American Economic Review 2022, 112(1): 304–342
https://doi.org/10.1257/aer.20201073

When Should There Be Vertical Choice in Health Insurance Markets?†

By Victoria R. Marone and Adrienne Sabety*

We study the welfare effects of offering choice over coverage levels—"vertical choice"—in regulated health insurance markets. We emphasize that heterogeneity in efficient coverage level is not sufficient to motivate choice. When premiums cannot reflect individuals' costs, it may not be in consumers' best interest to select their efficient coverage level. We show that vertical choice is efficient only if consumers with higher willingness to pay have a higher efficient level of coverage. We investigate this condition empirically and find that as long as a minimum coverage level can be enforced, the welfare gains from vertical choice are either zero or economically small. (*JEL* D82, G22, H75, I13, I21)

Choice over vertically differentiated financial coverage levels—which we term "vertical choice"—is widely available in US health insurance markets. A notable example is the metal-tiered plans (e.g., Gold, Silver, Bronze) offered on Affordable Care Act (ACA) exchanges. In contrast, national health insurance schemes often offer only a single level of coverage. For example, Britons are automatically enrolled in the level of coverage provided by the National Health Service, without a choice. In both contexts, regulation plays a central role in determining the extent of vertical choice. But to date, the economics literature has provided limited guidance to regulators on this topic. This paper aims to fill that gap.

The basic argument in favor of vertical choice is the standard argument in favor of product variety: with more options, consumers can more closely match with their socially efficient product by revealed preference (Dixit and Stiglitz 1977). This argument, however, relies critically on the condition that privately optimal choices align with socially optimal choices. In competitive markets in which costs are independent of consumers' private valuations, this alignment is standard. But in markets

* Marone: Department of Economics, University of Texas at Austin and NBER (email: marone@utexas.edu); Sabety: Department of Economics, University of Notre Dame (email: asabety@nd.edu). Liran Einav was the coeditor for this article. We are grateful to Vivek Bhattacharya, David Cutler, Leemore Dafny, David Dranove, Amy Finkelstein, Tal Gross, Igal Hendel, Gaston Illanes, Matthew Leisten, Thomas McGuire, Matt Notowidigdo, Chris Ody, Rob Porter, Elena Prager, Mar Reguant, Bill Rogerson, Mark Shepard, Amanda Starc, Bob Town, Tom Wiseman, and Gabriel Ziegler, as well as to the coeditor and three anonymous referees for advice and suggestions that greatly benefited this research. We also thank many participants in seminars at Northwestern, Kellogg, BI Norwegian Business School, University of Chicago, Princeton, MIT, Washington University in St. Louis, Yale SOM, Rochester Simon, NYU, MIT Sloan, Chicago Booth, Wisconsin, UT Austin, NBER Summer Institute, and the 2018 ASHEcon Conference for helpful comments. Finally, we thank Jason Abaluck and Jon Gruber for access to the data and for their support of this research project.

† Go to https://doi.org/10.1257/aer.20201073 to visit the article page for additional materials and author disclosure statements.

with selection, like health insurance markets, this alignment may not be possible. In these markets, costs are inextricably related to private valuations, and asymmetric information (or regulation) prevents prices from reflecting marginal costs (Akerlof 1970, Rothschild and Stiglitz 1976). We show that even if health insurance markets are competitive, regulated, and populated by informed consumers, whether vertical choice can increase welfare is theoretically ambiguous.

Our welfare metric derives from a seminal literature on optimal insurance, which holds that the efficient level of coverage equates the marginal benefit of risk protection and the marginal social cost of utilization induced by insurance (Arrow 1965; Pauly 1968, 1974; Zeckhauser 1970). We focus attention on how this central tradeoff between the "value of risk protection" and the "social cost of moral hazard" plays out on a consumer-by-consumer basis. The aim is to design a plan menu that reflects these social incentives by inducing consumers to select their efficient level of coverage. But in doing so, the designer must contend with private incentives: namely that consumers with higher willingness to pay for insurance will select higher coverage. The key challenge is that consumers with higher willingness to pay are not necessarily the consumers with a higher efficient coverage level. It is precisely this statement that captures the theoretical ambiguity of whether it is optimal to offer a vertical choice.

We consider the menu design problem facing a market regulator that can offer vertically differentiated plans and can set premiums.[1] The regulator's objective is to maximize allocative efficiency with respect to consumers and plans. As is standard in employer-sponsored insurance and national health insurance schemes, the regulator need not break even plan by plan, nor in aggregate. If more than one plan is demanded from the regulator's chosen menu, we say it has offered vertical choice. Extending the widely used graphical framework of Einav, Finkelstein, and Cullen (2010), we show that the key condition determining whether the optimal menu features vertical choice is whether consumers with higher willingness to pay have a higher efficient level of coverage. The principal empirical focus of this paper is to determine whether this is likely to be true.

We begin by presenting a model of consumer demand for health insurance, building on Cardon and Hendel (2001) and Einav et al. (2013). The model has two stages. In the first, consumers make a discrete choice over plans under uncertainty about their health. In the second, upon realizing their health, consumers make a continuous choice of health-care utilization. We use the model to show that willingness to pay for insurance can be partitioned into two parts: one that is both privately *and* socially relevant (the value of risk protection), and one that is only privately relevant (the expected reduction in out-of-pocket spending). Because a portion of the private value of insurance is a transfer, higher willingness to pay does not necessarily imply higher social surplus. For example, allocating higher coverage to a sick but risk-neutral consumer delivers her a private benefit, but generates no social benefit; more of her expected health-care spending is simply shifted to others. If

[1] By market regulator, we mean the entity that administers a particular health insurance market: in employer-sponsored insurance, this is the employer; in Medicare, it is the Centers for Medicare and Medicaid Services; under a national health insurance scheme, it is the government.

she consumes more health care than she values in response to higher coverage, the regulator would prefer she had lower coverage.

We estimate the model using data from the population of public school employees in Oregon. The data contain health insurance plan menus, plan choices, and the subsequent health-care utilization of nearly 45,000 households over the period 2008 to 2013. Crucially for identification, we observe plausibly exogenous variation in the plan menus offered to employees. The variation is driven by the fact that plan menus are set independently by each of 187 school districts in the state, which in turn select plans from a common superset determined at the state level. In addition, we observe several coverage levels offered by the same insurer with the same provider network, providing isolated variation along our focal dimension. Our model incorporates observed and unobserved heterogeneity across households along the key dimensions of health status, propensity for moral hazard, and risk aversion. We use the model to recover the joint distribution of household types in this population.

Modeling the structural primitives that underlie demand and costs in the market allows us to evaluate these objects for any level of coverage a regulator might wish to consider. The key advantage of this strategy is that we can look beyond the set of contracts observed in our particular setting (Einav, Finkelstein, and Cullen 2010). This flexibility also raises the distinct design question of how "closely spaced" to permit contracts to be. Should the regulator permit choice over contracts differentiated by only \$10 in deductible? We do not directly model the potential costs of such fine-grained choice (such as fixed costs of offering contracts), but we do evaluate the potential benefits. As discussed below, in practice we find that even under ideal conditions, the returns to allowing very closely spaced contracts are economically small.

Our estimates imply that all households have a fairly high efficient level of coverage, ranging between a high-deductible contract (with a \$10,000 deductible and full coverage thereafter) and full insurance. Contracts outside this range can be ruled out from the optimal menu (as they deliver lower social surplus for every household). Within this range, we find that households with higher willingness to pay are primarily motivated by a greater expected reduction in out-of-pocket spending, rather than by a greater value of risk protection. These households are highly likely to spend past \$10,000, and therefore face little out-of-pocket cost uncertainty under any contract in the relevant range of coverage levels. Although there are competing factors, this negative relationship between willingness to pay and "relevant risk" quantitatively dominates. As a result, we find at best only a weak relationship between willingness to pay and efficient coverage level.

We solve for the optimal menu under a baseline requirement that contracts be no closer than \$2,500 out-of-pocket maximum intervals. We find that the optimal menu consists of a single contract. Introducing any other contract, at any price, leads to over- or underinsurance (on average) among households that would choose the alternative. We then increase the permissible density of contracts by a factor of ten (to \$250 out-of-pocket maximum intervals). Here, we find that it is efficient to offer a vertical choice: the optimal menu features four contracts, clustered around the original optimal single contract. However, because social surplus is quite flat across coverage levels near the optimum, the welfare gains are small. Offering a choice increases welfare by only $5 per household per year relative to what is achieved by a single contract, a gain which may be quickly eroded by factors outside the scope of our model.

It is important to emphasize that these results may not reflect the value of vertical choice in all health insurance markets. Indeed, robustness analysis in Section IV reveals that doubling average risk aversion leads the optimal menu to feature vertical choice, even in the "sparse" contract space. When risk protection accounts for a larger part of the variation in willingness to pay for higher coverage, the revelation of private information (through choice) becomes more valuable. On the other hand, there are some reasons to think that our findings may be reflected in other settings. The negative relationship between willingness to pay and "relevant risk" is a central driver of our results, and this relationship follows from a pair of factors: (i) variation in willingness to pay is primarily driven by consumers' information about their upcoming health needs, and (ii) the lowest relevant level of coverage is reasonably high. The first factor implies that the highest willingness-to-pay consumers are the sickest, and the second implies that these consumers would face little out-of-pocket cost uncertainty even in the lowest relevant coverage level. These factors seem particularly plausible in settings in which contracts span a short time, and in which exposure to substantial out-of-pocket spending risk is not efficient for anyone.

In the last part of the paper (Section V), we evaluate the welfare and distributional implications of the optimal plan menu relative to a status quo with vertical choice. We find that the optimal menu (the single contract) increases welfare by $315 per household per year. But these gains are not shared evenly in the population. Sicker and larger households fare best under the single contract, while healthier and smaller households fare best under vertical choice. Our results suggest that one reason for the persistence of vertical choice in settings such as employer-sponsored insurance could be to limit redistribution across these groups.

Beyond the work noted above, our theoretical approach is most closely related to Azevedo and Gottlieb (2017), who also model demand for health insurance in a setting with vertically differentiated contracts and multiple dimensions of consumer heterogeneity. While their focus is on competitive equilibria, their numerical simulations also consider optimal pricing. They document that under certain distributions of consumer types, offering choice is optimal, while under others it is not. Our paper focuses directly on why this is the case, and brings to bear an empirical approach that permits substantially more flexibility in the distribution of consumer types.

Our paper also closely relates to work that evaluates allocative efficiency in health insurance markets (Cutler and Reber 1998; Lustig 2008; Carlin and Town 2008; Dafny, Ho, and Varela 2013; Kowalski 2015; Tilipman 2018), and more specifically to the growing literature on menu design in these markets.[2] In the context of insurer choice, Bundorf, Levin, and Mahoney (2012) investigate the optimal allocation of consumers to insurers, and find that it cannot be achieved by uniform pricing. Our paper is similar in spirit (and in findings), but focuses instead on the financial dimension of insurance. In this same context, Ericson and Sydnor (2017)

[2] We also note the close relationship between our paper and recent work by Landais et al. (2021) on unemployment insurance and Hendren, Landais, and Spinnewijn (2021) on social insurance more broadly. Like us, these papers consider the value of offering a choice from the perspective of a social planner that can set prices.

also consider the question of whether choice is welfare improving. A key difference of our work is that we consider a setting in which contract characteristics are endogenous and premiums are exogenous, as opposed to the reverse. In a similar spirit, Ho and Lee (2021) study optimal menu design from the perspective of an employer. Like us, they find that the gains from offering a choice over coverage levels are small. Our contribution relative to these papers is to provide a conceptual characterization of when choice over financial coverage levels is and is not valuable. We view this characterization as a tool that can be used to re-examine the design of existing health insurance markets through a new lens. Our empirical analysis demonstrates the relevance of the prediction that vertical choice may not be valuable, and links it to the distribution of fundamentals—risk aversion, propensity for moral hazard, and distributions of health outcomes—in a population.

Finally, we view our work as complementary to the large literature documenting the fact that consumers have difficultly optimizing over health insurance products (Abaluck and Gruber 2011, 2016; Ketcham et al. 2012; Handel and Kolstad 2015; Bhargava, Loewenstein, and Sydnor 2017), which has recently also focused on ways in which consumers can be nudged into doing so (Abaluck and Gruber 2016, 2017; Gruber et al. 2019; Bundorf et al. 2019; Samek and Sydnor 2020). Importantly, if privately and socially optimal allocations do not align, more diligent consumers may just as well lead to less desirable outcomes (as is found by Handel 2013). A central aim of the present paper is to inform the design of health insurance markets in such a way that better-informed consumers always lead to better allocations.

The paper proceeds as follows. Section I presents our theoretical model and derives the objects relevant to describe private and social incentives. Section II describes our data and the variation it provides. Section III presents the empirical implementation of our model. Section IV presents the model estimates and main results. Section V evaluates welfare and distributional outcomes. Section VI concludes.

I. Theoretical Framework

A. *Model*

We consider a model of a health insurance market in which consumers are heterogeneous along multiple dimensions and the set of traded contracts is endogenous. We assume that premiums may not vary with consumer characteristics, that claims may be contingent only on health-care utilization, and that each consumer will select a single contract.[3]

We denote a set of potential contracts by $X = \{x_0, x_1, \ldots, x_n\}$, where x_0 is a null contract that provides no insurance. Within X, contracts are vertically differentiated by the financial level of coverage provided. Consumers are characterized by type $\theta = \{F, \psi, \omega\}$, where F is a distribution over potential health states, $\psi \in \mathbb{R}_{++}$ is a risk aversion parameter, and ω is a parameter that governs consumer preferences

[3] It may not be possible to vary premiums with consumer attributes if consumers have private information (Cardon and Hendel 2001), or it may not be desirable to do so to prevent exposing consumers to reclassification risk (Handel, Hendel, and Whinston 2015). Otherwise, the market could be partitioned according to observable characteristics, and each submarket could be considered separately.

for health-care utilization (and ultimately captures the degree of moral hazard). A population is defined by a distribution $G(\theta)$.

Demand for Health Insurance and Health-Care Utilization.—Consumers are subject to a stochastic health state l, drawn from their distribution F. Given their health state, consumers decide the money amount $m \in \mathbb{R}_+$ of health-care utilization ("spending") to consume, a decision which in part depends on their insurance contract. Contracts are characterized by an increasing and concave out-of-pocket cost schedule $c_x : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, where $c_x(m) \leq m$.

Consumers value health-care spending m and residual income y. Preferences are represented by $u_\psi(y + b(m; l, \omega))$, where b is a money-metric valuation of health-care utilization, and u_ψ and $b(\cdot; l, \omega)$ are each strictly increasing and concave. Upon realizing their health state, consumers choose their health-care utilization by trading off its benefit with its out-of-pocket cost: $m^*(l, \omega, x) = \text{argmax}_m(b(m; l, \omega) - c_x(m))$. Privately optimal utilization implies indirect benefit $b^*(l, \omega, x) = b(m^*(l, \omega, x); l, \omega)$ and indirect out-of-pocket cost $c_x^*(l, \omega, x) = c_x(m^*(l, \omega, x))$. Before the health state is realized, expected utility is given by

$$U(x, p, \theta) = E\left[u_\psi(\hat{y} - p - c_x^*(l, \omega, x) + b^*(l, \omega, x)) \,|\, l \sim F\right], \tag{1}$$

where p is the contract premium and $\hat{y}$ is initial income.

