans.

I assume that consumers have a CARA utility function with a risk aversion coefficient of 0.0004 (the median and mean estimated by Handel (2013).) Under the CARA utility function, w is irrelevant. As a result, equation (7) can be simplified as $v_{ij} = -p_{ij}' + \int u\left(-OOP_j(x)\right) dF_i(x) + \beta\epsilon_{ij}$, where p_{ij}' is the net premium. I convert each plan's cost-sharing attributes into a simplified three-arm design (see Appendix E for details). With the three-arm design and individuals' loss distribution, I can calculate τ_{ij} and p_j. The net premiums, p_{ij}, for those who are eligible for subsidy, is p_j minus the subsidy amount.

Finally, I assume ϵ_{ij} to be i.i.d following the extreme value type one distribution. I vary β from 0 to some positive numbers. Under each β, I calculate the plan chosen by each individual type as the one maximizing v_{ij}. Let $j_i^*(\beta)$ denote the plan chosen by individual i under β. The total efficiency of the market under β is calculated as

$$ss(\beta) = \sum_c \sum_{i \in c} w_{ic} \left(\int u\left(-OOP_{j_i^*(\beta)}(x)\right) dF_i(x) - \theta\tau_{ij_i^*(\beta)} \right), \tag{9}$$

where $i \in c$ indicates individuals in county c, and w_{ic} are population weights of that type.

The consumer surplus of individual i under β is:

$$cs_i(\beta) = \int u\left(-OOP_{j_i^*(\beta)}(x)\right) dF_i(x) - p_{ij_i^*(\beta)}. \tag{10}$$

I elaborate more on calibration details in Appendix E and summarize the parameter sources in Appendix Table E1.

---

[15] https://www.cms.gov/data-research/statistics-trends-and-reports/marketplace-products/2017-marketplace-open-enrollment-period-public-use-files

Model caveat. The simulation makes a few simplification assumptions. First, I assume no insurance is not in the choice set of consumers. This assumption abstracts away from the extensive margin of the market. Second, I assume that all insurers have the same medical loss ratios and are perfectly competitive. Third, I focus on the variation in plans' financial attributes and ignore other plan characteristics. Fourth, I assume consumers have no moral hazard responses. These abstractions make the model tractable and highlight the key insights.

4.2 Results

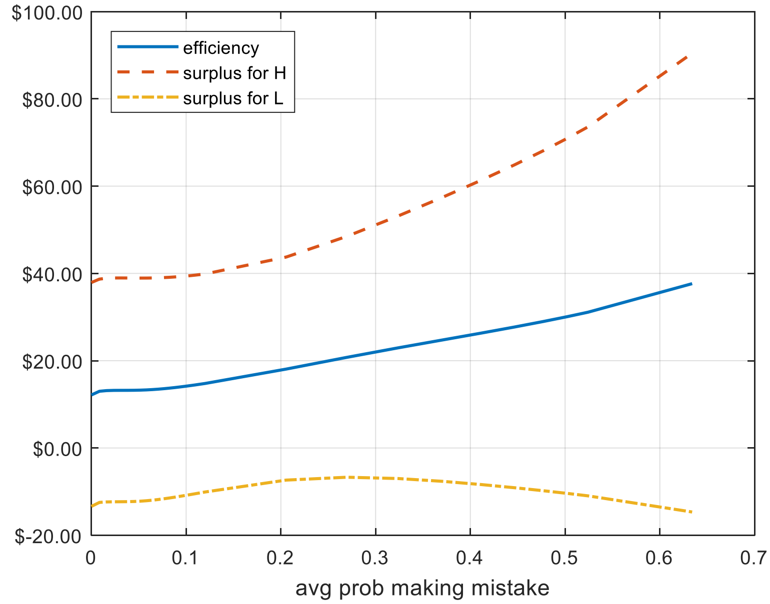
Figure 4 shows the results. The y-axis represents the social or consumer surplus difference between the hypothetical menu and the actual menu. A positive number means consumers are better off under the straight-deductible-only environment than facing the current ACA menu. To illustrate the distributional effects, I split consumers into those with above and below median expected medical expenditure. The solid line represents the overall social surplus, while the the two dashed lines represent consumer surplus for each type. The x-axis is the fraction of consumers choosing the non-optimal plan, an increasing function of β.

When there is no confusion, limiting plans to straight-deductible design increases overall welfare by $12 per person per year. The welfare gain comes from offering straight-deductible plans to high-risk types in places where these plans are not available, while such benefits are largely offset by the welfare loss from forcing the low-risk types to choose such plans.

When consumers in the market are more likely to make a mistake in choosing plans, both the overall efficiency and the surplus for higher-risk types increase. For example, when 50% of consumers sort into a wrong plan, the average efficiency is $30 higher with regulation per year, and the surplus for the higher-risk types is $70 higher per year with regulation. However, such a change is not a Pareto improvement: The lower-risk types are worse off under such regulation. At the 50% confusion level, they are worse off by about $10 per year under the design regulation.

The simulation illustrates two key points. First, standardizing plan designs has distributional impacts: limiting to straight-deductible designs is not a Pareto improvement because such restriction removes valuable options for low-risk types. Second, the level of confusion matters for the overall welfare gains of regulating plan designs. When the confusion level is high enough, standardizing plan designs to straight-deductible plans can create large welfare benefits for the high-risk types.

Figure 4. Efficiency Effects of Regulating Plan Designs in ACA



Note: The y-axis represents the difference between the value under and without the design regulation. The regulation replaces all current ACA plans with a straight-deductible plan of the same premium.

Ultimately, the welfare impacts of plan standardization policies depends on the extent to which consumers have confusion when choosing health insurance plans. Previous literature has documented that consumers respond strongly to plan standardization policies in marketplaces like the ACA exchange (Ericson and Starc 2007). There are also works showing that consumers choose ACA cost-sharing reduction variations properly (DeLeire et al. 2017). My calibration illustrates the importance in estimating consumer confusion in this market.

5 Conclusion

In this paper, I identify an understudied dimension of sorting in insurance markets: Sorting by plans' multi-dimensional cost-sharing attributes. I show that in a market with asymmetric information, lower-risk consumers will sort into designs with less coverage for larger losses in exchange for more coverage for smaller losses, while higher-risk consumers sort into straight-deductible plans. The framework extends the classic model considering binary losses and can rationalize the proliferation of plan designs in health insurance markets.

The framework provides a new perspective on the trade-offs introduced by plan standardization policies. Prior literature recognizes that simplifying insurance contract characteristics can make it easier for consumers to compare plans and promote competition and efficiency. I illustrate that in a market with asymmetric information, plan design variation can also serve as a tool to separate different risk types and support an equilibrium where lower-risk consumers suffer less distortion under asymmetric information. As a result, removing plan design variation may harm efficiency. The overall benefits of standardizing plan design thus depend on the relative importance of these concerns.

