Sorting on Plan Design: Theory and Evidence from the ACA

Chenyuan Liu[1]
University of Wisconsin – Madison
17 November 2019

Latest Version Available Here

Abstract

It is well known that under asymmetric information, adverse selection can lead to distortions in the level of insurance coverage purchased. This paper identifies an understudied dimension of sorting in insurance: sorting by plan design. In many markets, insurance plans have multi-dimensional cost-sharing attributes that can vary across plans with the same expected level of coverage. Classic theory suggests that plans with optimal risk protection have a straight-deductible design. I show that under asymmetric information, only high-risk individuals sort into such plans. Those with lower risk prefer plans that trade higher out-of-pocket limits for lower deductibles and coinsurance. Consistent with this theory, I find empirical evidence in the ACA Individual Exchange that (1) there is significant variation in plan design (i.e., level of out-of-pocket risk) across plans in the same coverage tier, and (2) the straight-deductible plans attract enrollees with higher risk. I consider the implications of sorting on plan design for efficiency and market regulation using simulations. I show that regulating away complex plan designs could have large negative effects in an otherwise unregulated competitive market. Further, in risk-adjusted markets such as the ACA Exchanges, limiting complex plan designs would likely not have large benefits unless consumer confusion plays an important role in choices.

---

[1] Contact information: chenyuan.liu@wisc.edu. I thank my advisor Justin Sydnor for his guidance throughout the development of the paper. I also thank Naoki Aizawa, Daniel Bauer, Thomas DeLeire, Yu Ding, Ty Leverty, Yue Li, John Mullahy, Margie Rosenberg, Casey Rothschild, Joan Schmit, Alan Sorensen and numerous seminar participants for helpful comments. Financial support from the Geneva Association is gratefully acknowledged. All errors are my own.

In the U.S., consumers often face options between health insurance plans with complicated financial designs. Plans can differ from each other along multiple financial dimensions, such as deductibles, coinsurance rates, and out-of-pocket limits. What economic forces lead to this variation in the financial design of insurance coverage? This question is important both for our understanding of how insurance markets operate and for policy discussions on regulations regarding insurance design and the standardization of insurance products.

The existing literature provides an incomplete explanation. It is well established that under asymmetric information, consumers with different risks will sort into different insurance coverage (Rothschild and Stiglitz 1976). However, the literature has focused on sorting on the *level* of coverage along a single dimension, such as deductibles. A given level of coverage can be achieved by many different insurance *designs* using different combinations of financial attributes. It is an open question why plans vary along multiple financial dimensions and which plan designs will emerge under different market conditions.

I study this issue both theoretically and empirically and demonstrate that asymmetric information can generate substantial variation in insurance plan designs. Theoretically, I show that sorting under asymmetric information leads different risk types to demand plans with different insurance coverage designs. In particular, under asymmetric information, lower-risk types will be attracted to plans with very different designs than they would in insurance markets with perfect information. Empirically, I analyze the financial attributes and enrollment patterns for plans in the Affordable Care Act (ACA) Individual Federal Exchange and document both substantial variation in plan designs and patterns of sorting consistent with my theoretical predictions.

I begin the paper with my theoretical analysis. I show that asymmetric information distorts not only the *levels* of coverage people with different risks seek but also the *plan designs* they will prefer. There is a classic result in insurance economics showing that when risk-averse consumers face premiums reflecting their level of risk, the optimal plan design is a “straight deductible”, in which consumers pay full losses below the deductible, and make no payment once they hit the deductible (Arrow 1963; Gollier and Schlesinger, 1996). Straight-deductible designs provide the best risk protection and the lowest variance of uninsured risk among all designs. However, I show that under asymmetric information, the classic result holds for those with higher risk but not for those with lower risk. Asymmetric information creates a force pushing lower-risk types to choose plan designs that have more coverage for smaller losses (in the form of coinsurance) while sacrificing coverage on larger losses (in the form of higher out-of-pocket limits).[2]

I show with formal proofs and simulation examples that this pattern holds both in unregulated competitive environments and in regulated markets with risk adjustment.[3] In unregulated competitive environments, lower-risk types want insurance coverage but also want to sort into a plan that higher-risk types find unattractive. Plan designs that offer relatively strong coverage for smaller losses but high out-of-pocket limits strike this balance. In regulated markets with risk adjustment, lower-risk types cross-subsidize higher-risk types and face prices reflecting market-average risk levels. They prefer plans covering more of their own losses than the other type's losses. Holding fixed the premium level, a plan offering more coverage for smaller losses and sacrificing coverage for larger losses will give the lower-risk types more coverage than a straight-deductible plan.

In the second part of the paper, I examine the empirical relevance of sorting by plan design in the ACA Federal Individual Exchanges (Healthcare.gov), a market with risk adjustment regulations. I combine publicly available data on the cost-sharing attributes, premium, enrollment, and claim costs for the universe of 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. Within these tiers though, insurers have significant latitude in designing the cost-sharing attributes of their plans in different

[2] Lower (higher) risk types are defined as individuals having a distribution of losses weighted more heavily towards smaller (larger) losses.

[3] In a competitive market, plan prices reflect the cost of consumers who actually enroll. Under perfect risk adjustment, plan prices reflect the average cost of coverage under that plan for the full population.

combinations. I document that, within each of these coverage levels, insurers use this latitude and offer plans varying substantially in their designs. For example, consider the Gold Tier, where plans should cover around 80% of costs for the benchmark average population. A consumer shopping for plans in this tier would face plans with out-of-pocket limits varying by over $2,000 on average. This variation in plan designs within a metal tier is stable over time in the data, with no clear trends toward more plans with higher out-of-pocket limits or toward straight-deductible designs. I also find that the variation in plan designs is widespread among different counties and insurers.

The variation in financial attributes translates to economically meaningful differences in the risk protection provided by different plans within the same coverage tier. I quantify the design variation across plans by calculating the expected utility of choosing each plan for an individual with a market-average distribution of health risk and a moderate degree of risk aversion. For such an individual, the straight-deductible plan offers the best risk protection. Sorting into the other plans available in each tier can have utility costs equivalent to as much as $1,000 per year.

This variation in plan design creates room for sorting by risk types in the ACA market. The theoretical prediction is that plans with straight-deductible designs will be favorable to those with average to above-average risk, but will be unattractive for the lower-risk participants. Using plan-level claim costs and insurer-level risk transfers collected from the Uniform Rate Review dataset, I find that within a metal tier, the straight-deductible plans are chosen by individuals with significantly larger ex-ante risk scores and ex-post medical expenditure. The theoretical prediction of sorting by risk type into different plan designs is clearly reflected in these measures.

In the absence of risk adjustment, straight-deductible plans attracting higher-cost enrollees would be expected to put pressure on the premiums of those plans. However, I find that premiums are very similar for different plan designs within the same metal tier. This suggests that the risk adjustment scheme and pricing regulations in the ACA Exchanges are blunting the pass-through of these selection differences to consumers. Ultimately, the patterns of sorting by risk type help explain why a wide range of plan designs can emerge in this market, while the effectiveness of the risk-adjustment scheme helps explain why the market does not converge to the designs that attract the lower-risk types.

In the last part of the paper, I examine the implications of variation in plan designs for market efficiency and regulation. To do so, I construct realistic distributions of health risks derived from Truven MarketScan data. I use the k-means clustering method to group people based on age, gender, prior expenditure, and pre-existing conditions and estimate separate spending distributions within these groups. I then simulate the change in market efficiency when only straight-deductible plans are allowed, using levels of risk aversion estimated in prior literature.

I show that the effects of variation in plan design on efficiency depend on the underlying market conditions. In unregulated competitive markets, plan design variation, specifically, the existence of high out-of-pocket limit plans, helps sustain a more efficient separating equilibrium. When people can sort along only one dimension of cost-sharing (i.e., deductibles), the lower-risk types end up sacrificing substantially more coverage to avoid pooling with the higher-risk types and may drop out of the market completely.

In a market with perfect risk adjustment, on the other hand, the existence of plan-design variation has more ambiguous and smaller impacts on market efficiency. Under risk adjustment, lower-risk types are inefficiently under-insured because they respond to market-average prices. For a given price, plans with non-straight-deductible designs provide lower-risk types more insurance. This can help offset their inefficiently low levels of coverage. On the other hand, on the margin, their decisions may be inefficient because the marginal price they pay for the additional coverage under these designs may be below the marginal cost of providing it to them. Ultimately, the overall impact of removing non-straight-deductible plans depends on the relative importance of the two effects and is ambiguous in risk-adjusted markets.

My simulations suggest that, for the ACA, the overall efficiency of the market may be similar with or without regulations on plan designs. I use a multinomial logit model to simulate a counterfactual scenario where the actual Exchange plans are replaced with straight-deductible plans of the same premium. In my baseline simulation, I estimate that the average efficiency would be slightly higher ($10 per person per year) under the design regulation than without. I also extend the model to allow for the possibility that some consumers might make mistakes when selecting plans, which has been shown to be an issue in health insurance choice in other settings (e.g., Abaluck and Gruber, 2011; Abaluck and Gruber, 2019; Bhargava, Lowenstein and Sydnor, 2017). The plans with high out-of-pocket limits that attract lower-risk individuals create the possibility of a costly mistake for lower-risk types. I show that the efficiency benefits of regulating plan designs in the ACA are significantly higher if a moderate share of consumers makes plan-choice mistakes.

My work contributes to three streams of literature. First, I build on and offer new insights to the literature on insurance coverage distortions under asymmetric information. Theoretical literature establishes that insurers distort contracts to screen lower-cost individuals (Rothschild and Stiglitz 1976, Crocker and Snow, 1985; Veiga and Weyl, 2016; Handel, Hendel, and Whinston, 2015; Azevedo and Gottlieb, 2017). My work extends this literature by showing that this sorting can happen along multi-dimensional cost-sharing features. Moreover, my results highlight that the availability of multi-dimensional cost-sharing designs could play a large role in improving the efficiency of perfectly competitive markets without risk adjustment. I also extend the literature by studying sorting into different plan designs in markets with perfect risk adjustment. My work shows that in risk-adjusted markets we can expect lower-risk individuals to sort into plans with both a distorted *level* of coverage and distorted cost-sharing *designs* relative to choices under perfect information.

Second, my analysis of the ACA exchanges also highlights the empirical relevance of adverse selection in the market with risk adjustment. I combine publicly available datasets in a novel way to show selection happens both across metal tiers and along different designs within a metal tier in the ACA Federal Exchange. Prior literature shows evidence that risk adjustment schemes are imperfect, and insurers exploit imperfect risk-adjustments by using non-price attributes to attract profitable consumers.[4] My work shows, however, that despite possible imperfections in the ACA risk adjustment scheme, it is effective at flattening the effect of adverse selection on the premiums across plans with different financial designs.

Finally, this paper identifies an understudied mechanism shaping the complex financial plan designs available in insurance markets. Prior literature identified how moral hazard (Pauly, 1968; Zeckhauser, 1970; Einav et al., 2013), nonlinear loading factors or risk-averse insurers (Raviv, 1979), background risk (Doherty and Schlesinger, 1983), and liquidity constraints (Ericson and Sydnor, 2018) can potentially lead people to select into these types of plan designs. This paper illustrates how asymmetric information can also rationalize the complex plan design variation. I find empirical evidence in the ACA Exchange that different risk types sort on different plan designs, suggesting this mechanism can play a role in rationalizing observed plan offerings and sorting in the ACA Exchange.

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 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

There is a long literature tracing at least to Arrow (1963) exploring the optimal design of insurance plans. A classic result from this literature is that straight-deductible plans offer optimal risk protection in the absence of moral hazard concerns and background risk (Arrow, 1963; Gollier and Schlesinger 1996). Under such plans, consumers pay full losses out-of-pocket before the deductible level and pay a fixed deductible level afterward. However, the optimality of the straight-deductible design hinges on the assumption that all plans are priced based on individual risk. With heterogeneous risk types and asymmetric information, theorems establishing the optimality of straight deductible plans do not apply. I show that in general, higher-

[4] Examples include managed care (Frank, Glazer, and McGuire, 2000), provider networks (Shepard, 2016), drug formulary (Geruso, Layton and Prinz, 2018), and advertisement (Aizawa and Kim, 2018).

risk types will sort into straight-deductible designs while lower-risk types trade more coverage for smaller events with higher worst-case risk.

