Hospital Network Competition and Adverse Selection:
Evidence from the Massachusetts Health Insurance Exchange

Mark Shepard*
Harvard Kennedy School and NBER

July 14, 2015

Abstract

Health insurers increasingly compete based on their networks of covered medical providers. Using data from Massachusetts' pioneer insurance exchange, I show evidence of substantial adverse selection against plans covering the most prestigious and expensive academic hospitals. These plans attract consumers with high costs in two ways: (1) they are sicker than average and (2) when sick, they tend to use expensive providers of care. Standard risk adjustment partly offsets the first but does not offset the second cost dimension, which is driven by heterogeneity in patients' preferences for and choices of providers. I study the implications of this selection using a structural model of insurer competition on hospital coverage and premiums. I find that adverse selection encourages plans to exclude the prestigious hospitals from network. But this outcome has two offsetting benefits: it reduces costs substantially and puts pressure on the expensive hospitals to lower prices. Simple modifications to risk adjustment to encourage coverage of the hospitals do little to improve welfare because they do not address the fundamental issue: efficiently sorting which patients should use expensive academic medical centers.

* Post-Doctoral Fellow in Aging and Health Care at NBER (2015-16) and Assistant Professor of Public Policy at Harvard Kennedy School (2016+). Email: mark_shepard@hks.harvard.edu. I thank my Ph.D. advisors David Cutler, Jeffrey Liebman, and Ariel Pakes for extensive comments and support in writing this paper. I thank the Massachusetts Health Connector (and particularly Michael Norton, Sam Osoro, Nicole Waickman, and Marissa Woltmann) for assistance in providing and interpreting the data. I also thank Katherine Baicker, Amitabh Chandra, Jeffrey Clemens, Keith Ericson, Amy Finkelstein, Jon Gruber, Kate Ho, Sonia Jaffe, Tim Layton, Robin Lee, Greg Lewis, Tom McGuire, Joe Newhouse, Daria Pelech, Amanda Starc, Karen Stockley, Rich Sweeney, Jacob Wallace, Tom Wollmann, Ali Yurukoglu, and participants in seminars at Boston College, the Boston Fed, CBO, Columbia, Chicago Booth, Duke Fuqua, Harvard, Northwestern, U Penn, Princeton, Stanford GSB, UCLA, UCSD, and Washington University for helpful discussions and comments. I gratefully acknowledge data funding from Harvard's Lab for Economic Applications and Policy, and Ph.D. funding support from National Institute on Aging Grant No. T32-AG000186 (via the National Bureau of Economic Research), the Rumsfeld Foundation, and the National Science Foundation Graduate Research Fellowship.

Introduction

Public programs increasingly use regulated markets to provide health insurance to enrollees. These types of markets now cover more than 75 million people and cost over $300 billion in U.S. programs including the Affordable Care Act (ACA), Medicare Advantage, and Medicaid managed care. Markets can improve welfare by giving consumers choice and encouraging plan competition to improve quality and reduce costs. But a perennial concern in insurance markets is adverse selection. When consumer costs vary in ways that cannot be priced, insurers may have inefficient incentives to cut benefits to avoid attracting high-cost customers.[1] Because of this concern, recent programs attempt to address selection by regulating plan benefits and using a tool called "risk adjustment" that compensates plans when they enroll sicker people. Whether adverse selection is a concern even with these policies is an important question and a matter of active research.

As the ACA has regulated plans' covered services and patient cost sharing, insurers have differentiated on an alternate dimension: covered networks of hospitals and other medical providers. In the first years of the ACA, almost half of all exchange plans have adopted "narrow network" designs (McKinsey 2014).[2] These plans have generated controversy, including calls for broader network requirements, partly because they tend to exclude the most prestigious academic hospitals. These "star" hospitals are known as centers of advanced medical treatment and research but, partly as a result, are quite expensive (Ho 2009). By excluding them, insurers limit access to top providers but also reduce costs by steering patients to cheaper settings. However, insurers' incentives to balance this cost-quality tradeoff may also be influenced by adverse selection. Whether selection is involved an important question for the policy debate and one on which there is little direct evidence.[3]

In this paper, I show evidence of adverse selection against plans covering star hospitals through a channel that is theoretically distinct from the usual selection story and therefore poses a challenge for standard policies like risk adjustment. Typically, adverse selection in health insurance is associated with *sicker* people choosing generous plans. But in addition to medical risk, consumers can be costly because of their *provider choices* – if for a given illness, they use more expensive doctors and hospitals. This second cost dimension may be important, since prices and costs vary widely across hospitals within areas. Importantly, insurers typically cover the bulk of these cost variations, with patients paying fees that differ

[1] The inability to charge higher prices to high-cost customers may occur either because of asymmetric information or because of regulations, like the ACA's prohibition on pricing based on pre-existing conditions. A large literature has studied selection starting with the seminal theoretical work of Akerlof (1970) and Rothschild and Stiglitz (1976).
[2] "Narrow networks" were defined as plans covering fewer than 70% of area hospitals. The McKinsey report documented a sharp rise in narrow network plans in 2014 relative to individual insurance markets in 2013.
[3] An older literature studying the rise of HMOs found that HMOs attracted healthier customers than traditional insurance (Miller and Luft 1997). But HMOs differed from both via limited networks and a variety of managed care restrictions. By contrast, today's competition involves exclusively managed care plans.

little across in-network providers.[4] As a result, patients who choose star hospitals when sick are more costly than patients who use less expensive alternatives. Risk adjustment is unlikely to offset these higher costs, which are driven by provider choices rather than purely medical factors. Moreover, people with strong preferences for using the star hospital – e.g., because they live nearby or because their regular doctor practices there – are also likely to select plans that cover their preferred provider. Thus, even idiosyncratic preferences for the star hospital (uncorrelated with medical risk) can lead to adverse selection against plans that cover them.

In some ways, the implications of this alternate channel for adverse selection are standard – inefficient sorting and potentially unravelling of generous coverage. Because regulations prevent plan premiums from varying based on consumer-specific use of star hospitals, plan sorting can be inefficient. For example, some people might want the option value of accessing a star hospital if they get a serious cancer. But to buy a plan covering it, they have to pool with people who use star providers for all their health care needs. Plans covering star hospitals differentially attract these high users, forcing them to raise premiums on all customers. The result is a cycle of adverse selection in which plans continually raise premiums and lose low-cost consumers, with this process either stabilizing or leading to full unravelling.