The framework abstracts away from certain market conditions and opens for future research. First, the baseline model does not consider moral hazard responses. Understanding how plan cost-sharing attributes induce expenditure *within* a coverage level and how that interacts with plan selection under an endogenous contract design framework is an important question. Second, the framework assumes perfect competition. More research is needed to understand how market power may complicate the welfare implications of plan standardization policies.

References

Abaluck, Jason, and Jonathan Gruber. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *The American Economic Review* 101, no. 4 (June 2011): 1180–1210.

Abaluck, Jason, and Jonathan Gruber. "When Less is More: Improving Choices in Health Insurance Markets." *The Review of Economic Studies*, 90 (3), (May 2023): 1011–1040.

Agterberg, Joshua, Fanghao Zhong, Richard Crabb, and Marjorie Rosenberg. "A Data-Driven Clustering Application to Identify Clusters of Individual Profiles with High Health Expenditures." Working paper (2019).

Arrow, Kenneth J. "Uncertainty and the Welfare Economics of Medical Care." *The American Economic Review* 53, no. 5 (1963): 941–73.

Azevedo, Eduardo M., and Daniel Gottlieb. “Perfect Competition in Markets With Adverse Selection.” *Econometrica* 85, no. 1 (2017): 67–105.

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor. “Choose to Lose: Health Plan Choices from a Menu with Dominated Option.” *The Quarterly Journal of Economics* 132, no. 3 (August 1, 2017): 1319–72.

Carey, Colleen. “Technological Change and Risk Adjustment: Benefit Design Incentives in Medicare Part D.” *American Economic Journal: Economic Policy* 9 no. 1 (2017): 38–73.

Chade, Hector, Victoria R. Marone, Amanda Starc, and Jeroen Swinkels. “Multidimensional Screening and Menu Design in Health Insurance Markets.” Working Paper Series. National Bureau of Economic Research (2022): 30542.

Crocker, Keith J., and Arthur Snow. “Multidimensional Screening in Insurance Markets with Adverse Selection.” *Journal of Risk and Insurance* 78, no. 2 (2011): 287–307.

Decarolis, Francesco, and Andrea Guglielmo. 2017. “Insurers’ Response to Selection Risk: Evidence from Medicare Enrollment Reforms.” *Journal of Health Economics* 56 (December): 383–96.

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* 56 (December 2017): 71–86.

Doherty, Neil A., and Harris Schlesinger. “Optimal Insurance in Incomplete Markets.” *Journal of Political Economy* 91, no. 6 (1983): 1045–1054.

Einav, Liran, Amy Finkelstein, Stephen P Ryan, Paul Schrimpf, and Mark R Cullen. “Selection on Moral Hazard in Health Insurance.” *American Economic Review* 103, no. 1 (February 2013): 178–219.

Ellis, Randall P., and Thomas G. McGuire. 2007. “Predictability and Predictiveness in Health Care Spending.” *Journal of Health Economics* 26 (1): 25–48.

Ericson, Keith M., and Amanda Starc. “How Product Standardization Affects Choice: Evidence from the Massachusetts Health Insurance Exchange.” *Journal of Health Economics* 50 (December 2016): 71–85.

Ericson, Keith Marzilli, and Justin Sydnor. 2017. "The Questionable Value of Having a Choice of Levels of Health Insurance Coverage." *Journal of Economic Perspectives* 31 (4): 51–72.

Ericson, Keith M. and Justin Sydnor "Liquidity Constraints and Insurance Demand" Working Paper 2018.

Ericson, Keith Marzilli, Philipp Kircher, Johannes Spinnewijn, and Amanda Starc. 2020. "Inferring Risk Perceptions and Preferences Using Choice from Insurance Menus: Theory and Evidence." *The Economic Journal*, no. ueaa069 (May).

Frank, Richard G, Jacob Glazer, and Thomas G McGuire. "Measuring Adverse Selection in Managed Health Care." *Journal of Health Economics* 19, no. 6 (2000): 829–54.

Finkelstein, Amy, James Poterba, and Casey Rothschild. "Redistribution by Insurance Market Regulation: Analyzing a Ban on Gender-Based Retirement Annuities." *Journal of Financial Economics* 91, no. 1 (2009): 38–58.

Geruso, Michael, Timothy Layton, and Daniel Prinz. "Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges." *American Economic Journal: Economic Policy* 11, no. 2 (May 2019): 64–107.

Geruso, Michael, Timothy J. Layton, Grace McCormack, and Mark Shepard. 2023. "The Two Margin Problem in Insurance Markets." *Review of Economics and Statistics* 105 (2): 237-257.

Gollier, Christian, and Harris Schlesinger. "Arrow's Theorem on the Optimality of Deductibles: A Stochastic Dominance Approach." *Economic Theory* 7, no. 2 (June 1, 1996): 359–63.

Handel, Benjamin R. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review* 103, no. 7 (2013): 2643–82.

Ho, Kate, and Robin S. Lee. Forthcoming. "Health Insurance Menu Design for Large Employers." *The RAND Journal of Economics*.

Kautter, John, Gregory Pope, Melvin Ingber, 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 and Medicaid Research Review* 4, no. 3 (2014): E1-E46.

Lavetti, Kurt, and Kosali Simon. 2018. “Strategic Formulary Design in Medicare Part D Plans.” *American Economic Journal: Economic Policy* 10 (3): 154–92.

Liu, Chenyuan, and Justin Sydnor. 2022. “Dominated Options in Health Insurance Plans.” *American Economic Journal: Economic Policy* 14 (1): 277–300.

Marone, Victoria R., and Adrienne Sabety. 2022. “When Should There Be Vertical Choice in Health Insurance Markets?” *American Economic Review* 112 (1): 304–42.

Pauly, Mark V. “The Economics of Moral Hazard: Comment.” *The American Economic Review* 58, no. 3 (1968): 531–537.

Polyakova, Maria. “Regulation of Insurance with Adverse Selection and Switching Costs: Evidence from Medicare Part D.” *American Economic Journal: Applied Economics* 8, no. 3 (July 2016): 165–95.

Raviv, Artur. “The Design of an Optimal Insurance Policy.” *The American Economic Review* 69, no. 1 (1979): 84–96.

Rothschild, Casey. “Nonexclusivity, Linear Pricing, and Annuity Market Screening.” *Journal of Risk and Insurance* 82, no. 1 (2015): 1–32.

Rothschild, Michael, and Joseph Stiglitz. “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information.” *The Quarterly Journal of Economics* 90, no. 4 (November 1976): 629.

Shepard, Mark. 2022. “Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange.” *American Economic Review* 112 (2): 578–615.

Spence, Michael. 1978. “Product Differentiation and Performance in Insurance Markets.” *Journal of Public Economics* 10, no. 3 (December 1, 1978): 427–47.

Veiga, André, and E. Glen Weyl. “Product Design in Selection Markets.” *The Quarterly Journal of Economics*, 131 no. 2 (2016): 1007-1056.