Private versus Social Incentives.—Absent insurance, consumers pay the full cost of health-care utilization, m. Socially optimal health-care utilization therefore coincides with privately optimal utilization absent insurance.[4] The difference between privately optimal spending $m^*(l, \omega, x)$ and socially optimal spending $m^*(l, \omega, x_0)$ determines the social cost of insurance. Since insurance reduces the price consumers pay for health care, $m^*(l, \omega, x)$ typically exceeds $m^*(l, \omega, x_0)$. We refer to this induced utilization as "moral hazard spending."[5] A consumer's net payoff from moral hazard spending is given by

$$v(l, \omega, x) = \underbrace{b^*(l, \omega, x) - b^*(l, \omega, x_0)}_{\text{Benefit of moral hazard spending}} - \underbrace{(c_x^*(l, \omega, x) - c_x^*(l, \omega, x_0))}_{\text{Out-of-pocket cost of moral hazard spending}},$$

where $b^*(l, \omega, x_0)$ is the indirect benefit of uninsured behavior, and $c_x^*(l, \omega, x_0)$ is the out-of-pocket cost of uninsured behavior at insured prices. Note that since any change in behavior is voluntary, $v(l, \omega, x)$ is weakly positive.

[4] Importantly, this is true only if m represents the true cost of health-care provision and if there are no externalities associated with health-care utilization, as we assume here.

[5] Following convention, we use the term "moral hazard" to describe the scenario at hand, in which there is elastic demand for the insured good and a state that is not contractible. Note that this is not a problem of hidden action, but rather of hidden information. A fuller discussion of this (ab)use of terminology in the health insurance literature can be found in Section I.B of Einav et al. (2013), as well as in the dialogue between Pauly (1968) and Arrow (1968).

Calculations in online Appendix A.1 show that if u_ψ features constant absolute risk aversion, willingness to pay for contract x relative to the null contract x_0 can be expressed as[6]

$$(2)\quad WTP(x,\theta) \;=\; \underbrace{E_l\left[c^*_{x_0}(l,\omega,x_0) - c^*_x(l,\omega,x_0)\right]}_{\text{Expected reduction in out-of-pocket cost holding behavior fixed}} + \underbrace{E_l\left[v(l,\omega,x)\right]}_{\text{Expected payoff from moral hazard spending}} + \underbrace{\Psi(x,\theta)}_{\text{Value of risk protection}}.$$

Willingness to pay is composed of three terms: the expected reduction in out-of-pocket cost holding behavior fixed (at uninsured behavior), the expected payoff from moral hazard spending, and the value of risk protection. The first term captures the transfer from the consumer to the insurer of the expected health-care spending liability that exists even absent moral hazard. It will be an equal and opposite cost to the insurer. The second and third terms, in contrast, depend on consumers preferences and are relevant to social welfare. Consumers partially value the additional health care they consume when they have higher coverage, as well as the ability to smooth consumption across health states.

Insurer costs are given by $k_x(m)$, where $m = k_x(m) + c_x(m)$. A reduction in out-of-pocket cost is an increase in insurer cost, so $c^*_{x_0}(l,\omega,x_0) - c^*_x(l,\omega,x_0) = k^*_x(l,\omega,x_0)$. The social surplus generated by allocating a consumer of type θ to contract x (relative to allocating the same consumer to the null contract) is the difference between $WTP(x,\theta)$ and expected insured cost $E_l\left[k^*_x(l,\omega,x)\right]$, which after simplifying is

$$(3)\qquad SS(x,\theta) \;=\; \underbrace{\Psi(x,\theta)}_{\text{Value of risk protection}} - \underbrace{E_l\left[k^*_x(l,\omega,x) - k^*_x(l,\omega,x_0) - v(l,\omega,x)\right]}_{\text{Social cost of moral hazard}}.$$

Because the insurer is risk neutral, it bears no extra cost from uncertain payoffs. If there is moral hazard, the consumer's value of her expected health-care spending falls below its cost, generating a welfare loss from insurance. The welfare loss equals the portion of the expected increase in health-care spending that is not valued.

The socially optimal contract for each type of consumer optimally trades off the value of risk protection and the social cost of moral hazard: $x^{eff}(\theta) = \text{argmax}_{x \in X} SS(x,\theta)$. Given premium vector $\mathbf{p} = \{p_x\}_{x\in X}$, the privately optimal contract optimally trades off private utility and premium: $x^*(\theta,\mathbf{p}) = \text{argmax}_{x\in X}\big(WTP(x,\theta) - p_x\big)$.

Supply and Regulation.—We suppose contracts are supplied by a regulator, which can observe the distribution of consumer types and can set premiums on all contracts except x_0, which has zero premium. The regulator need not break even on any given contract, nor in aggregate. It can effectively remove any nonnull contract from the set of contracts on offer by setting a premium of infinity. It can effectively remove x_0 from offer by setting the premium of any nonnull contract to zero. This simple model of supply is isomorphic to a more complicated model involving

[6]The single role of constant absolute risk aversion is to ensure that the value of risk protection, and thereby social surplus, is invariant to the contract premium.

perfect competition among private insurers and a regulator that can strategically tax or subsidize contracts. Precisely such a model is formalized in Section 6 of Azevedo and Gottlieb (2017).

The regulator sets premiums $\mathbf{p}$ in order to maximize social welfare, given by

$$W(\mathbf{p}) = \int SS\big(x^*(\theta,\mathbf{p}),\theta\big)\,dG(\theta).$$

Our question is whether, or when, the regulator's solution will involve vertical choice.

B. *Graphical Analysis*

We characterize the answer graphically for the case of a market with only two potential contracts. This case conveys the basic intuition and can be depicted easily using the graphical framework introduced by Einav, Finkelstein, and Cullen (2010).

First, it is useful to recognize that moral hazard, risk aversion, and consumer heterogeneity are necessary conditions for vertical choice to be efficient. If there were no moral hazard, the highest coverage contract would be socially optimal for all consumers, and the optimal menu would involve only this contract. If there were no risk aversion, the same would be true with the lowest coverage contract. If there were no consumer heterogeneity, all consumers would again have the same socially optimal contract, and the optimal menu would again feature a single contract. In the following, we explore the more interesting (and more realistic) cases in which consumers do not all have the same socially optimal contract.

Two Contract Example.—Suppose there are two potential contracts, x_H and x_L, where x_H provides higher coverage than x_L. Figure 1 depicts the market for x_H in two populations. If a consumer does not choose x_H, they receive x_L; x_0 is excluded by setting p_L to zero. As x_H provides higher coverage, $WTP(x_H,\theta) \geq WTP(x_L,\theta)$ for all consumers. Each panel shows the demand curve D for contract x_H, representing marginal willingness to pay for x_H relative to x_L. The vertical axis plots the marginal premium $p = p_H - p_L$ at which the contracts are offered. The horizontal axis plots the fraction q of consumers that choose x_H.

Each panel also shows the marginal cost curve MC and the marginal social surplus curve SS. The marginal cost curve measures the expected cost of insuring consumers under x_H relative to x_L: $E_l\big[k^*_{x_H}(l,\omega,x_H) - k^*_{x_L}(l,\omega,x_L)\big]$. Because consumers with the same willingness to pay can have different costs, MC represents the average marginal cost among all consumers at a particular point on the horizontal axis (a particular level of marginal willingness to pay). The social surplus curve SS plots the vertical difference between D and MC, or equivalently, the average value of $SS(x_H,\theta) - SS(x_L,\theta)$ among all consumers at a particular point on the horizontal axis.

Though vertical differentiation implies D and MC must be weakly positive, the presence of moral hazard means that SS need not be. It is possible for consumers to be overinsured. Moreover, our precondition that all consumers do not have the same socially optimal contract guarantees that in both populations, marginal social surplus

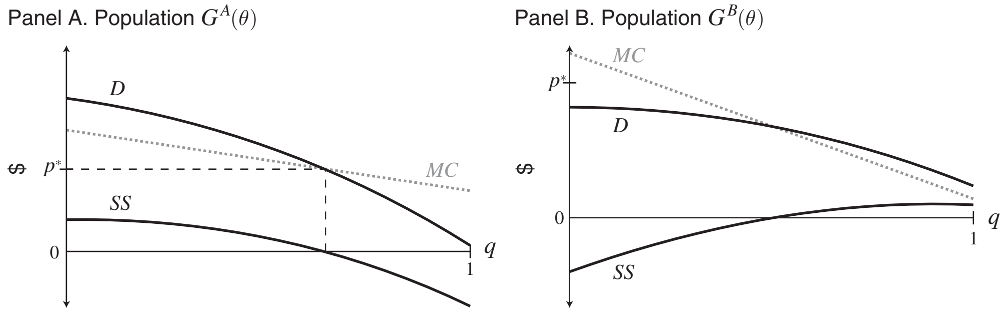

Figure 1. Examples in Which Vertical Choice (a) *Is* and (b) *Is Not* Efficient

Notes: The figure shows two health insurance markets in which there are two contracts available: x_H and x_L, where x_H provides higher coverage than x_L. Each panel shows the demand curve *D*, the marginal cost curve *MC*, and the social surplus curve *SS* for x_H relative to x_L. In the left panel, the regulator optimally offers vertical choice. In the right panel, the regulator optimally does not offer vertical choice, and all consumers choose x_L.

will be positive for some consumers and negative for others.[7] The key difference between populations $G^A(\theta)$ and $G^B(\theta)$ is whether high or low willingness-to-pay consumers have a higher efficient level of coverage. In population $G^A(\theta)$, marginal social surplus is increasing in marginal willingness to pay. The premium p^* can therefore sort consumers with on-average positive *SS* into x_H, and on-average negative *SS* into x_L. In population $G^B(\theta)$, meanwhile, such a premium does not exist.

In population $G^B(\theta)$, any interior allocation results in some amount of "backward sorting," meaning that there is a group of consumers enrolled in x_H who would be more efficiently enrolled in x_L, and vice versa. Consequently, any allocation with enrollment in both contracts is dominated by an allocation with enrollment in only one. No sorting dominates backward sorting because it is always possible to prevent "one side" of the backward sort. To see this, consider the allocation $\tilde{q}$ at the point where *SS* intersects zero. Any allocation to the right of $\tilde{q}$ strictly dominates, as more consumers with positive marginal social surplus now enroll in x_H. The same logic applies to the left of $\tilde{q}$. The only allocations that cannot easily be ruled out as suboptimal are the endpoints, at which all consumers enroll in the same contract. In the example shown, the integral of *SS* is negative, meaning that the population would on average be overinsured in x_H. p^* is therefore anything high enough to induce all consumers to choose x_L.

Remarks.—The limitation of choice as a screening mechanism is directly related to the idea that a single (community-rated) price may not be able to efficiently sort consumers that vary in cost (Einav, Finkelstein, and Levin 2010; Glazer and McGuire 2011; Bundorf, Levin, and Mahoney 2012; Shepard 2016; Geruso 2017). Consumers select a contract based on the available consumer surplus, $CS = WTP - p$, while

[7] Note that as *SS* represents an average, this condition does not itself guarantee that the social surplus curve will cross zero. Since it is necessary for *SS* to cross zero for vertical choice to be optimal, we focus our two examples on cases in which that occurs. If *SS* did not cross zero, a single plan would be on-average optimal at every level of willingness to pay, and the optimal menu would feature a single contract.

efficiency relies on a comparison with costs, $SS = WTP - MC$. When CS and SS diverge (when $p \neq MC$), the efficiency of choice turns on whether they are at least positively related. If they are not, choice can only result in some degree of "backward sorting."

In the simple case of two contracts and a social surplus curve that crosses zero at most once, vertical choice is efficient if and only if it crosses from above. In a more general case with multiple potential contracts and arbitrary social surplus curves, this necessary and sufficient condition is still directly informative. If consumers all have the same socially optimal contract (or more plausibly, if the same contract is socially optimal at all levels of willingness to pay), there will be no crossing in the upper envelope of social surplus curves, and the optimal menu will feature this single contract. If instead there *is* crossing in the upper envelope of social surplus curves, one must assess whether the higher-coverage contracts cross from above, or in other words, whether or not choice would lead to backward sorting.

Taken together, the procedure for evaluating the efficiency of vertical choice can be summarized by a test for the condition of whether consumers with higher willingness to pay have a higher efficient coverage level, where we emphasize that *higher*, in both instances, is to be evaluated strictly. This condition itself is complex. It is both theoretically ambiguous and, by our own assessment, not obvious. If healthy consumers change their behavior more in response to insurance, as is suggested by findings in Brot-Goldberg et al. (2017), this would tend toward positively aligning willingness to pay and efficient coverage level. If healthy consumers are more risk averse, as is suggested by findings in Finkelstein and McGarry (2006), this would tend toward negatively aligning them.

There is a question of what characteristics drive variation in willingness to pay, and in turn how those characteristics determine the efficient level of coverage. The net result depends on the joint distribution of expected health spending, uncertainty in health spending, risk aversion, and moral hazard in the population. Moreover, it depends on how these primitives map into marginal willingness to pay and marginal insurer cost across nonlinear insurance contracts, as are common around the world and present in the empirical setting we study. Ultimately, whether consumers with higher private valuations of higher coverage also generate a higher social value from higher coverage is an open empirical question.

II. Empirical Setting

A. *Data*

Our data are derived from the employer-sponsored health insurance market for public school employees in Oregon between 2008 and 2013 (OEBB 2018). The market is operated by the Oregon Educators Benefit Board (OEBB), which administers benefits for the employees of Oregon's 187 school districts. Each year, OEBB contracts with insurers to create a state-level "master list" of plans and associated premiums that school districts can offer to their employees. During our time period, OEBB contracted with three insurers, each of which offered a selection of plans. School districts then independently selected a subset of plans from the state-level menu and set an "employer contribution" toward plan premiums. Between 2008 and

2010, school districts could offer at most four plans; after 2010, there was no limit, but many still offered only a subset.

The data contain employees' plan menus, realized plan choices, plan characteristics, and medical and pharmaceutical claims for all insured individuals. We observe detailed demographic information about employees and their families, including age, gender, zip code, health risk score, family type, and employee occupation type.[8] An employee's *plan menu* consists of a plan choice set and plan prices. Plan prices consist of the subsidized premium, potential contributions to a health reimbursement arrangement (HRA) or a health savings account (HSA), and potential contributions toward a vision or dental insurance plan.[9]

The decentralized determination of plan menus provides a plausibly exogenous source of variation in both prices and choice sets. While all plan menus we observe are quite generous, in that the plans are generally high coverage and are highly subsidized, there is substantial variation across districts in the range of coverage levels offered and in the exact nature of the subsidies.[10] Moreover, school districts can vary plan menus by family type and occupation type, resulting in variation both within and across districts. Plan menu decisions are made by benefits committees consisting of district administrators and employees, and subsidy designs are influenced by bargaining agreements with local teachers' unions. Between 2008 and 2013, we observe 13,661 unique combinations of year, school district, family type, and occupation type, resulting in 7,835 unique plan menus.

