Health Insurance Menu Design with Multi-Dimensional Plans: Evidence from ACA Exchanges*

Myles Wagner†

March, 2023

Click here for most recent version

Abstract

Offering choice of health insurance plans is often a policy objective. But choice over a single dimension of plan quality, the focus of most past work, faces several challenges: selection on moral hazard can reduce or even reverse the efficiency gains from choice, and adverse selection can cause markets to unravel even when choice would be efficient. I study choice in the Affordable Care Act (ACA) setting, where plans have *multiple dimensions* of quality, including financial generosity and provider network breadth. I show theoretically that offering a tradeoff between quality types – e.g., by pairing a broad network with low financial generosity and a narrow network with high financial generosity – can reduce the scope for inefficient plan selection. Using empirical evidence from individual-level plan choices and medical claims in the Massachusetts ACA exchange, I find that offering choice between a less generous broad-network and more generous narrow-network plan substantially reduces the impacts of both selection on moral hazard and adverse selection. The resulting menu of plans produces greater efficiency gains than choice over either quality dimension on its own, and between \$190-\$475 of surplus per member year relative to the best single-plan option.

JEL Codes: D82, G22, I13, I18

*Acknowledgments: I thank my advisors David Cutler, Nathan Hendren, and Mark Shepard for their continual patience, guidance, and support. I am also grateful to Michael Chernew, Liran Einav, Oliver Hart, Tim Layton, Robin Lee, Jeremy Majerovitz, Victoria Marone, Ariel Pakes, Nicholas Tilipman, and Anthony Yu for helpful discussions, as well as to participants at the Harvard Labor/Public Finance Workshop, Harvard Industrial Organization Workshop, Harvard Department of Healthcare Policy Health Economics Seminar, and American Society of Health Economists meetings 2022. This research was supported by National Institute on Aging, Grant Number T32-AG000186.

†Department of Economics, Harvard University. Email: mnwagner@g.harvard.edu

1 Introduction

Many insurance markets offer consumers choice over plans. Because individuals may have heterogeneous tastes for different plan characteristics, sustaining a large number of choices is viewed as an explicit policy objective in a number of settings, such as the Medicare Advantage and Affordable Care Act (ACA) marketplaces. However, in selection markets such as health insurance, offering choice is not always socially efficient. When insurance issuers are constrained in their ability to price-discriminate on the basis of expected liability – either because of community-rating regulations (which require insurers to charge the same price to all members of a population) or because consumers have private information – consumers may select plans that are sub-optimal from a social planner's perspective.

The problem arises because the *relative* – or *incremental* – cost of changing insurance coverage varies across individuals, due to differences in the quantity of care individuals tend to use (health risk), and differences in how changes in coverage affect healthcare consumption behavior (so-called "moral hazard"). This variation in incremental costs opens the door for *selection on incremental costs* (also often called selection on moral hazard), wherein individual plan preferences are correlated with their incremental costs. As a result, choice has an ambiguous effect on social welfare. Indeed, recent work across a variety of settings has found the value of choice to be small or even negative relative to restricting choice to just a single option.[1]

Although health insurance plans can differ in many ways – including the extent of patient cost-sharing (e.g., deductible) and degree of doctor and hospital choice (provider network) – existing work has typically considered the problem of choice (both conceptually and empirically) in settings where plan differentiation is limited to a single dimension. The characteristics of the optimal menu in a setting with multiple plan dimensions are poorly understood.[2]

In this paper, I study the conditions under which multiple dimensions of plan differentiation make it optimal to offer choice. First, I develop a conceptual framework that generalizes the challenge of

[1]For example, recent work by Marone and Sabety (2022) has found that, when health insurance plans are *vertically differentiated* (more versus less generous out-of-pocket prices), offering choice is inefficient relative to offering just a single option because more generous plans are predominantly taken up by sicker patients who do so in order to get a larger transfer from the insurer, but benefit little from additional risk protection. Similarly, Shepard (2022) finds that when plans differ in whether they cover expensive "star" hospitals, selection into the star-hospital plan by sicker patients loyal to those hospitals creates additional costs for the insurer that exceed the value those patients place on the broader provider choice.

[2]In the general case, the sponsor faces a multi-dimensional screening problem with multiple instruments (plan features) and multiple dimensions of heterogeneity (consumer preferences and costs with respect to plan features). Problems of this type are notoriously intractable.

offering choice in selection markets. The key impetus for this paper is the observation that a market designer (in the insurance setting often called a market sponsor or regulator) can take advantage of the *multi-dimensional* nature of plans to ameliorate the challenge of offering choice by more efficiently screening patients. The idea is to offer a menu of plans that requires patients to *trade off* between various quality dimensions. Such menus use a form of multi-dimensional screening which I refer to as "diagonal differentiation," to reflect the fact that the respective plans are more generous on some dimensions but simultaneously less generous on others. An example would be a firm that offers employees a choice between an HMO plan with generous out-of-pocket costs but a narrow network of hospitals and physicians, and a PPO plan that contracts with a large set of providers but also requires patients to pay higher prices out-of-pocket at the point care via a higher deductible and/or co-pays.

I identify conditions under which a market sponsor would want to offer choice over a menu diagonally differentiated across two plan quality dimensions, e.g., financial coverage (out-of-pocket price schedule) and provider network breadth, even when it would not offer choice if plans could only be differentiated along a single quality dimension (as has been the finding in prior work). The conditions are statistical in nature and have an intuitive interpretation: the effectiveness of diagonal differentiation depends on the extent to which relative plan *preferences* (which govern willingness to pay and plan selection) and relative plan *costs* are explained by a patient's *health risk* – the quantity or severity of care the patient can be expected to use. When health risk explains a relatively large share of variation in costs across plans, but a smaller share of variation in plan demand (i.e., there is substantial preference heterogeneity conditional on health risk), diagonal differentiation improves screening efficiency through two channels. First, it reduces the extent to which plan selection is driven by health risk.[3] Second, it reduces the heterogeneity in incremental costs across individuals, and thereby reduces the *scope* for selection on incremental costs.[4]

I take this framework to data by estimating a structural model of plan preferences and enrollee costs in the Massachusetts ACA exchange (Connector). I observe enrollment and medical claims for enrollees in the Connector by drawing on the Massachusetts All-Payer Claims Database (APCD), and take advantage of several sources of price variation in the exchange in order to identify heterogeneity in

[3]Since selection on health risk is a primary driver of adverse selection in insurance markets, a diagonally differentiated menu may also be more stable under insurer competition.

[4]In the special case in which relative plan costs are *fully* explained by health risk, it is possible (under certain divisibility conditions on the plan characteristics space) to offer a menu of plans that reduces the variance in incremental costs to zero and therefore eliminate the potential for selection on incremental costs. In this case, it will always be preferable to offer some form of diagonally differentiated choice relative to a single-plan menu.

preferences for financial generosity and network breadth, and to estimate the effect of plan features on costs. The Connector is well suited to study the problem of multi-dimensional menu design because it features rich variation in health plan features. Patients buying insurance through the exchange choose between insurance carriers that cover different networks of providers, including a number of carriers that have narrow provider networks, and one that is vertically integrated with a large academic medical system (Partners Healthcare) and is the only carrier to cover providers in that system.[5] Furthermore, each insurance carrier is required to offer a standardized menu of plans with vertically differentiated out-of-pocket price schedules.

The key driver of heterogeneity in demand for broad-network coverage in my setting is whether the patient used Partners providers before becoming enrolled in the Connector. Patients with a history of Partners use, as well as sicker patients, are willing to pay the greatest premium for broad provider choice. However, sicker patients are also willing to pay more for generous out-of-pocket coverage relative to former Partners patients. Sick patients also generate the greatest costs for an insurer when covered under more generous coverage or broader provider networks. Given this distribution of preferences and costs, I examine whether a social planner constrained to set uniform prices, and attempting to maximize unweighted social surplus, would offer choice over various counterfactual menus of plans given the joint distribution of demand and costs estimated from the population of Connector enrollees. Because sicker patients tend to choose more generous coverage when choice is offered, the social planner would prefer to restrict choice to a single Silver-tier plan rather than offer choice over plans that are differentiated strictly in out-of-pocket coverage. Similarly, when patients are able to choose among plans that differ only in their coverage of Partners providers, sorting by sickness – a key driver of cost-sensitivity to Partners coverage – restricts a planner's ability to efficiently offer choice.

In my primary counterfactual analysis, I find the social planner's optimal two-plan menu from a space of plans that features narrow vs broad-network and more versus less generous financial coverage. The optimal menu features diagonal differentiation: offering patients choice between a less generous broad-network plan and a more generous narrow-network plan. Compared to the best single-plan menu, which offers a more generous broad-network plan, the optimal menu increases net social surplus

[5]Partners Healthcare, which has since the period studied in this paper been re-branded as Mass General Brigham, is a large Massachusetts-based healthcare system which includes Massachusetts General Hospital and Brigham and Women's Hospital, both prestigious teaching hospitals that are also two of the most expensive hospitals in the state.

by between \$190-\$475 per member-year.

From the perspective of the menu design problem, adverse selection acts as a constraint on the ability of the market sponsor to implement the efficient allocation (price) within a given menu, and can therefore affect optimal menu design. Even when the market sponsor wants to offer choice, adverse selection may distort the equilibrium pricing of plans from the efficient price (Akerlof, 1970; Einav, Finkelstein and Cullen, 2010) and in some cases can result in the total unraveling of choice (Cutler and Reber, 1998). While I find both vertically and diagonally differented menus suffer from adverse selection, the diagonally differentiated menu is more stable under adverse selection. Because the diagonally differentiated menu does not offer one plan that is unambiguously more generous (on any single dimension) than the other, riskier (sicker) patients do not sort as strongly as in a vertically differentiated menu. Allowing the market sponsor to use a risk adjustment scheme similar to that used in the ACA, I find that a diagonally differentiated menu leads to an equilibrium allocation that generates close to 100 percent of the surplus of the efficient allocation. However, the equilibrium allocation when plans are vertically differentiated in network alone generates only about a third of its potential surplus.

This paper relates to several strands of literature in the economics of selection markets. It complements a large literature on "managed competition" (Enthoven, 1988) in health insurance markets, including papers studying demand for health insurance in individual exchanges (e.g., Dafny, Ho and Varela, 2013; Dafny, Hendel and Wilson, 2015; Ericson and Starc, 2015a, 2016), and policies designed to promote efficiency in health plan provision and pricing under competition (e.g., Einav et al., 2010; Handel et al., 2015; Azevedo and Gottlieb, 2017; Geruso et al., 2021). It also relates to the literature on the effects of financial (Einav et al., 2013) and provider network (Shepard, 2022) health plan characteristics on selection and costs, and follows Finkelstein and Poterba (2004) and Finkelstein and McGarry (2006) in emphasizing the importance of multi-dimensional preferences and costs in selection markets.

Most directly, this paper contributes to a growing literature in the design and pricing of menus in selection markets. This includes Marone and Sabety (2022), Ho and Lee (2021) and Tilipman (2022) who study health insurance menu design in the employer context, and Landais et al. (2021) who study menu design in the context of unemployment insurance. This paper is also closely related to Bundorf, Levin and Mahoney (2012), who study efficient pricing in health insurance with plans that differ in multiple dimensions (notably, the setting studied by Bundorf et al. (2012) features a health plan menu that is diagonally differentiated.)

The rest of the paper proceeds as follows. Section 2 outlines my conceptual framework and uses it to illustrate how a regulator can use multiple plan dimensions to create menus where choice is efficient. Section 3 describes the Massachusetts ACA exchange setting and the enrollment and claims data I use in estimating the empirical models of demand and cost described in Section 4. I use these estimates to analyze a social planner's menu design problem in Section 5, and Section 6 considers additional extensions to the menu design problem where pricing is constrained by adverse selection. Finally, Section 7 concludes.

2 Conceptual Framework

2.1 Setup and Notation

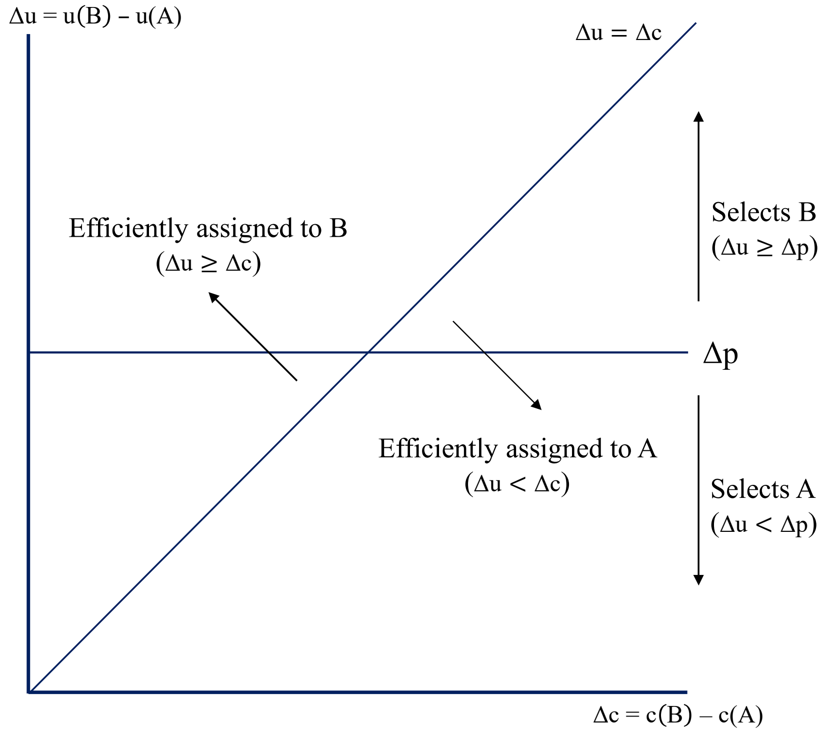
I consider a setting in which consumers in a given population, denoted $\mathbb{I}$, choose one of two health insurance plans $h \in \{A, B\}$. Consumers, whom I index by $i \in \mathbb{I}$, have demand for insurance, which I denote $u_i(h)$, which gives consumer i's willingness-to-pay (WTP) for plan h. Because there are only two plans available in my setting, it is convenient to define the *relative* or *incremental* willingness-to-pay (WTP) for plan B versus plan A. I denote this by $\Delta u_i \equiv u_i(B) - u_i(A)$ for consumer i. Demand is therefore a function of the relative price Δp of plan B versus plan A: consumer i chooses plan B if and only if $\Delta u_i \geq \Delta p$, and chooses plan A otherwise. Throughout the paper, I will assume that the relative price Δp is constant across all consumers. I follow the health insurance literature in calling this a "community-rating" constraint.

When consumer i chooses a plan, they generate costs for the insurer. I refer to the expected cost of consumer i on plan h by $c_i(h)$, and define the relative or incremental cost of consumer i choosing plan B versus plan A as $\Delta c_i \equiv c_i(B) - c_i(A)$.

2.1.1 Efficient Allocation

I consider the preferences of a social planner, which in the insurance market context is often called a "market sponsor," or simply "sponsor," which seeks to maximize (unweighted) social surplus: the difference between aggregate willingness-to-pay and insurer cost. Let $\mathbb{I}_B(\Delta p) \subset \mathbb{I}$ refer to the set of

Figure 1: Sponsor's Plan Pricing Problem



Notes: This figure illustrates the tradeoff the sponsor faces in setting the relative price (Δp) for plan B versus plan A. The sponsor wishes to assign any enrollee *above* the 45-degree line to plan B, since their incremental WTP exceeds their incremental cost, and any enrollees *below* the 45-degree line to plan A, since their incremental cost exceeds their incremental demand. However, the sponsor is restricted to assignment plans on the basis of incremental WTP alone using a uniform price Δp.

consumers who purchase B at price Δp, that is $\mathbb{I}_B = \{i : \Delta u_i \geq \Delta p\}$. Then social surplus can be written as

$$S(A, B; \Delta p) \equiv \int_{i \in \mathbb{I}_B(\Delta p)} u_i(B) - c_i(B) \, \mathrm{d}i + \int_{i \notin \mathbb{I}_B(\Delta p)} u_i(A) - c_i(A) \, \mathrm{d}i. \tag{1}$$

By subtracting $\int_{i \in \mathbb{I}} u_i(A) - c_i(A) \, \mathrm{d}i$ from the above equation, the sponsor's problem can be expressed as,

$$max_{\mathbb{I}_B \subset \mathbb{I}} E\left[\Delta u_i - \Delta c_i | i \in \mathbb{I}_B\right], \tag{2}$$

in terms of solely the incremental demand Δu_i and incremental cost Δc_i.[6]

Figure 1 portrays regions of $(\Delta u, \Delta c)$-space that illustrate the sponsor's tradeoff of changing plan assignments using a community-rated (uniform) price. The sponsor's optimal feasible plan assignment (i.e., the price the sponsor should set) thus depends on the joint distribution of incremental WTP and

[6]Importantly, for existence of a solution, I allow the price Δp to take on infinite values.

incremental cost – that is, where they fall in $(\Delta u, \Delta c)$-space.