2.1 Model Setup

Setting. Suppose individual i faces uncertainty in losses distributed over a set S of states with generic element s. Denote by x_s the size of the loss in state s. Individuals differ from each other by the probabilities of experiencing each loss state. Let f_s^i denote the probability of individual i being in state s.[5]

I consider a general state-dependent insurance plan that captures the wide range of potentially complex plan design consumers would desire. Specifically, an insurance plan is defined as a function $\boldsymbol{l}: s \rightarrow R^+$. I also define $l_s \equiv l(s)$ as the value of the function evaluated at s, so l_s represents the insurer indemnity in state s. The insurance payment l_s satisfies the condition $0 \leq l_s \leq x_s$, which implies the insurance payment is non-negative and no larger than the size of the loss. The expected insurer indemnity for individual i is $\sum_s f_s^i l_s$.

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 $\boldsymbol{l}$. Individuals have a concave utility function u_i over this financial outcome of each loss state: $u_i' > 0, u_i'' < 0$. Consumers choose insurance plans to maximize expected utility:

$$\max_{l} \sum_{s} u_i(w_i - x_s + l_s - p(\boldsymbol{l})) f_s^i, \tag{1}$$

subject to

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

and premium constraints. In general, if premiums are independent of type, then the optimal contract in (1) will differ across individuals: different individuals will choose different contracts.

Case 1. Single Risk Type/Risk-Based Pricing.

[5] In my setup, the loss states can be either discretely or continuously distributed. In the following notations I use summation over all loss states to represent the discretely distributed scenario. The notation can be adapted using integrals to represent the continuous case.

I first confirm that if there is a single risk type in the market (or, equivalently, if risk types are fully known and can be priced), then the classic results of Arrow (1963) and Gollier and Schlesinger (1996) apply given my model setup. That is, if all contracts have equal loading, then a straight deductible plan is the solution to (1) as long as it is an option in the choice set of possible plan designs.

For this single-risk-type case, I drop subscript i for simplicity of exposition. Assume perfectly competitive insurers set premium as a linear function of the expected covered expenditure: $p(\boldsymbol{l}) = \theta \sum_s f_s l_s + c$, where $\theta \geq 1$ is a proportional loading factor, and $c \geq 0$ is a fixed loading.[6] The premium of an insurance plan l is:

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

Suppose further that all possible insurance contracts are available and priced this way. Proposition 1 states the form of optimal insurance in this case:

Proposition 1. [Arrow: The Optimality of Straight-Deductible Plans under Risk-based Pricing] Suppose there is a single risk type in the market, and the premium is a linear function of the expected covered expenditure. For any fixed loading factors, the expected-utility-maximizing contract is a straight deductible plan.

Proof: See Appendix A.

One way to see the optimal insurance is to consider the first-order-condition:

$$u_s' \leq \theta \sum_\tau u_\tau' f_\tau \,, \forall s, \tag{2}$$

with equality if $l_s > 0$.

The left-hand side represents the marginal benefit of the reduction in out-of-pocket costs, and the right-hand-side represents the marginal costs (via the effects on premium) of increasing coverage for that loss state. The left-hand side is a decreasing function with regard to l_s because $u_s'' < 0$ for all loss states (consumers are risk averse). When $\theta = 1$, the left-hand side can be the same for all loss states. Essentially, individuals get full insurance. When $\theta > 1$, $u_s' < \theta \sum_\tau u_\tau' f_\tau$ for small loss states and thus $l_s = 0$ (no coverage for smaller losses). When x_s is large enough, $l_s > 0$ and u_s' is

[6] The loadings capture costs of operation for the competitive insurers.

the same across these loss states. This implies that for these loss states the consumption, $w - x_s + l_s - p(\boldsymbol{l})$, is constant. As a result, consumers pay a fixed deductible $(x_s - l_s)$ in these states. In summary, when $\theta > 1$, the optimal insurance is in the form of no coverage for small losses, and a deductible once the loss is large enough. Intuitively, a straight-deductible plan smooths consumption for large losses and has the lowest variance in uninsured risk, holding fixed the premium. The deductible level is determined by the loading factor and risk aversion level.

Case 2. Asymmetric Information with Separating Equilibrium

Now consider the case when there are two risk types (L and H) equally distributed in the market. Their respective probabilities of being in state s are f_s^L and f_s^H, and the utility functions are u_L and u_H respectively. Let l_{Ls}^* and l_{Hs}^* denote the utility-maximizing insurance payment in state s for each type.

To characterize the insurance contracts under this case, I need a model for how the premium is determined. I look at equilibria in which premiums are a mechanical function of the expected covered losses given who sort into that plan. Examples of such equilibria include Rothschild and Stiglitz (1976), and Azevedo and Gottlieb (2017). This rules out equilibrium concepts with cross-subsidization among plans (as in Spence 1978). Without delving into the details of the existence of such equilibrium, I characterize the equilibrium plans implied by the minimum necessary condition if they exist.

Analogous to Rothschild and Stiglitz (1976), I consider a separating equilibrium where one risk type (H) gets the optimal contract under full information, and the other type (L) distorts their coverage to prevent the higher-risk types from pooling with them.[7] A necessary condition for a plan being a separating equilibrium is that the contract purchased by the *lower*-risk type gives the higher-risk type no more utility than does the plan purchased by the higher-risk type (i.e., the contracts must be incentive compatible).

[7] Here, the H and L types are defined such that the incentive compatibility constraint for the H type is constrained while the incentive compatibility constraint for L is slack. Such an equilibrium may not exist. The exercise is to characterize the property of such an equilibrium if it exists.

In summary, the equilibrium plan chosen by the lower-risk types satisfy the following conditions:

$$0 \leq \ l_s \leq \ x_s,$$
$$p(\boldsymbol{l}) \ = \theta \sum_s f_s^L l_s + c,$$
$$\sum_s u_H(w - x_s + l_s - p(\boldsymbol{l}))f_s^H \ \leq \sum_s u_H(w - x_s + l_{Hs}^* - p(\boldsymbol{l}_H^*))f_s^H,$$

where $\boldsymbol{l}_H^*$ is the optimal plan H chooses under full information, and l_{Hs}^* denotes the insurance indemnity implied by that plan in loss state s. The only difference from the above case is the incentive compatibility constraint. This makes the plan desired by the lower-risk type not only depend on the loss distribution of the lower-risk type, but also the higher-risk type.

Proposition 2 shows the property of the plans desired by the lower-risk type:

Proposition 2. Suppose the market consists of two risk types, L and H. Assume that for any two loss states where $x_s \neq x_t$, we have $\frac{f_s^L}{f_s^H} \neq \frac{f_t^L}{f_t^H}$. Suppose all possible contracts are available and the premium is a function of the expected covered expenditure. Among all contracts giving H the same utility as the optimal contract for H under full information, the one that maximizes the utility of L has a non-straight-deductible design.

Proof: see Appendix A.

Proposition 2 states that if the ratio of the probabilities of the two types (f_s^L/f_s^H) is different across loss states, then the lower-risk types would sort into plans with a non-straight-deductible design. We can expect this condition to generally hold, for example, if both types have a log-normal distribution with different means. The intuition can be illustrated starting from the original straight-deductible plan the lower-risk type would choose under full information. Under such a plan, the consumption is the same across different loss states for all loss states above the deductible. Such a plan will not be incentive compatible, however, because it is priced based on the risk of the lower-risk types thus the higher-risk types will deviate. Thus, the lower-risk types need to change their coverage to prevent pooling with the higher-risk types. They could achieve this by either reducing coverage for larger loss states or reducing coverage for small loss states. Now if they reduce coverage of larger loss states and keep the same coverage of smaller loss states, they make the plan less attractive to the higher-risk types since larger losses are more likely to happen for higher-risk types. Sacrificing coverage for large losses and transferring to coverage for small losses is less problematic for the lower-risk types though, since most of their losses are likely to be small.

Case 3. Asymmetric Information with Perfect Risk Adjustment

In many markets, regulators impose risk adjustment regulations to flatten premium differences among different plans and to remove screening incentives for insurers. Such regulation enforces a cross-subsidization from lower-risk types to higher-risk types. 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 that *would* be obtained *if* all risk types enroll in the plan.[8] This approximates the regulatory environment in many U.S. health insurance markets, including Medicare Advantage, Medicare Part D, and the ACA Exchanges, and will be relevant for the empirical analyses in Section 3. [9]

The premium is:

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

The premium of insurance plans is a function of both types' risk distributions and is the same regardless of which type sorts into the plan.

[8] This definition is a special case of the Einav, Finkelstein and Tebaldi (2018) framework. In their definition, a risk adjustment is defined as a transfer r_i to the insurer if individual i enrolls in the plan. Insurers' profits are then defined as $p_j - (c_i - r_i)$. My setting is equivalent as setting r_i as the difference between c_i and the market average cost. In a perfectly competitive market, the insurer will then set the premium at the level of the market average cost. The premium formula can be thought as a reduced-form of the market equilibrium with risk adjustment. Geruso et al. (2019) also uses the same formula to define perfect risk adjustment.

[9] Except for its ability in capturing the key features of the regulatory environment, the perfect risk adjustment premium formula abstracts away from several institutional details. First, the formula abstracts away from the extensive margin of insurance demand. The premium formula is accurate only if $\overline{l_j}$ also incorporates risk types without any insurance. In practice, however, the risk adjustment transfers are typically based only on claim costs of people with a plan selection. Second, the premium formula implicitly assumes full community rating, while in many markets, including the ACA Exchanges, premiums can partially vary by age, household composition, etc.

In this environment, I consider a subset of plans with a non-increasing implied consumption: for any loss states where $x_z \leq x_t$, the implied consumption follows $c_z \geq c_t$. Almost all health insurance plans satisfy this property as the spending is typically accumulated annually. Proposition 3 characterizes the plans maximizing expected utility for each type:

Proposition 3. Suppose the market consists of two risk types, and the premium is a linear function of the expected covered expenditure of both types. Suppose $\frac{f_s^L}{f_s^H}$ is everywhere decreasing with regard to x_s. Among plans with non-increasing implied consumption, the higher-risk types will sort into a straight-deductible plan, but not the lower-risk type.

Proof: see Appendix A.

Under perfect risk adjustment, the premiums are effectively "shared" between the two types. The marginal cost of reducing out-of-pocket spending does not only depend on their own spending, but also the spending of the other type. Ideally, they want to have the premium covering more of their own spending than the spending of the other type. Holding fixed a premium level, the straight-deductible plans give the lower-risk types the lowest expected coverage relative to all other possible designs. This is because straight-deductible designs offer full coverage for larger losses, and it is that coverage that gets reflected in the premium for these plans. High-loss states though, are dominated by higher-risk individuals. Alternatively, a plan that gives up coverage of large losses for coinsurance rates of small losses is more beneficial to the lower-risk types because they are more likely to incur small losses, and thus get more expected coverage from such a plan. This suggests that the plan desired by the lower-risk types has coinsurance rates (if not full coverage) for smaller losses, and smaller or even no coverage for large losses. For the higher-risk types, the opposite is true. A straight-deductible plan provides them with higher consumption in larger loss states, and are thus are preferred.

Proposition 2 and 3 imply that under asymmetric information, different risk types will sort into different plan designs: the higher-risk type chooses a straight-deductible plan, but not the lower-risk type. I refer to this pattern as "sorting by plan design." Note that the classic adverse selection forces indicate that the coverage for the lower-risk type will be distorted downward (become less than the efficient level of coverage) in a single-loss state scenario. My results characterize the form of that distortion in a multi-loss-states environment when all designs are available: the design desired by the lower-risk types deviates from the first-best straight-deductible design that would hold under full information and has (partial) insurance for smaller losses.

2.2 Simulation

In this section, I present a numerical example illustrating the sorting pattern based on an empirically realistic set of risk distributions.

2.2.1 Simulation Setup

The Demand Side. In this simulation, I parameterize the consumer preference using the constant-absolute-risk-aversion (CARA) utility function. Consumers are risk averse with a risk-aversion coefficient $r_i > 0$.[10] For the following simulation, I also assume homogenous risk aversion for the two types and set $r = 0.0004$, which is the mean level of risk aversion estimated by Handel (2013) for a population of employees selecting among health insurance plans.

For my initial simulation, I continue to consider two risk types, but I also consider extensions to multiple risk types. In Section 2.2.2. below, I explain how I use data on health spending to model the ex-ante medical expenditure distributions of the different types in the simulation to reflect a realistic division of the population into a discrete set of types.

The Supply Side. For the premium, I assume insurers charge a premium 20% higher than the claim costs ($\theta = 1.2$).[11] I simulate the premiums of each plan, either under no risk adjustment or perfect risk adjustment. For the no risk adjustment case,

[10] This functional form removes income effects and is used in many prior works modeling insurance choice (e.g. Handel 2013; Abaluck and Gruber 2019).