Zeckhauser, Richard. “Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives.” *Journal of Economic Theory* 2, no. 1 (1970): 10–26.

<u>For Online Publication Only</u>

Appendix A. Proofs in Section 2

Proposition 1. Proof:

The optimization problem for the consumer is:

$$v = \max_{l} \sum_{s} u(w - x_s + l_s - p(\boldsymbol{l}))f_s, \qquad \forall \boldsymbol{l}$$

subject to:

$0 \leq l_s \leq x_s,$

$$p(\boldsymbol{l}) = \theta \sum_{s} f_s l_s + c.$$

$\boldsymbol{l} = (l_1, l_2, \dots, l_s, \dots, l_S)$ is the vector of the insurance payments in each state.

The first-order condition is:

$$\frac{\partial v}{\partial l_s} = u_s'(1 - \theta f_s)f_s - \theta f_s \sum_{\tau \neq s} u_\tau' f_\tau \leq 0, \forall s,$$

with equality if $l_s > 0$.

First, note that if $l_s = x_s$ is binding for state s, then it's binding for all other states. This corresponds to the case when $\theta = 1$ and the optimal insurance is full insurance. For $\theta > 1$, full insurance is no longer optimal because of loading. For all states, $l_s < x_s$.

Second, note that the FOC can be rewritten as $u_s' \leq \theta \sum_\tau u_\tau' f_\tau$. The right-hand side is the same for all states, which implies that once binding, $x_s - l_s$ is a constant. Since $u'' < 0$, FOC is binding when x_s is larger than a certain level. Suppose $x_d = d$ is the level where $u_s'(w - x_d) = \theta \sum_\tau u_\tau' f_\tau$. Then the optimal insurance plan has the following form:

$$l_s^* = \begin{cases} 0, & if\ x_s < d \\ x_s - d, & if\ x_s \geq d' \end{cases}$$

which is the straight-deductible design. ■

Proposition 2. Proof:

The optimization problem for L is:

$$\max_{l} \sum_{s} u_L(w - x_s + l_s - p(\boldsymbol{l}))f_s^L$$

subject to:

$$p(\boldsymbol{l}) = \theta \sum_{s} f_s^L l_s + c,$$

$$0 \leq l_s \leq x_s,$$

$$\sum_{s} u_H(w - x_s + l_s - p)f_s^H = A.$$

$\boldsymbol{l} = (l_1, l_2, \dots, l_s, \dots, l_S)$ is the vector of the insurance payments in each state. A represents the utility H gets from choosing their optimal contract under full information. The last condition thus represents the binding incentive compatibility constraint for H.

The Lagrange of the above optimization problem is:

$$\mathcal{L}(\boldsymbol{l}) = \sum_{s} u_L(w - x_s + l_s - p(\boldsymbol{l}))f_s^L - \lambda(\sum_{s} u_H(w - x_s + l_s - p(\boldsymbol{l}))f_s^H - A).$$

Let u'_{Ls} denote the derivative of lower-risk type utility function with regard to consumption in loss state s. The first-order condition is:

$$u'_{Ls} - \lambda \frac{f_s^H}{f_s^L} u'_{Hs} \leq \theta \left(\sum_{\tau} u'_{L\tau} f_\tau^L - \lambda \sum_{\tau} u'_{H\tau} f_\tau^H \right) \forall s, \tag{4}$$

with equality if $l_s > 0$. Note that since the right-hand side is a constant, $u'_{Ls} - \lambda \frac{f_s^H}{f_s^L} u'_{Hs}$ is the same across loss states with $l_s > 0$.

Now take two loss states s and t such that $x_s \neq x_t, l_s > 0$ and $l_t > 0$. $\frac{f_s^H}{f_s^L} \neq \frac{f_t^H}{f_t^L}$ by assumption. Suppose that a straight deductible is optimal, then $l_s - x_s = l_t - x_t$ (equal consumption when losses are larger than the deductible level). This then implies that $u'_{Ls} = u'_{Lt}$ and $u'_{Hs} = u'_{Ht}$. But since $\frac{f_s^H}{f_s^L} \neq \frac{f_t^H}{f_t^L}$, $u'_{Ls} - \lambda \frac{f_s^H}{f_s^L} u'_{Hs} \neq u'_{Lt} - \lambda \frac{f_t^H}{f_t^L} u'_{Ht}$, contradictory to (4). As a result, the optimal plan for the lower-risk type cannot be a straight-deductible plan. ■

Proposition 3 Proof:

Take any loss state s, and assume that the $\frac{f_s^L}{f_s^H} = \alpha$. The first-order conditions of the coverage in state s for H are:

$$u'_{sH} \leq \frac{\theta}{2}(1+\alpha) \sum_{\tau} u'_{\tau H} f_\tau^H, \forall l_s,$$

with equality if $l_s > 0$. Similarly, the first-order conditions for L are:

$$u'_{sL} \leq \frac{\theta}{2}(1+\frac{1}{\alpha}) \sum_{\tau} u'_{\tau L} f_\tau^L, \forall l_s,$$

with equality if $l_s > 0$.

For any two loss states $x_t > x_z$, we know that $\frac{f_t^L}{f_t^H} < \frac{f_z^L}{f_z^H}$. This means whenever $l_t > 0$ and $l_z > 0, u'_{tH} < u'_{zH}$ and $u'_{tL} > u'_{zL}$. Since $u''_H < 0$ and $u''_L < 0$, $l^*_{Ht} - x_t \geq l^*_{Hz} - x_z$ and $l^*_{Lt} - x_t \leq l^*_{Lz} - x_z$. That is, the consumption in state t is always no smaller than the consumption in z for the higher-risk type, and the consumption in state z is always no smaller than the consumption in t for the lower-risk type.

Note that among the plans with non-increasing consumption, the implied consumption in t cannot be larger than the implied consumption in z. This means the higher-risk type will either have zero indemnity at small loss states or a constant consumption c^* once the indemnity is positive. The higher-risk type would want larger consumption in state t than in z, but cannot because of the non-increasing consumption constraint.[16] This means the higher-risk type will desire a straight-deductible plan. The lower-risk type is not constrained and will desire a plan with larger consumption for smaller losses (x_s) than in larger losses (x_t), a non-straight-deductible design. ■

[16] The non-increasing consumption constraint implies that for any loss states z and t where $x_t > x_z$, the allowed plan must imply $c_t \leq c_z$, where c_s denote consumption in state s. Empirically, since most health insurance plans accumulate spending over a year, almost all comprehensive health insurance plans satisfy this condition.

Corollary 1 Proof:

The optimization problem for i is:

$$\max_{l} \sum_{s} u_i(w - x_s + l_s - \theta \sum_{s} f_s^L l_s - c) f_s^i$$

subject to:

$$\theta \sum_{s} f_s^L l_s + c = A,$$
$$0 \leq l_s \leq x_s.$$

$\boldsymbol{l} = (l_1, l_2, \ldots, l_s, \ldots, l_S)$ is the vector of the insurance payments in each state. A is the fixed premium level individuals are required to choose from.