I define the solution to the sponsor's problem as the *optimal*, or *efficient*, allocation. The only constraint in this definition of the sponsor's problem is the community-rating constraint, and the sponsor's only instrument is the price Δp. The health insurance literature has considered a number of extensions to this problem. For example, Bundorf et al. (2012) consider relaxing the community-rating constraint, while Einav et al. (2010) consider constraining the set of feasible allocations (prices) to that resulting from the equilibrium of a pricing game between competing insurers. This paper studies the sponsor's *menu design* problem, which considers the characteristics of the plans A and B, in addition to the relative price Δp, as policy instruments of the market sponsor.

2.1.2 Offering Choice

For given plans A and B, it is possible that the sponsor's constrained-efficient allocation involves assigning all consumers to one of the two plans, i.e., $\mathbb{I}_B = \mathbb{I}$ or $\mathbb{I}_B = \emptyset$. In other words, it is possible the sponsor *may not want to offer choice* over A and B, even when there is substantial heterogeneity in preferences.[7] To see why, consider the two cases presented in Figure 2 which portrays two potential joint distributions of $(\Delta u, \Delta c)$ for pairs of plans A and B.

In Panel 2, consumers with the highest incremental WTP for B have (on average) incremental costs that exceed their incremental WTP. In this case, any interior allocation would be characterized by "backwards sorting" (Marone and Sabety, 2022), which is reflected graphically in the fact that the Einav et al. (2010) incremental cost (MC) curve is shallower than the 45-degree line (where $\Delta u = \Delta c$). As a result, the sponsor's objective function is *convex* in the share of consumers choosing B, and the efficient allocation lies at one of the two end-points.

The graphical example can be used to motivate a formal characterization of conditions on the distribution of $(\Delta u, \Delta c)$ under which the sponsor offers choice. Under the assumption that the incremental cost (MC) curve is monotonic in incremental demand, the sponsor will offer choice only if

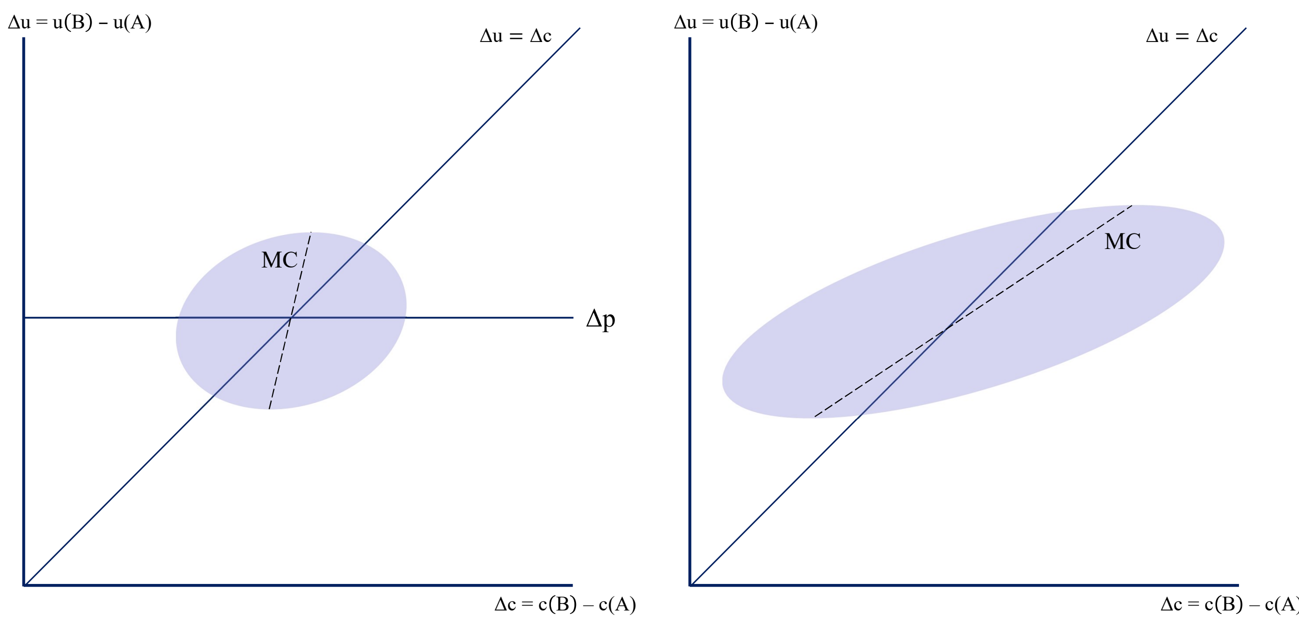
$$cov\left(\Delta u, \Delta u - \Delta c\right) \geq 0, \tag{3}$$

that is, if enrollees with greater incremental WTP for plan B are on average more efficiently enrolled in

[7] In this case, and when the distribution of Δu does not have bounded support, the efficient allocation corresponds to an infinite incremental price Δp.

Figure 2: Offering choice may be inefficient

Panel 1: Sponsor offers choice **Panel 2:** Sponsor does not offer choice



Notes: Each panel illustrates a joint distribution of $(\Delta u, \Delta c)$ represented by the shaded ovals. I have also drawn dashed lines representing the marginal incremental cost (MC) curve in each panel, where MC is defined as $E[\Delta c|\Delta u]$ as in Einav et al. (2010). The 45-degree line represents the demand curve. In Panel 1 on the left, the sponsor's (constrained) optimal allocation is achieved by setting a uniform price equal to Δp. However, in Panel 2, because MC rises in Δu more quickly than 1-for-1 (the marginal cost curve rises more quickly than the demand curve), any interior allocation would result in "backwards-sorting," where plan B is taken up by consumers who on average value it less than the additional costs they generate for the insurer.

plan B. Condition 3 is a necessary but not sufficient condition for the sponsor to offer choice; choice also requires that there exists a unique level of incremental demand , Δu^0, s.t. $E\left[\Delta u - \Delta c | \Delta u = \Delta u^0\right] = 0$ – that is, the MC curve crosses the 45-degree line exactly once (a standard single-crossing property).

2.1.3 Statistical Demand Decomposition

In order to explore when Condition 3 holds, it is useful to decompose demand Δu into components that are explained by heterogeneity in incremental cost Δc and factors that are not related to cost, which can be thought of as instead driven by heterogeneity in preferences for plan characteristics. Previous work has accomplished this by using standard functional form assumptions on utility. For example, Marone and Sabety (2022) and Finkelstein, Hendren and Luttmer (2019) each show that, when utility follows a standard CRRA (constant relative risk aversion) form, demand for (financial) insurance can be decomposed into the sum of a component that captures willingness-to-pay for the expected transfer (which may also include willingness-to-pay for moral hazard when applicable), and a component that captures willingness-to-pay for risk protection.[8] Because I study a setting in which plans can be differentiated in non-financial characteristics, rather than making assumptions about the functional form of preferences over those characteristics, I instead show that a general argument can still be made with an analogous *statistical* decomposition. For an arbitrary utility function and underlying distribution of primitives, Δu can be linearly decomposed into a component statistically predicted by incremental cost Δc, and an uncorrelated "preference" term,

$$\Delta u_i = E^L\left[\Delta u | \Delta c_i\right] + \psi_i = \underbrace{\alpha + \beta \Delta c_i}_{\text{Cost-related component of } \Delta u_i} + \underbrace{\psi_i}_{\text{Preference } (\psi \text{ uncorrelated with } \Delta c)} \tag{4}$$

where $E^L\left[\Delta u | \Delta c_i\right] = \alpha + \beta \Delta c_i$ is the linear conditional expectation of Δu for enrollees with incremental cost equal to Δc_i, and ψ is an individual-specific preference parameter that is (by construction) uncorrelated with Δc.

Substituting Equation 4 into Condition 3 yields that it is efficient for the sponsor to offer choice

[8] Landais, Hendren and Spinnewijn (2022) also perform a conceptually similar decomposition of demand for financial insurance with a more general setup.

over A and B only if

$$V(A,B) \equiv \underbrace{\beta(\beta-1)\,var(\Delta c)}_{\text{Selection on incremental cost}} + \underbrace{var(\psi)}_{\text{Preference heterogeneity}} \geq 0. \tag{5}$$

Condition 5 shows that the sponsor's decision whether to offer choice depends on three distinct features of the distribution of $(\Delta u, \Delta c)$, each of which is graphically reflected in the examples shown in Figure 2.

1. **The nature of selection on incremental cost** β, which is represented graphically in the slope of the best-fit line through the joint distribution of Δu and Δc.[9] The first term in equation 5 is weakly negative if and only if $0 < \beta < 1$, that is, when incremental WTP on average increases less than one-for-one with incremental cost. This is consistent with *selection on moral hazard* (Einav et al. (2013); Shepard (2022)), in which enrollees with a greater tendency towards moral hazard (and thus greater incremental cost) are more likely to choose a plan which enables moral hazard. In models of moral hazard, patients undertake costlier care (e.g., greater quantities of care, and/or more expensive forms of care) only because they do not pay its full marginal cost – consequently their *willingness-to-pay* for moral hazard is exceeded by the moral hazard's marginal cost.

2. **The scope for selection on incremental cost** $var(\Delta c)$, represented graphically in the horizontal dispersion of the distribution along the Δc-axis. Heterogeneity in incremental cost is the defining feature of selection markets. When $0 < \beta < 1$, greater heterogeneity in incremental costs weakens the case for offering choice, in the sense that greater preference heterogeneity is required in order for choice to be efficient (in order for $V(A,B)$ to be weakly positive). In the case without heterogeneity in incremental costs, the first term of Equation 5 resolves to zero and the sponsor unambiguously offers choice – the standard result in markets without selection.

3. **The degree of heterogeneity in incremental WTP** $var(\psi)$ *conditional* on incremental cost. More heterogeneity in plan preferences conditional on incremental cost can only strengthen the case for choice. In selection markets, whether a market sponsor would want to offer choice between two given plans comes down to a tension between two factors: the degree of heterogeneity in willingness-to-pay given by $var(\psi)$, and the degree of selection on incremental cost given by $\beta(\beta-1)\,var(\Delta c)$.

[9] Formally, $\beta = \frac{cov(\Delta u, \Delta c)}{var(\Delta c)}$.

2.2 Menu Design with Multi-Dimensional Plans

I now introduce the menu design problem, which considers the characteristics of plans A and B as instruments of the market sponsor. The key feature of my setting is that I allow plans to be characterized by two quality dimensions, which I denote m and x. To reflect my empirical setting, m parameterizes the generosity of the plan's out-of-pocket price schedule, and x parameterizes the quality of the plan's provider network. Let $\mathbb{P}$ be the space of all possible plans (m, x), and the set of two-plan menus thus includes all unordered pairs from $\mathbb{P}$. Consumers i have WTP for plan (m, x) denoted by $u_i(m, x)$, and generate expected insurer costs of $c_i(m, x)$.

2.2.1 Menu Design

For a given menu, I assume the market sponsor can achieve the efficient allocation described in Section 2.1.1 by regulating the relative price of plans, constrained only in that the price is community-rated. The sponsor's *menu design problem* thus consists in finding the pair of plans in $\mathbb{P}$ with the efficient allocation that generates the greatest surplus. The sponsor's problem can be written as

$$max_{A,B\in\mathbb{P}}\left[\underbrace{max_{\Delta p}\left\{S\left(A,B;\Delta p\right)\right\}}_{V_{eff}(A,B)}\right]. \tag{6}$$

Note that the sponsor's problem admits single-plan menus (by choosing a pair of plans with $A = B$) as well as two-plan menus where the optimal allocation involves assigning all consumers to a single plan. I therefore define the solution to the sponsor's menu design problem as the set of menus (A, B) that satisfy $V_{eff}(A, B) \geq V_{eff}(A', B')$, where $V_{eff}(A, B)$ gives the social surplus of the efficient allocation between A and B, although I will also refer to an element of this set as the optimal menu. I assume the underlying primitives imply that the solution set is non-empty.

2.2.2 Is Choice Offered in the Optimal Menu?

Characterizing the optimal menu(s) in general requires solving a multi-dimensional screening problem with multiple instruments and (possibly) multiple dimensions of heterogeneity. I provide a partial characterization which is intended to provide intuition regarding the more general problem in the setting with multi-dimensional plans. The result is that, when there is a single dimension of consumer heterogeneity in *cost*, the solution set to the sponsor's menu design problem includes a menu that offers choice – that is, there is a menu that offers choice which weakly dominates all other menus in terms of social surplus. This is because it is possible to reduce the scope for selection on moral hazard, $var\left(\Delta c\right)$, to zero, perfectly flattening the marginal cost curve.

The argument involves taking a local-linear approximation to the (heterogeneous) insurer cost function, $c_i\left(m, x\right)$. Let $A = \left(m_0, x_0\right)$ be any plan in $\mathbb{P}$ and consider menus consisting of A and $B = \left(m_0 + \Delta m, x_0 + \Delta x\right)$. As shown above in Section 2.1.3, whether the sponsor efficiently offers choice in the menu $\left(A, B\right)$ depends on the distribution of incremental costs

$$\Delta c_i\left(\Delta m, \Delta x\right) \equiv c_i\left(m_0 + \Delta m, x_0 + \Delta x\right) - c_i\left(m_0, x_0\right). \tag{7}$$

I take a local linear approximation to $\Delta c_i\left(\Delta m, \Delta x\right)$:

$$\begin{aligned}
\Delta c_i\left(\Delta m, \Delta x\right) &= c_i\left(m_0 + \Delta m, x_0 + \Delta x\right) - c_i\left(m_0, x_0\right) && (8)\\
&\approx c_i\left(m_0, x_0\right) + \frac{\partial c_i}{\partial m}\Delta m + \frac{\partial c_i}{\partial x}\Delta x - c_i\left(m_0, x_0\right) && (9)\\
&= \frac{\partial c_i}{\partial m}\Delta m + \frac{\partial c_i}{\partial x}\Delta x. && (10)
\end{aligned}$$

For notational brevity, I will refer to $c_m = \frac{\partial c_i}{\partial m}$ and $c_x = \frac{\partial c_i}{\partial x}$ and leave the subscript-i indicating heterogeneity across consumers i as implicit.

Cost Dimensionality Assumption I now introduce the assumption that there is a single dimension of cost heterogeneity. Formally, I assume $\left|corr\left(c_m\left(m, x\right), c_x\left(m, x\right)\right)\right| = 1$ at all $\left(m, x\right)$.[10] In the health insurance context, where plans are characterized by out-of-pocket price generosity m, and provider network quality x, it might be natural to think of *sickness* or the *quantity* of healthcare consumption as the common driver of cost responses to changes in either characteristic, since the cost to the insurer of covering more out-of-pocket expenses or covering a more or less expensive set of providers scales with the quantity of care each patient uses. In this case, the correlation between cost responses would be positive, so I proceed using $corr\left(c_m, c_x\right) = 1$, although the argument can be just as easily applied in the opposite case where the correlation is negative.