[11] 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 claim costs. When this regulation binds, it implies a loading factor of around 1.2.

I follow the equilibrium notion by Azevedo and Gottlieb (2017) to calculate the equilibrium plans chosen by each risk type. The details are in Appendix A.

The Choice Set. 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: a deductible, an out-of-pocket-limit (OOP-limit, the maximum of the out-of-pocket spending per year), a coinsurance rate (the share of medical expenditure paid by consumers) before the deductible, and a coinsurance rate after the deductible. To make the simulation tractable, I discretize the contract space and assume OOP-limit is no larger than $100,000. The second category consists of constant coinsurance plans, with a coinsurance rate ranging between 0 and 1.[12] Both full and no insurance are in the choice set.

2.2.2 Constructing Risk Distributions from Claim Data

To simulate plans chosen by different risk types, I need information about the ex-ante medical expenditure distributions. I derive such information using the Truven MarketScan database, a large claim database for the U.S. employer-sponsored plans. The Truven data have been used to benchmark health spending in many studies and was also used to calculate the actuarial value (AV)[13] for plans in the first two years of the ACA markets. I select a sample of individuals enrolled in a non-capitated plan in both 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 U.S. 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 together and puts individuals with dissimilar characteristics in different groups (Agterberg et al. 2019).[14] I use age, gender, employment status, pre-existing conditions constructed

---

[12] The three-arm design is popular in the markets such as ACA Exchange, Medicare Part C and D, and employer sponsored insurance plans. Traditional Medicare plans are (a variation of) constant coinsurance plans.

[13] Actuarial value is defined as the fraction of expenditure covered by a plan for an individual with market-average risk. Under perfect risk adjustment, the premium of a plan is a function of AV.

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

based on diagnosis codes and procedures performed, and medical expenditure in 2012 as inputs to the model. For illustrative purposes, I initially create 2 clusters and use these to separate the population into two risk types. I create more clusters for analyses in later sections. After obtaining the clusters, I fit a 3-parameter log-normal distribution with a mass at 0 to the 2013 medical expenditure for each group to get the risk distribution (Einav et al., 2013) and inflate the expenditure to 2017 dollars.

The resulting lower-risk type has an expected risk of $1,700 and a standard deviation of $7,100, representing 26% population in the sample.[15] The higher risk has an expected risk of $7,200 and a standard deviation of $21,600. Appendix Figure 1 plots the probability density function of the two distributions. From the graph, it is clear that the two probability density functions have different shapes: the lower-risk type has greater probability density on smaller losses while the higher-risk type has greater probability density on larger losses.

2.2.3 Simulation Results

As expected from the analysis above, when there is no asymmetric information and both types face type-specific premiums, the straight deductible plans are optimal. Table 1 rows (1) and (2) show this result. The lower-risk type desires a plan with a $933 straight-deductible, and higher-risk type desires a plan with an $1820 straight-deductible. Both are plausible and similar to the plans observed in the employer-sponsored market and the ACA Exchange. The deductible level is larger than 0 (they get less than full insurance) because there is a positive loading in insurance plans. On average, the higher-risk type has 83% of their expenses covered, and the lower-risk type has 77% of their expenses covered.

When there is asymmetric information and no risk adjustment, the lower-risk type sorts into a non-straight-deductible design with a constant coinsurance rate (paid by consumers) of 23%. The plan covers 77% of medical expenditures for the

[15] In this classification, the median individual is classified as higher-risk type. In Section 2.3.1 I show the results when there are more than 2 risk types and illustrate the plans chosen by the extremely high risk type (who may only take a small fraction of the population) are straight-deductible plans.

lower-risk type at all loss levels.[16] Even though the coverage levels are similar to those under perfect information, there is a large distortion in the design such that the lower-risk type is subject to a large variation in their out-of-pocket spending. The higher-risk type has no distortion and gets the same straight-deductible plan they would have chosen under full information.

Table 1. Plans Chosen Under Different Market Conditions

		Risk type	deductible	out-of-pocket limit	co1	co2	% losses covered
Case 1: Single Risk Type/Risk-Based Pricing							
Risk-based pricing	(1)	H	1,820	1,820	1	0	83%
	(2)	L	933	933	1	0	77%
Case 2: Asymmetric Information with Separating Equilibrium							
No risk adjustment	(3)	H	1,820	1,820	1	0	83%
	(4)	L	0	∞	0.23	0.23	77%
Case 3: Asymmetric Information with Perfect Risk Adjustment							
Perfect risk adjustment	(5)	H	677	677	1	0	92%
	(6)	L	7,300	20,500	1	0.2	28%

Note: Numbers from author simulation. co1 is the coinsurance rate before the deductible (paid by consumers). co2 is the coinsurance rate after the deductible (paid by consumers). For example, a coinsurance rate before the deductible being one means the consumer pays full expenditures before hitting the deductible. A straight-deductible plan has 1) the same deductible and out-of-pocket limit, 2) full consumer cost-sharing before hitting the deductible (co1=1), and 3) no cost-sharing after hitting the deductible (co2=0).

The design distortion persists when there is perfect risk adjustment. The lower-risk type still sorts into a non-straight-deductible design, though the OOP-limit is much smaller. In this case, the lower-risk type sorts into a plan with a $7300 deductible, a $20,500 OOP-limit, and a 20% coinsurance rate between the two. On average, only 28% of the expenses are covered. This reflects both a coverage level

[16] This extremely high out-of-pocket limit in this example may be driven in part by the CARA utility form, which ignores income effects.

distortion and a design distortion. The higher-risk type, on the other hand, sorts into a straight-deductible plan with a $677 deductible. Note that risk adjustment also creates deviation for the higher-risk type: the higher-risk type is cross-subsidized by the lower-risk type and sorts into plans with higher coverage.

2.3 Model Extensions

The pattern of sorting by plan design illustrated above is not specific to two risk types with a homogenous risk aversion level. In this section, I discuss the robustness of the pattern with more than two risk types, heterogeneity in risk aversion, and the existence of moral hazard responses.

2.3.1 More than Two Risk Types

The sorting on plan design mechanism can be easily extended to a market with more than two discrete risk types. The intuition is similar: non-straight-deductible designs provide more coverage in risk states where the lower-risk types are more likely to experience relative to the higher-risk types. Appendix Figure 2 shows an example with 15 risk distributions constructed from the Truven MarketScan data. Under perfect risk adjustment, the types choosing the straight-deductible design have significantly more probability mass on the right tail.

2.3.2 Heterogeneity in Risk Aversion

A growing literature documents the existence of advantageous selection and attributes one explanation as heterogeneity in risk aversion, or more specifically, risk aversion being negatively correlated to risk types (Finkelstein and McGarry 2006; Fang, Keane, and Silverman 2008). In my model, adding heterogeneity in risk aversion does not change the basic sorting pattern where different risk types prefer different designs. The proofs of Proposition 2 and 3 incorporate cases where the two risk types have different risk aversion levels, or even different utility functions. The sorting pattern persists as long as both types are risk averse.[17]

---

[17] Proposition 2 relies on the assumption that at least one risk type has a binding incentive compatibility constraint.

A negative correlation between risk aversion and risk distribution can create a sorting pattern where both types sort into plans with similar *coverage levels,* yet they would still sort into *different designs.* For example, under perfect risk adjustment, if the above higher-risk type and lower-riks type have risk aversion coefficients of 0.0005 and 0.002 respectively, then both types will desire a plan covering 85% of the market average risk. The higher-risk type with lower risk aversion desires a plan with a straight-deductible of $1234, and the lower-risk type with more risk aversion desires a plan with a $700 deductible, $3200 OOP-limit, and 10% coinsurance rate between the two.

2.3.3 Moral Hazard

Prior literature establishes that moral hazard responses play an important role in explaining the sorting and spending patterns observed in empirical data (Einav et al., 2013 and Brot-Goldberg et al. 2017). This raises the question of whether the pattern is robust when there is moral hazard or heterogeneity in risk types.

To answer that, I modify my model to incorporate moral hazard responses following the framework introduced by Einav et al. (2013). The model assumes after the health risk is realized, individuals endogenously pick the spending level and get utility from over-spending. I then simulate the first-best plans and plans chosen by different types when both types have moral hazard responses. Appendix B specifies the details of the model.

Adding moral hazard responses to the original model with two risk types gives a similar sorting pattern. I assume that both types over-spend by 40% when moving from no insurance to full insurance. The over expenditure is proportional to the ex-ante risk. Under full information, the plans desired by both types are straight-deductible plans. With asymmetric information and perfect risk adjustment, the plan desired by the higher-risk type has a straight deductible design, and the plan desired by the lower-risk type has a non-straight-deductible design. Appendix B specifies the details of the results.

3. Empirical Analysis

In this section, I show evidence that such a mechanism plays a significant role in the current U.S. health insurance market. I use the ACA Exchange data and document large variation in plan designs, with a sizable demand for non-straight-deductible plans. I then document a sorting pattern into plan designs consistent with the prediction of the conceptual framework.

3.1 Institutional Background

The Affordable Care Act Exchange (the Exchange henceforth) was launched in 2014 and provided health insurance plans for non-elderly individuals whose employers do not offer plans or who are self-employed. Each state can either join the Federal Exchange or establish its state Exchange. I focus on the federally administered Individual Exchange operated via Healthcare.gov as it covers most states (around 40 states) and has the following regulations suitable to study the plan design variation.

The Exchange regulates the actuarial value of plans but leaves insurers with latitude to offer a range of different plan designs. The Exchange has regulation on the market-average actuarial value: 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.[18] The Exchange also requires plans to have an OOP-limit no larger than some amount ($7150 in 2017). Insurers are otherwise free to offer any cost-sharing attributes. Some state Exchanges (such as California) further regulate the plan designs, so I only focus on the Federal Exchange.

There are also ACA regulations limiting 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 to equalize plan costs within insurers. The premium of plans offered by the same insurer will be set based on the overall risk pool of that insurer, not the risk

[18] Consumers with income level below a certain level are qualified for cost-sharing variation plans, which have a higher AV than standard Silver plan.

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 user status, and (partially) by age groups. In the following analysis, I provide new evidence on the risk adjustment scheme in the ACA Exchange and discuss how the perfect risk adjustment assumption likely holds along the plan design dimension.

3.2 Data

I use Health Insurance Exchange Public Use Files from 2014 to 2017, a publicly available dataset of the universe of plans launched through Healthcare.gov. Appendix Table 1 shows the states in the sample. I define a unique plan based on the plan ID administered by CMS, which is a unique combination of state, insurer, network, and cost-sharing attributes, and also is the level of choice in the menu faced by consumers. [19] For each plan, I observe the financial attributes (deductibles, coinsurance rates, copays, OOP-limits, etc.), the premium (which varies at plan-rating area level) and the enrollment number in that plan (at plan-state level). I focus on the 2017 year for the main analysis, but the results are similar for other years.

For a subset of plans with a premium increase of more than 10%, I can link them to the Rate Review Data and observe the average claim costs (including both insurer payments to providers and the consumer cost-sharing) at the plan-state level. I discuss these data in more detail in Section 3.4 when I examine the extent to which there is differential sorting across risk types into different plan designs.

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 2 shows the market share of straight-deductible plans over time. Take the year 2016 as an example. There are around 4000 unique plans offered

[19] As noted before, each Silver plan has three cost-sharing variations. In almost all cases, the straight-deductible design is true across the standard plans and the cost-sharing variations. In such analyses, I count a Silver plans along with its variations as one plan, because the enrollment and claim costs are at plan level, not plan-variation level. In the analysis of the risk premium, I use the standard Silver plans.

in this market.[20] Among them, 13% are straight-deductible plans. In total there are 9.7 million consumers purchased a plan in this market, and about 8.3% of them select a straight-deductible plan.

Table 2. Market Share of Straight-Deductible Plans

year	% plans that are straight-deductible	Total number of plans	% of consumers choosing straight-deductible	Total Number of consumers (mm)
2014	10.5%	2,864	5.7%	5.5
2015	9.6%	4,573	7.5%	9.2
2016	13.0%	3,966	8.3%	9.7
2017	11.1%	3,106	4.7%	9.0

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. Almost all Silver plans have either all of these variations with a straight-deductible design (with different deductibles) or have none of them being straight-deductible plans.

There is also substantial variation in plan designs within a metal tier. Figure 1 shows the deductible and the OOP-limits distribution by metal tier.[21] For example, the OOP-limits of Gold plans vary between $2,000 and $7,000. Since the OOP-limits capture the worst-case risk, such variation represents a substantial difference in risk protection against the catastrophic risk.