But selection on use of star hospitals also has non-standard implications for two reasons. First, it is fundamentally linked to the tradeoff between risk protection and moral hazard in health care. When insurers cover a star hospital, patients' costs increase because they can now use it (instead of cheaper alternatives) without paying the full bill. Coverage of star hospitals therefore generates a form of moral hazard.[5] What I find is that the people most likely to use star hospitals when covered (i.e. highest moral hazard) tend to select into plans that cover them. Thus, this pattern is an example of the idea of "selection on moral hazard" documented by Einav et al. (2013). Because moral hazard is involved, the welfare implications of unravelling are ambiguous and depend on whether the lost value of access exceeds or falls short of the cost savings. It also implies that policies to reduce moral hazard (e.g., higher "tiered" copays for expensive hospitals) may also reduce adverse selection. However, as long as the higher copays do not cover the full extra cost of star hospitals, some moral hazard and adverse selection is likely to remain as a tradeoff to providing risk protection.

A second difference from the standard analysis is that the selection is linked to a service (care at star hospitals) whose prices are not set competitively. Instead, hospital prices are set in negotiations with insurers driven by market power. This market power complicates policy responses to selection. For instance, a mandate to cover the hospitals – one standard response to selection – would be problematic

[4] In the Massachusetts exchange in particular, regulations mandated equal copays across covered hospitals.
[5] The moral hazard terminology can be confusing because contract theory typically refers to moral hazard as a hidden or non-contractible agent action that is costly to the principal. Here, the action is using a star hospital, and while it is not hidden, regulation prevents writing an insurance contract whose price varies with star hospital use.

because it would give star hospitals extreme power to raise prices. Other policies that subsidize plans covering star hospitals would have a conceptually similar (but less extreme) effect. A key question is whether government is willing to subsidize star hospitals' high prices to ensure that exchange enrollees have access to them. Because these academic hospitals' high prices partly fund teaching, medical research, and care for the poor, the answer to this question is not obvious.

To study these issues, I use data from a market that was a key precursor to the ACA: Massachusetts' subsidized insurance exchange.[6] The Massachusetts exchange provides a nice setting for studying networks and selection. Covered services and patient cost-sharing rules are fixed by regulation for all plans, so provider networks are the only significant benefit that differs across plans. Further, the exchange has excellent administrative data on all consumers' plan choices and insurance claims. These detailed data let me estimate a flexible model of demand and costs to capture heterogeneity driving adverse selection. Finally, Massachusetts has a clear set of star hospitals: the Partners Healthcare System. Partners is centered around Mass. General and Brigham & Women's hospitals, which are consistently ranked by *U.S. News & World Report* as the top two hospitals in Massachusetts and among the top 10 hospitals in the nation. Consistent with state reports (e.g., Coakley 2013), I find that Partners hospitals are extremely expensive. I estimate that their risk-adjusted prices per admission are almost twice the average of other hospitals and over $5,000 (or 33%) more than the average of other academic medical centers.

I start by testing for adverse selection against plans covering Partners using reduced form methods. I show that these plans attract a group who appear to strongly prefer Partners: people who have used Partners hospitals in the past, either for inpatient or (more often) outpatient care. Compared to an average enrollee, these past Partners patients are (1) 15% costlier even after risk adjustment, (2) 80% more likely to select a plan that covers Partners, and (3) more than twice as likely to use Partners for subsequent hospitalizations. These facts suggest that Partners patients are loyal to their preferred hospitals and select plans based on their desire to use Partners in the future.[7] I find that this loyalty to past-used hospitals is true more broadly across all hospitals in my data, suggesting that it is a general phenomenon likely to drive plan selection in health insurance markets.[8]

---

[6] Massachusetts' subsidized exchange (which I study) is distinct from its unsubsidized exchange that has been studied by Ericson and Starc (2013, *forthcoming*). The only research I am aware of using the subsidized exchange is by Chandra, Gruber, and McKnight (2010, 2011, *forthcoming*) studying the effect of the individual mandate on market enrollment and the effects of cost-sharing changes in 2008 on utilization of care.

[7] Past Partners patients also appear to be unobservably sick, since they have above-average risk-adjusted costs even in plans that do not cover Partners. However, their costs are much higher in plans that cover Partners.

[8] It is less clear how much of this loyalty is driven by state dependence (a preference for hospitals used in the past) versus more durable preference heterogeneity. Both are valid channels for the short-run adverse selection results I find. But state dependence implies lower long-run welfare impacts of unraveling of Partners coverage, since patients need only incur a one-time cost of switching providers. Disentangling the roles of state dependence versus heterogeneity in loyalty to providers is an important question for future research.

I next study how this selection played out in a case in 2012 when a large plan dropped Partners and several other hospitals from its network. This type of network change provides a natural source of evidence that past research has rarely studied. Consistent with selection, I find a sharp increase in enrollees switching away from the plan that dropped Partners, driven almost entirely by a nearly 40% switching rate by Partners patients. While exchange enrollees typically switch plans less than 5% of the time (c.f. Handel 2013), they are much more willing to switch plans when a preferred provider is dropped. Using my model (discussed below) to decompose the plan's large risk-adjusted cost reduction after dropping Partners, I find that selection can account for between 36% and 50% of the change.

The reduced form analysis suggests the importance of adverse selection based on hospital networks. However, both the welfare and policy implications of this selection are less clear. To investigate these issues, I estimate a structural model of consumer preferences and insurer costs. The model – which follows a structure used in past work (Ho 2009; Capps, Dranove, and Satterthwaite 2003) – consists of three pieces: (1) a hospital demand system capturing hospital choices under different plan networks, (2) an insurance demand system capturing plan choice patterns, and (3) a cost system estimated from the insurance claims data. Relative to past work, the main innovation is to allow for detailed preference heterogeneity and use the individual-level data to capture the correlations among hospital choices, plan preferences, and costs – which are critical for adverse selection. In addition, I pay special attention to the identification of the premium and network coefficients in plan demand, using only within-plan variation to identify them. For premiums, I use variation across consumers driven by Massachusetts' subsidy rules. For networks, I use variation in how different consumers value a given hospital network.

My demand estimates imply that individuals value both lower prices and better hospital networks, though with significant heterogeneity in this tradeoff. Consistent with the reduced form evidence, I find that past patients of a hospital are particularly likely to use it again and to select plans that cover it. These effects are particularly strong for past patients of Partners hospitals. Thus, the demand estimates are consistent with significant selection based on coverage of the prestigious Partners hospitals.