I now show that there exists $\left(\Delta m, \Delta x\right)$ such that $var\left(c_m \Delta m + c_x \Delta x\right) = 0$. As discussed above in Section 2.1.3, this is sufficient for choice to be (weakly) efficient. It is straightforward to show that,

[10] Where $corr\left(a, b\right) = \frac{cov(a,b)}{\sqrt{var(a)var(b)}}$ is the correlation coefficient.

when $corr\left(c_m, c_x\right) = 1$,

$$\left(\Delta m, \Delta x\right) = \left(\frac{-cov\left(c_m, c_x\right)}{var\left(c_m\right)}\Delta x, \Delta x\right)$$

solves $var\left(c_m \Delta m + c_x \Delta x\right) = 0$. Since this argument can be applied at any single-plan menu $A = (m_0, x_0)$, it follows that any single-plan menu is weakly dominated by a menu that offers choice.

Intuition The argument provides intuition for the general case in which costs are multi-dimensional. When incremental costs c_m and c_x are *closely (positively) correlated*, providing offsetting differentiation, which I call "diagonal differentiation" that requires patients to trade off between plan attributes reduces the variance in the relative/incremental costs between plans, and therefore *reduces the scope* for selection on incremental costs. Although in general the variance cannot be reduced to zero, in order for choice to be efficient it need only be reduced *sufficiently* relative to the degree of heterogeneity in preference, as shown in Equation 5. Of course, providing such offsetting differentiation may also reduce the heterogeneity in relative preference between the two plans if preferences for each plan dimension are similarly highly correlated. Offering choice over diagonally differentiated plans is most likely to yield efficient choice when cost responses to plan characteristics are highly correlated, but preferences for plan characteristics are relatively weakly (or even negatively) correlated. A key motivation for this paper is the observation that, in health insurance, this may be particularly likely to be the case. The reason is that a primary driver of cost and cost-response heterogeneity in health insurance is a patient's underlying sickness, which, through quantity of healthcare consumption, is a likely common driver of cost responses to a number of health plan characteristics.

3 Setting and Data

3.1 Massachusetts Connector

The Massachusetts Health Connector is the state's health insurance exchange – a regulated individual insurance market – established under the Affordable Care Act (ACA). The Connector opened in 2015, taking over for the pre-ACA Massachusetts exchange, which was established in 2006 under the state's "Romneycare" reform.[11] To the best of the author's knowledge, this paper is the first to study the Massachusetts ACA exchange using individual-level enrollment and claims data. Individuals purchas-

[11] For research on the pre-ACA Massachusetts health insurance exchange, see e.g., Shepard (2022); Ericson and Starc (2015a,b, 2016); Hackmann et al. (2015)

ing insurance through the Connector choose from plans offered by competing private insurers. By rule under the ACA, each insurer offers a standardized menu of plans, corresponding to four "metal tiers": Bronze, Silver, Gold and Platinum. These tiers range from higher cost-sharing/less generous (Bronze), to lowest cost-sharing/most generous (Platinum).[12] The Connector strictly regulates the financial features of each of these metal tiers for insurance carriers offering plans on the exchange, but within each metal tier, provider networks differ across insurance carriers. The Connector also offers premium and cost-sharing subsidies for lower income consumers through a program called ConnectorCare. Consumers falling below 300% of the federal poverty level and who are not eligible for employer insurance or other public programs (such as Medicaid) choose among generously-subsidized versions of the "baseline" Silver-tier plans, which means that, in effect, consumers below 300% of poverty choose only among insurance carriers, and not over coverage level (metal tier). Consumers above 300% of poverty do not receive state subsidies, and choose among plans that differ in coverage level (metal tier) and are offered by different insurance carriers.

The Connector setting is especially well suited to studying multi-dimensional menu design because of the rich variety of plans available. In addition to clearly defined and regulated variation in coverage level through the metal tier design, major carriers participating in the Connector differ substantially in the quality (breadth) of their provider networks. I focus particularly on three major health insurance carriers in the Connector during my period of study (2015-2018): Boston Medical Center (BMC) HealthNet, Network Health (also called Tufts Health Plan Direct), and Neighborhood Health Plan (NHP). During my period of study, NHP's provider network was the only one which covered Partners Healthcare, a large system that includes Massachusetts General Hospital (MGH) and Brigham and Women's Hospital (BWH).[13] Many Partners providers, in particular MGH and BWH, are known for both their prestige and high cost of care (Shepard, 2022). BMC Healthnet is a vertically-integrated plan operated by Boston Medical Center, a major Boston-based public hospital which covers a narrow provider and hospital network, while Network Health is the narrow-network option offered by Tufts Health Plan, a Massachusetts-based health insurance carrier which also offers broad-network commercial networks in the employer-sponsored insurance market. The variation in network quality between BMC, Network Health, and NHP, intersected with the variation in coverage tiers mandated by the

[12] Insurers in the Connector also have the option of offering a least-generous "Catastrophic" tier plan; in practice, few insurers offer Catastrophic coverage and those plans that are offered see very low enrollments.

[13] Partners Healthcare has since re-branded and is now called Mass General Brigham, and Neighborhood Health Plan is now called Mass General Brigham Health Plan.

ACA, creates a discrete approximation to the two-dimensional plan space $\mathbb{P}$ described in Section 2.2, and provides a natural setting to study menu design with multi-dimensional plans.

3.2 Data Sources

I draw on two primary administrative data sources that, combined, contain information on plan choice sets, prices, enrollment decisions and healthcare utilization for Connector enrollees over the period 2013-2018. The first is linked enrollment and claims data from the Massachusetts All-Payer Claims Database (APCD), which I supplement with administrative enrollment and pricing data from the Connector.

Massachusetts All-Payer Claims Database (APCD) I use individual-level claims and enrollment data for the years 2013-2018 from the Massachusetts APCD, which provides linked enrollment-claims data from public and private insurance payers in the state of Massachusetts, including all carriers participating in the Connector.

The APCD enrollment data is an individual-level panel describing each record of an individual enrolled in health insurance in Massachusetts, and includes variables describing the period of enrollment, the identity of the insurance carrier, whether the plan was purchased through the individual exchange, and features of the plan such as actuarial value and ACA metal tier (Catastrophic, Bronze, Silver, Gold or Platinum). The enrollment data also describe detailed individual-level demographics including gender, age, household composition and location of residence up to a 5-digit zip code. Using individual-level identifiers, enrollment data can be linked to APCD insurance claims, which describe insurer costs, medical diagnoses and procedures, as well as the identity of hospitals and physicians providing medical care for each patient.

Massachusetts Connector Pricing Data I use administrative pricing and enrollment data for the entirety of the Massachusetts Connector over the period 2015-2018, which allow me to observe plan availability and pricing by geographic region and member age. I combine this information on plan availability and pricing to construct a dataset of individual plan enrollment decisions.

3.2.1 Plan Choice and Cost Data

I construct a comprehensive dataset at the level of each individual choice instance over plans purchased through the Connector for the period 2015-2018. Individuals in the Connector make plan enrollment decisions at two key points on the calendar. The first is when *new enrollees* initially purchase insurance through the Connector, and the second is at an annual open enrollment period during January of each calendar year, when *continuing enrollees* can switch their enrollment to a different plan. For each instance of an individual making a choice between plans, I define variables describing the set of available plans, the characteristics of these plans including the insurance carriers, metal tiers, and premiums, and the plan enrollment decisions (which plan was chosen). For each individual-choice instance, I also define patient-level variables that describe each individual's demographics, their history of healthcare use, and measures of their concurrent health (based on subsequent medical claims).

Key individual-level variables include: patient-level risk scores and total medical costs paid by the insurer (calculated over the subsequent year before the next choice instance), the average distance between the patient's zip code and Partners-affiliated hospitals, and whether the patient used Partners (defined as outpatient, non-emergency care) *before* ever enrolling in the Connector.

3.3 Price Variation in the Connector

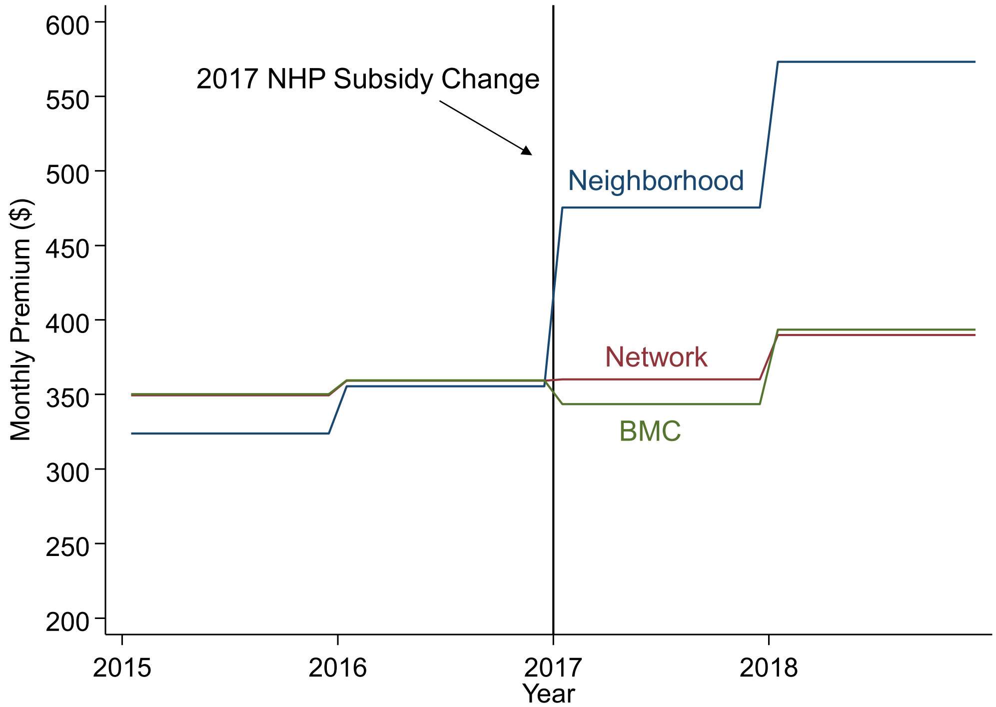
I draw on variation in plan prices over time to identify heterogeneity in plan preferences. Over the period 2015-2018, pricing by insurance carrier and by metal tier varies due to several institutional factors.

Reduced-Form Evidence: Variation in Carrier Prices

Between 2016-2017, the Connector changed the structure of its premium subsidies for plans in ConnectorCare, causing the price of NHP (the broad-network plan) to rise substantially relative to narrow-network options BMC and Network Health, and precipitating a "death spiral" (Cutler and Reber, 1998) pattern of adverse selection against NHP.[14] ConnectorCare prices are linked to each carrier's "baseline" premium in the individual market for consumers who do not receive the state (ConnectorCare) subsidies, meaning this policy change had spillover effects on the price of NHP's plans in the

[14] Prior to 2017, the Connector granted more generous premium subsidies to NHP because of its more expensive provider network. Citing concerns that this policy created an incentive for NHP to raise its premiums, the Connector changed this policy in 2017. For research on price-linked subsidies in health insurance exchanges, see Jaffe and Shepard (2020).

Figure 3: Price Variation: Carriers



Notes: Figure shows average monthly baseline (pre-subsidy) premium by year for the plans offered in the Connector by each of the three major carriers: Neighborhood (NHP), BMC Healthnet, and Network Health, over the first four years (2015-2018) of the Connector. Averages are computed by taking the mean of the monthly premium of the lowest-price plan within each metal tier (Bronze, Silver, Gold and Platinum) offered by each of the carriers. Premiums vary by geographic region and age; averages are weighted by the distribution of individual enrollees in the Connector in 2016.

individual market. Figure 3 plots the average monthly premium of plans offered by each of the three major carriers over time.

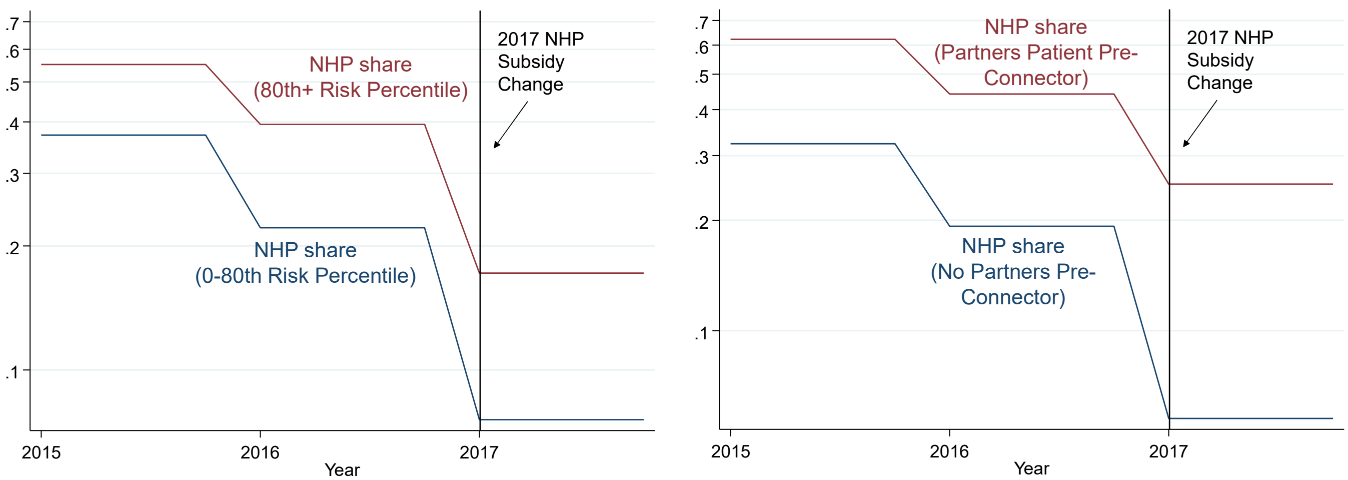
The change in market shares due to this price variation provides identification of the demand curve for NHP relative to BMC and Network. Figure 4 provides graphical analysis of heterogeneity in demand for NHP by sickness and history of Partners use, among first-time enrollees to the Connector. The results suggest that sickness and past use of Partners providers are both drivers of demand for the NHP plan.[15] Panel A shows that sicker patients 1) purchase NHP at greater rates, and 2) see a proportionally smaller decline in their likelihood of buying NHP plans in 2017 (when the price is higher) relative to 2016. Panel B shows a similar pattern for patients who have used Partners providers before entering the Connector, consistent with a preference for Partners coverage.

[15]The 2015-2016 price variation could in principle be used to identify another segment of the demand curve, under the assumption that unobserved carrier quality . In practice, I draw identification of demand for carrier networks from only the 2017 subsidy change, as I discuss in greater detail below in Section 4.1.