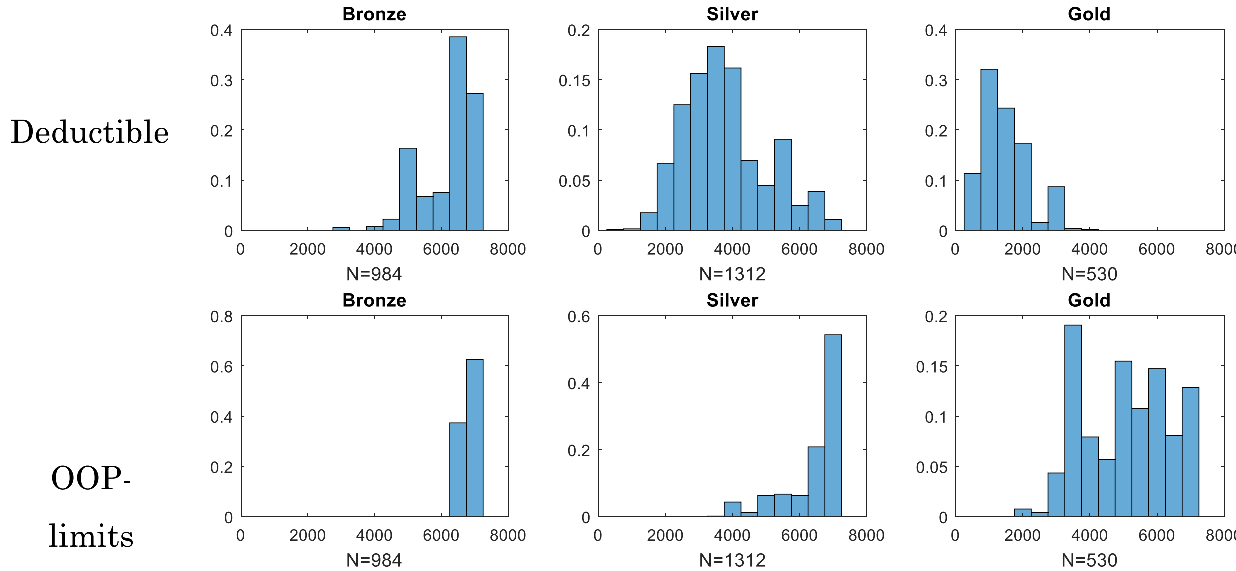
The histograms show the variation among all plans launched in the Federal Exchange. Any particular consumer in the market, however, only faces a subset of these plans. Consumers' choice set varies at the county level, so I also calculate the plan variation faced by a particular consumer, averaged across all counties. The

[20] A unique plan is defined as a unique combination of insurer, cost-sharing attributes, provider network, and geographic areas offered. Each consumer only face a fraction of these plans.

[21] The values are for a household with one individual, in-network, and first tier coverage. The deductible and the OOP-limits for families have a similar pattern. Most plans have almost all utilization coming from the in-network, first-tier providers. The Silver plans include only standard Silver plans and exclude cost-sharing variations available only to lower-income households.

variation is smaller but still sizable. For example, on average, the range in OOP-limits of Gold plans faced by a particular consumer is $2000.

Figure 1. Distribution of the Deductible and the OOP-Limits by Metal Tier



Note: Data from 2017 CMS Health Insurance Exchange Public Use Files. The sample includes all Exchange qualified health plans offered to individuals through the Health Insurance Exchange. For Silver tier, the sample excludes the cost-sharing reduction plans.

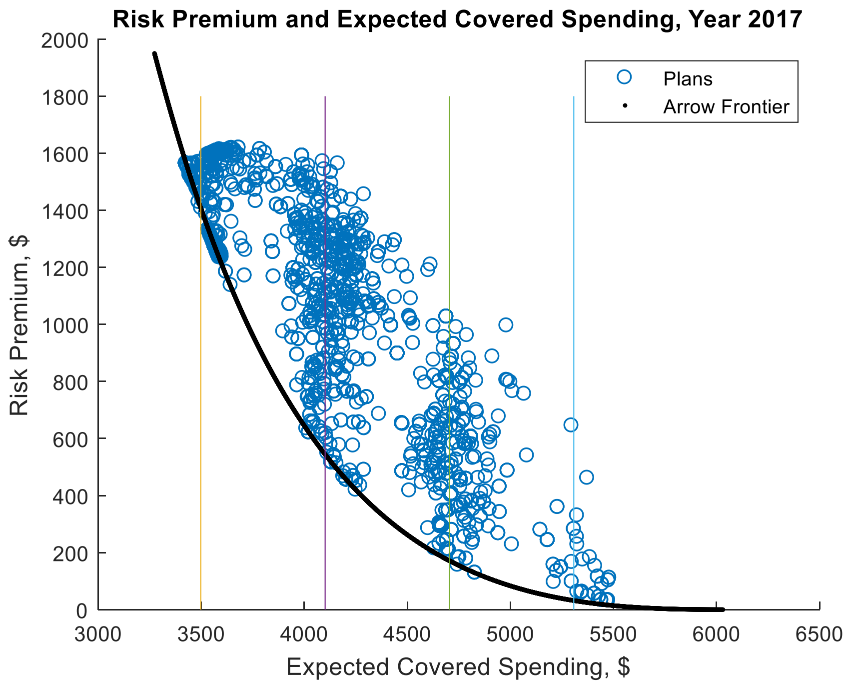
A natural question then is whether the variation in plan designs within a metal tier is economically important. To quantify this, I evaluate each plan available in the Federal Exchange for the individual facing the market-average risk. For this individual, plans in the same metal tier have similar AV and thus provide similar expected coverage. To capture the variation in plan design, I calculate the risk premium R, defined by the following formula:

$$EU(w-a) = U\,(w - E(a) - R),$$

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 for a plan is defined relative to a full-insurance benchmark. It represents an amount of money that a person would need to receive to be indifferent between enrolling in that plan versus a full-insurance plan when both plans are priced at their fair actuarial value. The risk premium is zero for a risk-neutral enrollee and is positive for risk-averse individuals, and increases with the level of uninsured risk in the plan.

To calculate the risk premium, I follow the literature in measuring the financial value of health insurance (for example, Handel 2013). I use the CARA utility model, so the wealth level is irrelevant. I set the risk-averse benchmark coefficient at 0.0004, which is the mean and median of the risk aversion estimated by Handel (2013) from health insurance plan choices after accounting for inertia. I also get the market-average risk in the ACA Exchange from the 2017 Actuarial Value Calculator.[22]

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



Note: Data from 2017 CMS Health Insurance Exchange Public Use Files. The sample includes all Exchange qualified health plans offered to individuals via Healthcare.gov. For Silver tier, the sample excludes the cost-sharing reduction plans. Risk premium and expected covered spending is calculated using the 2017 CMS AV Calculator Distribution. Each dot represents one unique plan. Arrow frontier shows the lowest possible risk premium conditional on expected covered spending level (achieved by straight-deductible plans) and is not actual plans. The vertical lines show the targeted actuarial value. Not all plans line up with the vertical lines perfectly, partially because the regulator allows for a 2% error margin, and partially because of measurement error in my calculation.

[22] Accessed from https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Final-2017-AVC-Methodology-012016.pdf

The variation in the risk protection provided by different plan options within a metal tier is sizable for individuals facing the market average risk. Figure 2 shows the risk premium and the expected covered spending for all plans in the four metal tiers for 2017. The expected covered spending is a scaled function of AV, and the target metal tier is represented by the vertical line. 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 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 OOP-limits closer to the maximum allowed by the regulation.

In Appendix C, I show that the variation is stable over time, and not correlated to aggregate demand differences and insurer characteristics. I examine whether the risk premium variation is different 1) among different network types, 2) in larger, more competitive markets, and 3) whether there is a difference in plan design offering by for-profit insurers versus non-profit insurers, or by insurer size (measured by total premium and enrollment). None of these factors appear to correlate with plan design variations. The variation in plan designs seems to be prevalent in many different markets. Most insurers also offer both straight-deductible plans and other designs.

3.4 Evidence of Sorting by Health into Different Plan Designs

The existence of the plan design variation creates room for selection. The theoretical analyses in Section 2 suggest that the straight-deductible designs are more attractive to the higher-risk types. In this section, I examine whether this is empirically true in the context of the ACA Federal Exchange.

3.4.1 Data

The data I use is the Uniform Rate Review data. The first part of the data contains premium and claim cost information at the plan level. The rate review regulation requires insurers operating on the Exchange to submit justification for any plan experiencing a premium increase of more than 10%. The justification includes detailed information on average premium and the average total medical expenditure.

The total medical expenditure is the ex-post expenditure incurred by enrollees in a plan, including the insurer's liability, consumers' cost-sharing, and any government payment towards medical treatment.[23] I can link about half of the plans in my sample with this information between the year 2014 and 2017 (the claim sample henceforth).

The second part of the data has premium and claim information at the insurer level. Unlike the plan-level information, all insurers are required to report the insurer-level data, giving a better representation of the overall sample. I can link 74% of the insurers with plan level information with the claim dataset (baseline insurer sample). For the rest of the insurers, I can match 98% of them in the Medical Loss Ratio filing, another insurer-level dataset reporting premium and insurer loss ratio information, but only for a limited number of variables of interest. Results using the Medical Loss Ratio data are similar and are presented in Appendix D.

3.4.2 Methodology

I measure the extent to which plans with a non-straight deductible design attract healthier enrollees than straight deductible designs in three ways. First, I examine the ex-post reported total medical expenditure between straight-deductible plans and the other designs within an insurer. There is a concern that the claim sample is biased because these plans are relatively "underpriced" and experience a higher premium increase. As a result, the sorting pattern may be confounded by other factors. To account for that, I leverage the fact that insurers are subject to the single risk pool requirement when setting the premium.[24] This means if one particular plan experiences unexpected medical expenditure increases, the insurers are required to spread out these costs among the premium of all plans offered, making all plans subject to reporting. In the claim sample, the majority (75%) of insurers report either all or none of their plans in this sample. As such, while the selection of insurers who have claim information at plan-level is a biased sub-sample, within each insurer,

[23] A fraction of low-income households are eligible for lower-cost-sharing plans, and the government pays a fraction of the out-of-pocket spending these households incurred.

[24] In my dataset, insurers are a unique combination of insurer-state. For simplicity, I use "insurer" henceforth to represent insurer-state.

there is a more representative set of the plans they offer. In my plan-level analysis, I include insurer-year fixed effects so that the differences in claim costs between different plan designs are identified off of within-insurer variation.

Second, I examine the average medical expenditure at the insurer level as a function of the enrollment share in straight-deductible plans. Since all insurers are required to report the insurer-level information, there is less concern that the sample is biased when using the insurer-level sample.

Third, I examine the level of ex-ante risk adjustment transfers given to insurers as a function of the share of their plans that are straight deductible. The risk transfers will disentangle the impacts of moral hazard from adverse selection. This is because risk adjustment payments are calculated based on the average risk score of an insurer's enrollees, which is a function of demographic information and pre-existing risk factors calculated from prior claim information.[25] The transfers thus reflect ex-ante risk other than moral hazard responses. This identification strategy is similar to Polyakova (2016).