Plan Characteristics.—During our sample period, OEBB contracted with three insurers: Kaiser, Moda, and Providence. Kaiser offered health maintenance organization (HMO) plans that require enrollees to use only Kaiser health-care providers and obtain referrals for specialist care. Moda and Providence offered preferred provider organization (PPO) plans with broad provider networks. Each insurer used a single provider network and offered multiple plans. Within insurer, plans were differentiated only by financial coverage level.

Table 1 summarizes the state-level master list of plans made available by OEBB in 2009. The average employee premium represents the average annual premium employees would have had to pay for each plan. The full premium reflects the per-employee premium paid to the insurer. This premium varies formulaically by family type; the one shown is for an employee plus spouse. The difference between

[8] Individual risk scores are calculated based on prior-year medical diagnoses and demographics using Johns Hopkins ACG Case-Mix software. This software uses diagnostic information contained in past claims data as well as demographic information to predict future health-care spending. See, for example, Brot-Goldberg et al. (2017), Carlin and Town (2008), or Handel and Kolstad (2015) for more in-depth explanation of the software and examples of its use in economic research.

[9] Decisions about HSA/HRA and vision/dental contributions are also made independently by school districts. An HRA is a notional account that employers can use to reimburse employees' uninsured medical expenses on a pretax basis; balances expire at the end of the year or when the employee leaves the employer. An HSA is a financial account maintained by an external broker to which employers or employees can make pretax contributions. Data on employer premium contributions and savings account contributions were hand collected via surveys of each school district. Additional details on the data collection process can be found in Abaluck and Gruber (2016).

[10] The majority of school districts used either a fixed dollar contribution or a percentage contribution, but the levels of the contribution varied widely. Other districts used a fixed employee contribution. In addition, the districts' policies for how "excess" contributions were treated varied; in some cases, contribution amounts in excess of the full plan premium could be "banked" by the employee in a HSA or HRA, or else put toward the purchase of a vision or dental insurance plan.

Table 1—Plan Characteristics, 2009

Plan	Actuarial value	Avg. employee premium ($)	Full premium ($)	Deductible ($)	OOP max. ($)	Market share
Kaiser–1	0.97	688	10,971	0	1,200	0.07
Kaiser–2	0.96	554	10,485	0	2,000	0.11
Kaiser–3	0.95	473	10,163	0	3,000	< 0.01
Moda–1	0.92	1,594	12,421	300	500	0.27
Moda–2	0.89	1,223	11,839	300	1,000	0.05
Moda–3	0.88	809	11,174	600	1,000	0.11
Moda–4	0.86	621	10,702	900	1,500	0.10
Moda–5	0.82	428	9,912	1,500	2,000	0.13
Moda–6	0.78	271	8,959	3,000	3,000	0.04
Moda–7	0.68	92	6,841	3,000	10,000	0.01
Providence–1	0.96	2,264	13,217	900	1,200	0.07
Providence–2	0.95	1,995	12,895	900	2,000	0.02
Providence–3	0.94	1,825	12,683	900	3,000	0.01

Notes: The table shows the state-level master list of plans available in 2009. Actuarial value is the ratio of the sum of insured spending across all households to the sum of total spending across all households. The average employee premium is taken across all employees, even those who did not choose a particular plan. The full premium reflects the premium negotiated by OEBB and the insurer; the one shown is for an employee plus spouse. The deductible and out-of-pocket (OOP) maximum shown are for in-network services for a family household.

the employee premium and the full premium is the contribution by the school district. Plan cost-sharing features vary by whether the household is an individual (the employee alone) or a family (anything else). The deductible and out-of-pocket maximums shown are for a family household.[11]

As a way to summarize and compare plan coverage levels, we construct each plan's actuarial value. This measure reflects the share of total population spending that would be insured under a given plan.[12] Full insurance would have an actuarial value of one; less generous plans have lower actuarial values. In later years, the distribution of coverage levels looks qualitatively similar, with the notable exception that Providence was no longer available in 2012 and 2013. Table A.1 provides corresponding information for the plans offered in other years.

Household Characteristics.—We restrict our analysis sample to households in which the oldest member is not older than 65, the employee is not retired, and all members are enrolled in the same plan for the entire year. Further, because a prior year of claims data is required to estimate an individual's prospective health risk score, we require that households have one year of data prior to inclusion; this means our sample begins in 2009. These restrictions leave us with 44,562 households, representing 117,934 individuals. Table A.2 provides additional details on sample construction.

There is a clear bifurcation of our sample between Kaiser and non-Kaiser households. That is, 78 percent of households always chose either Moda or Providence,

[11] Many other cost-sharing details determine plan coverage level. For the purposes of our empirical model, we estimate the coinsurance rate and out-of-pocket maximum that best fit the relationship between out-of-pocket spending and total spending observed in the claims data. This procedure is described in online Appendix B.1.

[12] We evaluate out-of-pocket spending for each household in each plan, and then divide average insured spending by average total spending across all households for each plan. We evaluate counterfactual out-of-pocket spending using the "claims calculator" developed for this setting by Abaluck and Gruber (2016).

Table 2—Household Summary Statistics

	Full sample	Excluding Kaiser
Number of households	44,562	34,606
Number of enrollees	117,934	92,244
Percent of households with children	0.49	0.49
Enrollees per household, mean (median)	2.57 (2)	2.60 (2)
Enrollee age, mean (median)	39.8 (37.8)	40.4 (38.7)
Premiums		
Employee premium ($), mean (median)	880 (0)	843 (0)
Full premium ($), mean (median)	11,500 (11,801)	11,582 (11,801)
Household health-care spending		
Total spending ($), mean (median)	10,754 (4,620)	11,689 (5,173)
Out-of-pocket ($), mean (median)	1,694 (1,093)	2,054 (1,540)
Switching (percent of household-years)		
Forced to switch plan	0.20	0.21
insurer	0.01	0.02
Unforced, switched plan	0.17	0.20
insurer	0.03	0.03

Notes: Enrollees are employees plus their covered dependents. Sample statistics are calculated across all years, 2009–2013. Premiums statistics are for households' chosen plans, as opposed to for all possible plans. Sample medians are shown in parentheses.

19 percent always chose Kaiser, and only 3 percent at some point switched between. This pattern is not necessarily surprising. Kaiser offers a substantially different type of insurance product, and persistent consumer preference heterogeneity along this dimension would be a reasonable expectation. That said, modeling the choice over insurer type somewhat distracts from our focus on choice over financial coverage level. We therefore take advantage of this division in the data and conduct our primary analysis on the set of households that never enrolled with Kaiser. We consider the full sample in a robustness analysis in Section IVC.

Table 2 provides summary statistics on our panel of households. The first column describes the full sample, while the second column describes the subset of households that never enrolled in a Kaiser plan. Focusing on the non-Kaiser sample, 49 percent of households have children, and 74 percent of households are "families" (anything other than the employee alone). The average employee is age 47.9, and the average enrollee (employees and their covered dependents) is age 40.4. Households on average have 2.6 enrollees.

Employees received large subsidies toward the purchase of health insurance. The average household paid only $843 per year for their chosen plan; the median household paid nothing. Meanwhile, the average full premium paid to insurers was $11,582, meaning that the average household received an employer contribution of $10,739. Households had average out-of-pocket spending of $2,054 and average total health-care spending of $11,689.

Households were highly likely to remain in the same plan and with the same insurer they chose the previous year. That said, OEBB could adjust the state-level master list of available plans, and school districts could adjust choice sets over time. Because their prior choice was no longer available, such adjustments forced 21 percent of household-years to switch plans, and 2 percent to switch insurers. When the

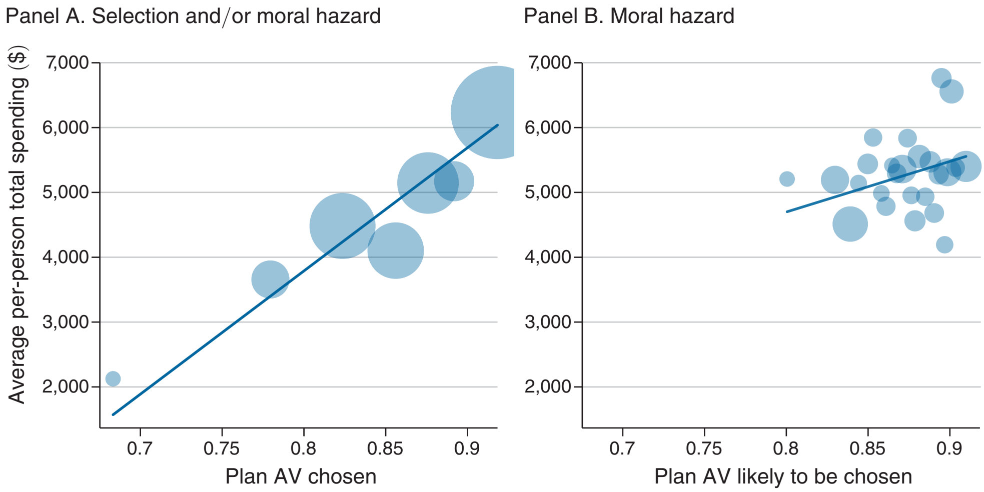

FIGURE 2. AVERAGE SPENDING BY PLAN COVERAGE LEVEL

Notes: The figure shows the relationship between average per-person total spending and plan actuarial value (AV) for households that selected Moda in 2009. In the left panel, households are grouped by their chosen plan, and each dot represents one of the seven Moda plans. In the right panel, households are grouped by the plan menu they were offered, and each dot represents a set of plan menus with the same predicted actuarial value. Predicted actuarial value is the AV most likely to be chosen if an average household were presented with that plan menu. The size of each dot indicates the number of households represented. Lines of best fit are weighted accordingly.

prior choice *was* available, 20 percent of household-years voluntarily switched plans and only 3 percent voluntarily switched insurers. The presence of both forced and unforced switching is important in our empirical model for identifying the extent of "inertia" in households' choice of plan and insurer.

B. *Variation in Plan Menus*

For the purposes of the present research, the two most important features of our setting are the isolated variation along the dimension of coverage level and the plausibly exogenous variation in plan menus. Variation in coverage level exists primarily among the plans offered by Moda. Variation in plan menus stems from the decentralized determination of employee health benefits. Both are central to identification of our empirical model.

To provide a sense of this variation, Figure 2 shows the relationship between health-care spending and plan actuarial value for households that chose Moda in 2009. In the left panel, households are grouped by their chosen plan. The plot shows average spending among households in each of the seven Moda plans, weighting each plan by enrollment. Unsurprisingly, households that enrolled in more generous plans had higher spending, reflecting adverse selection, moral hazard, or both.

The right panel groups households by their plan menu. It plots the actuarial value that an average household would be most likely to choose if offered a given plan menu, against the average spending of all households presented with that menu. This measure of plan menu generosity captures both the facts that a level of coverage can only be chosen if it is offered, and is more likely to be chosen if it is cheaper.[13] Each point on the plot represents the set of plan menus that share the same predicted actuarial value. Points are then weighted by the number of households represented. The resulting pattern indicates that households that were offered a more generous plan menu had higher spending. The patterns in both panels persist when we control for observables, suggesting the presence of adverse selection on unobservables, and of moral hazard.

Identification of our structural model proceeds in much the same way as the arguments above. A key identifying assumption is that plan menus are independent of household unobservables, conditional on household observables. An important threat to identification is that school districts chose plan menu generosity in response to unobservable information about employees that would also drive health-care spending. To the extent that districts with unobservably sicker households provided more generous health benefits, this would lead us to overstate the extent of moral hazard.[14]

We investigate this possibility by attempting to explain plan menu generosity with observable household characteristics, in particular health. We argue that if plan menus were not responding to *observable* information about household health, it is unlikely that they were responding to *unobservable* information. We find this argument compelling because we almost certainly have better information on household health (through health risk scores) than did school districts at the time they made plan menu decisions. Online Appendix Table A.5 presents this exercise. Conditional on family type, we find no correlation between plan menu generosity and household risk score. Online Appendix B.2 describes these results in greater detail. It also presents additional tests for what *does* explain variation in plan menus. We find that, among other things, plan menu generosity is higher for certain union affiliations, lower for substitute teachers and part-time employees, decreasing in district average house price index, and decreasing in the percentage of registered Republicans in a school district. None of these relationships are inconsistent with our understanding of the process by which district benefits decisions are made.

We exploit this identifying variation within our structural model, but can also use it in a more isolated way to produce reduced-form estimates of moral hazard. Online Appendix B.3 presents an instrumental variables analysis using two-stage least squares. The estimates yield a moral hazard "elasticity" that can be directly compared with others in the literature. We estimate that the elasticity of demand for health-care spending with respect to its average end-of-year out-of-pocket cost is -0.27, broadly similar to the benchmark estimate of -0.2 from the RAND experiment (Manning et al. 1987, Newhouse 1993). We also find suggestive evidence of heterogeneity in moral hazard effects, which is an important aspect of our structural model and of our research question.

[13] We construct this measure using a conditional logit model of household plan choice. This model and the resulting measure of plan menu generosity are described in detail in online Appendix B.2.

[14] The relationship could also run the other way: households could move across school districts, or select a district initially, based on the available health benefits. Such selection could again result in unobservably sicker households obtaining more generous health benefits. To the extent that observable health factors are correlated with unobservable factors that would drive this relationship, the analysis that follows is also relevant to this concern.

III. Empirical Model

A. *Parameterization*

We parameterize household utility and the distribution of health states, allowing us to represent our theoretical model fully in terms of data and parameters to estimate. We extend the theoretical model to account for the fact that in our empirical setting, there are multiple insurers, consumers are households consisting of individuals, a dollar in premiums may be valued differently than a dollar in out-of-pocket spending, and consumers make repeated plan choices over time.

Household Utility.—Following Cardon and Hendel (2001) and Einav et al. (2013), we parameterize the value of health-care spending to be quadratic in its distance from the health state. Household k's valuation of spending level m given health state realization l is given by

$$b(m; l, \omega_k) = (m - l) - \frac{1}{2\omega_k}(m - l)^2, \tag{4}$$

where ω_k governs the curvature of the benefit of spending and, ultimately, the degree to which optimal spending varies across coverage levels. Given out-of-pocket cost function $c_{jt}(m)$ for plan j in year t, privately optimal health-care spending is $m^*_{jt}(l, \omega_k) = \text{argmax}_m\big(b(m; l, \omega_k) - c_{jt}(m)\big)$.[15]

This parameterization is attractive because it produces reasonable predicted behavior under nonlinear insurance contracts, and is tractable enough to be used inside an optimization routine.[16] Additionally, ω_k can be usefully interpreted as the incremental spending induced by moving a household from no insurance to full insurance. Substituting for m^*, we denote the benefit of optimal utilization as $b^*_{jt}(l, \omega_k)$ and the associated out-of-pocket cost as $c^*_{jt}(l, \omega_k)$. Households face uncertainty in payoffs only through uncertainty in $b^*_{jt}(l, \omega_k) - c^*_{jt}(l, \omega_k)$.