Figure 4: Heterogeneity in Demand for Neighborhood Health Plan

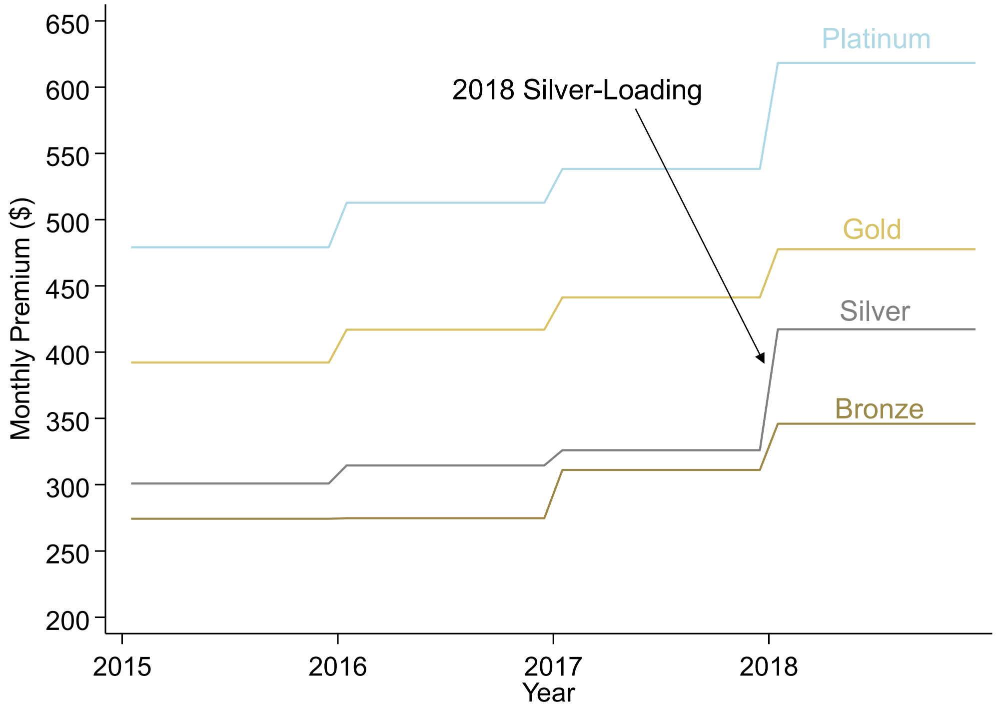
Panel A: Sick (80th+ Percentile Risk Score) vs. Healthy

Panel B: Partners Patient Pre-Exchange vs. Non-Partners Patient Pre-Exchange



Notes: Figure plots market share of NHP by years 2015-2017 among new (first-time) Connector enrollees. Panel A shows heterogeneity in NHP market shares by patient risk score, and Panel B shows heterogeneity in NHP market shares by whether individual enrollees used Partners outpatient (non-emergency) care within two years before enrolling in the Connector. Shares (y-axis) are shown in log-scale. NHP shares among new enrollees in 2018 are not shown because the APCD enrollment data only describe continuing enrollees in 2018. The sample excludes all former ConnectorCare enrollees.

Figure 5: Silver-loading Price Variation



Notes: Figure shows average monthly baseline (pre-subsidy) premium by year for plans of each primary metal tier (Bronze, Silver, Gold and Platinum) offered in the Connector by each of the three major carriers: Neighborhood (NHP), BMC Healthnet, and Network Health, over the first four years (2015-2018) of the Connector. Averages are computed by taking the mean of the monthly premium of the lowest-price plan within each metal tier for each carrier, and averaging across each of the three carriers. Averages are weighted by the distribution of individual enrollees in the Connector in 2016.

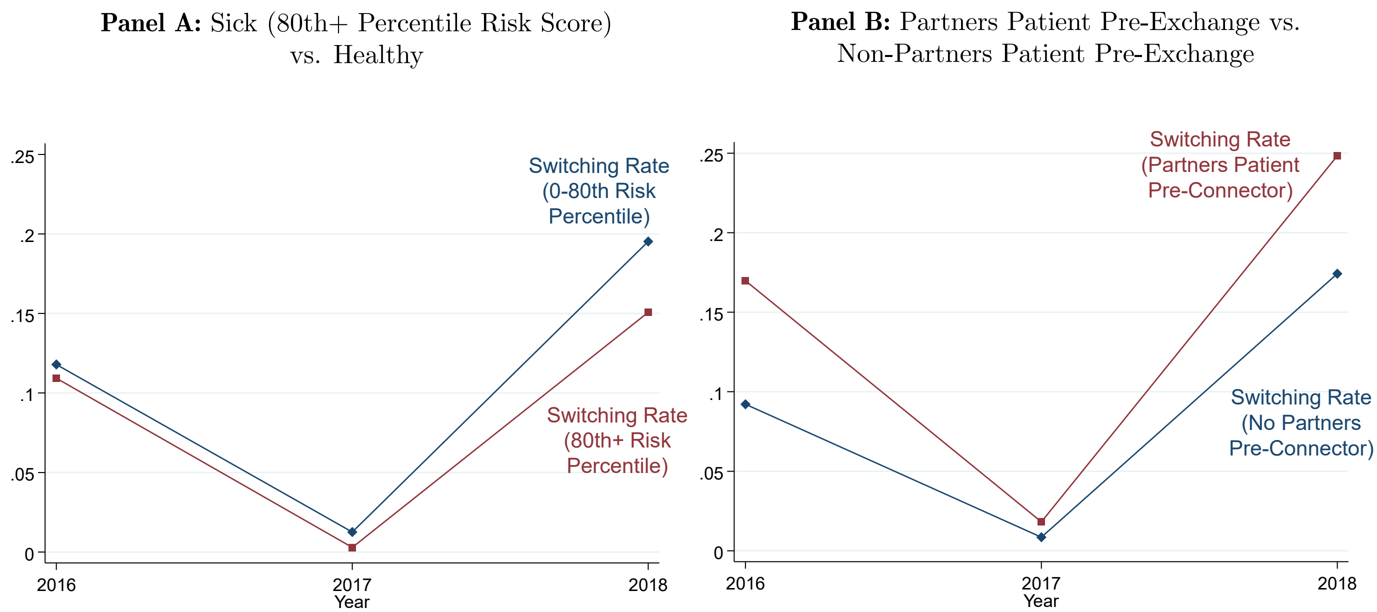
Reduced-Form Evidence: Variation in Silver-Tier Pricing

Between 2017-2018, the Connector allowed all carriers to raise the relative premiums of Silver-tier plans in order to make up for the loss of federal cost-sharing reduction subsidies, in a practice known as “Silver-loading.” Figure 5 shows the variation in average monthly premiums for plans in each of the four ACA metal tiers.

Variation in the relative price of Silver versus other metal tiers identifies demand for financial coverage. Because I only observe continuing enrollees in 2018, Figure 6 presents graphical evidence of heterogeneity in demand based on the rate at which individual enrollees *switch from Silver to Bronze* coverage in response to the price change. Panel A shows that sicker patients leave the more generous Silver plan at lower rates than healthier ones when the relative price of Silver rises, consistent with sicker patients having a preference for more generous coverage.[16] Panel B shows a different pattern

[16] An additional drawback of the APCD subset I use for this paper is that the data on medical claims for 2018 is

Figure 6: Silver-to-Bronze Switching Rates



Notes: For each year y shown on the x-axis, the figure plots the rate at which enrollees in a Silver-tier plan in year $y-1$ switch to a Bronze plan during open enrollment at the start of year y. Panel A shows heterogeneity in switching rate by whether a patient is sick (top 20% of risk scores) or healthy, and Panel B shows heterogeneity in switching rate by whether a patient used Partners (outpatient, non-emergency) providers within two years before enrolling in the exchange. The sample excludes all former ConnectorCare enrollees.

when switching rates are analyzed by whether the patient showed a preference for Partners providers before entering the exchange: previous Partners patients are more likely to switch from Silver to Bronze tier plans when the relative price of Silver increases.

The theory from Section 2.2 shows that quantifying heterogeneity in demand for plan characteristics (in the Connector setting, carrier network quality and metal tier), is a key input to the menu design problem. Given these patterns in plan choice along carrier and metal tier dimensions as a result of price variation, it is natural to ask if it is possible to quantify how much health risk, past use of Partners – and other factors such as distance to Partners hospitals, or demographics such as age – drive changes in plan shares when plans can differ in their carrier network or metal tier. Note that with sufficient price variation along each carrier-by-metal tier dimension, the demand function would be non-parametrically identified, and I could proceed in a similar fashion to Einav et al. (2010) in estimating

incomplete, meaning I do not observe healthcare utilization, and cannot construct risk scores, for patients based on their 2018 claims. I therefore use the 2017 risk score for continuing patients as a proxy for 2018 health risk.

demand curves. However, I have more limited price variation along each plan dimension *separately*, rather than independently along each carrier-by-metal tier. In order to quantify heterogeneity in demand, I therefore turn to a structural model in Section 4 and rely on parametric assumptions to bridge the gap between the price variation in my setting and the ideal price variation.

4 Empirical Framework

Section 3.3 uses price variation in the Connector to provide evidence of heterogeneity in demand for two plan features: carrier quality and metal tier, by a patient's sickness and preference for Partners healthcare providers. In Section 4.1, I develop a structural model of plan demand that allows me to leverage the price variation described above into an estimate of demand for plan features that varies by a rich set of consumer observables; this acts as the empirical analog to $u_i(m,x)$ from Section 2.2. Section 4.2 then describes how I estimate a model of expected insurer costs, including estimating causal effects of changes to plan features on expected costs; this provides the empirical analog to $c_i(m,x)$ from Section 2.2.

4.1 Plan Demand Model

I use the plan choice dataset to estimate a multinomial logit model of plan demand. My model takes the timing of an individual's participation in the exchange as exogenous, and models each individual's plan choice decision.[17] These decisions are made at two times: when the consumer first enters the exchange, and each year at annual open enrollment periods when enrollees are allowed to switch plans, for as long as the enrollee continues to participate in the exchange.[18] My model assumes that enrollee i choosing among plans j at time t does so in order to maximize the following utility function:

$$u_{ijt} = \underbrace{\alpha\left(z_{it}^{p}\right)p_{ijt}}_{\text{Price}} + \underbrace{\underbrace{f\left(Carrier_{jt};z_{it}\right)}_{\text{Carrier utility (network)}} + \underbrace{g\left(MetalTier_{jt};z_{it}\right)}_{\text{Coverage tier}}}_{\text{Empirical analog to } E[u_i(m,x)|z_{it}]} + \underbrace{\eta_{jt}\left(Age_{it}\right)}_{\text{Plan dummies}} + \underbrace{\gamma\left(z_{it}\right)1_{ijt}^{SamePlan}}_{\text{Plan inertia}} + \epsilon_{ijt}. \tag{11}$$

[17]This follows the standard assumption in the individual exchange setting, e.g. in Shepard (2022) who studies the pre-ACA Massachusetts individual exchange. The assumption is that changes in price variation over time in the Connector affects only the decision of which plan to choose, *conditional on* participating, but does not affect the decision of whether to purchase insurance through the Connector.

[18]I exclude any enrollees who re-enter the exchange after a period of absence, due to ambiguity in the default enrollment rules for returning enrollees.

In addition to an individual-plan-year-specific type-I extreme value taste shock ϵ_{ijt}, plan utility depends on prices p_{ijt}, which are observed and vary according to an individual's age and location, carrier (NHP, BMC or Network) and metal tier (Bronze, Silver, Gold or Platinum), plan inertia γ for current enrollees (those already enrolled in the Connector at the time they are making their choice), and a rich set of carrier-by-region-by-age and metal-tier-by-region-by-age fixed effects. I allow price sensitivity $\alpha(z_{it})$ to vary by individual observables z_{it}^p, where z_{it}^p includes the enrollee's health risk and interactions of the enrollee's sex and age. I also allow carrier, metal tier and plan inertia to vary by by observables z_{it}, which includes enrollee's health risk, interactions of sex and age, interactions of distance to Partners hospitals and use of Partners providers prior to entering the exchange. While the APCD does not allow me to observe patient-level socioeconomic status, I also include measures of the patient's neighborhood (ZIP code) level SES including percent of the population that is Black, that is Hispanic, and that is below poverty in z_{it}.

Identifying Price Variation Identification of premium coefficients requires isolated variation in prices that is orthogonal to changes in unobserved plan quality or unobserved shocks to demand. Section 3.3 describes variation in Connector prices over time, and I argue that the change in relative NHP prices from 2016-2017, and the change in relative Silver prices in 2017-2018, are each due to idiosyncratic policy shocks that are unlikely to be correlated with underlying changes in plan quality or demand. However, over the period 2015-2018, there are also other sources of price variation, which may bias demand estimates if they are correlated with unobserved plan quality. In order to ensure that my estimation uses only variation due to the aforementioned policy changes, I use a detailed set of plan dummies, $\eta_{jt}(Age_{it})$, to absorb all variation in prices over time *except* for changes in the relative price of NHP between 2016-2017, and changes in the relative price of Silver in 2017-2018.[19]

The intuition is analogous to identification of difference-in-differences treatment effects using fixed effects. The fixed effects $\eta_{it}(Age_{it})$ can be thought of as absorbing all variation in prices over time that may be correlated with changes in unobserved plan quality or demand (i.e., all price variation over time except that in the price of carriers between 2016-2017, and that in the price of metal tiers between 2017-2018). Because prices in the Connector vary by a consumer's age (in years) according to a fixed pricing schedule, I scale these fixed effects by this schedule as a function of age. The level of prices in the Connector also varies by geographic region, so I construct a series of region-age-scaled

[19] See e.g. Nevo (2000).

fixed effects. I then allow these fixed effects to vary flexibly by carrier-year and metal tier-year, with the exception of years 2016-2017 for the set of carrier fixed effects, and years 2017-2018 for the metal tier fixed effects, which are each constrained to be equal within region-age.

Another source of price variation comes from the variation in prices of carriers conditional on metal tier and variation in prices of metal tier conditional on carrier, *within* region-age-year. For example, the relative price of Silver versus Bronze may vary between NHP and BMC, or the relative price of NHP versus Network may differ between Silver and Gold within the same region-year. My identification strategy assumes this price variation is orthogonal to differences in unobserved plan quality not already captured by the fixed effects described above. This assumption seems reasonable because the Connector strictly regulates the financial features of plans in each metal tier, such that plans offered in the same tier by different carriers have the same out-of-pocket price schedules (Appendix Figures 13 and 14 give examples). Additionally, each carrier offers Bronze, Silver, Gold and Platinum plans on their standard commercial network. Instead, this variation in prices is likely due to differences in insurer *costs* due to the differences in each carrier's provider networks. In sum, the identifying assumption is that the relationship beween prices and unobserved plan quality/demand follows the structure of the fixed effects described above.[20]

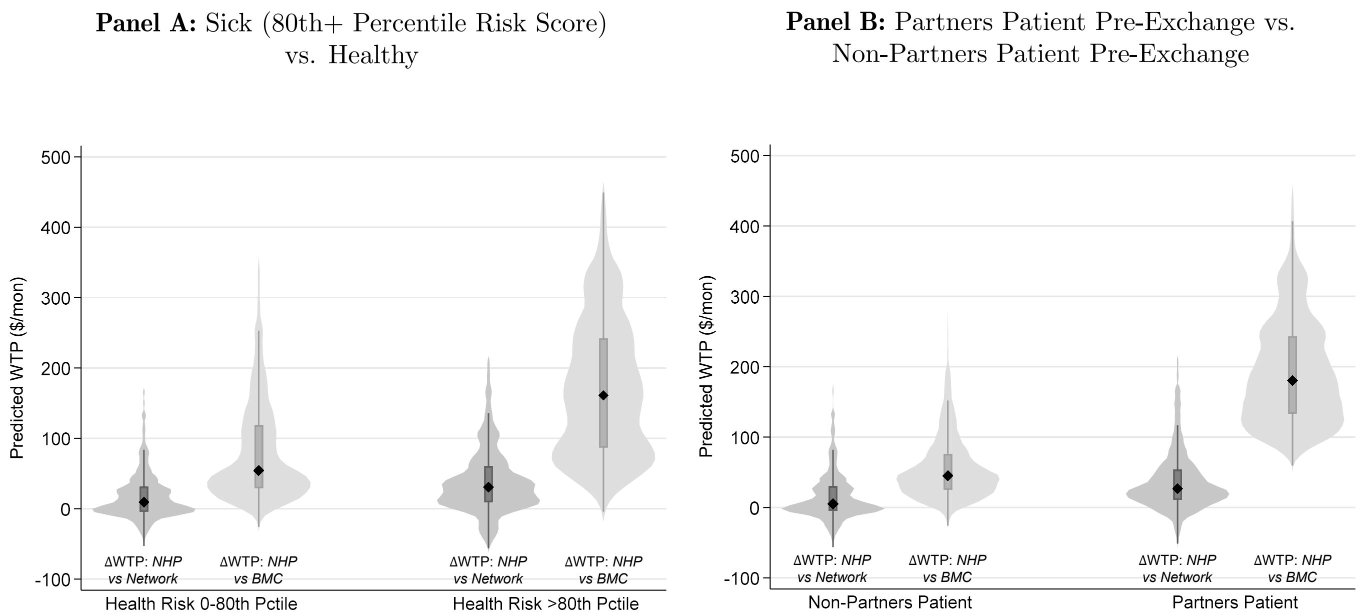
4.1.1 Plan Demand Estimation and Results

I estimate this demand model via maximum likelihood on the plan choice dataset described in Section 3.2.1. Plan premiums (in dollars per month) carry a negative coefficient across all enrollee types, but there is substantial heterogeneity in price sensitivity: the magnitude of the premium coefficient is smaller for older and sicker individuals, which is consistent with previous estimates in a similar context (see Shepard, 2022).