3.4.3 Selection Pattern in the Federal Exchange

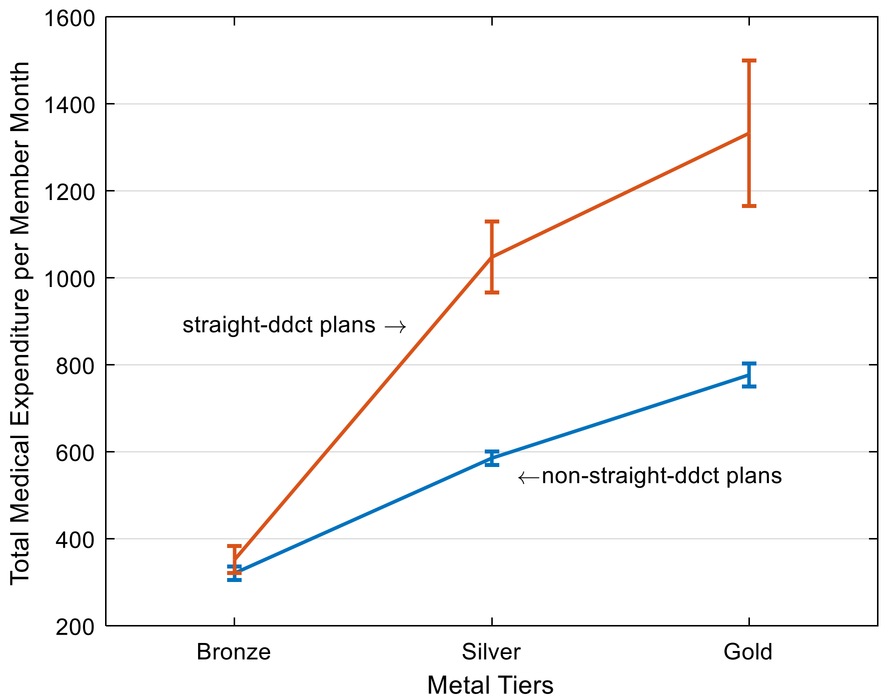
A comparison in unconditional means of the total medical expenditure illustrates that there is a strong correlation between average medical spending and plan designs, consistent with the theoretical predictions on sorting. Figure 3 shows the average monthly total medical expenditure (including insurer payment and consumer cost-sharing) for straight-deductible plans and the other designs across the four metal tiers for plans in the claim sample. Holding fixed the metal tier, the straight-deductible plans have higher medical expenditure than the other plans. The difference is more than $400 per month for Silver and Gold plans.[26]

[25] In the first two years of the ACA, risk transfers were calculated based on concurrent claim information, and as such did not reflect a true ex-ante risk measure. Limiting the analysis to later years where this was not an issue gives similar results.

[26] The difference in the Bronze is smaller as these plans are constrained by the OOP-limit regulation, and there is limited room for design difference. There are also very few Platinum plans in the market, so they are not shown in the graph.

The regression results controlling for potential confounding factors show a similar pattern. In the first model, I regress total medical expenditure per member month on whether the plan is a straight-deductible design ("straight-deductible"), the actuarial value calculated based on market-average risk ("AV"). I also control for metal tier, network type, insurer-year fixed effects. Occasionally, insurers may offer different plans in different counties. As such, I also control for the service areas where each plans are launched. The comparison captures the difference between straight-deductible plans and other designs offered by the same insurer and year and launched in the same set of counties. Standard errors are clustered at the insurer level.

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 the Federal Exchange in 2014-2017. Only plans with more than 10% premium changes are reported in the Uniform Rate Review data, accounting for about 50% of the universe of plans launched.

Table 3 column (1) shows the results. On average, straight-deductible plans attract individuals with significantly higher medical expenditure ($156 higher per month, and $1872 annually), relative to mean spending of $594 per month ($7128 annually).

I also run a regression at the insurer-year level. The independent variables represent the share of enrollees in the straight-deductible design ("straight-deductible"). I also control for the average AV, and the fraction of consumers in each metal tier and network type, along with state and year fixed effects.

Table 3. Average Monthly Total Medical Expenditure and Premium: Plan-Level

	(1)	(2)	(3)
	monthly total expenditure	monthly premium	
		collected	charged
straight-deductible	156.47***	-1.13	-0.46
	(23.46)	(3.14)	(1.62)
actuarial value	944.84***	226.95***	331.76***
	(288.00)	(81.45)	(69.89)
N	7369	7369	73102
R^2	0.30	0.28	0.57
y-mean	586.4	404.4	302.5
y-sd	593.6	233.5	105.0
Controls	metal tier, network type		
Fixed Effects	insurer-serviceArea-year		insurer-year, rating area

Note: Straight-ddct is a dummy variable indicating whether the plan is a straight-deductible plan. The actuarial value of a plan is the fraction of losses covered for the average population, which varies no more than 4% within a metal tier. y-mean is the mean of the dependent variable and y-sd is the standard deviation of the dependent variable. Column (1) and (2) include plans between 2014 and 2017 with a premium increase for more than 10%. I dropped those reporting non-positive total expenditure or premium and plans with the top and bottom 1% 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 of enrollees in a plan, including consumer cost-sharing and insurer payments. The dependent variable in (2) is the average premium per member month. Column (3) includes all standard plans between 2014 and 2017. The dependent variable is the per-month premium of the single coverage for a 21-old non-tobacco user. Since premium varies by rating area, each observation is a plan-rating area-year. All standard errors are clustered at the insurer level and shown in parenthesis. *p<10%, **p<5%, ***p<1%.

Table 4 column (1) shows the regression results. Similarly, there is a significantly higher total medical expenditure for insurers having a higher proportion of enrollees in the straight-deductible design.

The total expenditure difference mainly reflects selection rather than moral hazard responses. Table 4 column (3) shows the same regression with risk adjustment transfers per member month as the dependent variable. On average, the insurers with a higher share of straight-deductible plans receive significantly higher risk transfers ($143 per member month more transfers if moving from no straight-deductible enrollment to full straight-deductible enrollment.)

Table 4. Average Monthly Total Medical Expenditure and Premium: Insurer-Level

	(1) total expenditure	(2) insurer liability	(3) risk transfers	(4) average premium
share straight-ddct	275.25***	235.81***	142.93**	61.68
	(99.95)	(79.89)	(63.26)	(63.47)
average AV	1064.24*	1135.15**	571.61**	244.30
	(550.37)	(461.59)	(230.65)	(333.1)
N	617	617	617	617
R2	0.37	0.35	0.14	0.63
y-mean	474.7	357.1	-6.2	381.1
y-sd	124.1	102.5	66.0	97.4
Controls	metal tier, network type, state, year fixed effects			

Note: Each observation is an insurer-year. Straight-ddct is the share of enrollees choosing a straight-deductible design, and AV is the average AV weighted by enrollment share of each plan. Metal tier and network type is the fraction of enrollment in each metal tier/network type within an insurer. 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 transfers an insurer received. The dependent variable of (4) is the average premium. All dependent variables are per member month. The regressions are weighted by the enrollment at each insurer-year. *p<10%, **p<5%, ***p<1%.

In summary, there is a significant difference in the total medical expenditure of different designs with similar coverage levels. A large part of this difference, if not all of them, is driven by the selection force from different risk types.

3.4.4 Risk Adjustment Scheme in the ACA

The goal of the risk adjustment program in the ACA is, "to compensate health insurance plans for differences in enrollee health mix so that plan premiums reflect differences in scope of coverage and other plan factors, but not differences in health status" (Kautter et al. 2014). My empirical analyses offer some new insights on how well the risk adjustment scheme works in the Federal Exchange.

The different designs in the same metal tier offered by the *same* insurer have very similar premiums. I consider two premiums: the per month premium at plan-rating area level, set by insurers for single coverage of 21-year-old non-tobacco users, and the average premium per member-month collected by insurers, which takes into the different age, tobacco-using, geographic markets, and household composition of a plan. The first measure is constructed using the dataset with all plans. Since the premium varies at rating area level (typically a group of neighboring counties within a state), each observation is a plan-rating area. The second measure is only available for plans in the claim sample.

First, based on the single risk pool requirement, plans with different designs should have similar premiums if they have the same AV. Table 3 columns (2) and (3) show that in both regressions, the coefficient for the straight-deductible dummy is almost 0. This confirms that the single risk pool requirement is well enforced at the insurer level.

Second, the risk adjustment across different insurers flattens the premium among insurers with different proportions of straight-deductible plan enrollment. Table 4 column (4) shows that the coefficient of straight-deductible market share is around \$62, much smaller than the coefficient of the insurer liability regression (column (2)), and is not statistically significant. This means insurers with a higher fraction of straight-deductible plans have similar average premiums to insurers with a lower fraction of straight-deductible plans.

Though risk adjustment flattens the premiums across different designs for the most part, they may be imperfect and undercompensate the straight-deductible designs. In calculating the insurer's likely payment toward a plan, the current risk adjustment formula assumes a fixed plan design (non-straight-deductible) for each metal tier. Given that the higher-risk types have more of their expense covered under straight-deductible plans than non-straight-deductible plans in the same metal tier, the current formula underestimates the actual covered expenses of straight-deductible plans. This may partially explain the fact that the coefficient of risk transfers (Table 4 column (3)) is smaller than the coefficient of insurer liability (Table 4 column (2)).[27]

My analyses suggest that when determining the risk transfers across insurers, regulators should account for the fact that straight-deductible plans cover more expenses for the higher-risk types than other designs holding fixed a metal tier. A better formula to achieve the stated goal is to have the risk transfers depend on both the metal tier and the plan design consumers sort into.

4. Implications for Plan Design Regulation

The existence of sorting by plan design has important implications for regulating the health insurance markets. There is a wide public policy debate on whether and how to regulate the plan designs. A few markets in the United States (including the California State Health Insurance Exchange) and other OECD countries (including the Netherlands and Switzerland) have already adopted plan standardization regulations, mainly motivated by consumer confusion. However, there is no consensus on what the standardized plan design should look like: In the Netherlands, only straight-deductible plans are offered. In California, the standardized plans have a non-straight-deductible design.

In this section, I explore the issue using the conceptual framework developed in Section 2. The counterfactual regulation is that insurers can only offer straight-deductible plans. I show that under asymmetric information, removing plan design variation may harm market efficiency and make the market more likely to unravel. I illustrate that the impacts of such policies hinge on the existence of risk adjustment and consumer confusion. I also estimate the likely benefits of such regulation in the ACA Federal Exchange.

[27] These differences may also reflect moral hazard.

Before delving into details, I introduce two efficiency measures used throughout the section. The first is the overall efficiency, defined as the net value of providing the plan to a certain individual minus the true costs to the society providing that coverage, weighted across the population:

$$W = \sum a_i \left(v_{ij} - \theta l_{ij}\right), \tag{4}$$

where a_i is the relative weight of each type and v_{ij} is the value for individual i choosing plan j calculated based on out-of-pocket spendings (excluding premium). Under CARA utility, which I assume throughout the simulation and estimation, v_{ij} is the certainty equivalent of $\int u(-OOP_j)dF_i$ relative to full insurance. It is separable from the costs of providing insurance or premium because of CARA utility. θl_{ij} is the expected covered spending multiplied by a loading factor (θ =1.2). Note that the true costs to the society of providing the plan, θl_{ij}, is the same as the premium only under an unregulated competitive equilibrium, but not the same under perfect risk adjustment.

The second is the type i's consumer surplus for choosing a specific plan j:

$$S_{ij} = v_{ij} - p_j - tax, \tag{5}$$

where p_j is the premium of the plan, tax is a uniform tax imposed on the whole population if there is any difference in the overall premium collected and the claim costs plus loading incurred. Unregulated competitive markets have a neutral budget, so $tax = 0$. Under perfect risk adjustment, the budget may not be neutral. I assume the market finances the difference through a tax or subsidy equally imposed on everyone in the market.

For both measures, I report the difference from the plans they choose under full information for ease of interpretation. A positive number in the surplus means consumers are better off than in the full information scenario. The efficiency measures are always non-positive (because the full information is the first-best).

4.1 Design Regulation in A Competitive Equilibrium

The existence of design variation creates a mechanism for screening that allows lower-risk types to get greater coverage while avoiding pooling with higher-risks.

Table 5 shows the plans chosen by the lower-risk type constructed in Section 2.2.2 under different design regulations. In a separating equilibrium, the lower-risk types sort into a constant coinsurance plan where 77% of any loss is covered by insurance.

Removing the non-straight-deductible plans forces the lower-risk type to choose a plan with a very high deductible ($13,154) to avoid pooling with the higher-risk type. Under this plan, only 23% of the expenses are covered for the lower-risk type. The consumer surplus is $1,256 lower than choosing the full information optimal plan and almost $700 lower relative to no regulation. This amount is sizable given that the total expected medical expenditure of the lower-risk type is around $1,700. The market-average efficiency is reduced by about $180 per person annually if allocated between both types.

Table 5. Plans Chosen Under Design Regulation: No Risk Adjustment

	deductible	out-of-pocket limit	co1	co2	% losses covered	surplus
Risk-based pricing (1)	933	933	1	0	77%	0
No regulation (2)	0	∞	0.23	0.23	77%	-$561
Design regulation (3)	13,154	13,154	1	0	23%	-$1,256

Note: The table shows the plans chosen by the lower-risk type under three scenarios in a simulation. co1 is the coinsurance rate before the deductible paid by consumers. co2 is the coinsurance rate after the deductible paid by consumers. The consumer surplus (in this case, also the social surplus) is calculated relative to surplus achieved by the plan chosen under no asymmetric information of that type.