I use the model to study the competitive, welfare, and policy implications of network-based selection. I simulate equilibrium in a game where insurers first choose whether or not to cover the Partners hospitals (holding fixed other hospital coverage) and then compete in a static Nash-in-prices game. I model exchange policies similar to those in the ACA, which differ in several ways from those used in Massachusetts. The key limitation of these simulations is that they hold hospital prices fixed at their observed values, not modeling hospital-insurer price bargaining. At Partners' observed high prices, I find that adverse selection leads all insurers drop them from network. As in the reduced form results, a plan deviating to cover Partners loses money both through higher costs for its existing enrollees (moral hazard) and by attracting high-cost enrollees who particularly like Partners (adverse selection). I use the model to decompose the adverse selection into traditional selection on levels of cost and selection on use of Partners. Of the 27% higher risk-adjusted costs for the group that most highly values Partners, about 60% is a higher level of risk-adjusted costs even in a plan that does not cover Partners, and 40% is larger cost increases when a plan covers Partners. Thus, both selection on unobserved health risk and selection on likelihood to use Partners appear to be quantitatively important.

Finally, I use my model to analyze policy changes to address adverse selection. I find that modified risk adjustment and differential subsidies for higher price plans can reverse the unraveling. These policies give plans a greater incentive to cover Partners even though doing so requires raising prices and attracting high-cost enrollees. However, I highlight two tradeoffs. First, covering Partners creates moral hazard. When Partners is covered, not only people who value it highly can use it but also those who value it little. My model's estimates imply that past Partners patients have greater value of access than costs, but other enrollees on average do not. Because the latter group is much larger, I find a net decrease in social surplus when the government changes policy to encourage Partners coverage.

A second tradeoff of these policy changes is that they encourage both insurers and Partners hospitals to raise prices. My current model does not capture the higher Partners prices (which are held fixed). But I find important increases in insurance prices and markups, leading to a government-funded increase in insurer profits. This analysis aligns with recent work finding that adverse selection leads plans to reduce markups in imperfectly competitive markets (Starc 2014; Mahoney and Weyl 2014). Adverse selection gives insurers an incentive to keep prices low to attract low-cost consumers. Policies that offset this effect encourage plans to raise price markups. In exchanges, higher plan prices mean higher government subsidies, which are set based on these prices.

These results suggest that standard policies used to address adverse selection (e.g., risk adjustment and subsidies) are less effective at improving welfare with selection based on star hospital use. These policies compensate insurers for attracting high-cost enrollees but do not address the fundamental issue of efficiently sorting patients across hospitals. Policies that address this sorting challenge directly – e.g., higher "tiered" copays for high-price hospitals or payment incentives for doctors to steer patients to lower-cost hospitals (see Song, et al. 2011; Ho and Pakes 2014) – may be more effective and are a fruitful subject for future research.

The remainder of this paper is organized as follows. Section 1 outlines a simple model that captures the main intuition for network-based selection. Section 2 presents background on the Massachusetts exchange and hospital market and introduces the data. Section 3 shows reduced form results, and Sections 4-5 presents the structural model and estimates. Section 6 analyzes the model's implications for adverse selection, and Section 7 presents the equilibrium and counterfactual policy simulations. The final section concludes.

1 Basic Theory

Economists have long known that insurance markets are subject to adverse selection. When consumers' costs vary but firms cannot observe or price based on this variation, market competition can be inefficient. Insurers may have incentives to cut benefits to avoid high-cost types (Rothschild and Stiglitz 1976) and in the extreme, the market can unravel entirely (Akerlof 1970) or to the minimum quality option (Cutler and Reber 1998). Most past theories have focused on selection based on medical risk (e.g., having an expensive illness). As a result, the standard policy to address selection is risk adjustment – transfers to plans attracting sicker people to offset their higher expected costs.

In this section, I present a simple model to illustrate an alternate source of adverse selection. This selection arises because certain "star" hospitals (usually academic medical centers) negotiate high prices that insurers cover (i.e., do not pass onto sick patients) when they include the hospitals in network. Whether insurers *should* cover star hospitals depends on the classic tradeoff between insurance protection (against the risk of needing advanced care) and moral hazard. What I show is that whether profit-maximizing insurers *will* cover star hospitals is also influenced by adverse selection. Consumers who strongly prefer using star hospitals will both be high-cost and more likely to choose plans covering the hospitals. Standard risk adjustment is unlikely to capture this form of selection, giving insurers an incentive to drop the star hospitals. I use the model to analyze when this unravelling is inefficient and to discuss the policy implications, which are complicated by insurer and hospital market power.

Simple Model

Consider a model with two hospitals. One hospital *S* is a "star" hospital, which can use its reputation to bargain for a high price τ_S per visit. A second "non-star" hospital *NS* has a lower price, $\tau_{NS} < \tau_S$. I assume that the star hospital's price is above its marginal cost mc_S and that its markup above costs exceeds the markup for the non-star hospital: $\tau_S - mc_S > \tau_{NS} - mc_{NS}$. Consumers differ in two dimensions: (1) their *risk* (or probability) of hospitalization, $r_i \in [0,1]$, and (2) their relative *value* (in dollar terms) of using hospital *S* when sick, v_i^S. If patient copays are identical across hospitals (as in the Massachusetts exchange), a patient with access to both hospitals will choose *S* if $v_i^S \geq 0$ and *NS* if $v_i^S < 0$.[9] Let $s_i \equiv 1\{v_i^S \geq 0\}$ be an indicator of consumer *i*'s preference for *S*. Defining $\Delta\tau \equiv \tau_S - \tau_{NS}$, under these assumptions, an insurer's expected total cost for consumer *i* is:

[9] Negative values capture the fact that hospital *NS* may be closer and more convenient for many patients.

$$C_{ij} = \begin{cases} r_i \cdot \tau_{NS} & \text{if a plan does not cover } S \\ r_i\left(\tau_{NS} + s_i \cdot \Delta\tau\right) & \text{if a plan covers } S \end{cases}$$

Consider an insurance market with two plans (*A* and *B*). For simplicity, plan *B* is a non-strategic actor who covers only hospital *NS* and sets its premium at average cost.[10] Plan *A* also covers hospital *NS* but chooses whether to cover *S* and strategically sets its premium to maximize profits. I assume consumers' utility for a plan not covering *S* is normalized to 0. Their (monetized) utility for plan *A* if it covers *S* equals their expected value of access to *S*: $U_{iA}^{CoverS} = r_i \cdot s_i \cdot v_i^S$. The key assumption here is that consumers' utility for a plan covering the star hospital (U_{iA}^{CoverS}) is correlated with their likelihood of using that hospital ($= r_i \cdot s_i$).