The estimates imply substantial heterogeneity in demand for insurance carriers. Patients perceive BMC as the lowest quality carrier conditional on financial coverage. Although BMC and Network have similar hospital networks, Network is viewed as similar in quality to NHP by the average patient, although sicker patients and patients with a history of Partners use are willing to pay a premium for NHP coverage. Network's health plan is branded as "Tufts Direct," and owned by a large regional insurance company – Tufts Health Plan – that also offers broad-network plans (some of which cover

[20] Essentially, the identification assumes plan utility at each choice instance is *additively separable* in unobserved carrier and metal tier quality. Appendix B.1 discusses this assumption in more detail.

Figure 7: Distribution of WTP for NHP relative to BMC and Network



Notes: Figure plots densities of ΔWTP_{ijt} (where WTP_{ijt} is defined in Equation 12) for changes in plan carriers across various groups. Panel A compares ΔWTP for NHP versus Network and NHP versus BMC for healthy (risk score in the bottom 80%) and sick (risk score in the top 20%) patients in the Connector. Panel B compares ΔWTP for NHP versus Network and NHP versus BMC for Partners (before entering the Connector) versus non-Partners (before entering the Connector) patients. Densities are estimated on the distribution of Connector member-months 2015-2017.

Partners) in the employer-sponsored insurance market, which may play a role in driving Network's high valuation relative to BMC. Figure 7 plots heterogeneity in predicted willingness-to-pay (WTP) for different carriers by sickness and past Partners use, based on the distribution of enrollees in the Connector 2015-2017 (years for which I observe all new and continuing enrollees), where I define predicted WTP as

$$WTP_{ijt} = \frac{-1}{\alpha\left(z_{it}^{p}\right)} \times \left[\underbrace{f\left(Carrier_j; z_{it}\right) + g\left(MetalTier_j; z_{it}\right)}_{V_{ijt}}\right]. \tag{12}$$

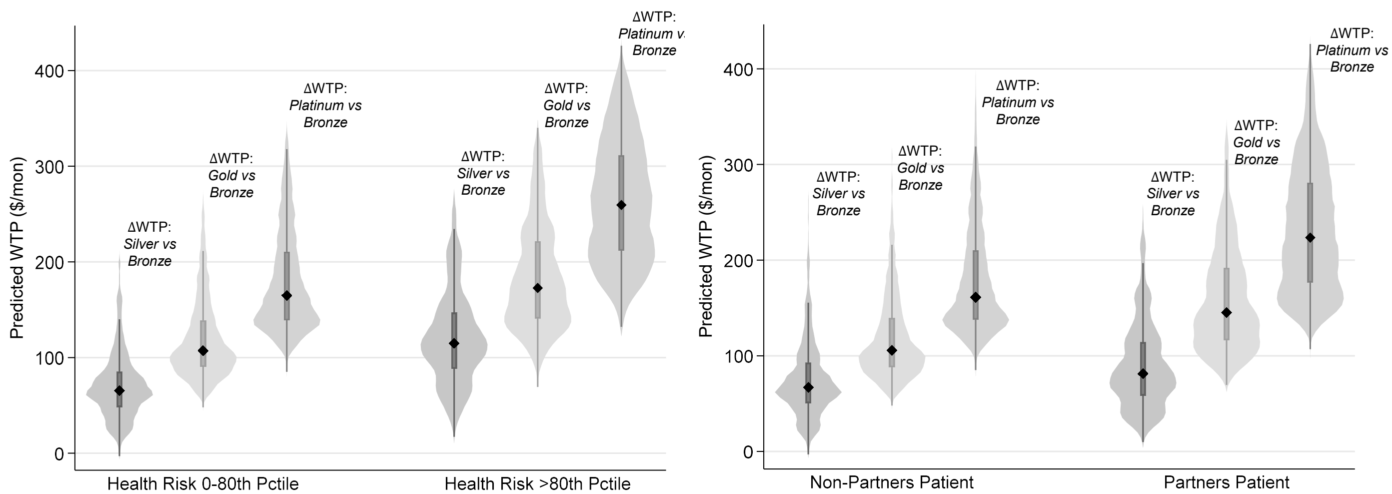
Multiplying by the inverse of the coefficient on monthly premium converts the predicted utility into units of dollars per month. Note that V_{ijt} excludes plan dummies, and so holds "unobserved" plan characteristics fixed across individuals making choices at difference points in time.[21]

[21]I normalize plan dummies to equal zero for 2017; the predicted WTP can thus be thought of as predicted WTP for 2017 plan characteristics (carrier quality and metal tiers).

Figure 8: Distribution of WTP for Metal Tiers

Panel A: Sick (80th+ Percentile Risk Score) vs. Healthy

Panel B: Partners Patient Pre-Exchange vs. Non-Partners Patient Pre-Exchange



Notes: Figure plots densities of ΔWTP_{ijt} (where WTP_{ijt} is defined in Equation 12) for changes in plan metal tier across various groups. Panel A compares ΔWTP for Silver versus Bronze, Gold versus Bronze, and Platinum versus Bronze for healthy (risk score in the bottom 80%) and sick (risk score in the top 20%) patients in the Connector. Panel B compares ΔWTP for Silver versus Bronze, Gold versus Bronze, and Platinum versus Bronze for Partners (before entering the Connector) versus non-Partners (before entering the Connector) patients. Densities are estimated on the distribution of Connector member-months 2015-2017.

Figure 8 plots heterogeneity in predicted WTP for metal tiers by sickness and past Partners use. Plan valuations are increasing in coverage tier across all groups, with sickness (top 20% of risk scores) being a strong driver of demand for more generous coverage. Being a past Partners patients is also a driver of demand for more generous coverage, especially for Gold and Platinum, but a relatively weaker driver of demand for Silver relative to Bronze.[22] This pattern, with sick and past Partners patients each preferring broad-network (NHP) coverage, while sick patients have a relatively stonger preference for Silver versus Bronze, is a key aspect of demand heterogeneity that will generate heterogeneity in relative preference between plans in a diagonally differentiated menu.

[22]Being a past Partners patient may be picking up heterogeneity in demand associated with unobserved socioeconomic factors that also drive demand for Gold and Platinum coverage, specifically.

4.2 Cost Model

The other input to the menu design problem is the distribution of insurer expected costs as a function of plan characteristics, denoted $c_i(m, x)$ in Section 2.2. I assume the distribution of expected insurer costs follows an affine transformation of the distribution of individual-level risk scores in the population of patients in the Connector.[23]

The other input required to define an empirical analog to $c_i(m, x)$ is the causal effect of changes in plan characteristics on an insurer's expected costs of insuring a given individual. Here, I assume a proportional model of insurer costs, to reflect the idea that the incremental cost *to the insurer* of a change in coverage level or provider network is proportional to the underlying sickness of each patient. This is captured by the empirical specification

$$EC_{ijt} = exp\left(\alpha_i + \beta r_{it} + \gamma_t + \delta^{covg} CS_{jt} + \lambda^{ntwk}(z_i)\, 1_{jt}^{NHP}\right), \tag{13}$$

for the insurer's expected cost of covering individual i on plan j at time t. The model includes an individual-specific (and time-invariant) individual fixed effect α_i, the individual's risk score r_{it}, and time-specific fixed effects γ_t.

The effect of an increase in coverage tier is captured by δ^{covg}; I assume that the proportional effect of an increase in coverage tier is linear in the change in cost-sharing tier CS_{jt}, that is, the proportional effect of Gold versus Silver is the same as the proportional effect of Platinum versus Gold, etc. The change in costs implied by δ^{covg} captures two separate cost effects of increased coverage generosity: the direct effect of the increase in the plan's actuarial value,[24] and "moral hazard," the increase in patient utilization due to a decrease in the out-of-pocket price of care. From the perspective of the menu design problem described in Section 2.2, these two channels enter the problem additively and so I treat them together.

Similarly, I assume that the broad network plan, NHP, has a proportional effect on expected insurer cost, captured by $\lambda^{ntwk}(z_i)$. Conceptually, $\lambda^{ntwk}(z_i)$ may reflect a combination of a price effect and a care intensity effect: patients on broad network plans may use a mix of providers with higher average

[23]These risk scores are constructed using diagnoses observed in concurrent medical claims, and represent a unitless scale measure of expected insurer costs; that is, the ratio of risk scores between any two patients approximates the ratio in expected costs given observed diagnoses. I construct these risk scores using the HSS-HCC method based on current year's claims data.

[24]The ACA's metal tiers correspond to actuarial values: Bronze ˜0.6, Silver ˜0.7, Gold ˜0.8 and Platinum ˜0.9, which represents the typical share of medical costs paid by the insurer.

prices, as well as providers that treat the patient *more intensively* (e.g., ordering more diagnostic tests), conditional on the patient's utilization (e.g., number of doctor's office visits). Because the cost effect of broad network coverage is likely to vary both by a patient's access to expensive/intensive providers, and the patients preference for those providers, I allow $\lambda^{ntwk}(z_i)$ to vary by observables z_i: interactions of the patient's distance from the nearest Partners hospital, and past use of outpatient Partners providers prior to enrolling in the Connector.

Estimation and Identifying Variation Identification of the key parameters δ^{covg} and $\lambda^{ntwk}(z_i)$ comes from within-individual variation in plan characteristics over time among *plan switchers*. The primary concern for this identification strategy is that plan switching in the Connector is not exogenous; enrollees who experience a large change in unobserved health state may be most likely to switch plans (along both cost-sharing and network breadth dimensions). It is therefore a natural concern that estimates of incremental cost are likely upwards-biased. To minimize concerns due to selection into plan switching, I restrict estimation of δ^{covg} and $\lambda^{ntwk}(z_i)$ to subsamples where plan switching is predominantly driven by (plausibly) exogenous changes in relative plan prices. These subsamples are different for network-switchers and coverage tier-switchers, respectively, and so in practice I estimate each set of parameters separately, using Poisson regression with individual fixed effects ("xtpossion, fe" in Stata).[25]

Network switchers (estimation of $\lambda^{ntwk}(z_i)$) I estimate $\lambda^{ntwk}(z_i)$ on the unbalanced panel of patients who are continuously enrolled from the end of 2016 to the beginning of 2017, and either

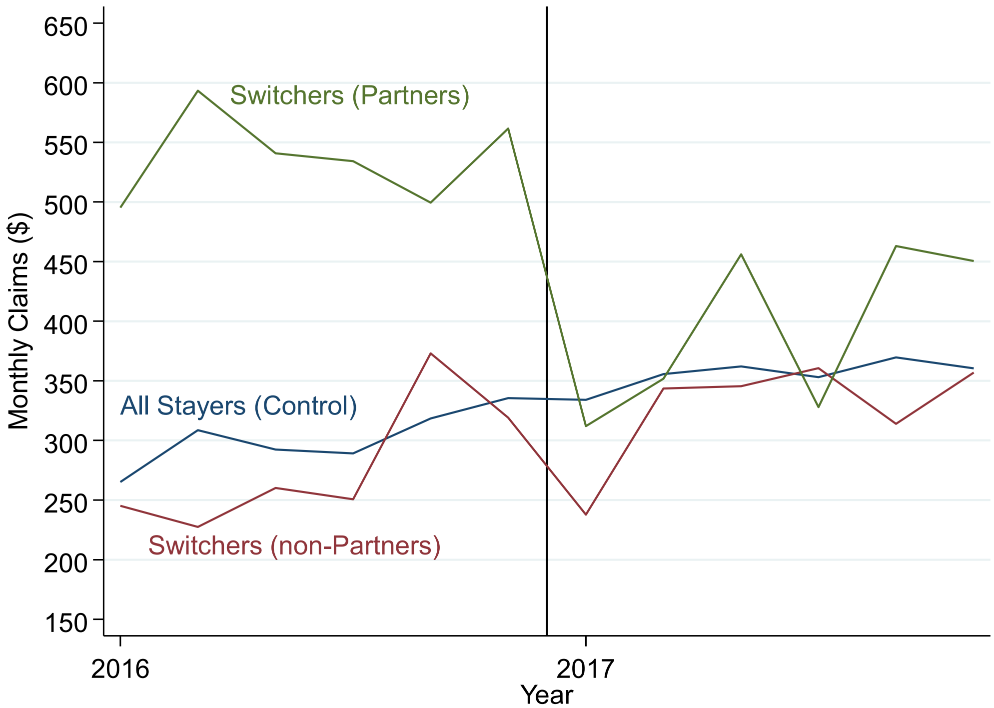
1. Switch from the broad-network plan (NHP) to narrow network plans (BMC or Network Health) when the price of the broad-network plan increases in 2017, or

2. Stay enrolled in the same plan (broad or narrow network) in both years.

I further restrict to ConnectorCare Plan Type III (200-300% FPL) enrollees to ensure that that there is no concurrent variation in cost-sharing plan features among either group. The predominance of plan switching in this sample is due to the change in price: 21.8% of broad-network enrollees switch to a narrow network plan in 2017, relative to an average plan-switching rate of 1.4% among enrollees

[25]Poisson regression estimators are consistent for multiplicative models such as equation 13, even if the underlying data (in this case, monthly medical spending) do not follow a Poisson distribution. I report heteroskedasticity-robust standard errors (clustered at the individual-level) to account for mis-specification of the conditional variance of medical spending. See Santos Silva and Tenreyro (2006) for a detailed discussion.

Figure 9: Identifying Variation for Network Cost Effects



Notes: This figure shows event study estimates of $exp\left(\lambda_t^{ntwk}\left(z_i\right)\right)$ for three groups z_i: switchers who used Partners in 2016 (green line), switchers who did not use Partners in 2016 (red line) and non-switchers (blue line). The event study coefficients have also been scaled by the 2016 monthly average of insurer cost of each group.

in other carriers. Identification of $\lambda^{ntwk}\left(z_i\right)$ comes from the year-on-year change in insurer costs for patients who switch from broad- to narrow-network coverage, relative to the control group of plan stayers; this is analogous to the identification of a difference-in-difference treatment effect.

I find that broad-network coverage increases expected insurer costs by an average of 28.5% (s.e. 9.78%). Consistent with predictions that $\lambda^{ntwk}\left(z_i\right)$ is driven by patients' use of expensive and intensive providers, I find that the network cost effect is driven almost exclusively by patients who used outpatient Partners providers in 2016, prior to switching to narrow-network coverage: Partners users see a 61.4% (s.e. 18.1%) increase in expected costs when enrolled in the broad-network plan relative to a narrow-network, while non-Partners patients see a statistically insignificant 3.4% (s.e. 10.7%) increase in expected cost.[26] Partners use on the exchange is only observable for patients already enrolled in NHP; I therefore estimate heterogeneity in $\lambda^{ntwk}\left(z_i\right)$ accordig to whether each patient used outpatient Partners providers before enrolling in the exchange. I find that pre-exchange Partners patients incur

[26]This finding also provides some reassurance that the estimates of $\lambda^{ntwk}\left(z_i\right)$ are not driven by selection on shocks to unobserved health – if that were the case, one should expect to see costs fall for non-Partners patients who switch plans.

a 35.8% (s.e. 15.9%) greater expected costs under broad-network coverage compared to 22.4% (s.e. 11.7%) for patient who did not use Partners prior to enrolling in the Connector. Appendix Table 3 shows the full Poisson regression results, including my final specification, which estimates the effect of NHP coverage separately by whether the patient used Partners before enrolling in the Connector, interacted with categorical measures of the patient's distance from the nearest Partners hospital.