This example illustrates the value of design variation in a market with heterogeneity in risk types and no price regulation: the existence of different designs helps sustain an equilibrium where there is much less coverage level distortion. Removing design variation will make the lower-risk types sort into plans with very little coverage. When the two risk types are different enough, the lower-risk types may exit the market completely under the straight-deductible only environment, but not without the design regulation.

4.2 Design Regulation under Risk Adjustment

In contrast to the prior case, regulating plan designs in a market with risk adjustment has ambiguous impacts on the overall efficiency. The key difference from the no-risk-adjustment case is that the premium does not reflect the costs of the type who actually enroll in the plan, but the risk of the overall market. The average cost pricing means that lower-risk types pay a marginal price above their marginal cost if they choose plans with higher population-average actuarial values. However, if they consider two plans with the same population-average actuarial value that have different plan designs, they may be able to get additional coverage under higher-risk plan designs with no increase in price.

To illustrate the mechanism, I use the same two risk types and simulate the plans they choose under full information, no regulation, and design regulation. The results are in Table 6. There are two impacts of restricting plans to straight-deductible plans. First, the straight-deductible plan has lower coverage for the lower-risk types (26% versus 28%). From this perspective, moving to the straight-deductible plan reduces efficiency because the lower-risk types move away from their full information plan.

Table 6. Plans Chosen Under Design Regulation: Perfect Risk Adjustment

		Risk type	deductible	OOP-limit	co1	co2	% losses covered	Consumer Surplus	Average Efficiency
Risk-based pricing	(1)	H	1,820	1,820	1	0	83%	0	0
	(2)	L	933	933	1	0	77%	0	
No Design Regulation	(3)	H	677	677	1	0	92%	$1107	-$304
	(4)	L	7,300	20,500	1	0.2	28%	-$4230	
Design Regulation	(5)	H	677	677	1	0	92%	$1110	-$318
	(6)	L	10,971	10,971	1	0	26%	-$4292	

Note: co1 is the coinsurance rate prior to the deductible paid by consumers. co2 is the coinsurance rate after the deductible paid by consumers. Consumer surplus is calculated as the net transfer to each risk type to make them indifferent from choosing that plan or the full information plan. The social surplus is calculated as the welfare of sorting into that plan from the social planner's perspective.

Second, the costs of choosing the non-straight-deductible plan relative to the straight-deductible plans are not fully borne by the lower-risk types. One way to see this is to compare the surplus for the higher-risk types. With or without regulation, the higher-risk types sort into the same straight-deductible plan with the same premium. However, their surplus increases by $3 under the straight-deductible regulation. This is because there is a smaller budget difference under the design regulation, which results in less tax imposed overall. This difference in tax means the lower-risk types do not fully bear the costs of making the choice and pass some of the costs to the higher-risk types. This is equivalent to taking on extra coverage without paying the fair price. Removing such options removes this externality.

The two forces move in opposite directions, and the overall efficiency depends on the relative size of the two. In this numeric example, the average efficiency is reduced by $14 under the regulation.

The regulation can also have an impact on the extensive margin. In this specific example, the two risk types choose some insurance over no insurance, so there is no interaction with the extensive margin. When there is more variation in the risk types in the market, some of the lowest risk types may choose to drop out of the market entirely. As they exit, the market premiums rise and there may be additional exits. In these cases, the existence of plan design variation helps limit market unraveling: when there are only straight-deductible plans, the lower-risk types cannot find the plans providing enough coverage at a reasonable price, and they are more likely to exit the market than if there are other designs. Given that the premium is based on the risk of those who choose a plan, this will increase the overall premium of the market, and create a feedback loop where the fraction insured decrease under the design regulation. In Appendix E, I provide a numeric example using 100 risk types created from the Truven MarketScan data. In the example, an additional 5% of the population will drop out of the market when there is a design regulation removing non-straight-deductible plans.[28]

4.3 Consumer Confusion

[28] The simulation caps the deductible at $100,000. The least insurance one could get other than no insurance is a straight-deductible plan with a deductible of $100,000. Theoretically it is possible that consumers may prefer a plan with even higher deductible. The point is to think of this “no insurance” as carrying very little coverage, if not truly no insurance.

Prior literature emphasizes the relevance of consumer confusion in this market (Abaluck and Gruber 2011, Abaluck and Gruber 2019; Bhargava et al. 2017). Under confusion, individuals may sort into the plan with a low value for them. Since in the full information case, all types want a straight-deductible plan, limiting plans to only these plans may help mitigate the consequence of sorting into the wrong plan.

The consequence of choosing the wrong plan is especially large for the higher-risk type. Take the two plans chosen by the two risk types in the perfect risk adjustment case as an example. The higher-risk types lose about $3,620 in their surplus from choosing the plan intended for the lower-risk type. For the lower-risk type, choosing the plan of the higher-risk type is also very bad for their surplus because that plan is too expensive for them. However, from the market efficiency perspective, the lower-risk type gets more coverage under the plan chosen by the high-risk type and thus incurs positive efficiency to the society (about $1,155). This suggests that restricting plans to straight-deductible plans might have sizable impacts on market efficiency with confusion and may have differential impacts on different risk types.

4.4 Estimating the Effects for the ACA Federal Exchange

In this section, I estimate the likely impacts of limiting plans to straight-deductible designs in the 2017 ACA Federal Exchange. Specifically, I compare the market outcome under two sets of menus: the actual 2017 plans offered in the ACA Exchange and a hypothetical choice set replacing all current options with a straight-deductible plan of the same premium. This new choice set has the same number of options and the same premiums as the existing one. The only difference is that all plans have a straight-deductible design.

To represent the likely risk distributions in the ACA Exchange, I create 100 risk types from the Truven MarketScan data using the algorithm specified in Section 2.2.2. I shift the means of these distributions so that overall, these 100 risk types have the same average medical expenditure as in the 2017 ACA Federal Exchange. I also assume that consumers have a constant absolute risk aversion utility function with a risk aversion coefficient of 0.0004.

To compare the market consequence under the actual choice set and this hypothetical one, I make the following assumptions. Since I do not have information about the risk distributions at each county, I assume that all counties have the same risk distributions. In each county, there are a fraction of consumers who enjoy a premium subsidy or are eligible for cost-sharing reduction plans. These plans have much more generous design but have the same premium as the associated standard Silver plans. I collect information on the fraction of each subsidized type and the amount of premium subsidy from the 2017 Open Enrollment Period County-Level Public Use File (OEP data). Finally, I assume that the individual mandate and premium subsidy lead everyone to choose an insurance plan, so no insurance is not in their choice set.

I assume that when selecting plans, individuals base their choice on a perceived utility from choosing plan j given by:

$$\int u(-OOP_j - p_j)dF_i + \beta\epsilon_{ij}.$$

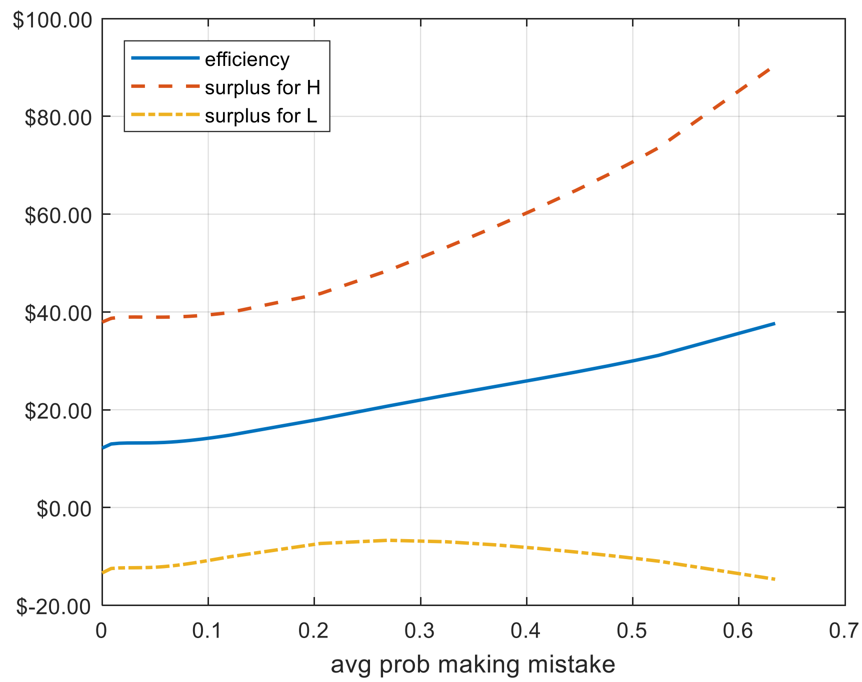
The deterministic part, $\int u(-OOP_j - p_j)dF_i$, is a function of the out-of-pocket spending and premium. For each plan, the premium p_j is calculated assuming perfect risk adjustment, such that it reflects the costs of all risk types enrolled in the plan, multiplied by a loading factor of 1.2, and minus any premium subsidy. I assume that all risk types are equally likely to receive a subsidy. The fraction of people with subsidy varies by county and is collected from the OEP dataset.

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. The error term allows me to incorporate the potential for consumer confusion into the simulations. Consumers in the ACA Exchanges often face a large choice set, typically around 20 options in each county, making confusion a likely concern. In almost all counties, there are at least two designs within a metal tier. This variation may allow different risk types to sort into plans that suit their needs, but at the same time, it may create room for consumer confusion. The larger the scaling parameter β, in the utility function, the more randomness there will be in plan choice, which I use to model a greater level of consumer confusion.

I calculate the impacts of limiting plans to straight-deductible plans on the overall efficiency, defined in (4), the consumer surplus for the higher-risk types (defined as the population with expected medical expenditure above the median), and lower-risk types (below median), defined in (5).

Figure 4 shows the difference between each value under the straight-deductible choice set and the current choice set. 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 welfare by about \$10 per person per year. This is slightly larger than the numeric example in the 2-risk type case. One possible reason is that in some counties, there are no straight-deductible plans, and having the straight-deductible plans offered will increase the overall efficiency.

Figure 4. Efficiency Effects of Regulating Plan Designs in ACA



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

When consumers in the market are more likely to make a mistake in choosing plans, both the overall efficiency and the surplus for the 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. Such improvement is not Pareto: the lower-risk types are worse off under such regulation. At the 50% confusion level, they are worse off by about $10 per year.

The simulation illustrates that the level of confusion matters for the welfare implication of simplifying plan designs. DeLeire et al. (2017) show that a very low percentage of enrollees who are eligible for cost-sharing reduction plans select dominated options. Their findings suggest that consumers can effectively sort into the optimal metal tiers in the ACA exchanges. However, there is also a broader literature documenting confusion in health insurance choices in other settings. It is an open question of how optimally consumers can choose between plan designs with different cost-sharing designs in the ACA exchanges. More research is needed on this front.

5 Conclusion

In this paper, I identify an underappreciated dimension of sorting in insurance markets: sorting by plan design. I prove that in a market with asymmetric information, the lower-risk type will sort into designs with less coverage for larger losses in exchange for more coverage for smaller losses, while the higher-risk types sort into straight-deductible plans. The sorting pattern exists in both a separating equilibrium without price regulation and a market with perfect risk adjustment.

This sorting pattern is empirically relevant in the United States health insurance markets. I document that in the ACA Federal Exchange, there is large variation in the cost-sharing features of plans launched and chosen within a similar coverage level. Such variation in the cost-sharing features translates to an economically significant difference in risk protection for consumers facing the market-average risk. I then show that within a coverage level, the straight-deductible plans have a similar premium as the other designs, but have significantly larger claim costs than the other plan designs.

More broadly, this paper argues for the importance of studying variation in plan designs. My framework illustrates that coverage design variation is an important dimension of sorting under asymmetric information. This work also highlights an underappreciated perspective in evaluating design regulations. Prior literature recognizes that simplifying insurance contract characteristics can make it easier for consumers to compare across plans and can promote competition and efficiency. I illustrate that in a market with asymmetric information, plan design variation can also serve as a tool to "screen" different risk types and support an equilibrium where the lower-risk types distort less in their coverage. As a result, removing plan design variation may make the market more likely to unravel and harm the market efficiency. The overall benefits of standardizing plan design thus depend on the relative importance of these concerns. This paper illustrates the trade-off in a market with risk adjustment and perfect competition. More research is needed to understand the implications of design variation when there is market power.

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. “Less Is More: Improving Choices by Limiting Choices in Health Insurance Markets.” Working paper (2019).

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).

Aizawa, Naoki, and You Suk Kim. “Advertising and Risk Selection in Health Insurance Markets.” *American Economic Review* 108, no. 3(March 2018): 828–67.

Akerlof, George. “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism.” *The Quarterly Journal of Economics* 84, no. 3 (Aug 1970) 175–188.