Household k in year t derives the following expected utility from plan choice j:

$$U_{kjt} = \int - \exp\big(-\psi_k z_{kjt}(l)\big)\, dF_{kft}(l), \tag{5}$$

where ψ_k is a coefficient of absolute risk aversion, z_{kjt} is the payoff associated with realization of health state l, and F_{kft} is the distribution of health states faced if the plan belongs to insurer $f(j)$. The payoff associated with health state realization l is given by

$$z_{kjt}(l) = -p_{kjt} + \alpha^{OOP}\big(b^*_{jt}(l, \omega_k) - c^*_{jt}(l, \omega_k)\big) + \delta^{f(j)}_{kj} + \gamma^{inertia}_{kjt} + \beta \mathbf{X}_{kjt} + \sigma_\epsilon \epsilon_{kjt}, \tag{6}$$

[15] Note that c_{jt} is indexed by t because cost-sharing parameters vary within a plan across years. It also varies by household type (individual versus family), but we omit an additional index to save on notation. With a linear out-of-pocket cost function with coinsurance rate c and nonnegative health states, $m^* = \omega(1 - c) + l$ and $b^* = \frac{\omega}{2}(1 - c^2)$. Online Appendix C.2 provides solutions when contracts are piecewise linear and negative health states are permitted.

[16] The model predicts, for example, that if a consumer realizes a health state just under the deductible, she will take advantage of the proximity to cheaper health care and consume a bit more (putting her into the coinsurance region). Online Appendix Figure A.2 provides a depiction of optimal spending behavior predicted by this model.

where p_{kjt} is the household's plan premium (net of the employer contribution); $b^*_{jt}(l, \omega_k) - c^*_{jt}(l, \omega_k)$ is the payoff from optimal utilization measured in units of out-of-pocket dollars; $\delta^{f(j)}_{kj}$ are insurer fixed effects interacted with household observables, $\gamma^{inertia}_{kjt}$ is a set of fixed effects for both the plan and the insurer a household was enrolled in the previous year, interacted with household observables; and $\mathbf{X}_{kjt}$ is a set of additional covariates that can affect household utility.[17] The payoff z_{kjt} is measured in units of premium dollars. Out-of-pocket costs may be valued differently from premiums through parameter α^{OOP}. Finally, ϵ_{kjt} represents a household-plan-year idiosyncratic shock, with magnitude σ_ϵ to be estimated. We assume these shocks are independent and distributed type I extreme value, and that households chose the plan that maximized expected utility from among the set of plans $\mathcal{J}_{kt}$ available to them: $j^*_{kt} = \operatorname{argmax}_{j \in \mathcal{J}_{kt}} U_{kjt}$.

Distribution of Health States.—We assume that individuals face a log normal distribution of health states, and that households face the sum of health state draws across all individuals in the household. Because there is no closed-form expression for the distribution of the sum of draws from log normal distributions, we represent a household's distribution of health states using a log normal that approximates. We derive the parameters of the approximating distribution using the Fenton-Wilkinson method. This novel means of modeling the household distribution of health states allows us to fully exploit the large amount of heterogeneity in household composition that exists in our data. It also allows us to closely fit observed spending distributions using a smaller number of parameters than would be required if demographic covariates were aggregated to the household level. Our method is to estimate individuals' health state distributions, allowing parameters to vary with individual-level demographics. Online Appendix C.1 provides additional details.

An individual i faces uncertain health state $\tilde{l}^i$, which has a shifted log normal distribution with support $(-\kappa_{it}, \infty)$:

$$\log\left(\tilde{l}^i + \kappa_{it}\right) \sim N\left(\mu_{it}, \sigma^2_{it}\right).$$

The shift is included to capture a mass of individuals with zero spending. If κ_{it} is positive, negative health states are permitted, which may imply zero spending. Parameters μ_{it}, σ_{it}, and κ_{it} are in turn projected onto individual demographics (such as health risk score), which can vary over time.

A household k faces uncertain health state $\tilde{l}$, which has a shifted log normal distribution with support $(-\kappa_{kt}, \infty)$: $\log\left(\tilde{l} + \kappa_{kt}\right) \sim N\left(\mu_{kt}, \sigma^2_{kt}\right)$. Under the approximation, household-level parameters μ_{kt}, σ_{kt}, and κ_{kt} are a function of individual-level parameters μ_{it}, σ_{it}, and κ_{it}. Variation in μ_{kt}, σ_{kt}, and κ_{kt} across households, as well as within households over time, arises from variation in household composition: the number of individuals and each individual's demographics. In addition to this observable heterogeneity, we incorporate unobserved heterogeneity in household

[17] $\mathbf{X}_{kjt}$ includes HRA or HSA contributions, HA_{kjt}; vision and dental plan contributions, VD_{kjt}; and a fixed effect $\nu^{NarrowNet}_{jt}$ for one plan (Moda–2) that had a narrow provider network in 2011 and 2012. The associated parameters for health account and vision/dental contributions are α^{HA} and α^{VD}, respectively.

health though parameter μ_{kt}. Households can in this way hold private information about their health that can drive both plan choices and spending outcomes.

Finally, we introduce an additional set of parameters ϕ_f to serve as "exchange rates" for monetary health states across insurers. These parameters are intended to capture differences in total health-care spending that are driven by differences in provider prices across insurers, conditional on health state.[18] For example, the same physician office visit might lead to different amounts of total spending across insurers simply because each insurer paid the physician a different price. We do not want such variation to be attributed to differences in underlying health. Our approach is to estimate insurer-level parameters that multiply realized health states, transforming them from underlying "quantities" of health-care utilization into the monetary spending amounts we observe in the claims data. We model a household's money-metric health state l as the product of an insurer-level "price" multiplier ϕ_f and the underlying "quantity" health state $\tilde{l}$, where $\tilde{l}$ is log normally distributed depending only on household characteristics. Taken together, the distribution F_{kft} is defined by

$$l = \phi_f \tilde{l},$$
$$\log\left(\tilde{l} + \kappa_{kt}\right) \sim N\left(\mu_{kt}, \sigma^2_{kt}\right).$$

B. *Identification*

Our aim is to recover the joint distribution across households of willingness to pay, risk protection, and the social cost of moral hazard associated with different levels of coverage. Variation in these objects arises from variation in either household preferences (the risk-aversion and moral-hazard parameters) or in households' distributions of health states. Our primary identification concerns are (i) distinguishing preferences from private information about health, (ii) distinguishing taste for out-of-pocket spending (α^{OOP}) from risk aversion, and (iii) identifying heterogeneity in the risk-aversion and moral-hazard parameters. We provide informal identification arguments addressing each concern.

We first explain how ω, which captures moral hazard, is distinguished from unobserved variation in μ_{kt}, which captures adverse selection on unobservables. In the data, there is a strong positive correlation between plan generosity and total health-care spending (see Figure 2 panel A). A large part of this relationship can be explained by observable household characteristics, but even conditional on observables, there is still residual positive correlation. This residual correlation could be attributable to either the effect of lower out-of-pocket prices driving utilization (moral hazard) or private information about health affecting both utilization and coverage choice (adverse selection). The key to distinguishing between these explanations is the variation in plan menus.

Both within and across school districts, we observe similar households facing different menus of plans. As a result, some households are more likely to choose higher

[18] Provider prices are a well-documented source of heterogeneity in total health-care spending across insurers (Cooper et al. 2018), and these differences are often modeled to be linear in utilization (Gowrisankaran, Nevo, and Town 2015; Ghili 2016; Ho and Lee 2017; Liebman 2018). Of course, ϕ_f may also capture other differences across insurers, such as care management protocols or provider practice patterns.

coverage only because of the plan menu they are offered. The amount of moral hazard is identified by the extent to which households facing more generous plan menus also have higher health-care spending. On the other hand, we also observe similar households facing similar menus of plans, but still making different plan choices. This variation identifies the degree of private information about health, as well as the magnitude of the idiosyncratic shock σ_ϵ. Conditional on observables and the predicted effects of moral hazard, if households that inexplicably choose more generous coverage also inexplicably realize higher health-care spending, this variation in plan choice will be attributed to private information about health. Any residual unexplained variation in plan choice will be attributed to the idiosyncratic shock.

Both risk aversion (ψ) and the relative valuation of premiums and out-of-pocket spending (α^{OOP}) affect households' taste for higher coverage, but do not affect health-care spending. To distinguish between them, we rely on cases in which observably different households face similar plan menus. Risk aversion is identified by the degree to which households' taste for higher coverage is positively related to uncertainty in out-of-pocket spending, holding expected out-of-pocket spending fixed. The parameter α^{OOP} is identified by the rate at which households trade off premiums with expected out-of-pocket spending, holding uncertainty in out-of-pocket spending fixed.

Unlike the preceding arguments, identification of unobserved heterogeneity in risk aversion and the moral hazard parameter relies on the panel nature of our data. Plan menus, household characteristics, and plan characteristics change over time. We therefore observe the same households making choices under different circumstances. If we had a large number of observations for each household and sufficient variation in circumstances, the preceding arguments could be applied household by household, and we could nonparametrically identify the distribution of ψ and ω in the population. In reality, we have at most five observations for each household. We ask less of this data by assuming that the unobserved heterogeneity is normally distributed. The variance and covariance of the unobserved components of household types are identified by the extent to which different households consistently act in different ways. For example, if some households consistently make choices that reflect high risk aversion and other (observationally equivalent) households consistently make choices that reflect low risk aversion, this will be interpreted as unobserved heterogeneity in risk aversion.

C. *Estimation*

We project the parameters of the individual health state distributions μ_{it}, σ_{it}, and κ_{it} onto time-varying individual demographics:

$$\begin{aligned} \mu_{it} &= \boldsymbol{\beta}^{\mu}\mathbf{X}^{\mu}_{it}, \\ \sigma_{it} &= \boldsymbol{\beta}^{\sigma}\mathbf{X}^{\sigma}_{it}, \\ \kappa_{it} &= \boldsymbol{\beta}^{\kappa}\mathbf{X}^{\kappa}_{it}. \end{aligned} \tag{7}$$

Covariate vectors $\mathbf{X}^{\mu}_{it}, \mathbf{X}^{\sigma}_{it}$, and $\mathbf{X}^{\kappa}_{it}$ contain indicators for the zeroth to thirtieth, thirtieth to sixtieth, sixtieth to ninetieth, and ninetieth to hundredth percentiles of individual health risk scores each year; $\mathbf{X}^{\mu}_{it}$ and $\mathbf{X}^{\kappa}_{it}$ also contain a linear term in risk score, separately for each percentile group; and $\mathbf{X}_{it}^{\mu}$ also contains an indicator for whether the individual is a female between the ages of 18 and 35 and for whether the individual is under 18 years old.

Using the derivations shown in online Appendix C.1, the parameters of households' health state distributions are a function of individual-level parameters:

$$
\begin{aligned}
(8)\quad \sigma_{kt}^2 &= \log\left[1+\left[\sum_{i\in\mathcal{I}_k}\exp\left(\mu_{it}+\frac{\sigma_{it}^2}{2}\right)\right]^{-2}\sum_{i\in\mathcal{I}_k}\left(\exp\left(\sigma_{it}^2\right)-1\right)\exp\left(2\,\mu_{it}+\sigma_{it}^2\right)\right],\\
\bar{\mu}_{kt} &= -\frac{\sigma_{kt}^2}{2}+\log\left[\sum_{i\in\mathcal{I}_k}\exp\left(\mu_{it}+\frac{\sigma_{it}^2}{2}\right)\right],\\
\kappa_{kt} &= \sum_{i\in\mathcal{I}_k}\kappa_{it},
\end{aligned}
$$

where $\mathcal{I}_k$ represents the set of individuals in household k. Private information about health is reflected in normally distributed unobservable heterogeneity in μ_{kt}. The household-specific mean of μ_{kt} is given by $\bar{\mu}_{kt}$, and its variance is given by σ_{μ}^2. A large σ_{μ}^2 means that households appear to have substantially more information about their health than the econometrician.

We assume that μ_{kt}, ω_k, and $\log(\psi_k)$ are jointly normally distributed:

$$
(9)\qquad \begin{bmatrix}\mu_{kt}\\ \omega_k\\ \log(\psi_k)\end{bmatrix} \sim N\left(\begin{bmatrix}\bar{\mu}_{kt}\\ \boldsymbol{\beta}^{\omega}\mathbf{X}_k^{\omega}\\ \boldsymbol{\beta}^{\psi}\mathbf{X}_k^{\psi}\end{bmatrix},\begin{bmatrix}\sigma_{\mu}^2 & & \\ \sigma_{\omega,\mu}^2 & \sigma_{\omega}^2 & \\ \sigma_{\psi,\mu}^2 & \sigma_{\omega,\psi}^2 & \sigma_{\psi}^2\end{bmatrix}\right).
$$

There is both observed (through the mean vector) and unobserved (through the covariance matrix) heterogeneity in each parameter. Covariates $\mathbf{X}_k^{\omega}$ and $\mathbf{X}_k^{\psi}$ include an indicator for whether the household has children and a constant.[19]

We model inertia at both the plan and insurer level: $\gamma_{kjt}^{inertia} = \gamma_k^{plan}\mathbf{1}_{kt,j=j(t-1)} + \gamma_k^{ins}\mathbf{1}_{kt,f=f(t-1)}$. We allow γ_k^{plan} to vary by whether a household has children. To capture whether sicker households face higher barriers to switching insurers (and therefore provider networks), we allow γ_k^{ins} to vary linearly with household risk score. Insurer fixed effects $\delta_k^{f(j)}$ can vary by household age and whether a household has children, and we allow the intercepts to vary by geographic region in order to capture the relative attractiveness of insurer provider networks across different parts of the state (as well as other sources of geographical heterogeneity in insurer preferences).[20] We normalize the fixed effect for Moda to be zero. As the parameters of individual health state distributions can vary freely, the "provider price" parameters require normalization: ϕ_{Moda} is normalized to one.

We estimate the model via maximum likelihood. Our estimation approach follows Revelt and Train (1998) and Train (2009), with the distinction that we model a

[19] Household-level covariates are fixed over time as follows. If a household has children in some years but not others, we assign it to its modal status. Household risk score is calculated as the mean risk score of all individuals in a household across all years. Household age is the mean age of all adults across all years.

[20] We divide the state into three regions, based on groups of adjacent hospital referral regions (HRRs): the Portland and Salem HRRs in northwest Oregon (containing 55 percent of households); the Eugene and Medford HRRs in southwest Oregon (32 percent of households); and the Bend, Spokane, and Boise HRRs in eastern Oregon (13 percent of households) (*Dartmouth Atlas Project* 2020). For more information and HRR maps, see http://www.dartmouthatlas.org/data/region.

discrete/continuous choice. Our construction of the discrete/continuous likelihood function follows Dubin and McFadden (1984). The likelihood function for a given household is the conditional density of its observed sequence of total health-care spending, given its observed sequence of plan choices. We use Gaussian quadrature to approximate the jointly normal distribution of unobserved heterogeneity, as well as to approximate the log normal distributions of household health states. Additional details on the estimation procedure are provided in online Appendix C.2.