The Lagrange of the above optimization problem is:

$$\mathcal{L}(\boldsymbol{l}) = \sum_{s} u_L(w - x_s + l_s - p(\boldsymbol{l})) f_s^L - \lambda(\theta \sum_{s} f_s^L l_s + c - A).$$

Take any loss state s, and assume that the $\frac{f_s^L}{f_s^H} = \alpha$. The first-order condition of the coverage in state s for H is:

$$u'_{sH} \leq \frac{\theta}{2}(1 + \alpha)(\sum_{\tau} u'_{\tau H} f_\tau^H + \lambda), \forall l_s,$$

with equality if $l_s > 0$. Similarly, the first-order conditions for the lower-risk type are:

$$u'_{sL} \leq \frac{\theta}{2}(1 + \frac{1}{\alpha})(\sum_{\tau} u'_{\tau L} f_\tau^L + \lambda), \forall l_s,$$

with equality if $l_s > 0$. The FOCs are the same as the FOCs in Proposition 3 except for adding a constant in the last term of the right-hand side. All the other arguments follow as the proof of proposition 3. ■

Appendix B. Numeric Example Details

B.1. Constructing Risk Distributions from Claims Data

I need information about the ex-ante medical expenditure distributions to calculate plans chosen by different risk types and simulate the welfare implications of plan standardization policy. I derive such information using the Truven MarketScan data, a large claims database for US employer-sponsored plans. The Truven data have been used to benchmark health spending in many studies (for example, Geruso, Layton, and Prinz 2019) and to calculate the AV for plans in the first two years of the ACA markets. I select a random 5% sample of individuals enrolled in a non-capitated plan in 2012 and 2013. In total, there are 190,283 unique individuals in the sample.

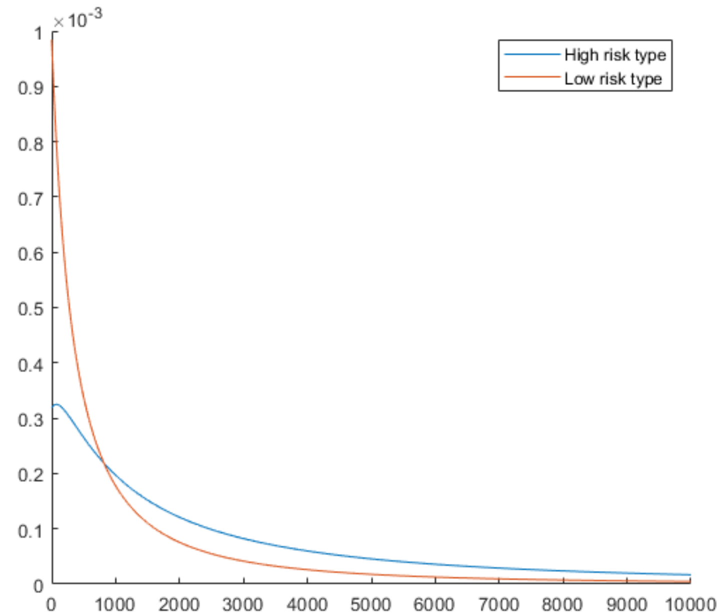
The goal is to construct a few ex-ante risk types representing the heterogeneity in medical expenditure in the US health insurance markets. I use the k-means clustering method to get these groupings. K-means clustering is a non-supervised learning algorithm that groups individuals with similar characteristics and puts individuals with dissimilar characteristics in different groups (Agterberg et al., 2019).[17] I use age, gender, employment

[17] This method is different from the supervised learning approach (such as regressions) to predict medical expenditure and construct risk scores (Kautter et al. 2014).

status, dummies for pre-existing chronic conditions (constructed based on diagnosis and procedure codes), and medical expenditure in 2012 as inputs to the model.

For illustrative purposes, I first create two clusters and use them to separate the population into two risk types. After obtaining the clusters, I fit a three-parameter log-normal distribution with a mass at zero to the 2013 medical expenditure for each group to get the risk distribution (Einav et al. 2013) and inflate the expenditure to 2017 dollars.

Appendix Figure B1. PDF of the Two Benchmark Risk Distributions



Note: Author estimation from Truven MarketScan data. Mass at zero was omitted for ease of exposition.

The resulting lower-risk type has an expected risk of \$1,843 and a standard deviation of \$7,414, representing 26% of the population in the sample. The higher risk has an expected risk of \$7,537 and a standard deviation of \$22,444. Appendix Figure B1 plots the probability density function of the two distributions. The lower risk has a 28.56% probability of incurring no losses, and the higher risk has a 4.53% probability of incurring no losses (not plotted in the graph.) The graph shows that the two probability density functions have different shapes: The low-risk type has greater probability density on smaller losses while the high-risk type has greater probability density on larger losses.

I then use similar methods to create 100 risk types and use them in the simulation exercise in Section 4.

B.2. Calculating Equilibrium Plans

I parameterize the consumer preference using the constant-absolute-risk-aversion (CARA) utility function. Consumers are risk averse with a risk-aversion coefficient γ = 0.0004, which is the mean level of risk aversion estimated by Handel (2013) for a population of employees choosing health insurance plans.

I consider a choice set with rich variation in the cost-sharing attributes. I allow for two broad categories of plan designs. The first category of plans has a three-arm design with four plan attributes: A deductible, a MOOP, a coinsurance before the deductible, and a coinsurance rate after the deductible. To make the simulation tractable, I discretize the contract space and assume the MOOP is no larger than $100,000. The second category consists of constant coinsurance plans, with a coinsurance rate ranging between zero and one. Both full insurance (in the form of zero constant coinsurance to consumers) and no insurance are in the choice set.

For the premium, I assume insurers charge a premium 20% higher than the claims costs (θ = 1.2).[18] I simulate the premiums of each plan, either under risk-based pricing (plan premiums differ by which type chooses the plan) or perfect risk adjustment. For the no-risk adjustment case, I follow the equilibrium notion by Azevedo and Gottlieb (2017) to calculate the equilibrium plans.

Appendix C. Supplementary Materials for the Empirical Analysis

In this section, I present supplementary tables and figures for the empirical analysis in Section 3.

Appendix Table C1. States in the Sample

2014	AK, AL, AR, AZ, DE, FL, GA, IA, ID, IL, IN, KS, LA, ME, MI, MO, MS, MT, NC, ND, NE, NH, NJ, NM, OH, OK, PA, SC, SD, TN, TX, UT, VA, WI, WV, WY
2015	AK, AL, AR, AZ, DE, FL, GA, , IA, IL, IN, KS, , LA, ME, MI, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NV, OH, OK, OR, PA, SC, SD, TN, TX, UT, VA, WI, WV, WY
2016	AK, AL, AR, AZ, DE, FL, GA, HI, IA, IL, IN, KS, , LA, ME, MI, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NV, OH, OK, OR, PA, SC, SD, TN, TX, UT, VA, WI, WV, WY
2017	AK, AL, AR, AZ, DE, FL, GA, HI, IA, IL, IN, KS, KY, LA, ME, MI, MO, MS, MT, NC, ND, NE, NH, NJ, NM, NV, OH, OK, OR, PA, SC, SD, TN, TX, UT, VA, WI, WV, WY

Appendix Table C2. Data Source of Empirical Analysis

Panel A.