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.

Brot-Goldberg, Zarek C., Amitabh Chandra, Benjamin R. Handel, and Jonathan T. Kolstad. “What Does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics.” *The Quarterly Journal of Economics* 132, no. 3 (August 1, 2017): 1261-1318.

Brown, Jason, Mark Duggan, Ilyana Kuziemko, and William Woolston. “How Does Risk Selection Respond to Risk Adjustment? New Evidence from the Medicare Advantage Program.” *American Economic Review* 104, no. 10 (October 2014): 3335–64.

Crocker, Keith J., and Arthur Snow. “A Simple Tax Structure for Competitive Equilibrium and Redistribution in Insurance Markets with Asymmetric Information.” *Southern Economic Journal* 51, no. 4 (April 1985): 1142.

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, and Mark R. Cullen. “Estimating Welfare in Insurance Markets Using Variation in Prices.” *The Quarterly Journal of Economics* 125, no. 3 (August 1, 2010): 877–921.

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.

Einav, Liran, Amy Finkelstein, and Pietro Tebaldi. “Market Design in Regulated Health Insurance Markets: Risk Adjustment vs. Subsidies.”Working paper 2018.

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 M. and Justin Sydnor “Liquidity Constraints and Insurance Demand” Working Paper 2018

Frank, Richard G, Jacob Glazer, and Thomas G McGuire. “Measuring Adverse Selection in Managed Health Care.” *Journal of Health Economics* 19, no. 6 (November 1, 2000): 829–54.

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 Layton, Grace McCormack, and Mark Shepard. “The Two Margin Problem in Insurance Markets.” Working paper (2019).

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.

Handel, Ben, Igal Hendel, and Michael D. Whinston. “Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk.” *Econometrica* 83, no.4 (2015): 1261–1313.

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.

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, 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. “Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange.” Working Paper (2016).

Spence, Michael. “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 (February 24, 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.

Appendix A. Proofs in Section 2

Proposition 1. Proof:

The optimization problem for the consumer is:

$$v = \max_{\boldsymbol{l}} \sum_s u(w - x_s + l_s - p(\boldsymbol{l})) f_s.$$

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:

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

with equality if $0 < l_s$. 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. [QED]

Proposition 2. Proof:

The optimization problem for the lower-risk type is:

$$v = \max_{\boldsymbol{l}} \sum_s u_L(w - x_s + l_s - p(\boldsymbol{l})) f_s^L.$$

subject to:

$$0 \leq l_s \leq x_s,$$
$$p(\boldsymbol{l}) = \theta \sum_s f_s^L l_s + c,$$
$$\sum_s u_H(w - x_s + l_s - p) f_s^H = A.$$

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

The Lagrange of the above optimization problem is:

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

Let u_{Ls}' denote the derivative of lower-risk types 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. [QED]

Proposition 3 Proof:

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 the higher-risk type is:

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

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

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

with equality if $l_s > 0$.

For any two loss states $x_t > x_z$, $\frac{f_t^L}{f_t^H} < \frac{f_z^L}{f_z^H}$ (because $\frac{f_s^L}{f_s^H}$ is decreasing with regard to x_s). This means whenever $l_t > 0$ and $l_z > 0, u'_{tH} < u'_{zH}$ and $u'_{tL} > u'_{zL}$. Since $u''_H < 0$ and $u''_L < 0$, $l^*_{Ht} - x_t \geq l^*_{Lz} - x_z$ 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. (They would want to have larger consumption in state t than in z, but are not able to because of the non-increasing consumption constraint.) 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. [QED]

Procedure calculating the separating equilibrium

To simulate the equilibrium plans chosen by the two risk types when there is no risk adjustment, I use the equilibrium concepts developed by Azevedo and Gottlieb

(2017). The equilibrium is defined as a set of plans chosen by each risk type, and the premium schedule for both the traded and non-traded plans satisfying the following proposition: any non-traded plan has a premium such that some consumers are indifferent from buying or not, and that premium is no larger than the costs of that consumer.

To calculate the plans chosen by each type, I use the following procedure:

1. Calculate the optimal plan for the higher-risk type when premium reflects their own risk. Label this plan as H^*.
2. Find the subset of plans which makes the higher-risk type no better off if priced based on the risk of the lower-risk type. Label these plans as in set C. Label the rest plans as in set D.
3. Among plans in C, find the plan to maximize the utility of the lower-risk type. Label this plan as L^*.
4. Assume (H^*, L^*) are the separating equilibrium plan. Each is priced based on the risk of the type who sorts into that plan.
5. The next step is to verify there exists a pricing scheme for all the other non-traded plans satisfying the above proposition. This is equivalent to assigning a premium for each plan assuming that they make at least one risk type indifferent to deviate and make the other type no better off.
    a. To do that, I first assign plans in set D a premium to make the higher-risk type indifferent from buying or not: p'_H.
    b. I then assign plans in set C a premium to make the lower-risk type indifferent from buying or not (p'_L). Among them, some will make the higher-risk type want to deviate. Label them as in set E. For these plans, reassign them a premium to make the higher-risk type indifferent from buying or not (p'_H).
    c. Verify that (H^*, L^*) and ($p'_{jH} \mid j \in D | j \in E$, $p'_{jL} | j \in C$) is an equilibrium by checking whether the incentive compatibility constraints hold. In the numeric example with two risk types in Section 2.3, it holds.

Appendix B Moral Hazard Responses

Model

Following Einav et al. (2013), I assume that individuals have moral hazard responses in medical spending. Individuals now will decide how many medical services to use after observing the health risk draw. Let λ denote the negative health shock, measured as the medical expenditure when individuals face no insurance. Their utility under plan j when they spend m in health care given the negative health draw λ, moral hazard level ω is:

$$\mathrm{u}(m; \lambda, \omega, \mathrm{j}) = \left[(m-\lambda) - \frac{1}{2\omega(\lambda)}(m-\lambda)^2\right] + \left[\mathrm{y} - \mathrm{p_j} - \mathrm{c_j}(m)\right],$$

where the item in the first bracket represents the utility from extra health (referred to as the health component henceforth) and the second bracket indicates the monetary utility (referred to as cost component henceforth). The moral hazard level, ω, is the extra medical expenditure one will consume if one goes from no insurance to full insurance. The larger the ω, the higher the moral hazard response. In the extreme case where ω is zero, the second term in the health component goes to negative infinity, so individuals will set $m = \lambda$, under which case the model reduces to the previous one. The cost component is consists of income (y), the premium of plan j (p_j), and the out-of-pocket costs ($c_j(m)$ represents the cost-sharing rule of plan j. It specifies when the medical expenditure is m, how much the out-of-cost will be). In this model, the value of a plan comes from both the health and cost component.

I then apply CARA utility function over u to get the final utility of choosing a plan:

$$v_{ij} = -E(u) - \frac{1}{2} rVar(u).$$

Whether and how moral hazard changes the incentive to insurers depends on whether ω is known to regulators when setting the risk adjustment formula. If the regulator can perfectly predict the moral hazard response and use it in the risk adjustment formula, then this goes back to the perfect risk adjustment case. The premium will reflect the average market expenditure. For the following simulation, I assume perfect risk adjustment.

This model provides a straightforward way to take account for moral hazard responses: the lower the cost-sharing attributes, the more the medical spending. There is a growing literature arguing that such extra spending may not be beneficial, and there is also cases of under-spending (Baicker et al. 2015; Starc and Town 2015). The model does not account for these factors.

Simulation Results

To get a reasonable size of ω, I follow Brot-Goldberg et al. (2017) and assume that individuals will overspend about 40% moving from no insurance to full insurance. Appendix Table B1 shows the plans each risk type sorts into when they have moral hazard responses. As in the case where there is no moral hazard, the lower-risk types sort into a non-straight deductible design while the higher-risk type sorts into a straight-deductible plan.

Appendix Table B1 Plans Chosen with Moral Hazard

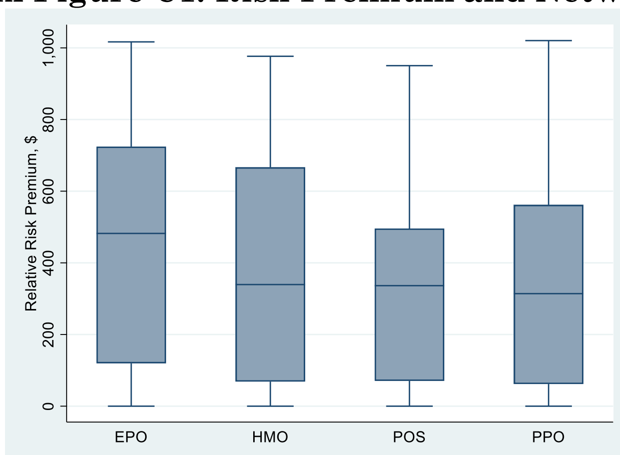
scheme	Risk Type	ddct	Out-of-pocket limit	co1	co2
Full information	H	5,176	5,176	1	0
	L	2,856	2,856	1	0
Perfect risk adjustment	H	3,081	3,081	1	0
	L	4,900	10,000	1	0.1

Appendix C. Variation in Plan Designs by Regions and Insurers

The plan design variation may reflect variation in aggregate demand or supply factors. To examine the issue, I document the correlation between design variation and the following factors.

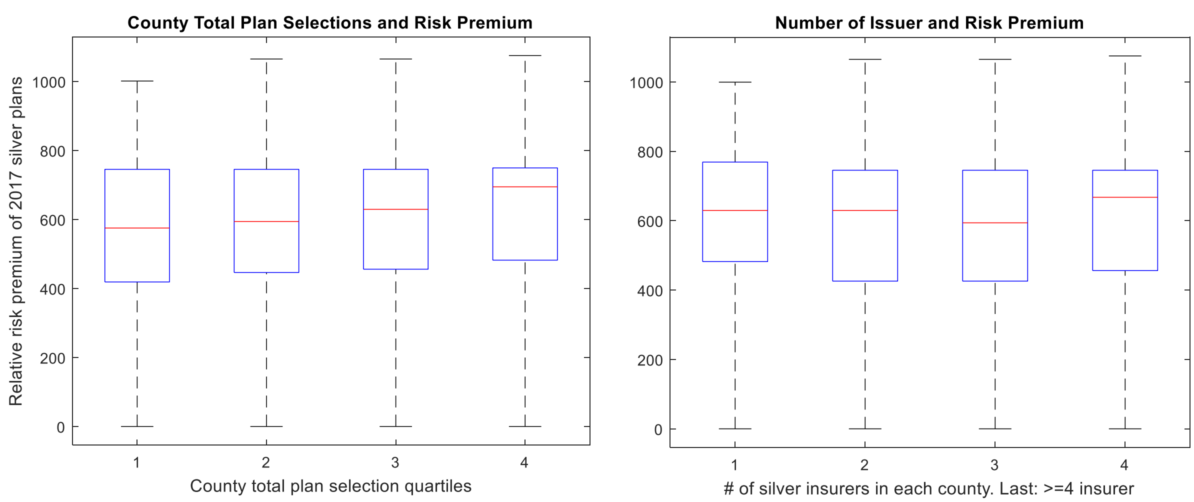
First, the variation in plan designs is prevalent for different network types. To measure plan design variation, I calculate the relative risk premium of a plan as the risk premium minus the lowest risk premium of the plan with the same market-average AV. Appendix Figure C1 shows the distribution of relative risk premium of 2017 plans by four network types: HMO, EPO, POS, and PPO.

Appendix Figure C1. Risk Premium and Network Type



Appendix Figure C2 Risk premium and county characteristics

Panel A: County Total Plan Selections Panel B: Number of Insurers

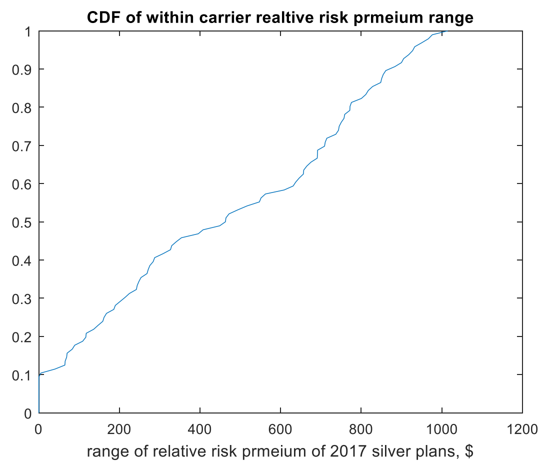


Second, the variation in risk premium is stable across markets with different size and competitiveness. For all plans in the 2017 Silver tier, I construct a new dataset where each observation represents one county-plan. I then classify plans based on their counties' market size, measured as total plan selections, and competition intensity, measured by the number of insurers. I plot the relative risk premium by quartiles of county market size, and by the number of insurers. Appendix Figure C2 shows that the distribution of risk premium is very similar for plans launched in counties with different market size, and different numbers of insurers.