Metal tier switchers (estimation of δ^{covg}) I focus on an unbalanced panel of broad-plan (NHP) enrollees continuously enrolled from the end of 2016 to the beginning of 2017 in order to estimate δ^{covg}. In 2017, Neighborhood Health Plan raised the price of its plans across all metal tiers, but raised its price *less* for its Silver-plan option relative to Bronze, Gold and Platinum. This was likely for strategic reasons, since the price of its ConnectorCare option (which accounts for a large share of its overall enrollment) is linked to the price of its Silver plan. The price of NHP Gold also fell relative to NHP Platinum. As a result, a number of NHP enrollees switch from non-Silver NHP to Silver NHP plans, and from NHP Platinum to NHP Gold: NHP enrollees switch across metal tiers at a 7.1% rate in 2017, compared to a rate of 1.1-1.5% among other carriers. I therefore construct an unbalanced panel consisting of

1. Neighborhood (NHP) enrollees who switch from Bronze, Gold or Platinum to Silver in 2017, as well as NHP enrollees who switch from Platinum to Gold, and

2. NHP enrollees who stay in the same coverage tier in 2017.

My estimate of δ^{covg} implies an increase in one coverage tier (e.g., Silver to Gold, or Gold to Platinum) increases insurer expected costs by 28.1% (s.e. 9.86%). For comparison, the ACA's risk adjustment formulas, including assumed "induced demand factors" that account for moral hazard, assume a proportional effect that ranges between 19.8-20.2% cost increases increase in metal tier. I also compute back-of-the-envelope implied arc elasticities of total claims to out-of-pocket price, approximating out of pocket price by $1 - AV$, and total claims by $\frac{1}{AV} \times \{\text{insurer cost}\}$, where AV represents the actuarial value of the metal tier. For example, the implied arc elasticity of demand of going from Platinum ($AV = 0.9$) to Silver ($AV = 0.7$) is -0.243.[27]

[27] The arc elasticity is given by the ratio of the percent change in total claims relative to the average, $\frac{\frac{1}{0.7}1.281 - \frac{1}{0.9}1.281^2}{\frac{1}{2}\left(\frac{1}{0.7}1.281 + \frac{1}{0.9}1.281^2\right)} \approx -0.243$ to the percent change in out-of-pocket price, $\frac{(1-0.7)-(1-0.9)}{\frac{1}{2}((1-0.7)+(1-0.9))} = \frac{0.3-0.1}{\frac{1}{2}(0.3+0.1)} = 1$.

5 Analysis: Menu Design

I study the menu design problem by taking the conceptual framework outlined in Section 2 to the distribution of preferences and costs estimated in Section 4. I analyze a counterfactual setting in which a market sponsor faces a two-step problem. In the first step, the sponsor admits up to two insurance constracts.[28] In the second step, the sponsor sets the relative price of the two plans in order to achieve the optimal allocation of patients to plans, under the constraint that this price must be uniform: all individuals pay the same price. First, I define the sponsor's objective function, which depends on both the characteristics of the plans in the sponsor's design space, and conditional on the design space, the allocations (constrained to be feasible with uniform plan prices) of enrollees to plans. Under a given set of plans and at a given price Δp, I evaluate the sponsor's objective defined in Equation 1 (unweighted social surplus). My framework identifies *relative* preference between plans, but does not identify absolute willingness-to-pay for insurance relative to being uninsured. I therefore report plan valuations relative to a baseline plan: the Silver-tier broad-network (NHP) plan.

Interpretation of Logit Term ϵ_{ijt} The unobserved logit error term is typically interpreted as capturing unobserved preference heterogeneity. Because the specification assumes draws of the unobserved preference term are i.i.d. across plans j for each choice instance, the logit specification is notorious for implying "new product" welfare effects which may overstate the value of expanding the choice set. This provides additional motivation to restricting analysis to two-plan menus, where the size of choice sets is held constant across counterfactual menus. Another challenge for interpretation of the logit distribution, which is specific to the ACA marketplace setting, is that there may be unobserved differences in the *price* faced by consumers due to external premium subsidies, e.g. employer reimbursements.[29] Variance in unobserved premium reimbursements would tend to make interpreting the logit distribution as unobserved preferences overstate the extent of preference heterogeneity. I therefore report results using two welfare objectives. The first has the sponsor interpret the logit term as unobserved preference and include the distribution of ϵ_{ijt} in its objective, and another in which the planner only uses the predicted (average) WTP from Equation 12 to evaluate welfare.[30] I consider these approaches

[28] Restricting to two-plan menus allows me to evaluate the efficiency of choice with straightforward graphical and computational analysis. Larger menus with potentially $N \geq 3$ plans are computationally more cumbersome without providing additional intuition.

[29] Such reimbursements were initially prohibited under the ACA, but became legal for small employers in 2017.

[30] Notably, the predicted WTP excludes preference heterogeneity attributed to predictable plan inertia. Inertia plays an important role in identifying preferences for plan characteristics, but interpreting it as preference heterogeneity runs

as giving upper and lower bounds on the value of choice.[31]

For my main results, I restrict the space of possible menus to *two-plan* subsets of

$$\mathbb{P} \equiv \{\text{BMC (Narrow), NHP (Broad)}\} \times \{\text{Bronze, Silver}\}.$$

Restricting to BMC versus NHP most closely approximates the conceptual setting where provider networks are vertically differentiated, and I restrict differentiation in coverage to Bronze and Silver because those are the most popular metal tiers in the Connector.

5.1 Vertical Differentiation

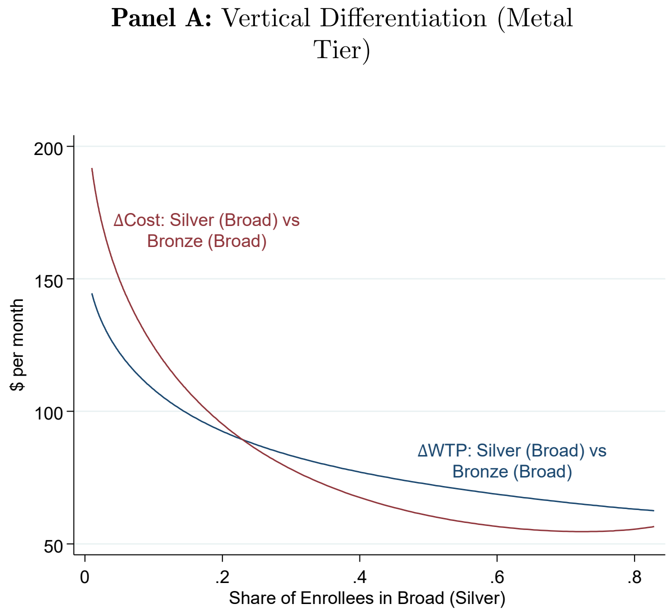
Suppose, counterfactually, that all plans had the same provider network, but could vary in their out-of-pocket price schedule. Would a social planner offer choice between vertically differentiated plans? Panel A of Figure 10 shows incremental willingness-to-pay and incremental cost curves for vertical differentiation in coverage level. These curves are constructed by tracing out the set of allocations between two plans, in this case Bronze-Broad and Silver-Broad, that are created by varying the relative price of the two plans. The curves plot the average incremental cost, and average incremental WTP of Silver relative to Bronze for the population of patients on the margin at each allocation. Several things are apparent from the figure. The steep downward slope of the incremental cost ($\Delta Cost$) curve indicates high selection on incremental cost. In fact, at higher prices, Silver is taken up by patients for whom the relative cost of providing them with Silver on average exceeds their relative willingness-to-pay for Silver, which is shown graphically by the $\Delta Cost$ curve crossing the ΔWTP curve from above. As a result, there is no price at which the market sponsor efficiently offers choice between Bronze and Silver plans; the sponsor's best allocation is achieved by restricting enrollees to Silver.

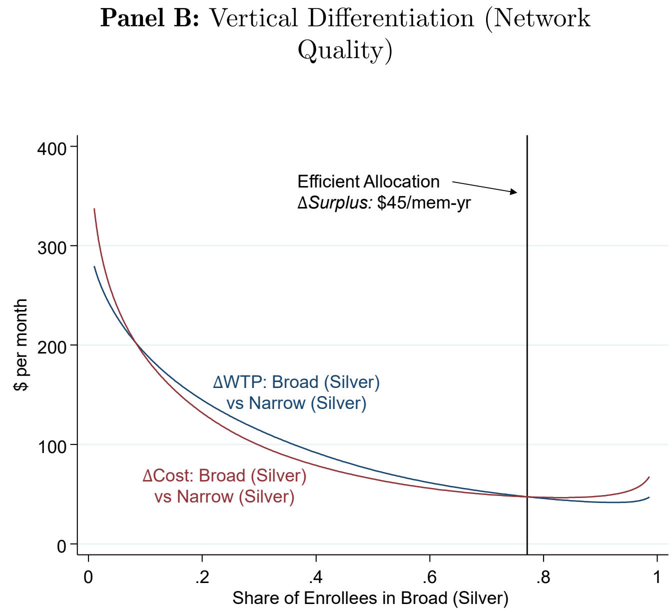
Alternatively, I can consider the counterfactual world in which plans can vary in their network quality, but all have the same out-of-pocket price schedule. Panel B of Figure 10 plots $\Delta Cost$ and ΔWTP for vertical differentiation in network quality. Again, differentiation in network quality induces selection on incremental costs. However, in this case, there is an interior allocation that maximizes the sponsor's objective while offering choice over plans with different networks: this can be achieved

into the well-known problem of distinguishing preference heterogeneity from state dependence. I therefore remove the plan inertia term from my welfare calculations, in effect treating all enrollees as if they are first-time enrollees.

[31]Non-welfare-relevant logit errors effectively act as noise in the assignment of plans; in the extreme case where plan assignment is entirely uncorrelated with predicted WTP, offering choice would be weakly worse than only offering the plan that is most efficient on average.

Figure 10: Vertical Differentiation: Incremental Benefit and Cost Curves





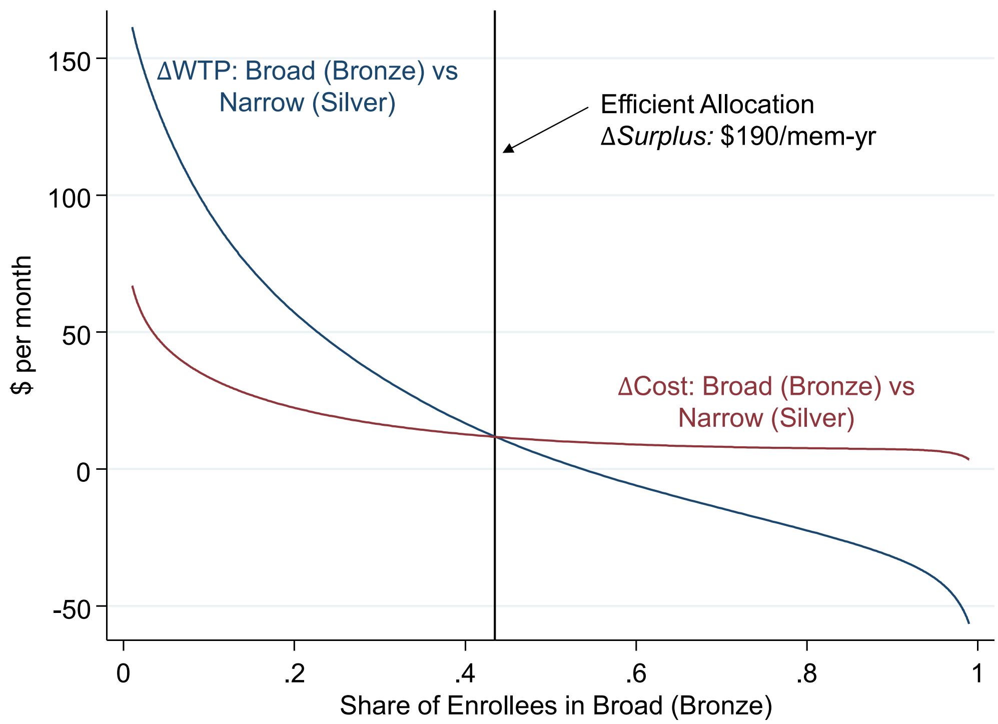
Notes: The figures plot the average $\Delta Cost$ (red) and ΔWTP (blue) of marginal enrollees against the allocations generated by varying the relative price of two vertically differentiated plans. ΔWTP excludes the logit term from the sponsor's objective. Panel A gives relative cost and relative demand curves for allocations between Silver-Broad and Bronze-Broad generated by $p_{Silver} - p_{Bronze} \geq 0$. Even at $p_{Silver} = p_{Bronze}$, a non-zero share of enrollees choose Bronze; this is a consequence of the logit demand specification. Panel B plots analogous curves for allocations between Silver-Broad and Silver-Narrow.

with a relative price of \$12 per month (\$144 per year), which induces 77.1 percent of enrollees to take up the Silver-Broad plan.

5.2 Diagonal Differentiation and the Optimal Two-Plan Menu

I now consider menus that are differentiated along both coverage and network dimensions simultaneously. Figure 11 plots $\Delta Cost$ and ΔWTP curves for a diagonally differentiated menu consisting of a Bronze-tier broad-network plan and a Silver-tier narrow-network plan. The key qualitative difference, compared with the examples of vertical differentiation shown above, is that this menu generates less selection on incremental cost, which can be seen from the relatively shallow slope of the $\Delta Cost$ curve. At the same time, this menu offers variety that appeals differently to different patient types. The planner's best allocation offers choice between these two plans with a relative price of \$24 per month (\$288 per year) resulting in 43.5 percent of enrollees choosing the Bronze-Broad plan.

Figure 11: Diagonal Differentiation



Notes: This figure plots $\Delta Cost$ and ΔWTP (excluding logit term in sponsor's objective) for marginal enrollees against the allocations generated by varying the relative price of Bronze-Broad and Silver-Narrow plans. ΔWTP excludes the logit term from the sponsor's objective.

Diagonal differentiation can reduce selection on incremental costs through two conceptual channels, highlighted in Section 2.1.3: reducing the *scope* of selection on incremental costs by reducing the variance in Δc, or by reversing selection by changing the nature of selection, given by β. In this case, the effect comes through the former channel – the diagonally differentiated menu results in substantially smaller *variance* in the incremental costs compared to the vertically differentiated menu. The mechanism is the offsetting plan features: broad networks and more generous cost-sharing both cost more to provide to sick patients, so the difference between healthy and sick patients in the relative cost of providing the Bronze-Broad compared to the Silver-Narrow plan is reduced.

Optimal menu I evaluate the sponsor's optimal allocation under all possible two-plan menus to find the sponsor's optimal menu. Table 1 gives the results. Panel A characterizes the efficient allocations when the sponsor's objective excludes the logit term. The diagonally differentiated menu described above (Bronze-Broad and Silver-Narrow) is the optimal two-plan menu, generating social surplus of approximately \$190.3 per member-year more than a single-plan menu consisting of only a Silver-Broad plan, and roughly \$145.3 per member-year more than the best vertically differentiated menu, which offers choice between Narrow and Broad Silver plans.

6 Constraints on Offering Choice: Adverse Selection

In the baseline counterfactual analysis, I have assumed the market sponsor faces no constraints on the relative price it sets between plans, other than that the price must be uniform. However, in some settings, plan prices may be subject to market forces. In insurance markets, adverse selection is a first-order concern, and in the context of the menu design problem can be considered as a constraint on the sponsor's ability to set prices (and thus allocations) in the second stage. It is therefore interesting to consider whether accounting for an adverse selection constraint would change the sponsor's menu design decision. To simplify analysis, I focus on the first- and second-best menus from Section 5.2: a diagonally differentiated menu offering choice between a Bronze-Broad and Silver-Narrow plan, and a vertically differentiated menu offering choice between a Silver-Broad and Silver-Narrow plan.