Data Source	Link	Unit of Observation	Key Variables	Analysis Using the Data
Health Insurance Exchange Public Use Files: 2014-2017	https://www.cms.gov/CCIIO/Resources/Data-Resources/marketplace-puf; https://www.cms.gov/CCIIO/Resources/Data-Resources/issuer-level-enrollment-data	Plan ID by year	Deductible, MOOP, coinsurance rates, AV, enrollment, HSA-eligibility	Figure 1-2, Table 3, Appendix Figure C2
		Plan ID by rating area by year	Premiums	Table 4 Column (3)
Uniform Rate Review Data: 2016 - 2019	https://www.cms.gov/CCIIO/Resources/Data-Resources/ratereview	Plan ID by year	Total expenditure and collected premiums per	Figure 3, Table 4 Column (1) and (2),

[18] Regulations adopted as part of the Affordable Care Act require insurers to have at least 80% or 85% (depending on the size) of their premium used to cover claims costs. When this regulation binds, it implies a loading factor of around 1.2.

			member per month (PMPM)	Appendix Figure C3, Appendix Table C4-C5
		Insurer by year	Total expenditure, insurer liability, risk transfers, and average premium PMPM	Table 5, Appendix Table C5- C7
Medical Loss Ratio filings: 2014-2017	https://www.cms.gov/CCIIO/Resources/Data-Resources/mlr	Insurer by year	Risk transfers PMPM	Appendix Table C7
Open Enrollment Period County-Level Public Use File	https://www.cms.gov/data-research/statistics-trends-and-reports/marketplace-products/2017-marketplace-open-enrollment-period-public-use-files	County by year	Enrollment share in CSR and premium subsidies	Figure 4

Panel B. Matching across Different Datasets

Datasets	# of insurer-year: 2014-2017	% matched
Insurer-year with plan information	821	100%
Uniform Rate Review Data	619*	75.4%
Medical Loss Ratio filings	796	97.0%
Combined – risk transfers	746	90.9%

Note: Each plan ID represents a unique combination of cost-sharing structure, plan type, drug formulary, and insurer. Cost-sharing variations are dropped for Silver plans, so only standard Silver plans are included in the sample. The Uniform Rate Review Data have a two-year lag, so the 2016 - 2019 reports match the 2014 - 2017 plan information. *The numbers are slightly larger than the sample size reported in Table 3 and Appendix Table C4 because two observations are absorbed by fixed effects.

Appendix Table C3. List of Essential Health Benefits

Category	Benefit Name
Medical Services	Emergency Room Services, Inpatient Physician and Surgical Services, Imaging (CT/PET Scans, MRIs), Laboratory Outpatient and Professional Services, Outpatient Surgery Physician/Surgical Services, Mental/Behavioral Health and Substance Use Disorder Outpatient Services, Outpatient Facility Fee (e.g., Ambulatory Surgery Center), Occupational and Physical Therapy, Primary Care Visit to Treat an Injury or Illness (exc. Preventive, and X-rays), Specialist Visit, Skilled Nursing Facility, Speech Therapy, X-rays and Diagnostic Imaging.
Drug Tiers	Generics, Preferred Brand Drugs, Non-Preferred Brand Drugs, Specialty Drugs (i.e. high-cost).

Appendix Table C4. Plan Design and Other Plan Characteristics

	Non-Straight-Deductible Plans	Straight-Deductible Plans	Difference
HMO	0.502	0.504	0.002
	(0.500)	(0.500)	(0.018)
National Network	0.327	0.322	-0.005
	(0.469)	(0.467)	(0.017)
HSA Eligible	0.130	0.723	0.593***
	(0.336)	(0.448)	(0.012)
New Plan	0.506	0.488	-0.017
	(0.500)	(0.500)	(0.018)
Fraction of Launched	0.366	0.365	-0.001
Counties That Are Rural	(0.311)	(0.336)	(0.011)
N	6,931	911	7,842

Note: The sample includes plans between 2014 and 2017 with a premium increase of more than 10%. I dropped those reporting non-positive total expenditure or premium and plans with the top and bottom one percent of either value to avoid impact from extreme values. Each observation is a plan-state-year. Means and standard errors in parenthesis. *: p<0.1, **: p<0.05, ***: p<0.01. "HMO" stands for health maintenance organization, as opposed to other managed care plan types, including preferred provider organization (PPO), exclusive provider organization (EPO), or point of service (POS) plans.

Appendix Table C5. Robustness Check: HSA Eligibility

Panel A. Plans with HSA

	(1)	(2)	(3)	(4)
	Dependent Variable: total medical expenditure per member month			
Straight-Deductible Plan	189.81 (27.86)			
% losses covered for first $2,000		-12.43 (4.21)		
Risk premium, $100			-78.85 (19.69)	
Deductible to MOOP ratio				442.22 (68.46)
N	1,441	1,441	1,441	1,441
R^2	0.66	0.65	0.66	0.67
y-mean		570.98		
y-sd		425.44		
Controls	metal tier, network type, HSA-eligibility, insurer FE, year FE, service area FE			

Panel B. Plans without HSA

	(1)	(2)	(3)	(4)
	Dependent Variable: total medical expenditure per member month			
Straight-Deductible Plan	58.17 (31.92)			
% losses covered for first $2,000		-3.22 (0.63)		
Risk premium, $100			-13.68 (3.13)	
Deductible to MOOP ratio				33.62 (18.32)
N	6,261	6,261	6,261	6,261
R^2	0.57	0.57	0.57	0.57
y-mean		552.31		
y-sd		371.77		
Controls	metal tier, network type, HSA-eligibility, insurer FE, year FE, service area FE			

Note: The sample includes plans between 2014 and 2017 with a premium increase of more than 10%. I dropped those reporting non-positive total expenditure or premium and plans with the top and bottom one percent of either value to avoid impact from extreme values. Each observation is a plan-state-year. "Straight-deductible plan" is a dummy variable indicating whether the plan has a straight-deductible design. "% losses covered for first $2,000" measures each plan's fraction of losses covered for the first $2,000 total medical expenditure evaluated for the individual with market-average risk. "Risk premium" measures the difference in risk premium of choosing the plan relative to the straight-deductible plan with the same actuarial value. The dependent is the average total medical expenditure per member per month. Standard errors are clustered at the insurer level and shown in parentheses.