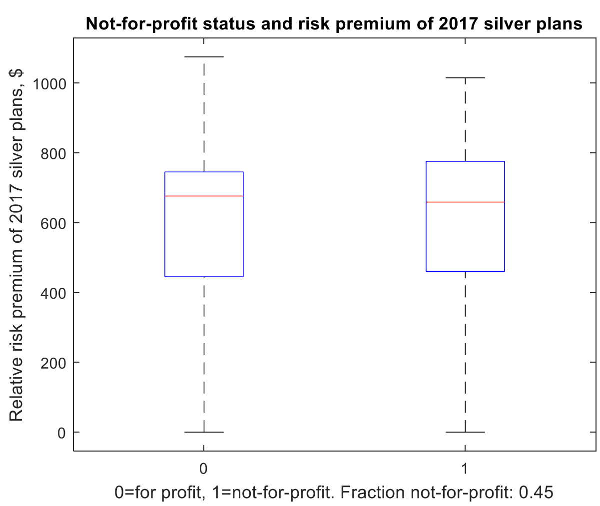
Finally, I examine whether the plan design variation is prevalent among different insurers. I first calculate the variation in risk premium within Silver plans offered by the same insurer. Appendix Figure C3 Panel A shows that all insurers have some variation in plan designs, and more than half insurers have the range in risk premium larger than $400. I then examine whether the variation in plan designs are correlated with the for-profit status of an insurer, or the insurer size measured by the total member-months in the individual market. Panel B shows that the distribution of risk premium of 2017 Silver plans is very similar among different insurers.

Appendix Figure C4. Risk Premium Variation among Insurers

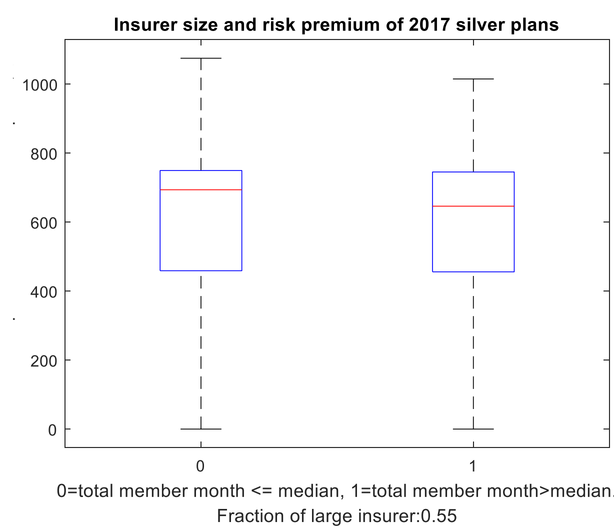
Panel A: Distribution of within insurer relative risk premium range of 2017 Silver plans



Panel B: Not-for-profit status



Panel C: insurer size



Notes: Not-for-profit status classifies insurers based on whether they are for-profit (0) or not-for-profit (1). Insurer size is based on the total member month underwritten in the 2015 individual market, where 1 represents insurers with the member months larger than the median. Data from the 2015 Medical Loss Ratio filing.

Appendix D. Empirical Analysis with Medical Loss Ratio Data

In the main text, I show that the straight-deductible plans have higher total medical expenditure and also receive more risk adjustment payments than other designs, both at plan level or aggregated at insurer level. In this section, I present the robustness check of these results using 1) a larger sample from Medical Loss Ratio filings 2) use risk premium for the market average population as a measure of plan design.

Appendix Table D1 shows the number of observations in different samples. Among all the insurer-years with plan information, 73.6% are in the baseline sample presented in Section 4. About 62.3% of them are in the Medical Loss Ratio fillings (MLR), which have information on average premium, insurer claim costs (the amount insurers are liable to pay per member month, not including consumer cost-sharing nor government payments). Some of them also have risk adjustment transfers. But there is no total medical expenditure information. I use this information to impute the missing information from the baseline sample as much as possible and resulting in two other datasets: the 815-sample accounting for 97.3% insurers with plan information. This sample has average premium and insurer claim costs information; the 744-sample accounts for 88.8% observations with risk adjustment transfer information.

Appendix Table D1 Matching across Different Datasets

Datasets	# of insurer-year: 2014-2017	% matched
Insurer-year with plan information	838	100%
Uniform Rate Review (benchmark insurer sample)	617	73.6%
Medical Loss Ratio filings	522	62.3%
Combined – premium, insurer claim costs	815	97.3%
Combined – risk adjustment	744	88.8%
Combined – total medical expenditure	617	73.6%

Appendix Table D2 shows that incorporating information from the MLR dataset is similar to the results in Section 4: the risk transfers and insurer claim costs are much higher for insurers with a higher proportion of enrollees in the straight-deductible plans, but the premiums are similar.

In Section 3.3, I show that within a metal tier, there is a range of plan design differences. Some plans, even though not straight-deductible, may have a design very close to a straight-deductible (for example, have a deductible very close to the OOP-limit). Such plans are also likely to attract higher expenditure individuals. To account for this, I use the relative risk premium calculated in Section 3.3 as a measure of plan design. This number reflects the risk premium of a design for an individual with a market average risk and is rescaled to 0 for straight-deductible designs. The higher the risk premium, the less risk protection it provides to the average risk, and the more attractive it should be for the lower-risk types.

Appendix Table D2. Sample with Missing Values Imputed from MLR

	(1)	(2)	(3)
Per month:	insurer liability	risk transfers	average premium
straight-ddct	161.85**	114.75**	8.30
	(69.25)	(48.35)	(53.91)
AV	1520.32***	549.33***	607.09**
	(417.41)	(194.71)	(295.92)
N	815	744	815
R2	0.17	0.12	0.54
y-mean	359.3	-6.1	379.8
y-sd	141.4	63.4	101.8
Controls	metal tier, network type, state fixed effects		

Note: Each observation is an insurer-year. Straight-ddct is the share of enrollees in the straight-deductible design, and AV is the average AV weighted by enrollment share of each plan. All dependent variables are per member month. The regressions are weighted by the total enrollment at each insurer-year level. *p<10%, **p<5%, ***p<1%.

Appendix Table D3. Risk premium as design measure

	(1)	(2)	(3)	(4)
Per month:	total expenditure	insurer liability	risk transfers	average premium
relative risk premium	-24.96***	-21.01***	-18.34***	-8.27
	(6.77)	(5.62)	(3.89)	(4.02)
AV	-1279.37**	-1308.85***	-762.29***	-357.09
	(552.75)	(468.51)	(209.85)	(350.69)
N	617	617	617	617
R2	0.66	0.43	0.23	0.64
y-mean	381.1	357.1	-6.2	381.1
y-sd	97.4	102.2	66.0	97.4
Controls	metal tier, network type, state fixed effects			

Appendix Table D3 shows that the key results are the same using relative risk premium (in $100) as independent variables. A one deviation change in the relative risk premium is around $200 and it will lead to a total expenditure change of $34 per member month and $31 decrease in risk transfers, both are significant. On the other hand, the coefficient for the average premium is almost zero.

Appendix E. Extensive Margin

In this simulation, I assume there is no individual mandate or premium subsidy. Consumers have the option to choose no insurance. The premiums are set based on the risk pool of individuals who participate in the market. If some risk types choose no insurance, then the premium will adjust to reflect only the risks of those who participate in the market, until the risk pool is stable.

I created 100 risk types using the method specified in Section 2.2.2 from the Truven MarketScan dataset. There is large heterogeneity in the risk distributions among these 100 types: the most healthy types have a mean medical expenditure of $620 and the least healthy types have a mean of $43,380.

I simulate the insurance choice when the 100 risk types face two choice sets: all designs are available; only straight-deductible plans are available. Table 8 shows the results. Without regulation, about 84% population in the market choose some insurance. The number reduced to 78% under the design regulation. The design regulation also has small impacts on the intensive margin: among those choosing an insurance plan, the average effective coverage level reduces from 71.35% to 70.56%.

Table 8. The Intensive and Extensive Margin under Design Regulation

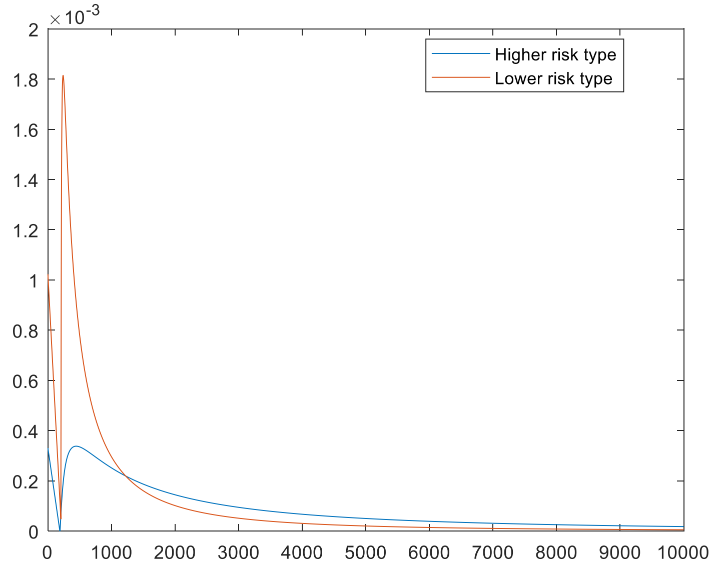
	% with insurance	% losses covered among insured	Overall % losses insured	Average efficiency
No design regulation	83.75%	71.35%	59.76%	-1497
Only straight-deductible plans	78.23%	70.56%	55.20%	-1506

This analysis illustrates how the design variation interacts with the extensive margin. With design variation, lower-risk types are more likely to stay in the market. Restricting plans to only straight-deductible design will drive the lower-risk types out of the market. This will, in turn, creates a feedback loop for the premiums and lower the average enrollment. On average, imposing the regulation reduces the market-average efficiency by $9.

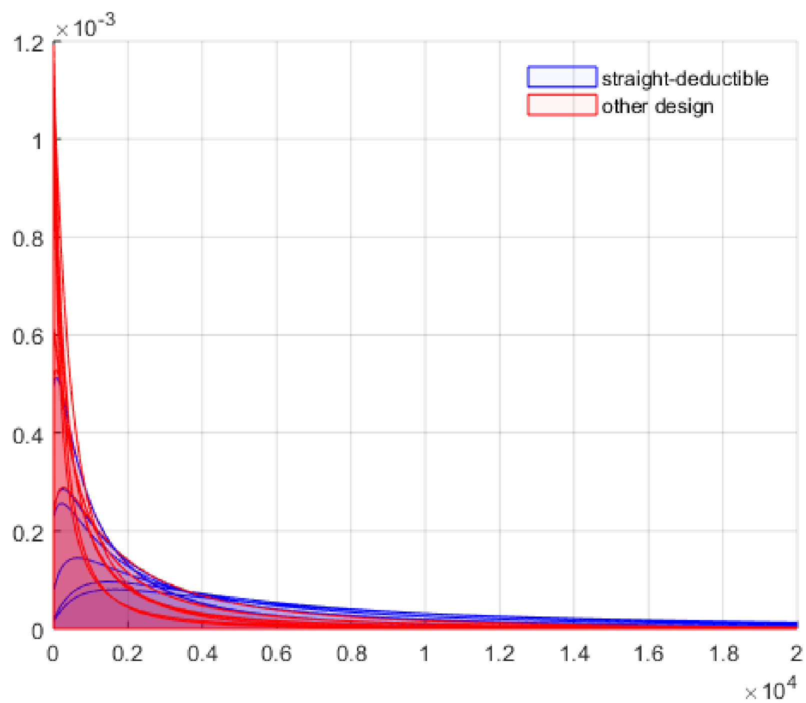
Appendix Table 1. States in the Sample

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 Figure 1. Probability Density Functions of the Two Benchmark Risk Distributions



Appendix Figure 2. Probability Density Functions of Risk Types Choosing Different Designs



Note: Distributions colored in blue (with relatively more weights on larger losses) are those choosing a straight-deductible plan, and that colored red (with relatively more weights on smaller losses) choose a non-straight-deductible design. The probability density function at 0 is not plotted in the graph to make it more readable.