Following Massachusetts' rules, assume that each plan *j* sets a single premium P_j that cannot vary across consumers.[11] Although prices cannot vary, the exchange risk adjusts payments based on consumer observables Z_i so a plan in total receives $P_j + RA(Z_i)$ for consumer *i*.[12] The risk adjustment function is set to offset a consumer's expected extra costs, so $RA(Z_i) = E(c_{ij} \mid Z_i) - \overline{c}$ (where $\overline{c}$ is overall average cost). If risk adjustment captured costs perfectly, a plan's profit margin would be a constant $P_j - \overline{c}$ for all consumers. However, heterogeneous preferences for hospitals can confound risk adjustment in two ways. First, consumers selecting a plan covering *S* may have higher *unobserved risk* – a problem of imperfect risk adjustment. Second, consumers may select based on *preference* for hospital *S*, an independent source of cost variation that standard risk adjustment does not capture.

To see this, suppose that the exchange observes risk perfectly ($Z_i = r_i$) and consider a simple example with two risk types and specific parameters shown in the table on the next page. The table shows costs for each enrollee type in plan *A* if it covers hospital *S*. Because risk adjustment only captures average costs for each risk, the costs of consumers who prefer *S* exceed the risk-adjusted payment, and they are therefore less profitable to cover. Notice also that the interaction between risk and preference is important. For *S*-preferring types, risk-adjusted payments are too low by \$250 for low risks and by \$2,500 for high risks. This follows from high risks using the expensive hospital more and therefore having a *differential* cost increase in a plan covering *S*. I argue below that these differential costs induce inefficient

[10] If instead plan *B* set its premium to maximize profits, all of the basic adverse selection intuition of the model would carry through, but I would need to consider how the policies I discuss below would affect its markup.

[11] Assume that any subsidies are a flat amount for both plans so that the difference in consumer premiums equals the difference in the prices plans receive.

[12] Risk adjustment methods vary, and in general, the exchange could also make risk adjustment a function of prices. This was done in Massachusetts so that $RA_i^{Mass} = \left(\varphi\left(Z_i\right) - 1\right)P_j$, where $\varphi\left(Z_i\right)$ was a risk score and the plan's total payment was $\varphi\left(Z_i\right)P_j$. The ACA's risk adjustment is closer to the simple model, since its transfer is based on an enrollee risk score and the average price in the market.

Table: Simple Example

	Expected Enrollee Costs		Avg. Cost (for risk adj.)
	Prefer Hospital *NS*	Prefer Hospital *S*	
Low Risk	$r_L \cdot \tau_{NS}$ **\$500**	$r_L \cdot \tau_S$ **\$1,000**	$r_L\left(\tau_{NS} + \overline{s}_L \cdot \Delta\tau\right)$ **\$750**
High Risk	$r_H \cdot \tau_{NS}$ **\$5,000**	$r_H \cdot \tau_S$ **\$10,000**	$r_H\left(\tau_{NS} + \overline{s}_H \cdot \Delta\tau\right)$ **\$7,500**

Parameter Assumptions: $r_L = 0.05,\ r_H = 0.50,\ \tau_{NS} = \$10\text{k},\ \tau_S = \$20\text{k},\ \overline{s}_L = \overline{s}_H = 0.5$

sorting, since a premium increase sufficient to pay for high risks' extra costs may cause low risks to leave the plan. This inability of a homogenous premium to efficiently sort consumers with varying differential costs is related to a point made by Bundorf, Levin, and Mahoney (2012), Handel, Hendel, and Whinston (2013), and Glazer and McGuire (2013).

Equilibrium and Welfare Implications

To study the equilibrium, I simplify by assuming the exchange risk adjusts based on the cost of *NS*, so $RA(Z_i) = \left(E(r_i \mid Z_i) - \overline{r}\right)\tau_{NS}$.[13] Notice that I again allow for risk to be observed imperfectly. Define the risk adjustment error as $e_i \equiv \left(r_i - E(r_i \mid Z_i)\right)\tau_{NS}$, which is positive for the unobservably sick and negative for the unobservably healthy. Under these assumptions, plan *B* always charges a price (assumed to equal average costs) of $P_B = \overline{r} \cdot \tau_{NS} + \overline{e}_B$, where $\overline{e}_B$ is plan *B*'s enrollees average risk adjustment error. If plan *A* does not cover hospital *S*, the two plans are undifferentiated. Both charge $P_A^0 = P_B^0 = \overline{r} \cdot \tau_{NS}$, attract a random sample of enrollees, and earn zero profits. If plan *A* does cover *S*, it can raise its price and potentially earn profits. However, by doing so, it also increases its costs. Defining the price difference as $\Delta P_A \equiv P_A - P_B$, plan *A* will cover *S* if:

$$\text{Cover } S \text{ if:} \quad \max_{\Delta P_A} \left\{ \underbrace{\Delta P_A}_{\text{Price Difference}} - \underbrace{E\left[r_i \Delta\tau \mid r_i v_i^S > \Delta P_A\right]}_{\text{Insurer Cost Increase}} - \underbrace{\left(\overline{e}_A - \overline{e}_B\right)}_{\text{Unobs. Risk Selection}} \right\} \geq 0 \tag{1}$$

Adverse selection on use of the star hospital shows up in (1) in the conditional expectation – the people who choose *A* tend to be those who most value (and use) hospital *S*. More traditional selection on unobserved risk occurs through the $\overline{e}_A - \overline{e}_B$ term. If people preferring hospital *S* are unobservably sicker $(\overline{e}_A - \overline{e}_B > 0)$ – which I find to be true empirically – this is an additional channel for adverse selection.[14]

[13] This assumption does not change any of the intuition of the results but makes the math much simpler.
[14] In theory, these individuals could be unobservably healthier, creating a source of advantageous selection that offsets some of the adverse selection on use of the star hospital.

Now compare this equilibrium to the efficient outcome. For now, I assume that hospital markups are pure transfers, so the true social cost of using each hospital is mc_S and mc_{NS}.[15] The first-best outcome is for patients to choose the plan covering *S* if and only if $v_i^S \geq \Delta mc$. This first-best is unlikely to be attainable. Even if plan *A* covers *S*, consumers will choose plan *A* and get access to *S* only if $r_i v_i^S \geq \Delta P_A$. These conditions cannot coincide as long as there is consumer risk heterogeneity. With heterogeneous risks, some low-risk types who would highly value star hospital access if they did become sick will not choose plan *A* because doing so involves pooling with high-risk types who frequently use hospital *S*. Conversely, some high-risk types will inefficiently select into plan *A*. These errors reflect inefficient sorting with a single premium and heterogeneous costs.