Appendix Table C6. Straight-Deductible Offering and Insurer Characteristics

	Insurers offering no straight-deductible plans	Insurers offering at least one straight-deductible plan	Difference
Offering HMO	0.546	0.560	0.014
	(0.499)	(0.497)	(0.035)
Offering National Network	0.240	0.254	0.014
	(0.428)	(0.436)	(0.030)
Fraction Enrolled in HSA-eligible Plans	0.134	0.182	0.048***
	(0.195)	(0.170)	(0.013)
Operating in rural areas	0.223	0.183	-0.039
	(0.417)	(0.386)	(0.028)
Above Median Enrollment	0.460	0.526	0.066
	(0.499)	(0.500)	(0.041)
Observations	252	367	619

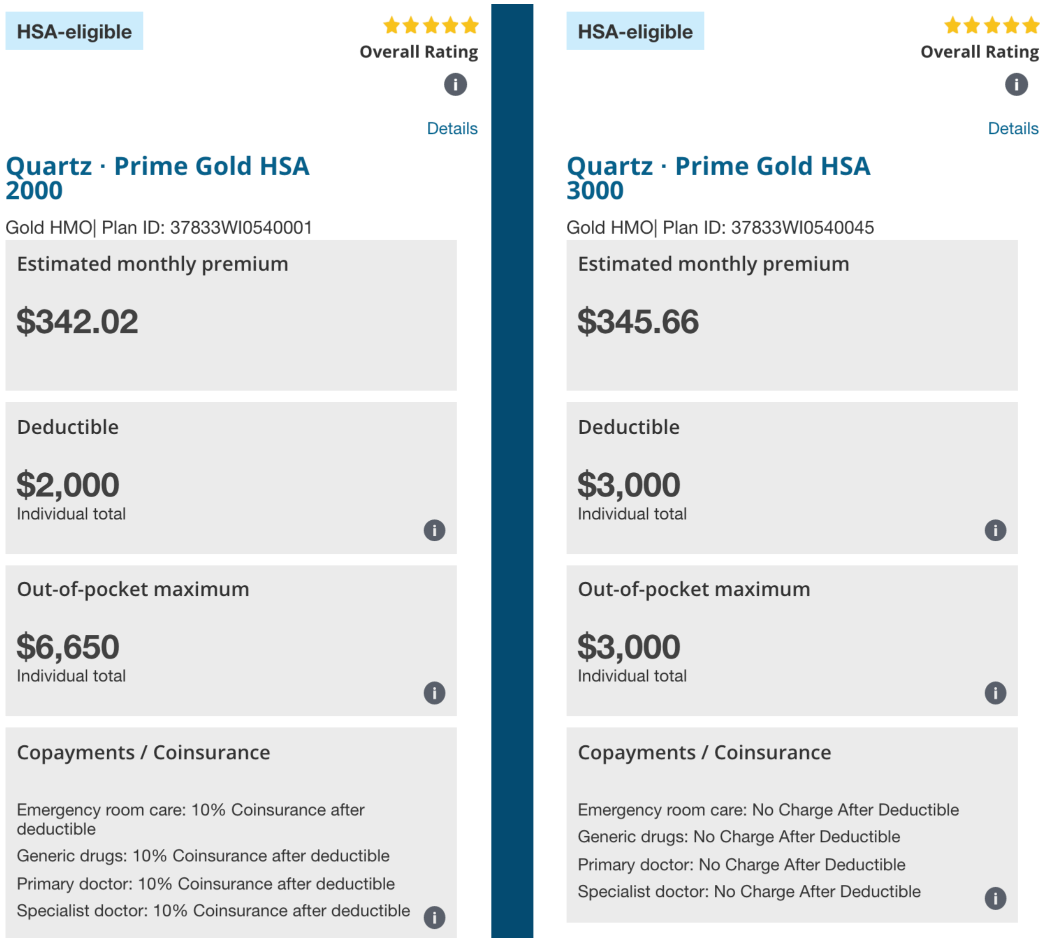
Note: Means and standard errors are in parentheses. *: p<0.1, **: p<0.05, ***: p<0.01. "HMO" stands for health maintenance organization, as opposed to other managed care plan types, including preferred provider organization (PPO), exclusive provider organization (EPO), or point of service (POS) plans.

Appendix Table C7. Robustness Check: Risk Transfers at the Insurer-Level

Dep. Var. = Risk Transfers Per Member Month	(1) Baseline	(2) More Controls	(3) MLR Sample
Offer Straight-Deductible Plan	40.58	33.89	34.93
	(13.51)	(12.19)	(11.99)
N	617	617	744
R^2	0.144	0.222	0.122
Dep. Var. Mean	-6.201	-6.201	-6.125
Dep. Var. SD	66.03	66.03	63.41

Note: Each observation is an insurer-year. The dependent variable is risk transfers received per member month. "Offer Straight-Deductible Plan" is a dummy variable indicating whether an insurer offers any straight-deductible plan. All models include year fixed effects and state fixed effects. The fraction of enrollees in health savings account is controlled for all columns. Column (2) further controls for each insurer's fraction of enrollees in different metal tiers and network types. Column (3) impute the missing values using the Medical Loss Ratio Files. Combining information from the Medical Loss Ratio reports, over 90% of insurers that launched a plan have the risk transfer information. Standard errors are clustered at the insurer level.

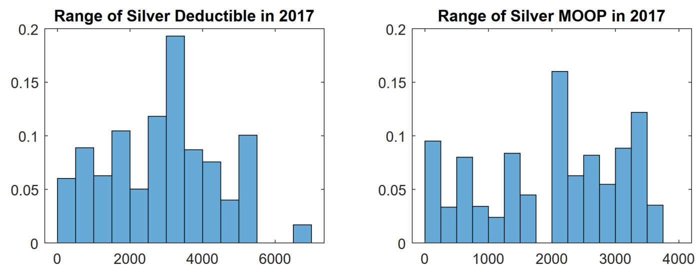
Appendix Figure C1. Illustration of Multiple Financial Attributes of ACA Plans



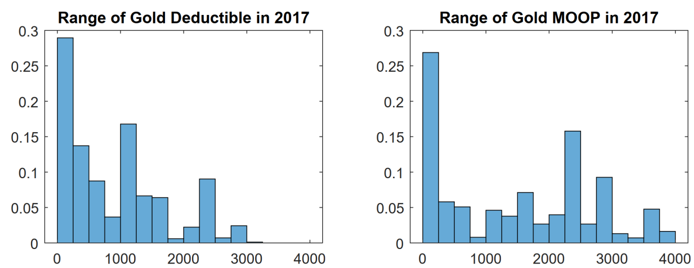
Note: Screenshots from healthcare.gov.