IV. Results

A. *Model Estimates*

Table 3 presents our parameter estimates. Column 3 presents our primary specification, as described in the previous section. Columns 1 and 2 present simpler specifications that are useful in understanding and validating the model. The table excludes insurer fixed effects and health state distribution parameters; these can be found in online Appendix Table A.8.

Column 1 presents a version of the model in which there is no moral hazard and no observable heterogeneity in individuals' health; that is, ω is fixed at zero, and we do not allow μ_{it}, σ_{it}, or κ_{it} to vary with individual demographics. Unobservable heterogeneity in household health (through σ_{μ}) is still permitted. In column 2, we introduce observable heterogeneity in health. A key difference across columns 1 and 2 is the magnitude of the adverse selection parameter σ_{μ}, which falls by more than half. When rich observable heterogeneity in health is introduced to the model, the estimated amount of unobservable heterogeneity in health falls substantially. In column 3, we introduced moral hazard. Here, an important difference is the increase in the estimated amount of risk aversion. With moral hazard as an available explanation, the model can explain a larger part of the dispersion in spending for observably similar households. This implies that households are facing less uncertainty in their health state than previously thought, and that more risk aversion is necessary to explain the same plan choices. Because estimated risk aversion increases, the relative valuation of premiums and out-of-pocket costs (α^{OOP}) falls.

Using column 3, we estimate an average moral hazard parameter ω of \$1,001 among individual households and \$1,478 among families.[21] Recall that ω represents the additional total spending induced by lowering marginal out-of-pocket cost from one to zero. Our estimates imply that moving a household from a plan where their health state was below the deductible to a plan where their health state would put them past the out-of-pocket maximum would increase total spending by 15.8 percent of mean spending for individuals and 11.4 percent for families.

Our estimates imply a mean (median) coefficient of absolute risk aversion of 0.92 (0.85).[22] Put differently, to make households indifferent between (i) a payoff of zero and (ii) an equal-odds gamble between gaining \$100 and losing \$X, the

[21] For comparison, the average ω estimated by Einav et al. (2013) is \$1,330, in a sample of households with average total health-care spending of \$5,283. In our sample, average total spending is \$6,339 for individuals and \$12,954 for families.

[22] Note that we measure monetary variables in thousands of dollars. Dividing our estimated coefficients of absolute risk aversion by 1,000 makes them comparable to estimates that use risk measured in dollars.

Table 3—Parameter Estimates

Variable	(1) Parameter	(1) SE	(2) Parameter	(2) SE	(3) Parameter	(3) SE
Employee premium ($000s)	−1.000[a]		−1.000[a]		−1.000[a]	
Out-of-pocket spending, $-\alpha^{OOP}$	−1.628	0.023	−1.661	0.024	−1.469	0.019
HRA/HSA contributions, α^{HA}	0.255	0.021	0.259	0.020	0.259	0.020
Vision/dental contributions, α^{VD}	1.341	0.024	1.302	0.022	1.209	0.021
Plan inertia intercept, γ^{plan}	4.763	0.060	4.431	0.056	4.630	0.063
Plan inertia × 1[Children], γ^{plan}	−0.129	0.039	−0.102	0.037	−0.138	0.038
Insurer inertia intercept, γ^{ins}	2.605	0.107	2.509	0.102	2.413	0.097
Insurer inertia × risk score, γ^{ins}	−0.074	0.083	−0.120	0.078	−0.037	0.080
Narrow network plan, $\nu^{NarrowNet}$	−2.440	0.155	−2.286	0.145	−2.334	0.151
Providence utiliz. multiplier, ϕ_P	1.022	0.018	1.072	0.017	1.063	0.002
Risk aversion intercept, β^{ψ}	−0.706	0.046	−1.018	0.059	−0.251	0.052
Risk aversion × 1[Children], β^{ψ}	0.005	0.031	−0.367	0.083	−0.361	0.050
Moral hazard intercept, β^{ω}					1.028	0.038
Moral hazard × 1[Children], β^{ω}					0.671	0.008
Std. dev. of private health info., σ_μ	0.683	0.002	0.331	0.064	0.225	0.005
Std. dev. of log risk aversion, σ_ψ	0.701	0.062	1.140	0.012	0.833	0.021
Std. dev. of moral hazard, σ_ω					0.281	0.013
Corr(μ, ψ), $\rho_{\mu,\psi}$	0.130	0.018	−0.365	0.049	0.227	0.005
Corr(ψ, ω), $\rho_{\psi,\omega}$					−0.137	0.042
Corr(μ, ω), $\rho_{\mu,\omega}$					0.062	0.017
Scale of idiosyncratic shock, σ_ϵ	2.313	0.025	2.160	0.023	2.116	0.024
Insurer × {Region, Age, 1[Child.]}	Yes		Yes		Yes	
Observable heterogeneity in health			Yes		Yes	
Number of observations	451,268		451,268		451,268	

Notes: The table presents estimates for selected parameters; online Appendix Table A.8 presents the remaining estimates. Standard errors are derived from the analytical Hessian of the likelihood function. Column 3 presents our primary estimates, while columns 1 and 2 present alternative specifications. All models are estimated on an unbalanced panel of 34,606 households, 11 plans, and 5 years. The utilization multiplier for Moda (ϕ_M) is normalized to one.
[a]By normalization.

mean (median) value of $X in our population is $91.7 ($92.1).[23] We note, however, that our estimates of risk aversion are with respect to both financial risk *and* risk in the value derived from health-care utilization (through b^*_{jt}), so they are not directly comparable to estimates that consider only financial risk. The standard deviation of the uncertain portion of payoffs $\left(b^*_{jt} - c^*_{jt}\right)$ with respect to the distribution of health states is $1,152 on average across household-plan-years. The standard deviation of out-of-pocket costs alone $\left(c^*_{jt}\right)$ is on average $1,280. To avoid a normally distributed lottery with mean zero and standard deviation $1,152 ($1,280), the median household would be willing to pay $489 ($544).

The importance of unobserved heterogeneity varies for health, risk aversion, and moral hazard.[24] Once we account for the full set of household observables and moral hazard, the estimated amount of private information about health is fairly small: Unobserved heterogeneity in μ_{kt} accounts for only 11 percent of the total

[23] A risk-neutral household would have $X equal to $100, and an infinitely risk-averse household would have $X equal to $0. Using the same example, Handel (2013) reports a mean $X of $91.0, Einav et al. (2013) report a mean $X of $84.0, and Cohen and Einav (2007) report a mean $X of $76.5.

[24] Following Revelt and Train (2001), we derive each household's posterior type distribution using Bayes' rule, conditioning on their observed choices and the population distribution. For the purposes of examining total variation in types across households (accounting for both observed and unobserved heterogeneity), we assign each household the expectation of their type with respect to their posterior distribution. This procedure is described in detail in online Appendix C.3.

variation in μ_{kt} across household-years. On the other hand, unobserved heterogeneity in risk aversion accounts for 93 percent of its total variation across households. Unobserved heterogeneity in the moral hazard parameter accounts for 18 percent of its total variation.

Conditional on observables, we find that households that are idiosyncratically risk averse are also idiosyncratically less prone to moral hazard $(\rho_{\psi,\omega} < 0)$ and also tend to have private information that they are unhealthy $(\rho_{\mu,\psi} > 0)$. We find that households with private information that they are unhealthy are also idiosyncratically more prone to moral hazard $(\rho_{\mu,\omega} > 0)$. Accounting for both unobservable *and* observable variation, our estimates imply that households' expected health state $E[\tilde{l}]$ has a correlation of 0.22 with risk aversion, and a correlation of 0.25 with the moral hazard parameter. The risk aversion and moral hazard parameters have a correlation of only 0.01. Online Appendix Figure A.3 plots the full joint distribution of these three key dimensions of household type.

Our estimates imply substantial disutility from switching insurer or plan. The average disutility from switching insurer is \$2,408, and from switching plans (but not insurers) is \$4,562. We estimate that insurer inertia is decreasing in household risk score, and that plan inertia is on average \$138 lower for households with children. The exceptionally large magnitudes of our inertia coefficients reflect, in large part, the infrequency with which households switch plan or insurer, as shown in Table 2. Only 3 percent of household-years ever voluntarily switch insurer, and only 20 percent of household-years ever voluntarily switch plan.

Finally, the estimates in column 3 indicate that households weight out-of-pocket expenditures 46.9 percent more than plan premiums. We believe this could be driven by a variety of factors, including (i) household premiums are tax deductible, while out-of-pocket expenditures are not; and (ii) employee premiums are very low (at the median, zero), perhaps rendering potential out-of-pocket costs in the thousands of dollars relatively more salient. We also find that households value a dollar in HSA/HRA contributions on average 75 percent less than a dollar of premiums. This is consistent with substantial hassle costs associated with these types of accounts, as documented by Reed et al. (2009) and McManus et al. (2006).

Model Fit.—We conduct two procedures to evaluate model fit, corresponding to the two stages of the model. First, we compare households' predicted plan choices with those observed in the data. Figure 3 displays the predicted and observed market shares for each plan, pooled across all years in our sample. Shares are matched exactly at the insurer level due to the presence of insurer fixed effects, but are not matched exactly plan by plan. Predicted choice probabilities over plans within an insurer are driven by plan prices; inertia; and households' valuation of different levels of coverage through their expectation of out-of-pocket spending, their value of risk protection, and their value of health-care utilization. Given the relative inflexibility of the model with respect to household plan choice within an insurer, the fit is quite good.

Second, we compare the predicted distributions of households' total health-care spending to the distributions of total health-care spending observed in the data. In a given year, each household faces a predicted distribution over health states and, due to moral hazard, a corresponding plan-specific distribution of total health-care

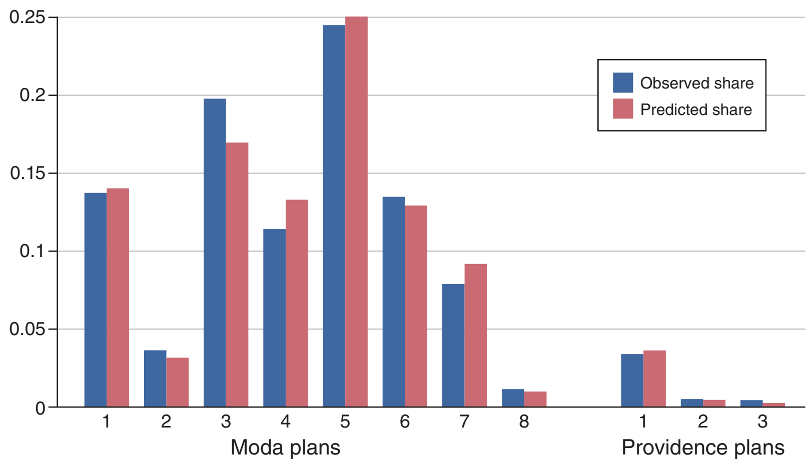

Figure 3. Model Fit: Plan Choices

Notes: The figure shows predicted and observed market shares at the plan level. All years are pooled, so an observation is a household-year. Predicted shares are calculated using the estimates in column 3 of Table 3 and online Appendix Table A.8.

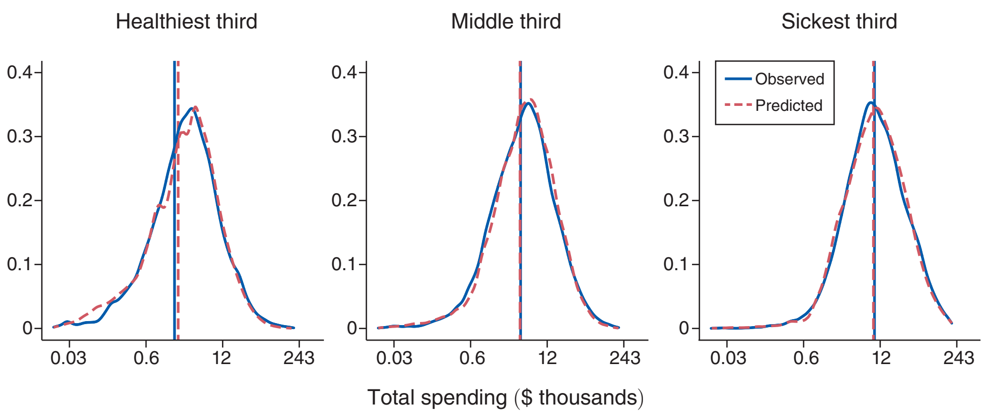

Figure 4. Model Fit: Health-Care Spending, by Tertile of Households by Risk Score

Notes: The figure shows kernel density plots of the predicted and observed distributions of total health-care spending on a log scale, separately by tertile of household health risk score, conditional on predicted/observed spending greater than $10. All years are pooled, so an observation is a household-year. Vertical lines represent the mean of the respective distribution. Predicted distributions are based on estimates from column 3 in Table 3 and online Appendix Table A.8. Overall, the observed probability of household spending less than $10 is 2.9 percent, and the predicted probability is 2.8 percent.

spending. To construct the predicted distribution of total spending in a population of households, we take a random draw from each household's predicted spending distribution corresponding to their chosen plan. Figure 4 presents kernel density plots of the predicted and observed distributions of total health-care spending. We assess fit separately by tertile of household risk score. Vertical lines in each plot represent the mean of the respective distribution. Overall, average total health-care spending is observed to be $11,689 and predicted to be $11,632. The standard deviation of total health-care spending is observed to be $20,803 and predicted to be $20,174.

The spending distributions fit well both overall and in subsamples of households, reflecting our flexible approach to modeling household health state distributions.

B. *Evaluating Vertical Choice*

We now construct the ingredients needed to evaluate the optimal plan menu: each household's willingness to pay for different levels of coverage, and the social surplus generated by allocating each household to different levels of coverage. We first specify the contracts under consideration.

Potential Contracts.—We consider concave, piecewise linear contracts that are vertically differentiated and well-ordered by coverage level.[25] While our numerical simulations consider all coverage levels between the null contract and full insurance, we limit attention in our graphical analysis to the range of coverage levels that are ultimately relevant given our parameter estimates. The lowest level of coverage we consider graphically is a contract with a deductible and out-of-pocket maximum of $10,000. The highest level of coverage remains full insurance. We begin by considering five contracts spanning this range, and refer to them as Catastrophic, Bronze, Silver, Gold, and full insurance. The contracts' actuarial values are 0.53, 0.61, 0.72, 0.86, and 1.00. Their out-of-pocket cost functions are depicted in online Appendix Figure A.4 panel (a).[26] We revisit the specification of potential contracts in Section IVC.

Willingness to Pay.—We make several simplifications to our empirical model in order to map it from the setting in Oregon back to our theoretical model, maintaining parameterizations and the estimated distribution of consumer types. To start, we put aside intertemporal variation in household health and focus on the first year each household appears in the data. We also use the provider price parameter $\phi = 1$ (corresponding to that of Moda). This leaves each household with a single type $\theta_k = \{F_k, \psi_k, \omega_k\}$, where F_k is a shifted log normal distribution described by parameters $\{\mu_k, \sigma_k, \kappa_k\}$. With respect to payoffs (equation (6)), we (i) hold all nonfinancial features fixed, so any insurer fixed effects cancel; (ii) suppose households choose from the new menu of contracts for the first time, removing any effects of inertia; (iii) set α^{OOP} to one so that premiums and out-of-pocket costs are valued one for one; and (iv) assume the idiosyncratic shock is not welfare relevant.[27]

[25] Coverage level ordering requires that contracts are well-ordered in the amount of risk protection provided. See online Appendix A.2 for a definition of coverage level ordering and the conditions on contracts that imply this ordering.