Equilibrium I consider adverse selection resulting from competition between insurers in the second stage, after the set of contracts has been determined by the market sponsor in the first stage of the

Table 1: Efficient Allocations in Two-Plan Menus

	Two-Plan Menus			Characteristics of Efficient Allocation			
				(1)	(2)	(3)	(4)
	Narrow (Tier)	vs	Broad (Tier)	Offers Choice?	Broad Plan Share	Rel. Price Broad ($/yr)	Rel. Surplus ($/member-yr)
A. Planner's Objective Excludes Logit							
i	Narrow (Bronze)		Broad (Bronze)	No	1	N/A	-180.0
ii	Narrow (Silver)		Broad (Silver)	Yes	0.771	144	45.0
iii	Narrow (Bronze)		Broad (Silver)	No	1	N/A	0
iv	Narrow (Silver)		Broad (Bronze)	Yes	0.435	288	190.3
B. Planner's Objective Includes Logit							
i	Narrow (Bronze)		Broad (Bronze)	Yes	0.661	480	257.1
ii	Narrow (Silver)		Broad (Silver)	Yes	0.595	672	310.8
iii	Narrow (Bronze)		Broad (Silver)	Yes	0.811	756	156.1
iv	Narrow (Silver)		Broad (Bronze)	Yes	0.487	132	474.6

Notes: Table describes characteristics of the optimal allocation within two-plan menus described in each row. Columns 1-3 describe the efficient allocation: whether it offers choice (Column 1), what share of enrollees are assigned the Broad-network plan (Column 2), and the Δp for the Broad-network plan which achieves the allocation. Column 4 gives the relative social surplus of the optimal allocation, evluated relative to a single-plan menu consisting of only a Silver-tier Broad-network plan. Panel A characterizes the efficient allocation when the sponsor's objective excludes the logit error term, while Panel B characterizes the efficient allocation in the same set of menus under an objective that includes the logit error term.

menu design problem. I compute equilibrium following the notion of price competition of Handel, Hendel and Whinston (2015) between insurers offering one of two pre-determined insurance contracts. Insurers earn zero profits in equilibrium, meaning the equilibrium is characterized by the break-even price, where the relative price between plans equals the difference in average insurer costs.

Risk Adjustment Market sponsors may also use policies in the second stage to counteract the price distortions resulting from adverse selection. I consider one such policy tool: risk adjustment. Under risk adjustment, the sponsor compensates plans that enroll costlier (sicker) patients. Risk adjustment reduces the difference in profitability between sicker and healthier patients, and in a setting with price competition and adverse selection, reduces the relative price between plans that results in equilibrium. I allow for a simple form of risk adjustment, which is similar to that used in the ACA: the risk adjustment payments to plans are proportional to the patient's concurrent (ex-post) risk score based on observed diagnoses on medical claims. However, the sponsor does not scale risk adjustment payments variably based on the insurance contract's provider network. Instead, I assume the sponsor uses scales risk adjustment payments scaled to the expected cost (risk score) of patients under the narrow-network plan. Under this risk adjustment scheme, the broad-network insurer's profits essentially internalize the savings of the narrow-network that result from changes in the relative price of the broad-network plan.

6.1 Competitive Equilibrium with Risk Adjustment

I find both the diagonal and vertical menus suffer from adverse selection, with the broad-network plan attracting sicker patients than the narrow-network plan in both cases. However, adverse selection is more severe for the vertical menu than the diagonal menu. Table 2 describes the equilibrium allocation in two-plan menus with ACA-like risk adjustment. The diagonal menu (row iv of Table 2) achieves an equilibrium allocation nearly as efficient as the optimal allocation, generating close to 100 percent of the surplus of the optimal allocation.

However, the vertical menu (row ii) generates a much smaller surplus in equilibrium than at the efficient allocation, due to adverse selection. When the sponsor's objective does not include the logit term (shown in Column 3), the vertical menu actually generates *less* surplus than restricting choice to the baseline Broad (Silver) plan. When the sponsor's objective includes logit, the equilibrium allocation generates only $\frac{107.5}{310.8} \approx 34.6$ percent of the surplus at the efficient allocation.

Table 2: Efficiency of Equilibrium Allocations in Two-Plan Menus

	Two-Plan Menus			Characteristics of Equilibrium Allocation			
				(1)	(2)	(3)	(4)
	Narrow (Tier)	vs	Broad (Tier)	Broad Plan Share	Rel. Price Broad ($/yr)	Rel. Surplus ($/member-yr) excl. Logit	Rel. Surplus ($/member-yr) incl. Logit
i	Narrow (Bronze)		Broad (Bronze)	0.328	1572	-103.0	138.0
ii	Narrow (Silver)		Broad (Silver)	0.107	2988	-27.9	107.5
iii	Narrow (Bronze)		Broad (Silver)	0.003	8208	-264.3	-255.2
iv	Narrow (Silver)		Broad (Bronze)	0.427	312	190.3	469.4

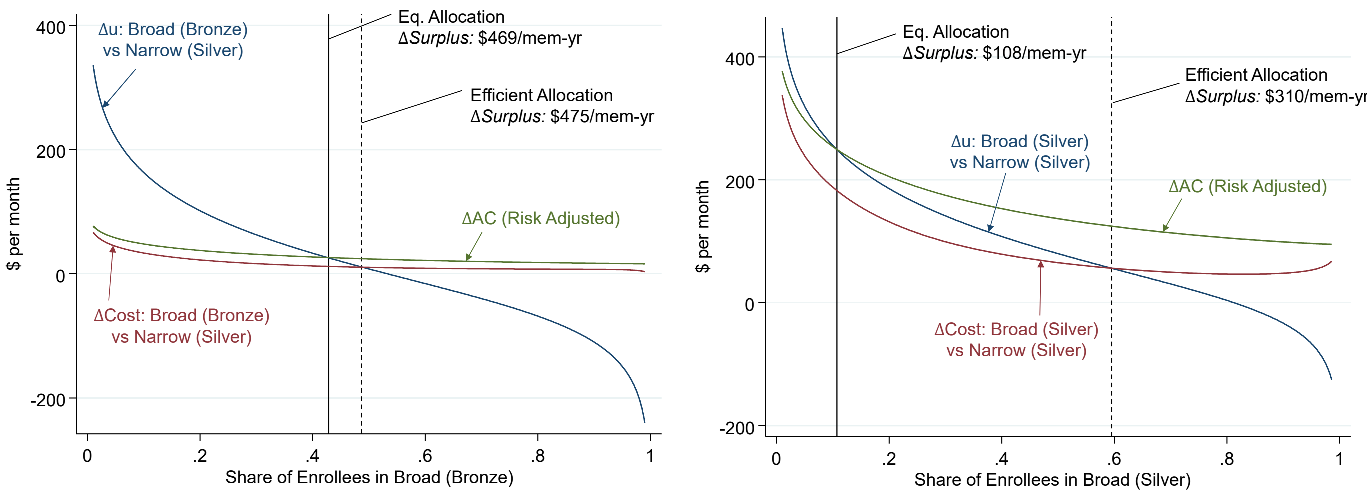
Notes: Table describes characteristics of the equilibrium allocation (with ACA-like risk adjustment) within two-plan menus described in each row. Columns 1 and 2 describe the allocation and equilibrium price (in $/year), respectively. Column 3 gives the relative social surplus of the optimal allocation, compared to a Silver-tier Broad-network plan, when the sponsor's objective is evaluated excluding the logit distribution term. Column 4 evaluates the *same allocation* relative to a Silver-tier Broad-network plan under a sponsor's objective that includes the logit term.

The diagonal menu has two advantages that contribute to its stability under adverse selection, which are analogous to the channels through which diagonal differentiation reduces selection on incremental cost. The first is that offsetting choice between two quality dimensions may reduce the extent of selection: if broad networks and generous coverage are both attractive to sicker (costlier) patients, then a diagonal menu may not suffer as much from adverse selection as a vertically differentiated menu. In this case, however, sicker patients tend to prefer network breadth relative to the difference in coverage between Silver and Bronze coverage. The diagonal menu has another advantage, which is that sick patients, although costlier, cost *less* to cover under a Bronze-Broad plan than under a Silver-Broad plan. This is clearly illustrated in Figure 12, which plots incremental demand (Δu, including the logit distribution term), incremental cost, and risk-adjusted Δ-average cost curves for the diagonally differentiated menu and the compares efficient and equilibrium allocations under diagonal differentiation and under vertical differentiation in network quality. Under identical (imperfect) risk adjustment schemes, the diagonal menu experiences much less adverse selection because offsetting plan features result in a much flatter incremental (marginal) cost curve.

In this sense, one can think of the diagonal differentiation as acting as a built-in form of risk adjustment, by compensating the broad-network provider indirectly (by allowing it to cover a smaller share of costs) rather than directly (through more sophisticated risk adjustment). The idea of combating adverse selection through menu design in this way dates back at least to Enthoven (1988).

Figure 12: Adverse Selection in Two-Plan Menus

Panel A: Diagonal Differentiation **Panel B:** Vertical Differentiation (Network)



Notes: This figure plots $\Delta Cost$ (red), Δu (blue) – the estimated (incremental) demand curve including the logit distribution term – and ΔAC (green), the risk-adjusted difference in average costs between plans, for two menus. Panel A shows a diagonally differentiated menu, and Panel B shows a vertically differentiated menu in network quality (both Silver-tier plans). Each Panel illustrates the *efficient* allocation, where incremental demand equals $\Delta Cost$, and the *equilibrium* allocation, where incremental demand intersects the risk-adjusted ΔAC curve. The value of choice is larger when welfare is evaluated using Δu (inclusive of the unobserved logit distribution term) than when using ΔWTP (using only predicted demand as in Figures 10 and 11), due to the additional preference heterogeneity implied by the logit distribution.

7 Conclusion

The economics of choice in selection markets such as health insurance are not straightforward. Offering choice may be inefficient when patients do not internalize the costs of providing them with different forms of coverage. Furthermore, market failures such as adverse selection may interfere with a market sponsor's ability to offer choice. This paper considers the menu design problem in a setting with multiple plan quality dimensions. Regulating plan features along multiple dimensions may allow market sponsors to offer more efficient choice, and to do so despite adverse selection and imperfect risk adjustment. These findings suggest potential avenues for menu design in a number of important policy settings, such as Medicare Advantage and the ACA exchanges, as well as providing potential economic explanations for the characteristics of menus in employer-sponsored health insurance settings. However, there are many areas where additional study may prove fruitful. Key questions include how menu design policies interact with market entry decisions of insurers when plan features are endogenous, as in the literature on provider network formation. When a health plan's provider network is a key quality attribute, this especially includes effects of menu design on the negotiation between insurers and providers.

An interesting implication of this paper is that it provides two economic explanations for (anecdotal) evidence that diagonally differentiated menus are common in the employer-sponsored insurance setting, which may have different implications for efficiency. The first is that diagonal differentiation is an effective way to provide efficient choice, and that employers are sophisticated maximizers of surplus. A second is that employers are unsophisticated risk adjusters, and diagonal differentiation is an effective way to combat adverse selection. The second explanation may hold even in situations where diagonal differentiation is less efficient than other kinds of menus. I view disentangling these two explanations as interesting ground for future research.

References

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

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

Bundorf, M. Kate, Jonathan Levin, and Neale Mahoney, "Pricing and Welfare in Health Plan Choice," *American Economic Review*, 2012, *102* (7), 3214–3248.

Cutler, David M. and Sarah J. Reber, "Paying For Health Insurance: The Trade-off Between Competition and Adverse Selection," *The Quarterly Journal of Economics*, May 1998.

Dafny, Leemore, Igal Hendel, and Nathan Wilson, "Narrow Networks on the Health Insurance Exchanges: What Do They Look Like and How Do They Affect Pricing? A Case Study of Texas," *American Economic Review: Papers and Proceedings*, 2015, *105* (5), 110–114.

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

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

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

Enthoven, Alain C., *Theory and Practice of Managed Competition in Health Care Finance*, Elsevier, 1988.

Ericson, Keith Marzilli and Amanda Starc, "Measuring Consumer Valuation of Limited Provider Networks," *American Economic Review*, 2015, *105* (5), 115–119.

_ and _ , "Pricing Regulation and Imperfect Competition on the Massachusetts Health Insurance Exchange," *Review of Economics and Statistics*, 2015, *97* (3), 667–682.

— and — , "How Product Standardization Affects Choice: Evidence from the Massachusetts Health Insurance Exchange," *Journal of Health Economics*, 2016, *50*, 71–85.

Finkelstein, Amy and James Poterba, "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market," *Journal of Political Economy*, 2004, *112* (1).

— **and Kathleen McGarry**, "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market," *American Economic Review*, September 2006, *96*, 938–958.

— **, Nathaniel Hendren, and Erzo F.P. Luttmer**, "The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment," *Journal of Political Economy*, 2019, *127* (6).

Geruso, Michael, Timothy Layton, Grace McCormack, and Mark Shepard, "Two Margin Problem in Insurance Markets," *Review of Economics and Statistics*, 2021, *Forthcoming.*

Hackmann, Martin B., Jonathan T. Kolstad, and Amanda E. Kowalski, "Adverse Selection and an Individual Mandate: When Theory Meets Practice," *American Economic Review*, 2015, *105* (3), 1030–1066.

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

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

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

Landais, Camille, Arash Nekoei, Peter Nilsson, David Seim, and Johannes Spinnewijn, "Risk-Based Selection in Unemployment Insurance: Evidence and Implications," *American Economic Review*, 2021, *111* (4), 1315–1355.

— **, Nathaniel Hendren, and Johannes Spinnewijn**, "Choice in Insurance Markets: A Pigouvian Approach to Social Insurance Design," *Annual Review of Economics*, 2022.

Marone, Victoria and Adrienne Sabety, "When Should There Be Vertical Choice in Health Insurance Markets?," *American Economic Review*, 2022, *112* (1), 304–342.

Nevo, Aviv, "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics*, 2000, *31* (3), 395–421.

Santos Silva, J.M.C. and Silvana Tenreyro, "The Log of Gravity," *The Review of Economics and Statistics*, November 2006, *88* (4), 641–658.

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

Tilipman, Nicholas, "Employer Incentives and Distortions in Health Insurance Design: Implications for Welfare and Costs," *American Economic Review*, 2022, *112* (3), 998–1037.

A APCD Details and Data Construction

A.1 Enrollment and Claims Data

The APCD allows me to observe or infer a number of key variables, including:

- **Enrollee Demographics:** the APCD includes variables describing each individual's household membership, age, sex, and location (zip code). Based on each enrollee's zip code, I construct measures of the average distance to a number of key Partners hospitals.

- **Plan Characteristics:** based on an enrollee's household income, he/she may be eligible for premium and cost-sharing subsidies through the ConnectorCare program. The APCD does not include information on individual- or household-level income, but does include information on plan characteristics including each plan's actuarial value. Based on the exchanges income-based subsidy schedule, I infer subsidy eligibility from the actuarial value of the each enrollee's plan.

- **Individual Health Risk:** sickness is likely a key driver of plan selection. I construct a measure of each enrollee's health ("risk score") using medical diagnoses listed on medical claims in the APCD, following the HHS-HCC risk score definition. The HCC risk score is a *concurrent* measure of risk, based on each patient's current-year medical diagnoses. While I observe plan enrollment for continuing enrollees in 2018, the APCD claims I have include only 2013-2017. For plan choices made in 2018, I use each patient's 2017 risk score as a proxy for health at the time of choice.

- **Provider Usage History:** an advantage of the APCD is that it allows me to observe both enrollment and claims data for individuals *outside* of the exchange. I use claims for enrollees before they first enroll in the exchange to create a measure of loyalty to Partners Healthcare providers, based on whether each patient used outpatient Partners up to 2 years before entering the exchange.[32] Unlike any measure of Partners use after enrolling in the exchange, this is not endogenous to plan choice on the Connector. I identify Partners-affiliated physicians and outpatient facilities using the Massachusetts Provider Database from Massachusetts Health Quality Partners (MHQP).

[32]Since I observe claims as early as 2013, and the exchange first opens in 2015, two years is the greatest window within which I can observe pre-exchange Partners use for all enrollees.

A.2 Construction of Plan Choice Dataset

I make a number of sample restrictions in the final plan choice dataset.

1. Restrict sample to Connector QHP individual (household size of one) enrollees not eligible for ConnectorCare plans (>300% of FPL), for which I directly observe relative plan prices.