Even if the first-best is unattainable, we can ask how plan *A*'s incentives compare to second-best efficiency given that consumers sort based on a single premium difference. It is socially optimal for plan *A* to cover hospital *S* (at some differential premium ΔP) if:

$$\text{Efficient to Cover } S \text{ if:} \quad \max_{\Delta P} \left\{ \underbrace{E\left(r_i v_i^S \mid r_i v_i^S > \Delta P\right)}_{\text{Cons. Value of Star Hospital}} - \underbrace{E\left(r_i \cdot \Delta mc \mid r_i v_i^S > \Delta P\right)}_{\text{Social Cost Increase}} \right\} \geq 0 \qquad (2)$$

Comparing conditions (1) and (2), we can say that *the plan has too little incentive to cover the star hospital* due to three factors:

(a) Selection on unobserved risk ($\bar{e}_A - \bar{e}_B > 0$), which the plan treats as a disincentive to cover *S* even though it is not a social cost;

(b) The star hospital's higher markup ($\Delta \tau > \Delta mc$), which makes the private cost of covering *S* exceed the social cost; and

(c) Consumer surplus for the plan when it covers hospital *S*: Plan *A*'s price increase equals the value to the marginal consumer (for whom $r_i v_i^S = \Delta P_A$), not the larger gain to the average consumer. This corresponds to the standard monopoly quality problem of Spence (1975).[16]

Thus, the plan will sometimes fail to cover the star hospital when it would be socially efficient to do so. This is the first inefficiency. Notice that selection on preference for using hospital *S* enters this inefficiency indirectly – by exacerbating the extra cost associated with the star hospital's markup.

A second inefficiency is that even if the plan covers hospital S, *its premium will be inefficiently high*. To maximize social welfare defined by (2), the optimal premium difference is $\Delta P^* = \Delta c \cdot E\left(r_i \mid r_i v_i^S = \Delta P^*\right)$. But for the plan to not lose money when covering *S*, we know by (1) that

[15] As I discussed in the introduction, how to value the star hospital's markup is not obvious. If it funds socially valuable services like teaching and medical research, it might have more than the valuation at cost I assume here.

[16] Unlike Spence (1975), the monopoly quality distortion can be signed here because coverage of hospital *A* is a discrete good.

$\Delta P_A \geq \Delta\tau \cdot E\left(r_i \mid r_i v_i^S > \Delta P_A\right) + \left(\overline{e}_A - \overline{e}_B\right)$. The plan's premium will be too high because of selection on unobserved risk ($\overline{e}_A - \overline{e}_B > 0$), the star hospital's markup ($\Delta\tau > \Delta c$), and selection on use of the star hospital (since $E\left(r_i \mid r_i v_i^S > \Delta P\right)$ exceeds $E\left(r_i \mid r_i v_i^S = \Delta P\right)$).

To summarize, under standard risk adjustment, several forces combine to give the plan a greater incentive to exclude the star hospital than is socially optimal. And if the plan does cover the star hospital, its price will be too high partly because it selects the highest-cost enrollees. Both inefficiencies suggest the potential for policy changes to improve the equilibrium.

Market Power and Policy Implications

How should exchange policy respond to these inefficiencies? Two natural policies are to modify subsidies or risk adjustment to encourage plan *A* to cover the star hospital and to reduce its relative premium if it does. In doing this, policymakers should consider potential negative competitive side effects. This consideration is important because when insurers have market power, adverse selection can have the positive effect of lowering price markups (Starc 2014; Mahoney and Weyl 2014). Policies used to offset selection may therefore lead insurers to raise prices and profits. In exchanges, higher prices are a public policy problem because subsidies are set based on prices[17] – so higher prices raise government costs, with an associated excess burden of taxation.

To see how this works, consider the condition for plan *A*'s profit-maximizing price:

$$P_A^* = \underbrace{\left[\overline{r}\tau_{NS} + \overline{r}_A \Delta\tau + \overline{e}_A\right]}_{\text{Risk-Adjusted Avg. Cost}} + \underbrace{\frac{1 - \frac{d\overline{r}_A}{dP_A}\Delta\tau - \frac{d\overline{e}_A}{dP_A}}{\eta_{D_A}}}_{\text{Plan Price Markup}} \tag{3}$$

where $\overline{r}_A$ is the average risk of people enrolling in plan *A* and $\eta_{D_A} \equiv -\frac{1}{D_A}\frac{dD_A}{dPrem_A}$ is the premium semi-elasticity of demand. As is standard with imperfect competition, the plan's price equals its risk adjusted costs plus a markup. Adverse selection (after risk adjustment) affects both of these terms. It first implies that plan *A* attracts worse risks (higher $\overline{r}_A$ and $\overline{e}_A$), raising its risk-adjusted costs. But it also implies that the plan has an incentive to hold down its markup, since by raising price it attracts even worse risks ($\frac{d\overline{r}_A}{dP_A} > 0$ and $\frac{d\overline{e}_A}{dP_A} > 0$). Policy changes that offset adverse selection push the other way: lowering risk-adjusted costs for the adversely selected plan and raising its price markups. The net effect on its price is

[17] Specifically, ACA subsidies are linked to the price of the second-cheapest silver tier plan. In Massachusetts, subsidies were linked to the price of the cheapest plan. These policies ensure that the cheapest plans will be affordable to consumers even if plan prices are higher than expected.

ambiguous. In Section 7, I analyze the effect of specific risk adjustment and subsidy changes to offset selection using my empirically estimated model.

A second competitive effect comes from the star hospital's market power. So far, I have been assuming that hospital prices are exogenous, but now suppose that the price of *S* can adjust, either by adjusting costs or markups. Another way to interpret the hospital coverage condition (1) is as a *maximum price* that *S* can charge while still being covered:[18]

$$\tau_S^{Max} = \tau_{NS} + \frac{P_A^* - \overline{r}\,\tau_{NS} - \overline{e}_A}{\overline{r}_A} \tag{4}$$

Adverse selection raises $\overline{r}_A$ and $\overline{e}_A$, both of which lower the maximum price the star hospital can charge, although increases in plan *A*'s price push the other direction. If the first effects dominate, adverse selection will spur the star hospital to cut its price, while policies to offset selection will do the opposite.