[26] The contracts' deductibles, coinsurance rates, and out-of-pocket maximums are respectively $10,000, –, $10,000 for Catastrophic; $5,846, 40%, $7,500 for Bronze; $3,182, 27%, $5,000 for Silver; and $1,125, 15%, $2,500 for Gold.

[27] As our model allows for rich heterogeneity in preferences over financially differentiated contracts, we are comfortable with the interpretation that any remaining determinants of plan choice contained in ϵ can be considered "mistake-making" (e.g., Handel and Kolstad 2015) or "monkey-on-the-shoulder tastes" (Akerlof and Shiller 2015), and so can be omitted from the social welfare calculation. In our counterfactuals, we suppose consumers have access to a tool that perfectly aids them in expressing their true preferences. Our question is whether, for this dimension of choice, such a tool is needed.

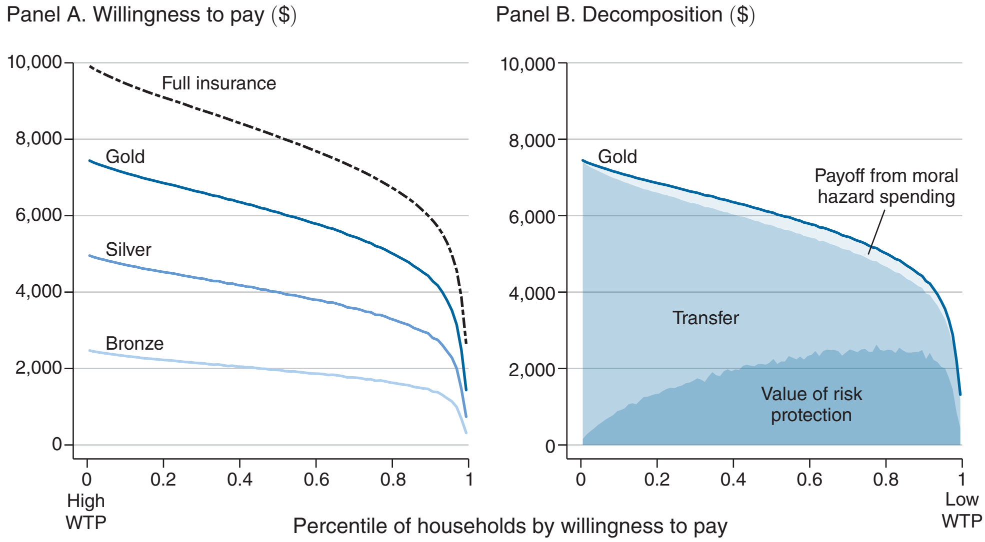

Figure 5. Willingness to Pay

Notes: The figure shows the distribution across households of (i) willingness to pay for each contract and (ii) the decomposition of willingness to pay for the Gold contract. The left panel consists of four connected binned scatterplots, with respect to 100 bins of households ordered by willingness to pay. The right panel consists of three connected binned scatterplots, with the area between each line shaded to indicate the component represented. Both willingness to pay and its components are measured relative to the Catastrophic contract.

With attention restricted to the dimension of coverage level, we can use equation (3) to express willingness to pay under our parameterization:[28]

$$WTP(x, \theta_k) = \underbrace{E_l\left[c_{x_0}(l) - c_x(l)\right]}_{\substack{\text{Expected reduction in out-of-pocket} \\ \text{cost holding behavior fixed}}} + \underbrace{E_l\left[\frac{\omega_k}{2}\left(1 - c_x'\left(m^*(l, \omega_k, x)\right)\right)^2\right]}_{\text{Expected payoff from moral hazard spending}} + \underbrace{\Psi(x, \theta_k)}_{\substack{\text{Value of risk} \\ \text{protection}}}.$$

As before, willingness to pay is composed of three parts: the "transfer" of expected out-of-pocket costs holding behavior fixed (at uninsured behavior), the expected payoff from moral hazard spending, and the value of risk protection. Recall that only the latter two components are relevant to social welfare.

Figure 5 presents the distribution of willingness to pay among family households.[29] Whereas our point of reference in the two-contract example was x_L, our reference contract now is the Catastrophic contract. We hereinafter refer to "willingness to pay" for a given contract, but emphasize that this is *marginal* willingness to pay with respect to this particular reference. Figure 5, as well as the figures that follow, is composed of connected binned scatterplots: households are ordered on the horizontal axis according to their willingness to pay, those at each percentile

[28] See online Appendix A.2 for the expression of the value of risk protection.

[29] We focus on family households because families make up 75 percent of the sample and because our set of potential contracts is chosen to mimic the coverage levels typically offered to families. Our results among individual households are qualitatively unchanged.

are binned together, and the average value of the vertical axis variable is plotted for each bin.[30] These 100 points are then connected with a line. The left panel shows the willingness to pay curves for our candidate contracts. As contracts are vertically differentiated, all households are willing to pay more for higher coverage. The highest willingness-to-pay households are willing to pay \$10,000 more for full insurance than for the Catastrophic contract. Online Appendix Figure A.5 provides demographic information about households across the distribution of willingness to pay. Higher willingness-to-pay households tend to be older, have more family members, be more risk averse, and most strikingly, have higher expected health-care spending.

The right panel shows, for one contract, the decomposition of willingness to pay. We find that the transfer represents the majority of willingness to pay for most households, but that this varies across the distribution of willingness to pay. For households with low willingness to pay, about one-half is made up by the transfer. For households with high willingness to pay, nearly all is made up by the transfer. The highest willingness-to-pay households are willing to pay \$7,500 more for Gold than for Catastrophic only in order to avoid paying an additional \$7,500 in out-of-pocket costs. Importantly, this means that allocating them to higher coverage generates almost no additional social surplus.

Social Surplus.—As in equation (3), the social surplus generated by allocating a household to a given contract is the difference between willingness to pay and expected insurer cost, which under our parameterization is equal to

$$(10) \qquad SS(x,\theta_k) \;=\; \underbrace{\Psi\left(x,\theta_k\right)}_{\text{Value of risk protection}} - \underbrace{E_l\left[\frac{\omega_k}{2}\left(1-c_x'\left(m^*\left(l,\omega_k,x\right)\right)\right)^2\right]}_{\text{Social cost of moral hazard}}.$$

The value of risk protection varies in the population to the extent there is variation in risk aversion and in the amount of uncertainty about out-of-pocket costs. The social cost of moral hazard varies in the population to the extent there is variation in the moral hazard parameter and in consumers' expected marginal out-of-pocket costs.

To understand the contribution of each of these components to the overall relationship between willingness to pay and social surplus, we first plot them separately. Figure 6 panel A shows the distribution across households of the marginal value of risk protection generated by each contract, relative to the Catastrophic contract. We find that the majority of the social welfare gains from more generous coverage are driven by households with intermediate levels of willingness to pay. This pattern is driven by the concavity of the contracts we consider. High willingness-to-pay households are likely to realize health states that put them above the out-of-pocket maximum of every contract, leaving them little uncertainty about out-of-pocket costs. Among the top fifth of households by willingness to pay, the probability of spending more than \$10,000, even without moral hazard, is 65 percent. Online Appendix

[30] Households are in fact ordered by willingness to pay for full insurance, but the ordering is nearly identical across contracts. The consistent willingness-to-pay ordering of households across contracts is what permits a graphical analysis of multiple contracts analogous to the two-contract example in Figure 1. See Geruso et al. (2019) for a detailed discussion of this point.

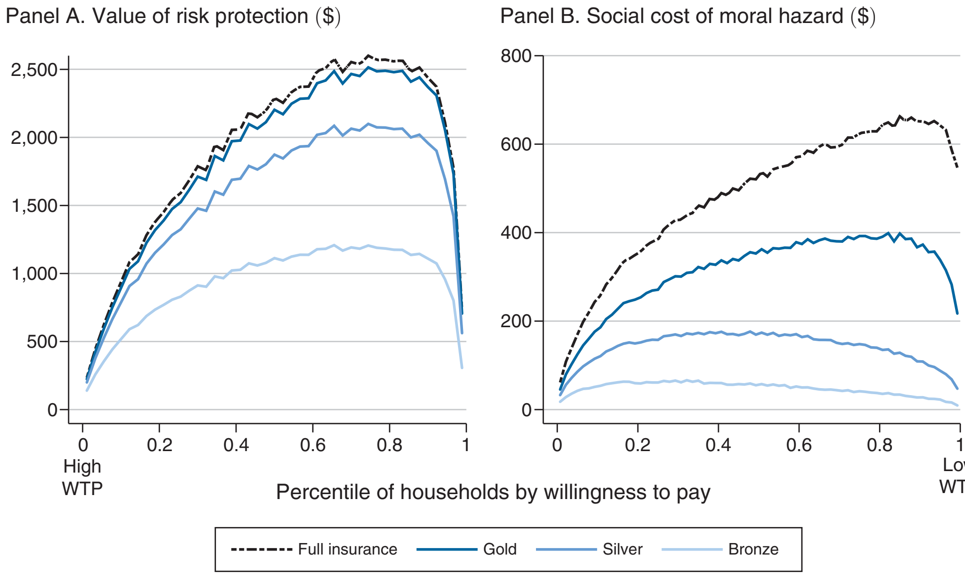

Figure 6. Components of Social Surplus

Notes: The figure shows the distribution across households of (i) the value of risk protection and (ii) the social cost of moral hazard, for each contract. Both are measured relative to the Catastrophic contract. Each panel is composed of four connected binned scatterplots, with respect to 50 (to reduced noise) bins of households ordered by willingness to pay.

Figure A.6 shows the spending distributions faced by households at different levels of willingness to pay. Variation in out-of-pocket uncertainty only becomes meaningful for households for whom much of the density of their spending distribution lies in the range \$0–\$10,000, within which marginal out-of-pocket cost varies across contracts.

Figure 6 panel B shows the distribution of the marginal social cost of moral hazard. It provides two important insights. First, high willingness-to-pay households on average barely change their behavior across this range of coverage levels.[31] For similar reasons as with risk protection, the majority of the social cost of more generous coverage is driven by households with lower willingness to pay. The second insight is that the Gold contract recovers about one-half of the social cost of moral hazard induced by full insurance. The \$1,125 deductible is high enough to deter excess spending, but low enough to sacrifice only a small amount of risk protection.

Finally, Figure 7 shows the resulting social surplus curves, equal to the vertical differences between the curves in Figure 6 panels A and B. The social surplus curves represent the average social surplus achieved by allocating all households at a given

[31] We note that this finding is closely related to the embedded assumption that moral hazard will not be expressed as long as end-of-year marginal out-of-pocket cost does not vary across contracts. While there is substantial empirical evidence that consumers do respond to spot prices (e.g., Aron-Dine et al. 2015; Dalton, Gowrisankaran, and Town 2020), here we do not find evidence of moral hazard among high-risk households (see online Appendix Table A.7). If the data did suggest a moral hazard response among these households, the model would load the effect onto the moral hazard parameter ω, compensating a weak treatment with a strong treatment effect.

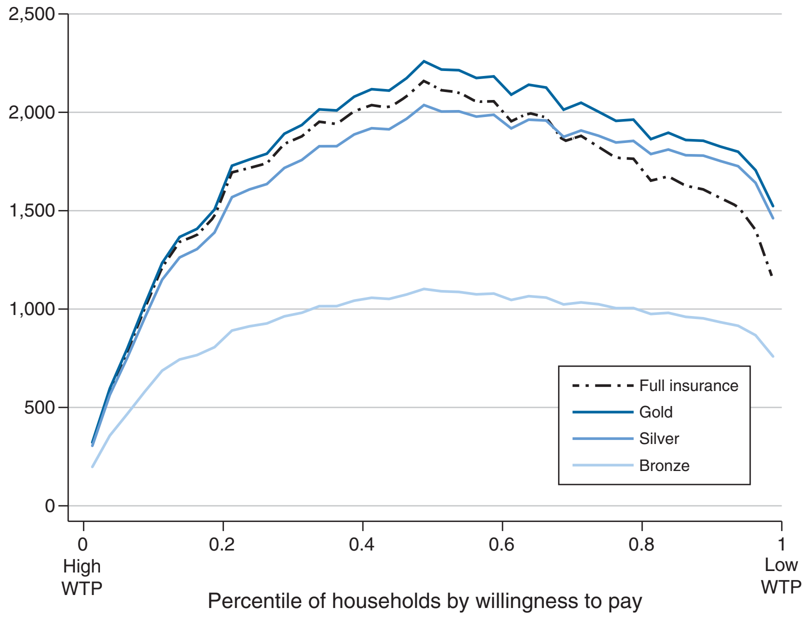

Figure 7. Social Surplus ($)

Notes: The figure shows the distribution across households of social surplus relative to the Catastrophic contract. The figure is composed of four connected binned scatterplots, with respect to 50 (to reduce noise) quantiles of households by willingness to pay.

percentile of willingness to pay to a given contract, relative to allocating them to the Catastrophic contract. Since households can be screened only by their willingness to pay, the figure permits a direct assessment of the optimal menu.

First, note that all curves lay everywhere above zero, meaning the Catastrophic contract (and any lower level of coverage) should be unambiguously excluded from the optimal menu. Any lower level of coverage can be ruled out because its social surplus curve will lay everywhere below that of the Catastrophic contract (c.f. Proposition 3 in online Appendix A.2). Among the remaining contracts, the social surplus curves of Bronze, Silver, and full insurance lay everywhere below that of the Gold contract, which delivers higher average social surplus at *every* level of willingness to pay. Households with higher willingness to pay should therefore not have a higher level of coverage than households with lower willingness to pay: they should, on average, have the same level of coverage.[32] It follows that offering choice over these contracts is not efficient in this population. Numerical optimization confirms this result. The optimal menu consists of only the Gold contract, and this allocation achieves social surplus (relative to allocating all households to Catastrophic) equal to the integral of the Gold social surplus curve: $1,514 per household.

[32] Although Gold is the efficient contract at every level of willingness to pay, it is not the efficient contract for every household. Online Appendix Figure A.7 shows the heterogeneity in households' efficient contracts.

C. *Robustness*

More Contracts.—A natural question is how the optimal menu would change if more contracts were available. Figure 7 indicates that the Silver contract is everywhere too little coverage, and that full insurance is everywhere too much coverage, but it says nothing about the potential gains of offering additional contracts within this range. We explore this question by expanding the number of contracts in the Silver-to-full insurance range from 1 (the Gold contract) to 20. The out-of-pocket cost functions for this denser set of potential contracts are depicted in online Appendix Figure A.4 panel (b).

We find that when efficient coverage level can be measured more finely, high willingness-to-pay households do have a slightly higher efficient level of coverage. In a small neighborhood of the Gold contract, it is therefore efficient to offer a choice. The optimal menu features four contracts.[33] This allocation achieves social surplus, relative to allocating all households to the Catastrophic contract, of $1,528 per household. This represents a gain of $14 over what is achieved by the Gold contract alone, and of only $5 over what can be achieved by a single contract in the denser set.