Appendix Figure C2. Deductible and MOOP Variation Within A County

Standard Silver Plans

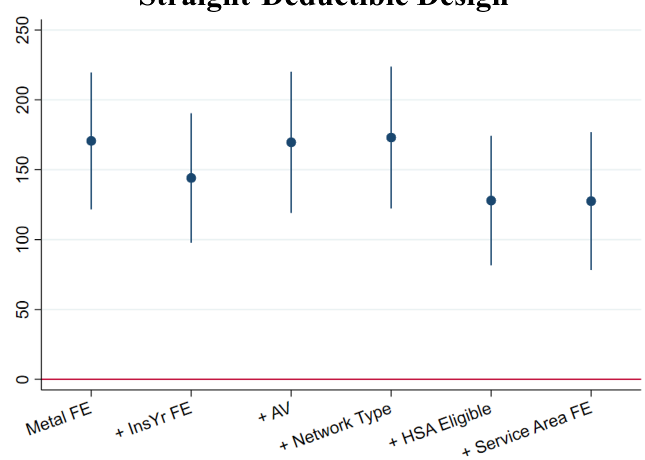


Gold Plans



Note: Data from 2017 CMS Health Insurance Exchange Public Use Files. I calculate the range of deductible (MOOP) for standard Silver and Gold plans within each county, then plot the distribution of all counties participating in the Federal Health Insurance Exchange. Plans include all exchange-qualified health plans offered to individuals through the Health Insurance Exchange. The deductible and MOOP refer to tier-one in-network coverage for an individual, and are cumulative over a year.

Appendix Figure C3. Total Medical Expenditure per Member Month and Straight-Deductible Design



Note: The figure shows the slope coefficient of the plan-level regression of average claim costs on whether the plan has a straight-deductible design. The sample includes all plans launched through HealthCare.gov between 2014 and 2017. Each observation is a plan by state. In each line, control variables are added on top of the left model, so for example, in the second line, both the metal fixed effects and insurer by year fixed effects are controlled. "Metal" represents metal-tier fixed effects. "InsYr FE" is insurer-by-year fixed effects; "AV" represents the actuarial value of a plan; "Network Type" includes three dummy variables indicating HMO, EPO, POS, and PPO (the baseline); "HSA eligible" is a dummy variable indicating whether a plan has a health savings account available; "Service Area FE" include dummy variables indicating the set of counties a plan is launched.

Appendix D. Estimating ACA Plans' Risk Premium

To quantify the economic value of different plan designs, I estimate each ACA plan's value to the market-average risk. Let a denote the stochastic out-of-pocket spending implied by the plan, with a distribution H. Define risk premium, R, as follows:

$$E[u(w-a)] = u\,(w - E(a) - R),$$

where w represents the wealth level, a represents the stochastic out-of-pocket spending and $E(a)$ represents its expected value, and $u(\cdot)$ is the utility function. The risk premium represents the sure amount the individual needs to receive to be indifferent between enrolling in that plan and a full-insurance plan, when both are priced at their fair AV. I now specify how H and $u(\cdot)$ are calculated.

Step 1. Calculating Out-of-Pocket Distributions for All Plans

I calculate H for each plan by applying each plan's cost-sharing rules on the market average risk distribution.

First, I collect the cost-sharing features of a plan's first-tier in-network coverage for essential health benefits. The utilization rate of the first-tier in-network coverage is 94.59% on average for the sample plans, and 99.47% of the total premium is contributed to cover the essential health benefits on average. I exclude preventive care because all plans must cover it with no cost-sharing. I collect each plan's deductible, MOOP, and copay/coinsurance rates for each benefit.

Second, I retrieved the ACA market representative distribution from the AV calculator, a tool created by CMS to compute the AV of each plan. The calculator contains a continuation table of the representative individual's medical expenditure distribution. The table is organized as follows: the overall distribution is discretized into 84 cells, each representing a range of total expenditure levels (e.g., 0, 0-100, 100-200, etc.). In each cell i, the table reports the average total expenditure level, x^i, the probability of being in that cell, p^i, and the expenditure amount and utilization frequency of each of the 17 benefits. Appendix Table C3 shows the list of benefits.

Third, I apply each plan's cost-sharing rules to the representative individual's expenditure distribution and calculate the out-of-pocket spending for each plan in each cell. Suppress the notation for each plan. Let i denote the i-th smallest total expenditure cell, x^i denote the total expenditure level in the cell, o^i denote the out-of-pocket spending level. The goal is to construct a mapping from x^i to o^i for all cells. Let x_s^i denote the expenditure amount for benefit s in that cell, n_s^i denote the utilization frequency of the benefit, c_s denote the coinsurance rate, q_s denote the copayment amount, d denote the deductible level, m denotes the MOOP, and o_s^i denote the out-of-pocket spending level in cell i. Let G denote the set of benefit subject to the deductible.

Fix a cell i. If $s \in G$, $o_s^i = x_s^i$. Otherwise, I apply the copay and coinsurance rules. If the benefit has coinsurance rate, $o_s^i = c_s x_s^i$. If the service has copays, $o_s^i = n_s^i q_s$. Among benefits subject to the deductible level, calculate the total amount subject to the deductible level: $g^i = \sum_G x_s^1$. Examine whether g^i succeeds the deductible level. If $g^i > d$, then allocate $g^i - d$ among all services subject to the deductible level proportionally to x_s^i, and then apply the copay and coinsurance rules. Next, check if $o^i \geq m$. If so, then replace $o^i = m$.

The result is the discrete version of the out-of-pocket spending distribution, $\{o^i, p^i\}_{i=0,1,\dots,83}$, for each plan i.

Step 2. Specifying the Utility Function.

In the calculation, I assume a CARA utility function with a risk-averse coefficient of 0.0004. I assume the risk aversion coefficient is 0.0004, the median and mean estimated by Handel (2013.) Under the CARA utility function, w is irrelevant.

Step 3. Calculating Risk Premiums

Given $a \sim H(\cdot)$ and $u(\cdot)$, I plug in them into $E[u(w-a)] = u\,(w - E(a) - R)$ to calculate R for each plan.

Appendix E. Calibration Details.

Step 1. Create Simplified Three-Arm Designs of ACA Plans

For each plan in the ACA market, I observe its deductible, MOOP, and cost-sharing rules for different benefits. The design of a plan is rather complicated because it involves cost-sharing rules for many different benefits. To make computation tractable and consistent with the theoretical analysis, I convert each plan's cost-sharing rules into a simplified three-arm design (Ericson et al., 2019; Liu and Sydnor, 2022).

Define a plan design as a mapping from total medical expenditure level to the out-of-pocket spending: $g: x \to R_0^+$. The simplified plan design is a piece-wise linear version of this function. The design contains three legs: a leg before the deductible level, a leg after hitting the deductible and before the MOOP, and a flat leg after hitting the MOOP. The design is represented by four parameters: the deductible level d, the MOOP, m, the coinsurance rate before the deductible, c_1, and the coinsurance rate after the deductible and before the MOOP, c_2.[19] The simplified design of a plan has the same fraction of losses covered for the average consumer in the ACA market for each leg as the original design.