2 Massachusetts Exchange Background and Data

I study the subsidized Massachusetts health insurance exchange – called Commonwealth Care, or CommCare. Created in Massachusetts' 2006 health insurance reform, it operated from November 2006 to December 2013, after which it shifted form to comply with ACA rules. Like the ACA exchanges, CommCare offered subsidized coverage to low-income people (below 300% of poverty) not eligible for employer-sponsored insurance or other public programs. Higher-income and otherwise ineligible people could buy unsubsidized plans in a separate exchange ("CommChoice") – a market studied by Ericson and Starc (2012, 2013).[19] In CommCare, enrollees could choose among competing private plans in a centralized marketplace. Over the 2010-2013 period I focus on, the exchange had five competing insurers and about 170,000 enrollees per month. This size makes it comparable to a very large employer insurance plan but still small relative to Massachusetts' overall population of 6.6 million.

CommCare's regulator was aware of the possibility of adverse selection and took several steps to counteract it. Generous subsidies (covering over 90% of premiums on average) and an insurance mandate encouraged broad participation to avoid just the sick from selecting into the market. To address adverse selection within the market, CommCare used several policies. First, it standardized nearly all benefits, requiring plans to follow pre-specified covered service and cost-sharing rules. The only major benefit left

[18] This would be the hospital's price in a Nash bargaining model in which the star hospital had all of the bargaining power. In this simple model, the hospital might want to charge a lower price because by doing so, plan *A* can charge a lower premium and attract more consumers who use hospital *S*. In a more realistic model with many hospitals and non-hospital costs, this effect would be sufficiently small that the maximum bargaining price would hold.

[19] The ACA differs from Massachusetts by pooling subsidized and unsubsidized enrollees into the same market. The ACA also differs in subsidizing people up to 400% of poverty but excluding people below 133% of poverty, who will be eligible for Medicaid in states expanding the program.

flexible were provider networks.[20] Second, the exchange used risk adjustment to compensate plans for attracting the sick. This worked by assigning each enrollee a risk score φ_i based on his expected costliness. A plan with price P_j would receive $\varphi_i P_j$ for covering enrollee *i*. CommCare used sophisticated diagnosis-based methods to set risk scores, similar to those in the ACA and Medicare.[21] Notably, risk adjustment does not incorporate people's past provider choices, since it is not intended to capture people's likelihood to use expensive hospitals. Although there is conflicting evidence on how well risk adjustment has worked in Medicare (see Brown et al. 2011; Newhouse et al. 2015), these methods nonetheless represent the state of the art.[22]

A final policy affecting adverse selection was Massachusetts' subsidy rules. Unlike the ACA which applies a flat subsidy to all plans, CommCare differentially subsidized higher-price plans. In theory, these "marginal" subsidies address selection by narrowing price differences and encouraging plans to compete less on price and more on quality. In practice, CommCare's subsidies were not intended to address selection nor were they well designed to do so. The main marginal subsidies were for enrollees below poverty, for whom federal Medicaid rules required all plans to be fully subsidized (i.e., zero consumer premiums). Because of this, the exchange had to set a maximum price to prevent arbitrary price increases. When binding, this maximum price in turn discourages plans from covering expensive hospitals, since they cannot increase their price when they do so.

Whether adverse selection is a significant concern, despite these policy steps, is an unknown but important question. Like CommCare, the ACA has generous subsidies and a mandate, risk adjustment and reinsurance, and benefit standardization relative to the pre-ACA market (though not as stringent as CommCare). A key question is whether provider networks, one of the few flexible benefits, create residual selection problems. Because of its network variation in isolation of other benefit differences, CommCare is a nice setting to study the effects of limiting provider networks.

Provider Networks and Plan Background

The Massachusetts provider market in which CommCare operates includes a large number of providers – about 80 hospitals and 28,000 active doctors (AAMC 2013). But these providers are

[20] In addition, insurers could flexibly set two more minor benefits: (1) a few "extra benefits" like gym membership discounts, and (2) prescription drug formularies, provided they covered at least two drugs per class.

[21] One difference was that Massachusetts used *prospective* risk adjustment (which uses only past observed diagnoses), whereas the ACA uses a *concurrent* method incorporating diagnoses observed in the current year. Massachusetts' method meant that risk adjustment for new enrollees could only be based on age and sex, since no past diagnoses information was available. In practice, I find the selection results hold just as strongly even for the subsample for whom diagnosis-based risk adjustment was used.

[22] Like the ACA, CommCare also used reinsurance for very high-cost enrollees (above $150,000 per year), but this accounted for just 0.03% of enrollees and about 1% of costs.

organized into a few dominant provider systems, most centered around large academic hospitals. Table 1 shows statistics on general acute care hospitals and hospital systems based on my CommCare data.[23] The top panel lists the most expensive hospitals based on the average per-admission cost to insurers. I also show a risk-adjusted price measure (which adjusts for patient severity, as I describe in Section 5.1) and the average severity of their patients (where the mean across all patients in all hospitals is normalized to 1.0). The bottom panel shows the most used hospital systems.

All of the five most expensive hospitals are academic medical centers – a designation given by the Massachusetts government to the state's six most significant academic hospitals. These academic hospitals both have high prices and treat sicker than average patients (severity > 1.0). But even among academic hospitals, the Partners Healthcare System is unique. Partners is both the most-used system and includes the two most expensive hospitals – Mass. General Hospital and Brigham & Women's Hospital, both in Boston.[24] These two are clear examples of what Ho (2009) called "star hospitals." *U.S. News & World Report's* rankings list them as the top two hospitals in Massachusetts, and they have consistently been listed among the top 10 hospitals in the U.S. (with Mass. General ranking #1 in 2013-14). The Partners hospitals' high prices have been repeatedly documented (see Coakley 2013; CHIA 2013) and have sparked anti-trust investigations by the U.S. Department of Justice and the Mass. Attorney General.

Table 2 shows hospital network coverage for each of the five exchange plans from 2009-2013, based on information the exchange posted online for enrollees to review. Because of its unique status, I focus attention on coverage of Partners, listing only the statewide share of all other hospitals covered (weighted by number of beds). Plans always cover/exclude the two Partners academic medical centers in tandem. Coverage of Partners' five community hospitals is more flexible but correlates strongly with coverage of the academic medical centers. As of 2011, three plans covered Partners – Network Health, Neighborhood Health Plan (NHP), and CeltiCare. The other two plans had specific reasons not to cover Partners: Boston Medical Center (BMC) plan is owned by a competitor academic hospital, and Fallon is centered in central-Massachusetts and does not offer coverage in most of Boston where Partners is based.