Different Contracts.—We next explore whether our results are robust to alternative contract designs. We have so far considered one particular design, as depicted in panels (a) and (b) of online Appendix Figure A.4. We now consider three alternatives, as depicted in panels (c)–(e). These are (c) removing deductibles, (d) removing the coinsurance region, and (e) extending the coinsurance region. Within each alternate contract design, we consider a set of five vertically differentiated contracts.

We solve for the optimal menu within each new set of contracts. These results are presented in online Appendix Table A.9. We find that among contracts without a deductible and without a coinsurance region, the optimal menu again features a single contract. For much the same reasons that this was true among the original contracts, higher willingness-to-pay consumers do not have a higher efficient level of coverage. We also find that among contracts with an extended coinsurance region, the optimal menu does feature vertical choice. Because it takes longer to reach the out-of-pocket maximum, households are less likely to hit it, and high willingness-to-pay households face much more variation in risk across contracts.

Our findings suggest that the contract dimension most relevant to the question of vertical choice is the stop-loss point, i.e., the level of total spending at which the out-of-pocket maximum is reached. Namely, our results suggest that if a regulator wanted consumers to pay on the margin for only a short time (a low stop-loss point), vertical choice may not offer meaningful welfare gains. If instead a regulator wanted consumers to pay on the margin for a long time (a high stop-loss point), our results suggest it may be useful to offer consumers a choice.[34]

[33] The four contracts are the Gold contract (actuarial value 0.86) and the three next-less-generous contracts (actuarial values 0.84, 0.83, and 0.81). At the optimal feasible allocation, 28 percent of households choose Gold; and 34 percent, 37 percent, and 1 percent of households choose the next three contracts respectively. The optimal single contract in the dense set is the 0.83 actuarial value contract.

[34] Evaluating a regulator's choice between these options is no longer a question of vertical choice. Though our estimates suggest that a lower stop-loss point is more efficient, we acknowledge that there are important

Different Consumers.—We next explore the robustness of our findings to different populations of consumers. We do this in two ways: (i) by re-estimating our model in the full sample of households that includes Kaiser enrollees, and (ii) by adjusting individual parameter estimates to establish their isolated effects on results.

As discussed in Section IIA, we exclude Kaiser enrollees from our primary analysis sample in order to focus on the vertical choice across coverage levels, as opposed to the horizontal choice across plan types (HMO versus PPO). Kaiser enrollees are on average slightly younger and healthier, and 3 percent of households did at one point switch between a Kaiser and non-Kaiser plan. We investigate how these factors impact our results by re-estimating our model using the full sample of households. Online Appendix Table A.10 presents these parameter estimates. Online Appendix Figure A.8 presents the corresponding willingness to pay and social surplus curves. Though the shapes and levels of the resulting social surplus curves are slightly different than under our original estimates, our qualitative results and the underlying mechanisms are unchanged. The optimal menu remains the Gold contract alone.

Second, we explore how specific perturbations of our parameter estimates affect our results. We explore nine cases, including raising and lowering the moral hazard parameter, raising and lowering risk aversion, and increasing heterogeneity in risk aversion and the moral hazard parameter. We also present three cases in which households vary only in their preferences: risk aversion and/or the moral hazard parameter. Our findings are summarized in online Appendix Table A.11.[35] For each case, the table shows the percent of households enrolled in each contract under the optimal menu. Intuitively, we find that the optimal menu is more likely to feature a choice when risk aversion plays a larger role in driving variation in willingness to pay. At the extreme, when households vary only in their risk aversion, nearly perfect screening is possible as private and social incentives are directly aligned.

Online Appendix Table A.11 also reports the welfare gains available from a denser contract space, and whether or not a choice would be efficient in that context. We find that while choice is almost always efficient in the denser contract space, the welfare gains available are consistently small (never exceeding \$16 per household per year). In the extreme case in which households vary only in their moral hazard parameter, private and social incentives are directly misaligned, and choice is not efficient even among the dense set of contracts. Across all nine cases, the welfare gains from vertical choice relative to what can be achieved by a single contract do not exceed \$10 per household. In a broad neighborhood of our parameter estimates, the efficiency loss from forgoing vertical choice is therefore either zero or economically small.

considerations our model may not capture. For example, consumers may inefficiently restrict utilization in response to even moderate marginal out-of-pocket costs (Aron-Dine et al. 2015; Dalton, Gowrisankaran, and Town 2020), which may pull in favor of a lower stop-loss point, or they may benefit from smoothing out-of-pocket spending within a year (Ericson and Sydnor 2018, Hong and Mommaerts 2021), which may pull in favor of a higher stop-loss point.

[35] We present fairly large perturbations, changing our estimates by a factor of two, in order to show cases in which our results do vary. Smaller changes to our parameter estimates, e.g., raising and lowering mean risk aversion by up to 30 percent, do not affect our results.

V. Counterfactual Pricing Policies

Returning to our focal set of metal-tier contracts and the estimated distribution of consumer types, we compare outcomes under five pricing policies: (i) regulated pricing with community rating, (ii) regulated pricing with type-specific prices, (iii) competitive pricing with community rating and a mandate, (iv) competitive pricing with type-specific prices and a mandate, and (v) premiums to support vertical choice. Under regulated pricing, premiums are set to maximize social surplus. Under competitive pricing, premiums are endogenous and must equal average costs on a plan-by-plan basis. A mandate enforces a minimum level of coverage at the Catastrophic contract. Under premiums to support vertical choice, premiums are set to support the availability of (read: enrollment in) every contract.

We consider two pricing policies, (ii) and (iv), in which premiums can vary by consumer attributes. If observable dimensions of household type are predictive of efficient coverage level, tailoring plan menus to observables may improve allocations. We divide households into four groups: childless households under age 45, childless households over age 45, households under age 45 with children, and households over age 45 with children.[36] We use age and whether the household has children because these are used in practice on ACA exchanges and are also important observables with which the parameters of our model may vary.

A. *Welfare Outcomes*

Table 4 summarizes outcomes under each of these five pricing policies. It shows the percent of households Q enrolled in each contract at the optimal feasible allocation, the percent of first-best social surplus achieved, and the expected per-household insurer cost AC among households enrolled in each contract (in thousands of dollars). We benchmark outcomes against the first best allocation of households to contracts (as depicted in online Appendix Figure A.7), which cannot be supported by prices unless premiums can vary by all aspects of consumer type. The first best allocation generates \$1,542 in social surplus per household, relative to allocating all households to the Catastrophic contract. Expected total health-care spending per household at the first best allocation is \$12,400, and expected insurer cost per household is \$10,351.

Policy (i) is the baseline policy considered in this paper, in which the regulator can set premiums but is restricted to community rating. As indicated by Figure 7, it is welfare maximizing to offer only the Gold contract.[37] Interestingly, although 25 percent of households are misallocated, this policy is almost equally as efficient as the first best allocation; that is, the ability to perfectly discriminate among

[36] Among family households, 6 percent are childless and under age 45, 27 percent are childless and over age 45, 52 percent have children and are under age 40, and 15 percent have children and are over age 45.

[37] This allocation is implementable because the regulator need not break even in aggregate. The Gold contract can be provided for free, and the deficit of \$10,619 per household can be funded by taxing incomes (here, at zero cost of public funds). We note that if the regulator *did* need to break even in aggregate, vertical choice would likely be efficient. The focus would shift to ensuring low willingness-to-pay consumers were not left out of the market entirely, even if that induced some high willingness-to-pay consumers to select lower-than-efficient coverage. See Azevedo and Gottlieb (2017) and Geruso et al. (2019) for a full treatment of a setting in which the regulator must break even in aggregate.

Table 4—Outcomes under Alternative Pricing Policies

	Policy	Percent of first best surplus		Potential contracts: Full	Gold	Silver	Bronze	Ctstr.
*	First best	1.000	*Q:*	0.06	0.75	0.19	< 0.01	–
			AC:	18.35	9.43	11.48	39.18	–
(i)	Regulated pricing with community rating	0.982	*Q:*	–	1.00	–	–	–
			AC:	–	10.62	–	–	–
(ii)	Regulated pricing with type-specific prices	0.989	*Q:*	–	0.98	0.02	–	–
			AC:	–	10.71	0.75	–	–
(iii)	Competitive pricing with community rating	0.000	*Q:*	–	–	–	–	1.00
			AC:	–	–	–	–	6.30
(iv)	Competitive pricing with type-specific prices	0.075	*Q:*	–	–	0.05	–	0.95
			AC:	–	–	4.95	–	6.41
(v)	Premiums to support vertical choice	0.796	*Q:*	0.01	0.07	0.63	0.28	0.01
			AC:	61.04	31.91	8.47	1.75	0.28

Notes: The table summarizes outcomes under five pricing policies as well as the first best allocation, among the 25,636 family households. *Q* represents the percent of households enrolled in each contract. *AC* represents average expected insurer cost (in thousands of dollars) among households enrolled in each contract. Social surplus is measured relative to the Catastrophic contract. At the first best allocation, social surplus is $1,542 per household and expected insurer cost is $10,351 per household.

consumers would increase welfare by only $28 per household per year. Driving this result is the fact that social welfare is quite flat across the top contracts, and particularly so among the households who are misallocated under policy (i). Among these households, the social surplus at stake between the Silver and Gold contracts is on average only $26; among all households, it is $112.

Because pricing policy (i) is almost as efficient as the first best outcome, there is little scope for improvement by varying prices by consumer type. Even so, under policy (ii) we do find that allowing the regulator to discriminate can improve allocational efficiency by a small amount. To young households under with children, it is efficient to offer a choice between Gold and Silver. To the other three sets of households, it is still efficient to offer only Gold. It becomes possible to productively offer lower coverage to young households with children because the other households, to whom it is not efficient to provide such low coverage, can now be excluded.

Policy (iii) considers competitive pricing with community rating and a mandated minimum level of coverage at the Catastrophic contract. We calculate the competitive equilibrium proposed by Azevedo and Gottlieb (2017). We find that a separating equilibrium above minimum coverage cannot be supported in this population, and the market unravels. Though choice is permitted, the market cannot deliver it. Policy (iv) allows the market to be segmented. We find that among young households with children, a separating equilibrium between the Silver and Catastrophic contracts can be supported. The other three market segments unravel.

The first four policies are natural benchmarks, but none turn out to feature the same degree of vertical choice that is observed in many US health insurance markets, including the market we study. A major difference between these real markets and our benchmark policies is that the former feature a complex set of taxes and subsidies that affect consumer premiums in ways not replicated by our benchmark policies. To mimic this status quo outcome, policy (v) implements premiums that can support enrollment in every contract. We target enrollment shares that match the true metal-tier shares observed on ACA exchanges in 2018.[38] Because households with intermediate willingness to pay (for whom social surplus increases steeply at low coverage levels; see Figure 7) now choose Silver instead of Bronze or Catastrophic, this allocation substantially increases welfare relative to the competitive outcome.

B. *Distributional Outcomes*

The population faces an unavoidable health-care spending bill of \$11,723 per household. It is unavoidable because it arises even if all households have the minimum allowable coverage (Catastrophic). While full insurance provides the benefit of additional risk protection, it also raises the population's health-care spending bill by 8 percent due to moral hazard, to \$12,695 per household.

The spending bill is funded by a combination of out-of-pocket costs and insured costs. Insured costs are in turn funded by premiums and, to the extent optimal premiums imply an aggregate deficit, by taxes. Different coverage levels imply different divisions between out-of-pocket costs and insured costs. For example, if all households had Catastrophic coverage, in expectation 47 percent of the spending bill would be paid out-of-pocket, and 53 percent would be insured. If all households had full insurance, 100 percent of the spending bill would be insured. There are therefore large differences across policies in the source of funding for the population health-care spending bill, and thereby in how evenly the spending bill is shared in the population.

Figure 8 shows distributional outcomes under three of our candidate policies: (i) regulated pricing ("All Gold"), (iii) competitive pricing ("All Catastrophic"), and (v) premiums to support vertical choice ("Vertical choice"). The left panel shows the distribution of health-care spending bills across households. Each household's health-care spending bill equals the premium plus expected out-of-pocket cost associated with their chosen contract, plus any tax assessed on all consumers. For simplicity (and because we lack information on income), we assess taxes equally across households. Households are again ordered on the horizontal axis according to their willingness to pay. Under "All Catastrophic," there is a premium-plus-tax of \$6,298. The highest willingness-to-pay households then also pay an expected out-of-pocket cost of \$9,708, implying a health-care spending bill of \$16,006. The lowest willingness-to-pay households pay an expected out-of-pocket cost of only \$1,500, implying a health-care spending bill of only \$7,798. When the population has higher coverage, as under the other two pricing policies, the health-care spending bill is shared more evenly in the population.

The right panel evaluates the distribution of consumer surplus, incorporating preferences over risk and health-care utilization in addition to just spending outcomes. In typical markets, consumer surplus is measured relative to the absence of a product. As we enforce a minimum level of coverage, here it is measured relative to the absence of a better product. Under each policy, consumer surplus is the difference between a household's marginal willingness to pay for their chosen

[38] Shares are from Kaiser Family Foundation (2018) and are available at https://www.kff.org/health-reform/state-indicator/marketplace-plan-selections-by-metal-level. We map Platinum coverage to full insurance. Premiums that can support these shares are \$7,059 for full insurance, \$4,594 for Gold, \$2,173 for Silver, \$375 for Bronze, and \$0 for Catastrophic; resulting in an aggregate deficit of \$6,856 per household.

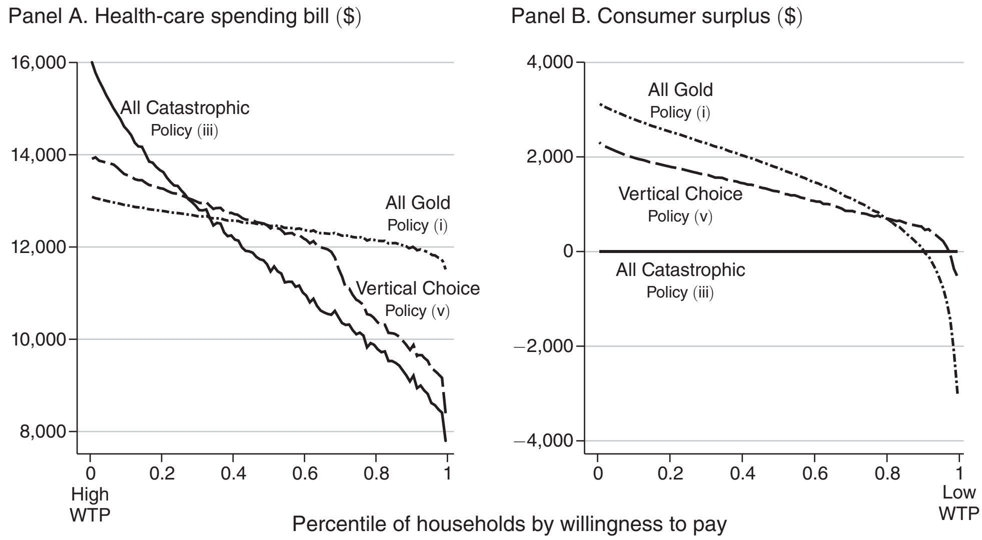

Figure 8. Distributional Outcomes

Notes: The figure shows the distribution across households of (i) expected health-care spending bill (premium plus tax plus expected out-of-pocket cost), and (ii) consumer surplus, under three of the policies considered in Table 4. Households are arranged on the horizontal axis according to their willingness to pay. As there may be variation in expected out-of-pocket costs among households with the same willingness to pay, panel A is composed of connected binned scatterplots, as in the previous figures. Consumer surplus equals marginal willingness to pay less marginal premium-plus-tax, relative to the allocation of all households to Catastrophic coverage.

plan, and the marginal premium-plus-tax associated with that choice. The integral of each consumer surplus curve equals the social welfare generated by that policy, relative to the "All Catastrophic" outcome. The difference between the "All Gold" consumer surplus curve in Figure 8 panel B and the Gold contract's social surplus curve in Figure 7 is that the former shows who receives the surplus, while the latter shows who generates it. The integrals of the two curves are the same.