I take the stochastic out-of-pocket spending distribution implied by each plan estimated in Appendix D Step 1. I then construct the simplified plan design by calculating the four parameters. The deductible and MOOP are directly observed in the data using the first-tier in-network values for single coverage.[20] Define two related values: the expenditure level when hitting the deductible, l_1, and the expenditure level when hitting the MOOP, l_2. Given $\{x^i, o^i\}_{i=0,1,\dots,83}$, calculate l_1 as the smallest expenditure level such that the out-of-pocket expenditure succeeds the deductible: $l_1 = argmin_{x_i}\{o^i - d : o^i \geq d\}$. l_2 is calculated as the smallest expenditure level such that the out-of-pocket expenditure succeeds the MOOP: $l_2 = argmin_{x_i}\{o^i - m : o^i \geq d\}$. Given l_1 and l_2, I then calculate $c_1 = d/l_1, c_2 = (m-d)/(l_2 - l_1)$.

Under the three-arm design, each plan design f is defined as:

$$g(x) = \begin{cases} c_1 x, & if\ x \leq \dfrac{d}{c_1}. \\ d + c_2(x-d), & if\ \dfrac{d}{c_1} < x \leq \dfrac{m-d}{c_2} + \dfrac{d}{c_1} \\ m, & if\ x > \dfrac{m-d}{c_2} + \dfrac{d}{c_1} \end{cases}$$

[19] In the ACA, all plans are required to have a MOOP, and plans offer full coverage on covered benefits after the cumulative out-of-pocket spending succeeds the MOOP. Some plans have certain benefits covered even before hitting the deductible level, thus I allow a coinsurance rate before the deductible level.

[20] Some plans have separate deductible or MOOP for drugs and medical services. I aggregate them into a single value.

Step 2. Create individual types

I create 100 risk types estimated using the Truven MarketScan data using the k-means clustering method in Appendix B. I scaled up the parameters such that the average market risk is the same as the 2017 ACA average level (obtained from the AV Calculator).

In the simulation, each market is a county. Let m denote each market. I assume that all counties have the same risk distributions. In each county, for each risk type, I then create five subtypes, representing consumers facing non-CSR plans and no premium subsidy, non-CSR plans and premium subsidy, and both (3 CSR plan types). This implicitly assumes that the medical expenditure risk is independent of income level. I use the 2016 CSR total enrollment share and premium subsidy share to get the relative weights of each type in a county. This is the only year where such data are available. I also use the 2017 county-level average premium subsidy and assume that it is the subsidy received by those eligible for premium subsidy. When aggregating to the national level, each county is weighted by the overall ACA enrollment share in 2017. The sub-types face different choice sets (the CSR population can choose the CSR-variation Silver plans instead of Standard Silver plans) and different net premiums.

Finally, consumers are assumed to have a CARA utility function with a risk aversion coefficient of 0.0004, the average and median value estimated by Handel (2013).

Step 3. Create the choice sets and premiums

I get the plans launched in each county in 2017 and fix them as the baseline. Let $C_{m,1}$ denote the sets of plans available in county m. For each plan in $C_{m,1}$, create a counterfactual plan with the same AV and has a straight-deductible design. The AVs are evaluated using the CMS 2017 AV Calculator distribution as in Step 1. Let $C_{m,2}$ denote the collection of plans.

For all plan j in both sets of plans, I calculate the expected covered losses to risk type i, τ_{ij} using the simplified three-arm design. I then calculate the premium p_j as the perfectly competitive, perfect risk adjusted premium with a loading factor of θ:

$$p_j = \theta \sum_i \tau_{ij} w_i,$$

where w_i is the weight of each type, τ_{ij} is the expected covered losses. In the baseline simulation, I set $\theta = 1.2$, the value is implied from the medical loss ratio regulation. The formula is only applied to non-CSR plans. By regulations, the cost-sharing reduction variation plans have the same premium as the associated non-CSR Silver plans. Further, I calculate τ_{ij} as

$$\tau_{ij} = \int \left(x - g_j(x)\right) dF_i,$$

where $g_j(x)$ is plan j's three-arm design estimated in step 1, $x \sim F_i$ is individual i's shifted log-normal distribution estimated in step 2.

The net premium, p_{ij}, for individuals eligible for the premium subsidy i is

$$p_{ij} = p_j - s_i,$$

where s_i is the subsidy level, varies at the county level.

Step 4. Calculate the chosen plan and welfare

Consumers are modeled as expected utility maximizers choosing plans based on the perceived utility:

$$v_{ij} = \overbrace{\underbrace{\int u\big(-g_j(x) - p_{ij}\big)dF_i(x)}_{welfare-relevant\ utility} + \beta\epsilon_{ij}}^{decision\ utility}.$$

The deterministic part, $\int u\big(-g_j(x) - p_j\big)dF_i$, is a function of the out-of-pocket spending, $g_j(x)$, and net premium, p_{ij}. It determines the welfare-relevant value of each plan j for individual i. The second component of the choice utility is an error component, ϵ_{ij}, that is assumed to be i.i.d following the extreme value type one distribution. It affects the choice of each consumer but is not relevant to welfare. The larger the scaling parameter β, the more randomness there will be in plan choice. $\beta = 0$ represents the case where all consumers choose optimally.

Under CARA, $v_{ij} = \int u\big(-g_j(x)\big)dF_i(x) - p_{ij} + \beta\epsilon_{ij}$. $\int u\big(-g_j(x)\big)dF_i(x)$ is calculated using the CARA utility functional form and the shifted log-normal distributions. I vary β from 0 to some very large positive number.

Consumers choose the plan maximizing v_{ij}. Let $j_i^*(\beta)$ denote the plan chosen by individual i under β. Define $t_i(\beta) = \int u\left(-g_{j_i^*(\beta)}(x)\right)dF_i(x)$. The total efficiency of the market under β is calculated as

$$ss(\beta) = \sum_c \sum_{i \in c} w_i\big(t_i(\beta) - \theta\tau_{ij_i^*(\beta)}\big).$$

where $i \in c$ indicates individuals in county c, and w_i are population weights of type i in county c, as a fraction of the total population. $\tau_{ij_i^*(\beta)}$ is the social costs of providing insurance $j_i^*(\beta)$ to individual i.

The consumer surplus of individual i under β is:

$$cs_i(\beta) = t_i(\beta) - p_{ij_i^*(\beta)}.$$

Appendix Table E1 summarizes the parameters used in the calibration and the source of data:

Appendix Table E1. Summary of Parameters

Calibration parameters	Meaning	Source
u_i	Individuals' utility function. Assumed to be CARA and identical for all individuals	The risk aversion coefficient is from Handel (2013)
F_i	Loss distributions	Estimated from Truven Market Scan data and scaled to match the 2017 ACA average
g_j	Plan design, a mapping from total loss to out-of-pocket spending	Estimated from the ACA data
$C_{m,1}$	Choice set of county m	Obtained from the ACA data
$C_{m,2}$	Counterfactual choice set of county m	Created by the author
θ	Loading factor	Assumed to be 1.2
w_i	Population weights of each individual	Obtained from the ACA data

s_i	Premium subsidies	Obtained from the ACA data
β	The standard deviation of the error term	Varies by the author