Over time, several changes occurred in Partners coverage. In 2012, the exchange introduced a limited choice policy that gave the lowest-price plans exclusive access to new below-poverty enrollees (for whom all plans were free). In response to this incentive to price low, CeltiCare and Network Health cut prices by 11% and 15%, respectively, to become the two cheapest plans. While CeltiCare already had low costs (partly through its very limited coverage of non-Partners hospitals), Network Health needed to cut costs to make possible this price reduction. To do so, it dropped almost all of the Partners system plus

[23] Because I exclude specialty hospitals and merge together different campuses of the same hospital (which are often not separately identified in the claims data), my data has a total of 64 unique hospitals.

[24] Partners also included (as of 2012) five community hospitals in Eastern Massachusetts and network including more than 1,100 primary care physicians (BCBS Foundation of Massachusetts 2013).

eight other hospitals from network.[25] As I show in Section 3, these changes helped Network Health reduce per-member costs by 21%. But a significant portion came from high-cost enrollees leaving Network Health and shifting to NHP and CeltiCare. In response, CeltiCare excluded the main Partners primary care physicians at the start of fiscal 2014, though it continued to cover the hospitals for care if patients were referred through an outside physician.[26] By contrast, NHP continued to cover Partners but had special reason to do so: Partners purchased it effective in fiscal year 2013. Thus, as CommCare transitioned into the ACA mid-way through fiscal 2014, only one plan fully covered Partners and that through a vertical ownership relationship.

Data: Plan Choices and Insurance Claims

To study CommCare, I use a comprehensive administrative dataset on plan enrollment and insurance claims for all plans from 2006-2013.[27] For each de-identified enrollee, I observe demographics, plan enrollment history, and claims for health care services while enrolled in the market. The claims include information on patient diagnoses, services provided, the identity of the provider, and the actual amounts the insurer paid for the care.

I use the raw data to construct two datasets for model estimation. The first dataset is for hospital demand and costs. From the claims, I pull out all instances of inpatient stays at general acute care hospitals in Massachusetts during fiscal years 2008-2013, the period over which I have data on plan networks. I add on hospital characteristics from the American Hospital Association (AHA) Annual Survey and define patient travel distance using the Google Maps driving distance from the patient's home zip code centroid to the hospital.[28] For each hospitalization, I sum up the insurer's total payment while the patient was admitted (including both the hospital facility fees and physician professional service fees) and use this to estimate the hospital price model described in Section 5.1.

The second dataset is for insurance plan demand and costs. Using the enrollment data, I construct a dataset of available plan choices, plan characteristics (including premium), and chosen options during

[25] Network Health had covered all of Partners in 2011 and dropped all except two small hospitals on the islands of Nantucket and Martha's Vineyard. It also excluded the affiliated Partners physician groups. The eight non-Partners hospitals dropped were mostly community hospitals but included one less prestigious academic medical center (Tufts), one teaching hospital (St. Vincent).

[26] Testimony from CeltiCare's CEO to the Massachusetts Health Policy Commission supports this interpretation: "For the contract year 2012, Network Health Plan removed Partners hospital system and their PCPs [primary care physicians] from their covered network. As a result, the CeltiCare membership with a Partners PCP increased 57.9%. CeltiCare's members with a Partner's PCP were a higher acuity population and sought treatment at high cost facilities. … A mutual decision was made to terminate the relationship with BWH [Brigham & Women's] and MGH [Mass. General] PCPs as of July 1, 2013." (HPC 2013)

[27] The data was obtained via a data use agreement with the Massachusetts Health Connector, the exchange regulator. To protect enrollees' privacy, the data was purged of all identifying variables.

[28] I thank Amanda Starc and Keith Ericson for sharing Google Maps distance data between hospitals and zip codes.

fiscal years 2008-2013. I consider plan choices made at two distinct times: (1) when an individual initially enrolls in CommCare or re-enrolls after a gap in coverage, and (2) at annual open enrollment when premiums change and current enrollees are allowed to switch plans. A key difference between these two situations is their default choice. New and re-enrollees must make an active choice to receive coverage – if they do nothing, they are not enrolled and do not appear in the data.[29] By contrast, current enrollees who do not respond to open enrollment mailings and emails are defaulted to their current plan. Over 95% of current enrollees do not switch, suggesting the likely importance of the default. For the remainder of the year following each plan choice, I sum up each individual's total costs in the claims data, distinguishing between inpatient hospital costs (included in the hospital cost model) and non-inpatient costs. I use the non-inpatient costs to estimate an additional plan cost model described in Section 5. The tables in Appendix A show summary statistics for both the hospital and plan choice samples.

3 Reduced Form Adverse Selection Evidence

In this section, I present reduced form evidence of adverse selection against plans covering the Partners hospitals. I also show evidence consistent with the key predictions of the theory in Section 1. Specifically, I document a factor that seems to capture enrollee preference for Partners – past use of care at a Partners facility. These past users are (1) higher cost (even after risk adjustment), (2) more likely to select a plan that covers Partners, and (3) more likely to use Partners for subsequent hospitalizations. These facts suggest that past users *select* plans based on their desire to *use* Partners, whose high prices mean higher costs for insurers. I show that these enrollees' loyalty gives Partners substantial market power, since these enrollees will switch plans to retain access.

I present two types of evidence to show these facts. First, I use the standard positive correlation test for asymmetric information. Second, I examine evidence around a large plan's exclusion of Partners from network in 2012.

Positive Correlation Test

My analysis starts with the positive correlation test (Cardon and Hendel 2001; Chiappori and Salanie 2000). This method tests whether individuals who select generous plans (here, plans covering Partners) are more costly in ways not captured by factors that prices can vary on. This is a joint test for adverse selection and moral hazard, since either factor may drive the correlation. Indeed, in my setting, the selection and moral hazard effects of covering Partners are connected.

[29] An exception to this rule prior to fiscal year 2010 was that the exchange auto-assigned the poorest new enrollees who failed to make an active choice. I observe these auto-assignees and exclude them from the plan demand dataset.