Figure 8 panel B depicts a classic feature of insurance markets with adverse selection. The optimal feasible allocation ("All Gold") results in higher coverage and greater social welfare gains, while the competitive outcome ("All Catastrophic") results in lower coverage but a more even distribution of welfare gains. At the competitive outcome, no one is made worse off than they were absent the market. Regulatory intervention can offer substantial efficiency gains, at the cost of making some households worse off.

Dynamic Considerations.—These static gains from trade, and the distribution thereof, are evaluated at a point in time at which households are aware of their endowed type, θ. In the spirit of Hendren (2021) and Handel, Hendel, and Whinston (2017), we can also consider welfare from the perspective of an "unborn" consumer, who, prior to participating in our spot insurance market, faces a lottery over types.[39]

[39] The interpretation that consumers face a lottery over all elements of type, including preferences, is consistent with Harsanyi (1953, 1955). We take this approach because it permits a simple informal analysis, but refer the reader to Eden (2020) for an alternative potential approach.

Note that a consumer's type θ uniquely determines their willingness to pay, and thus their position on the horizontal axes of Figure 8. Instead of considering a lottery over types, we can therefore directly consider the lottery over levels of willingness to pay. Under each policy, the lottery over types faced by an unborn consumer is equivalent to the uniformly distributed lottery over consumer surplus outcomes shown in Figure 8 panel B.

The question then becomes where to normalize utility across consumers. In Figure 8 panel B, we have assumed consumers are equally well off absent the market. But from "behind the veil of ignorance," it may be more natural to assume they are equally well off when fully insured. Such a renormalization would be reflected in Figure 8 panel B by rotating the consumer surplus curves counterclockwise, until an "All full insurance" consumer surplus curve was horizontal (as depicted in online Appendix Figure A.9). In this case, it becomes clear that the "All Gold" policy delivers the least risky distribution of surplus in the population, consistent with the intuition that higher coverage provides greater dynamic risk protection (Handel, Hendel, and Whinston 2015). Among the three candidate policies, the "All Gold" policy therefore delivers the most efficiency *and* the most equity in the spot market.[40]

VI. Conclusion

This paper presents a framework for evaluating optimal menus of coverage levels in regulated health insurance markets. Our framework incorporates consumer heterogeneity along multiple dimensions, endogenous health-care utilization, and menus of nonlinear insurance contracts among which traded contracts are endogenous. We show how willingness to pay for insurance can be decomposed into a component that is only privately relevant and a component that is also socially relevant, the latter of which gives insurance value beyond as a redistributive tool. We emphasize how the privately relevant, redistributive component plays a central role in determining feasible allocations. When premiums must be uniform, it may not be possible to align the private incentive to maximize one's own transfer with the social incentive to mitigate residual uncertainty. The presence of moral hazard means that the problem is more complicated than simply mandating full insurance for all.

We show that the efficiency of vertical choice hinges on whether consumers with higher willingness to pay have higher efficient levels of coverage. In reverse, this condition implies that a lowest-coverage plan should only be offered if the lowest willingness-to-pay consumers are the intended recipients. In the setting we study, we find that the lowest willingness-to-pay consumers are sufficiently risk averse, and facing sufficient risk, to warrant coverage as least as high as the Silver contract. At the other end, our key condition implies that a highest-coverage plan should only be offered if the highest willingness-to-pay consumers are the intended recipients. The highest coverage we consider is full insurance, and we find that it would be more efficient for the highest willingness-to-pay consumers to have lower coverage. Between these bounds, we find that private values for higher coverage are not

[40] Given the chosen normalization, maximal equity is achieved by allocating all households to full insurance. Maximal (static) efficiency, meanwhile, is achieved by allocating all households to Gold. A regulator placing some weight on each of these objectives may want to offer a vertical choice between the Gold contract and full insurance.

strongly positively correlated with social values, and thus that offering a choice cannot provide economically meaningful welfare gains. We also find that the welfare stakes of misallocation are low. Relative to what can be achieved by a single contract, the ability to perfectly screen consumers would increase welfare by only $28 per household per year.

An important limitation of this paper is that we assume the socially optimal level of health-care utilization is the level a consumer would choose absent insurance. If health-care providers charge supracompetitive prices, or if there are positive externalities of health-care utilization, it may well be the case that using insurance to induce additional utilization is desirable. In addition, important considerations our model does not address arise when consumers face liquidity constraints (Ericson and Sydnor 2018) or are protected from large losses by limited liability in addition to by insurance (Gross and Notowidigdo 2011, Mahoney 2015). Finally, a central simplification in our model is that health care is a homogeneous good over which consumers choose only the quantity to consume. In reality, health care is multidimensional, and the time and space over which utilization decisions are made is complex. We see the extension of our model in these directions to be a fruitful direction for future research.

REFERENCES

Abaluck, Jason, and Jonathan Gruber. 2011. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review* 101 (4): 1180–210.

Abaluck, Jason, and Jonathan Gruber. 2016. "Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance." *American Economic Review* 106 (8): 2145–84.

Abaluck, Jason, and Jonathan Gruber. 2017. "Improving the Quality of Choices in Health Insurance Markets." NBER Working Paper 22917.

Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84 (3): 488–500.

Akerlof, George A., and Robert J. Shiller. 2015. *Phishing for Phools: The Economics of Manipulation and Deception*. Princeton: Princeton University Press.

Aron-Dine, Aviva, Liran Einav, Amy Finkelstein, and Mark Cullen. 2015. "Moral Hazard in Health Insurance: Do Dynamic Incentives Matter?" *Review of Economics and Statistics* 97 (4): 725–41.

Arrow, Kenneth J. 1965. "Uncertainty and the Welfare Economics of Medical Care: Reply (The Implications of Transaction Costs and Adjustment Lags)." *American Economic Review* 55 (1): 154–8.

Arrow, Kenneth J. 1968. "The Economics of Moral Hazard: Further Comment." *American Economic Review* 58 (3): 537–39.

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

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

Brot-Goldberg, Zarek C., Amitabh Chandra, Benjamin R. Handel, and Jonathan T. Kolstad. 2017. "What Does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics." *Quarterly Journal of Economics* 132 (3): 1261–318.

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

Bundorf, M. Kate, Maria Polyakova, Cheryl Stults, Amy Meehan, Roman Klimke, Ting Pun, Albert Solomon Chan, and Ming Tai-Seale. 2019. "Machine-Based Expert Recommendations and Insurance Choices among Medicare Part D Enrollees." *Health Affairs* 38 (3): 482–90.

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

Carlin, Caroline, and Robert Town. 2008. "Sponsored Health Plans." *ReVision*, 0–67.

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

Cooper, Zack, Stuart V. Craig, Martin Gaynor, and John Van Reenen. 2018. "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured." *Quarterly Journal of Economics* 134 (1): 51–107.

Cutler, David M., and Sarah J. Reber. 1998. "Paying for Health Insurance: The Trade-off between Competition and Adverse Selection." *Quarterly Journal of Economics* 113 (2): 433–66.

Dafny, Leemore, Kate Ho, and Mauricio Varela. 2013. "Let Them Have Choice: Gains from Shifting Away from Employer-Sponsored Health Insurance and toward an Individual Exchange." *American Economic Journal: Economic Policy* 5 (1): 32–58.

Dalton, Christina M., Gautam Gowrisankaran, and Robert J. Town. 2020. "Salience, Myopia, and Complex Dynamic Incentives: Evidence from Medicare Part D." *Review of Economic Studies* 87 (2): 822–69.

Dartmouth Atlas Project. 2020. "Interactive Maps: Hospital Referral Regions."http://www.dartmouthatlas.org/data/region (accessed October 2019).

Dixit, Avinash K., and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67 (3): 297–308.

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

Eden, Maya. 2020. "Welfare Analysis with Heterogeneous Risk Preferences." *Journal of Political Economy* 128 (12): 4574–613.

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

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

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

Ericson, Keith Marzilli, and Justin R. Sydnor. 2018. "Liquidity Constraints and the Value of Insurance." NBER Working Paper 24993.

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

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

Geruso, Michael. 2017. "Demand Heterogeneity in Insurance Markets: Implications for Equity and Efficiency." *Quantitative Economics* 8 (3): 929–75.

Geruso, Michael, Timothy J. Layton, Grace McCormack, and Mark Shepard. 2019. "The Two Margin Problem in Insurance Markets." NBER Research Working Paper 26288.

Ghili, Soheil. 2016. "Network Formation and Bargaining in Vertical Markets: The Case of Narrow Networks in Health Insurance." Unpublished.

Glazer, Jacob, and Thomas G. McGuire. 2011. "Gold and Silver Health Plans: Accommodating Demand Heterogeneity in Managed Competition." *Journal of Health Economics* 30 (5): 1011–19.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town. 2015. "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry." *American Economic Review* 105 (1): 172–203.

Gross, Tal, and Matthew J. Notowidigdo. 2011. "Health Insurance and the Consumer Bankruptcy Decision: Evidence from Expansions of Medicaid." *Journal of Public Economics* 95 (7–8): 767–78.

Gruber, Jonathan, Benjamin R. Handel, Jonathan T. Kolstad, and Samuel H. Kina. 2019. "Managing Intelligence: Skilled Experts and AI in Markets for Complex Products." Unpublished.

Handel, Benjamin R. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review* 103 (7): 2643–82.

Handel, Benjamin R., Igal Hendel, and Michael D. Whinston. 2015. "Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk." *Econometrica* 83 (4): 1261–313.

Handel, Benjamin R., Igal Hendel, and Michael D. Whinston. 2017. "The Welfare Effects of Long-Term Health Insurance Contracts." NBER Working Paper 23624.

Handel, Benjamin R., and Jonathan T. Kolstad. 2015. "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare." *American Economic Review* 105 (8): 2449–500.

Harsanyi, John C. 1953. "Cardinal Utility in Welfare Economics and in the Theory of Risk-taking." *Journal of Political Economy* 61 (5): 434–35.

Harsanyi, John C. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy* 63 (4): 309–21.

Hendren, Nathaniel. 2021. "Measuring Ex Ante Welfare in Insurance Markets." *Review of Economic Studies* 88 (3): 1193–223.

Hendren, Nathaniel, Camille Landais, and Johannes Spinnewijn. 2021. "Choice in Insurance Markets: A Pigouvian Approach to Social Insurance Design." NBER Working Paper 27842.

Ho, Kate, and Robin S. Lee. 2017. "Insurer Competition in Health Care Markets." *Econometrica* 85 (2): 379–417.

Ho, Kate, and Robin Lee. 2021. "Health Insurance Menu Design for Large Employers." Unpublished.

Hong, Long, and Corina Mommaerts. 2021. "Time Aggregation in Health Insurance Deductibles." NBER Working Paper 28430.

Kaiser Family Foundation. 2018. "Marketplace Plan Selections by Metal Level." https://www.kff.org/health-reform/state-indicator/marketplace-plan-selections-by-metal-level. (accessed October 2019).

Ketcham, Jonathan D., Claudio Lucarelli, Eugenio J. Miravete, and M. Christopher Roebuck. 2012. "Sinking, Swimming, or Learning to Swim in Medicare Part D." *American Economic Review* 102 (6): 2639–73.

Kowalski, Amanda E. 2015. "Estimating the Tradeoff between Risk Protection and Moral Hazard with a Nonlinear Budget Set Model of Health Insurance." *International Journal of Industrial Organization* 43: 122–35.

Landais, Camille, Arash Nekoei, Peter Nilsson, David Seim, and Johannes Spinnewijn. 2021. "Risk-Based Selection in Unemployment Insurance: Evidence and Implications." *American Economic Review* 111 (4): 1315–55.

Liebman, Eli. 2018. "Bargaining in Markets with Exclusion: An Analysis of Health Insurance Networks." Unpublished.

Lustig, Joshua. 2008. "The Welfare Effects of Adverse Selection in Privatized Medicare." Unpublished.

Mahoney, Neale. 2015. "Bankruptcy as Implicit Health Insurance." *American Economic Review* 105 (2): 710–46.

Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, Arleen Leibowitz, and M. Susan Marquis. 1987. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *American Economic Review* 77 (3): 251–77.

Marone, Victoria R., and Adrienne Sabety. 2022. "Replication Data for: When Should There Be Vertical Choice in Health Insurance Markets?" American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E148941V1.

McManus, Margaret A., Stephen Berman, Thomas McInerny, and Suk-fong Tang. 2006. "Weighing the Risks of Consumer-Driven Health Plans for Families." *Pediatrics* 117 (4): 1420–24.

Newhouse, Joseph P. 1993. *Free for All? Lessons from the RAND Health Insurance Experiment*. Cambridge, MA: Harvard University Press.

Oregon Educators Benefit Board (OEBB). 2018. "Health Insurance Benefits Data."

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

Pauly, Mark V. 1974. "Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection." *Quarterly Journal of Economics* 88 (1): 44–62.

Reed, Mary, Vicki Fung, Mary Price, Richard Brand, Nancy Benedetti, Stephen F. Derose, Joseph P. Newhouse, and John Hsu. 2009. "High-Deductible Health Insurance Plans: Efforts to Sharpen a Blunt Instrument." *Health Affairs* 28 (4): 1145–54.

Revelt, David, and Kenneth Train. 1998. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *Review of Economics and Statistics* 80 (4): 647–57.

Revelt, David, and Kenneth Train. 2001. "Customer-Specific Taste Parameters and Mixed Logit: Households' Choice of Electricity Supplier." Unpublished.

Rothschild, Michael, and Joseph E. Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90 (4): 630–49.

Samek, Anya, and Justin R. Sydnor. 2020. "Impact of Consequence Information on Insurance Choice." NBER Working Paper 28003.

Shepard, Mark. 2016. "Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange." NBER Working Paper 22600.

Tilipman, Nicholas. 2018. "Narrow Physician Networks, Switching Costs, and Product Variety in Employer Markets." Unpublished.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. 2nd ed. Cambridge, MA: Cambridge University Press.

Zeckhauser, Richard. 1970. "Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives." *Journal of Economic Theory* 2 (1): 10–26.