2. Restrict choice set to Bronze, Silver, Gold and Platinum plans offered by Neighborhood Health Plan, BMC Healthnet, and Network Health. These three carriers collectively make up the large majority of enrollment in ConnectorCare, but there are two major carriers (Harvard Pilgrim and Blue Cross Blue Shield) that do not participate in ConnectorCare but do offer QHPs in the individual market. I exclude these carriers because while they are broad network plans, I do not directly observe their hospital networks, and because there is almost no margin of substitution between them and the three main carriers I study.

3. Restrict sample to enrollees who live in Massachusetts Connector rating areas 3-6. The Connector allows carriers to vary the price of their plans according to an individual's age and the "rating area" in which they live. Massachusetts is divided into seven rating areas, characterized by the first 3 digits of the location's zip code. I exclude rating areas 1 (3-digit zips 010, 011, 012 and 013) and 2 (014, 015 and 016) which capture Western Massachusetts, as well as rating area 7 (3-digit zip codes 025 and 026) which covers Cape Cod and Martha's Vineyard/Nantucket. The reason for excluding these areas is that BMC's provider network has very sparse coverage in these regions and, while I do observe non-zero shares of some BMC plans among enrollees in those zip codes, it is unclear whether those choices should inform demand estimates.

4. Collapse plan choices to one option per metal tier per carrier, for a total of up to 12 available choices. While the Connector requires each carrier to offer a standardized cost-sharing plan in each metal tier, in practice each carrier in my sample offers more than one plan in some metal tiers in some years. These additional plans are allowed to be "non-standard," with slightly different cost-sharing and prescription drug formularies than the exchange's standard requirements, although they are subject to the regulatory approval of the exchange. Unfortunately, I do not observe the contract characteristics of the non-standard plans, and so cannot control for differences in contract features across different plans within carrier-metal tier. Instead, I collapse all choices within carrier-metal tier-year into a single choice, and assign it the price of the cheapest such

plan. Likely as a result of regulatory oversight restricting the carriers' abilities to differentiate non-standard plans, the prices of all plans within carrier-metal tier-year do not differ widely.

B Identification Strategy

B.1 Additive Separability of Plan Utility

This section discusses the additive separability assumption underlying the identification strategy described in Section 4.1. One way to found additive separability is to begin with a two-period model as in Einav et al. (2013). In the second period, patients have already been enrolled in a plan j, characterized by generosity and network (m_j, x_j). They receive a health shock and make a healthcare utilization decision that takes into account the features of their health plan: the plan's financial generosity and provider network. In particular, I model care as consisting of the *quantity* of care (consisting of a number and type of visits) q, and the *identity* of the healthcare providers ϕ. At the beginning of the coverage period (one year in the case of the MA Connector) patients draw a health shock λ and a provider preference shock γ from distributions known by the individual.

I assume the second-period utility function over health is additively separable in the quantity of care and the identity of providers, so that these decisions are made independently: patients trade off the value of healthcare against the value of non-healthcare consumption in order to determine the quantity of care they consume, and choose providers in order to maximize their utility from provider identity. This is a natural assumption if, as is the case in the Connector, plans of the same coverage level m all have the same out-of-pocket price schedule. Formally, second period utility is given by

$$u(q, \phi; \lambda, \gamma, j) = v(q; \lambda, m_j) + h(\phi; \gamma, x_j),$$

where utility in healthcare quantity depends on both the health shock λ and the plan generosity m_j, and utility in provider identity depends on the provider preference shock γ and the plan's provider network x_j. It thus follows that plan utility, given by the expectation (over the joint distribution of health and provider preference shocks) of second-period utility, is also additively separable in the plan characteristics m and x.

Figure 13: Connector Standardized Metal Tier Designs, 2017

Plan Feature/ Service *A check mark (✓) indicates that this benefit is subject to the annual deductible*		Platinum	Gold	Silver	Bronze
Annual Deductible – Combined		N/A	$1,000	$2,000	N/A
		N/A	$2,000	$4,000	N/A
Annual Deductible – Medical		N/A	N/A	N/A	$2,750
		N/A	N/A	N/A	$5,500
Annual Deductible – Prescription Drugs		N/A	N/A	N/A	$250
		N/A	N/A	N/A	$500
Annual Out-of-Pocket Maximum		$3,000	$5,000	$7,150	$7,150
		$6,000	$10,000	$14,300	$14,300
Primary Care Provider (PCP) Office Visits		$25	$30	$30	$25 ✓
Specialist Office Visits		$40	$45	$50	$40 ✓
Emergency Room		$150	$150 ✓	$700 ✓	$500 ✓
Urgent Care		$40	$45	$50	$40 ✓
Inpatient Hospitalization		$500	$500 ✓	$1,000 ✓	$1,000 ✓
Skilled Nursing Facility		$500	$500 ✓	$1,000 ✓	$1,000 ✓
Durable Medical Equipment		20%	20% ✓	20% ✓	20% ✓
Rehabilitative Occupational and Rehabilitative Physical Therapy		$40	$45	$50	$40 ✓
Laboratory Outpatient and Professional Services		$0	$20 ✓	$25 ✓	$50 ✓
X-rays and Diagnostic Imaging		$0	$20 ✓	$25 ✓	$175 ✓
High-Cost Imaging		$150	$200 ✓	$500 ✓	$1,000 ✓
Outpatient Surgery: Ambulatory Surgery Center		$500	$250 ✓	$750 ✓	$750 ✓
Outpatient Surgery: Physician/Surgical Services		$0	$0 ✓	$0 ✓	$0 ✓
Prescription Drug	Retail Tier 1	$15	$20	$20	$25 ✓
	Retail Tier 2	$30	$30	$60	$75 ✓
	Retail Tier 3	$50	$50	$90	$100 ✓
	Mail Tier 1	$30	$40	$40	$50 ✓
	Mail Tier 2	$60	$60	$120	$150 ✓
	Mail Tier 3	$150	$150	$270	$300 ✓
2017 Final FAVC		91.73%	81.43%	71.84%	61.86%

31

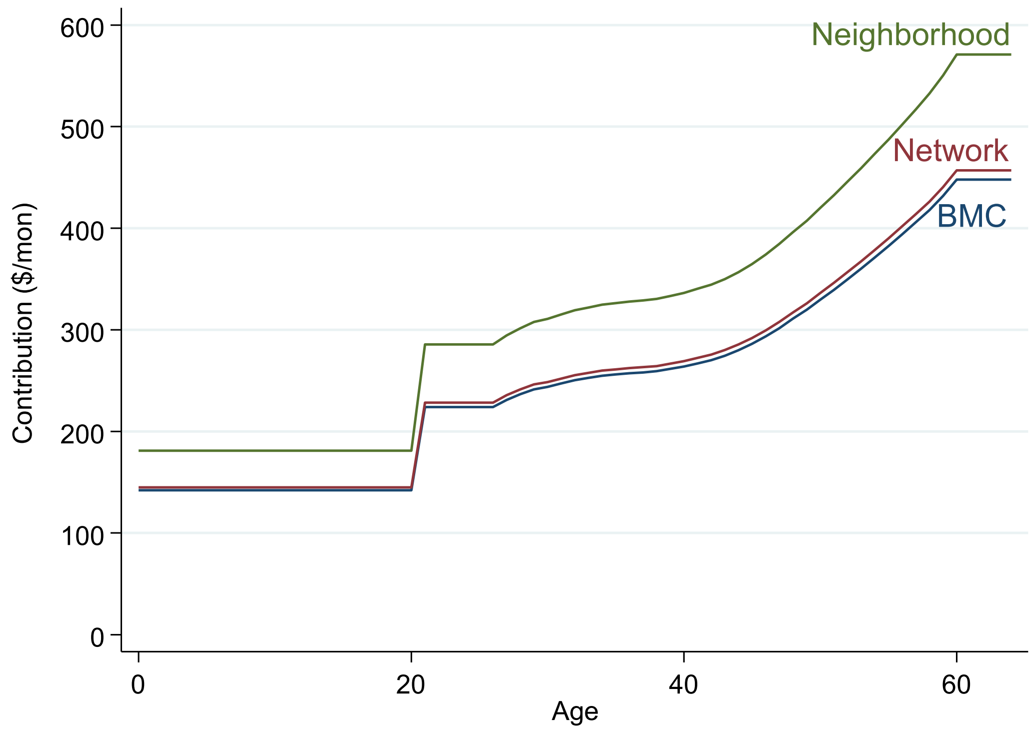
Taken from: MA Connector Final Award of 2017 Seal of Approval meeting slides.

Figure 14: Connector Standardized Metal Tier Designs, 2018

Plan Feature/ Service *A check mark (✓) indicates that this benefit is subject to the annual deductible*		Platinum	Gold	Silver	Bronze #1	Bronze #2 (HSA)
Annual Deductible – Combined		$0	N/A	$2,000	$2,500	$3,000
		$0	N/A	$4,000	$5,000	$6,000
Annual Deductible – Medical		N/A	$1,000	N/A	N/A	N/A
		N/A	$2,000	N/A	N/A	N/A
Annual Deductible – Prescription Drugs		N/A	$0	N/A	N/A	N/A
		N/A	$0	N/A	N/A	N/A
Annual Out-of-Pocket Maximum		$3,000	$5,000	$7,350	$7,350	$6,650
		$6,000	$10,000	$14,700	$14,700	$13,300
Primary Care Provider (PCP) Office Visits		$20	$30	$30	$30 ✓	$20 ✓
Specialist Office Visits		$40	$45	$50	$50 ✓	$40 ✓
Emergency Room		$150	$150 ✓	$700 ✓	$700 ✓	$250 ✓
Urgent Care		$40	$45	$50	$50 ✓	$40 ✓
Inpatient Hospitalization		$500	$500 ✓	$1,000 ✓	$1,000 ✓	$750 ✓
Skilled Nursing Facility		$500	$500 ✓	$1,000 ✓	$1,000 ✓	$750 ✓
Durable Medical Equipment		20%	20% ✓	20% ✓	20% ✓	20% ✓
Rehabilitative Occupational and Rehabilitative Physical Therapy		$40	$45	$50	$50 ✓	$40 ✓
Laboratory Outpatient and Professional Services		$0	$20 ✓	$25 ✓	$25 ✓	$25 ✓
X-rays and Diagnostic Imaging		$0	$20 ✓	$25 ✓	$25 ✓	$25 ✓
High-Cost Imaging		$150	$200 ✓	$500 ✓	$500 ✓	$500 ✓
Outpatient Surgery: Ambulatory Surgery Center		$250	$250 ✓	$750 ✓	$750 ✓	$500 ✓
Outpatient Surgery: Physician/Surgical Services		$0	$0 ✓	$0 ✓	$0 ✓	$0 ✓
Prescription Drug	Retail Tier 1	$10	$20	$20	$20	$20 ✓
	Retail Tier 2	$25	$30	$60	$60 ✓	$40 ✓
	Retail Tier 3	$50	$50	$90 ✓	$90 ✓	$60 ✓
	Mail Tier 1	$20	$40	$40	$40	$40 ✓
	Mail Tier 2	$50	$60	$120	$120 ✓	$80 ✓
	Mail Tier 3	$150	$150	$270 ✓	$270 ✓	$180 ✓
2018 Final FAVC		88.24%	79.69%	71.40%	64.84%	64.88%

Taken from: MA Connector Final Award of 2018 Seal of Approval meeting slides. New in 2018, the Connector provided an option for carriers to offer a second standardized Bronze plan option that was HSA compatible (Bronze #2). None of Neighborhood (NHP), BMC Healthnet or Network Health offered a Bronze #2 plan design in 2018.

Figure 15: Massachusetts Connector Pricing Function



Notes: Figure plots the ACA plan pricing function by age ("age curve"), illustrated by each carrier's Silver-tier plan offered in 2016 in Cambridge, MA (rating area 5). Insurance carriers have discretion over the level of price for each plan they offer within rating area × year, but the shape of the pricing schedule as a function of age is fixed across plans (and set by state regulation).

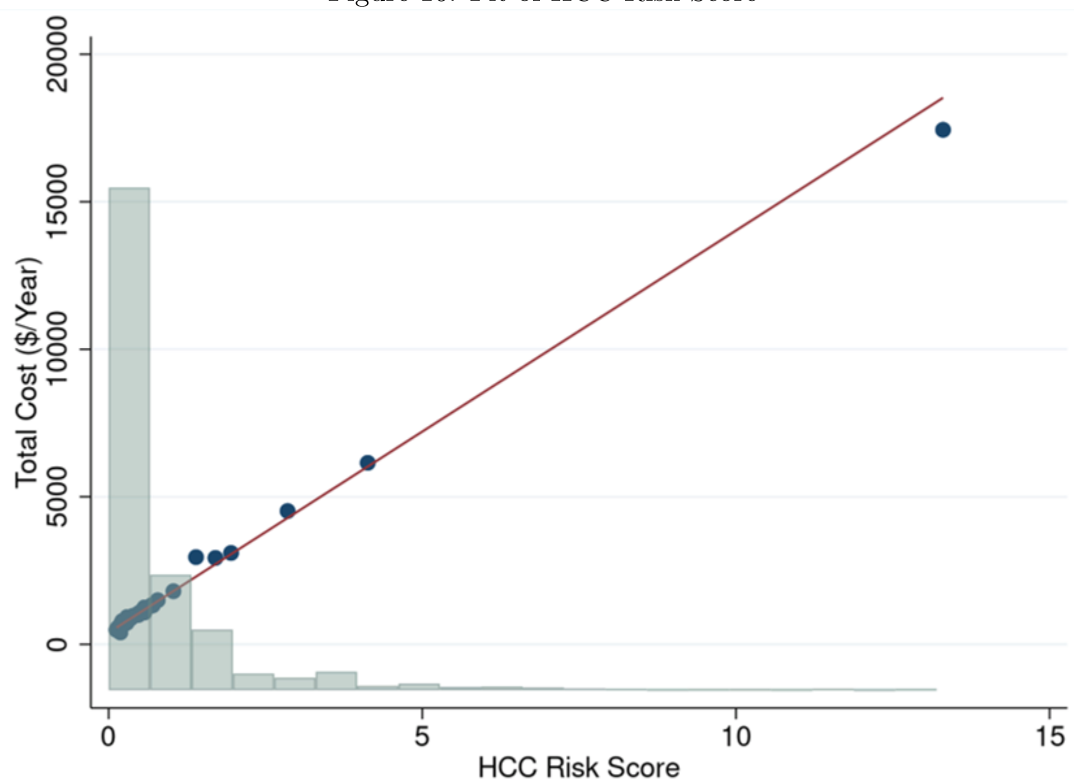
C Cost Effect Estimates

Table 3: Network Cost Effect Estimates

Coefficient	Network Cost Specification (1)	(2)	(3)	(4)
Proportional Effect of NHP	1.285			
	(0.098)			
x Non-Partners (pre-Switch)		1.034		
		(0.107)		
x Partners (pre-Switch)		1.613		
		(0.181)		
x Non-Partners (pre-Connector)			1.224	
			(0.117)	
x Distance > 25 mi.				1.212
				(0.183)
x Distance 5-25mi.				1.161
				(0.144)
x Distance <5 mi.				1.605
				(0.468)
x Partners (pre-Connector)			1.358	
			(0.159)	
x Distance > 25 mi.				1.307
				(0.358)
x Distance 5-25mi.				1.342
				(0.173)
x Distance <5 mi.				1.521
				(0.262)
Time FEs	Yes	Yes	Yes	Yes
N	47,460	47,460	47,460	47,460
N (enr. NHP in 2016)	10,271	10,271	10,271	10,271
N (switch from NHP in 2017)	2,240	2,240	2,240	2,240

Notes: Table shows estimates (and standard errors in parentheses) of the proportional effect of broad-network (NHP) coverage on average insurer costs for four poisson regression specifications. Column (4) estimates the effect of NHP coverage separately by whether the patient used Partners outpatient (non-emergency) care within two years before entering the Connector, interacted with categories of distance to the nearest Partners hospital.

Figure 16: Fit of HCC Risk Score



Notes: Figure plots the fit of HCC risk score (constructed based on concurrent medical claims) with realized costs, along with line of best fit in red. Histogram bars plot the density of the HCC risk score distribution. Sample is the population of Connector enrollees 2015-2017.