I follow the "unused observables" approach of Finkelstein and Poterba (2013) for the correlation test. This procedure works by separately regressing plan selection and cost on factors on which prices vary plus "unused" factors that I observe but were not included in pricing or risk adjustment. In the 2011-2013 years I analyze, insurers were not allowed to vary prices *at all* across enrollees (i.e. full community rating) but their payment for an enrollee equaled price times the enrollee's risk score. Therefore, the only "used" factor in the regression is the risk score.[30]

I include an unused factor that is likely to capture individuals' preferences for Partners: whether they have used a Partners hospital in the past, either for inpatient or outpatient care. Outpatient care – which includes visits to a doctor whose practice is in a hospital – is much more common and accounts for most of the past use. Past use is, of course, not an immutable characteristic but a proxy to pick up *future* preference for Partners. It may do so both through time-invariant heterogeneity and provider loyalty (state dependence), and I do not attempt to separate the two. The positive correlation test requires only a correlation between cost and plan choice, and I use past Partners patient status to test for this correlation.

Table 3 shows the test results. The first two columns show regressions for plan choice (where the outcome is a dummy for choosing a plan covering Partners), and the remaining columns show regressions for costs. Among all consumers (column 1), past Partners users are 32.8% points more likely to choose a plan that covers it, a very large effect relative to the mean probability of 41%. One concern is that this estimate partly reflects consumer inertia in plans covering Partners. To address this, Column 2 restricts the sample to re-enrollees who I know are making an active choice.[31] Based on Partners use in their earlier enrollment spell, these enrollees are still 26.8% more likely to choose a plan covering Partners.

The next set of columns shows the results for costs. For the full sample, past Partners use is significantly associated with \$56.2 higher risk-adjusted monthly costs, a 15% increase relative to mean costs of \$387.[32] As the theory suggested, past Partners users may be more expensive either because they are unobservably sicker or because for the same sickness, they use expensive providers. I examine this in several ways. First, columns 4 and 5 separate the sample by whether the enrollee's plan covers Partners. In plans not covering Partners the estimate is a smaller but still significant \$36.5, while in plans covering Partners, the estimate is a much larger \$62.3 (and the difference between these is statistically significant). This pattern is consistent with *both* unobserved sickness and use of Partners driving the higher costs.

[30] I also include plan-year dummies (for costs) and year dummies (for plan choice) to net out any constant effects across plans or time. All results are robust to excluding these.

[31] I restrict the sample to people with a break of at least three months to rule out cases where the break represents a short lapse in paying premiums. This estimate is robust to limiting to longer breaks. Even among enrollees gone for more than two years, past Partners users are 21% points more likely to choose a Partners-covering plan.

[32] I perform a robustness check to test whether a limitation in the risk adjuster – its lack of diagnosis information for new enrollees – is driving these results. Limiting the sample to individuals with full diagnosis risk adjustment, past Partners use still significantly predicts higher risk-adjusted costs, with a magnitude similar to the full sample.

Enrollees who prefer Partners are more costly even in plans that do not cover it. But they are particularly costly when their plan gives them access to Partners facilities.[33]

To further interpret these results, the bottom half of Table 3 shows other attributes of past Partners patients in plans that cover Partners. The estimates show that controlling for risk score, they are 46.3% points more likely to choose a Partners hospital when hospitalized and $1,953 more costly per hospitalization. Again, these results are consistent with individuals' decisions to use Partners directly contributing to their higher costs. However, their risk-adjusted hospitalization rate is no higher – although their absolute rate is higher, the risk score picks up the difference.

Evidence from Plan Network Changes in 2012

A second way to test for selection is to study plan network changes. I focus on changes in 2012 that were both the largest in CommCare's history and the only time when the main Partners hospitals were dropped. As discussed in Section 2, this change occurred after the exchange introduced new incentives rewarding the lowest-price plans. In response, Network Health cut its price by about 15% and, to cut costs, excluded the Partners system and eight other hospitals from its network. Other plans also changed prices but did not make significant network changes at the time.

I start by studying plan choice patterns, again using past use as a proxy for Partners preference. Figure 1 shows the share of current Network Health enrollees who switched plans just before the start of each plan year. The average switching rate is usually very low (about 5%), but it spikes in 2012 to just over 10%. All of this spike is driven by patients of the hospitals Network Health dropped. Almost 40% of past Partners patients switched away from Network Health in 2012, a more than *seven-fold* increase from adjacent years. One factor behind this increase may be that Partners providers encouraged their patients to switch plans.[34] Most of these switchers moved to CeltiCare and Neighborhood Health Plan (NHP), the two remaining plans covering Partners. Patients at the eight other dropped hospitals were also more likely to switch but only about half as frequently as Partners patients. This is consistent with Partners' star power giving it much greater ability to influence plan choices than non-star hospitals.

Because the Partners patients are a high-cost group, these switching patterns had important cost implications. Table 4 shows the change in unadjusted and risk-adjusted costs for Network Health between 2011 and 2012. Overall, its per-member-month costs fell by 21% (or 15% after risk adjustment), a huge

[33] An alternate interpretation is that only unobserved sickness is driving the results but that past Partners users who do not choose a plan covering Partners are less unobservably sick. Against this interpretation, however, I show below that these individuals are much more likely to choose Partners hospitals when hospitalized.

[34] By chance, I observed Partners doing so during a tour of Brigham & Women's Hospital in late 2013 when a Medicaid managed care plan was about to drop Partners from its network. The finance department was calling past patients to let them know they needed to switch plans to maintain access to Partners providers.

decline in the health insurance industry where costs rarely fall. However, among a fixed population of "stayers" in Network Health in both years, costs fell by just 6%. The remainder of the change came through selection of enrollees leaving and joining the plan. The most expensive group was those who switched away from Network Health in 2012 – their 2011 risk-adjusted costs were $509 per month, almost 40% above the plan's average.[35]

The bottom panel of the table breaks down costs for switchers and stayers into past Partners patients (as of the start of 2012) and all others. It makes clear that Partners patients drove the high costs among switchers away from Network Health. They represented 68% of all switchers and had risk-adjusted costs of $571 per month in 2011 (54% above the plan average), whereas all other switchers had below-average costs. In comparison, the Partners patients who stayed with Network Health were somewhat less expensive – only $472 per month in 2011. They also experienced a substantial 26% decline in costs in 2012, accounting for *all* of the cost declines among stayers. Although mean reversion surely plays a role, this pattern is also consistent with dropping Partners having a *differential* cost impact for the people most likely to use it, as in the model.[36]

These network changes and selection patterns had important impacts on hospital choice patterns and costs. Figure 2 shows the impact on admission shares at the Partners hospitals (top panel) and the other eight dropped hospitals (bottom panel). In both cases, admission shares among Network Health enrollees fell sharply.[37] But for Partners hospitals, admissions also rose sharply at other plans – despite their Partners coverage not changing. Selection by enrollees most likely to use Partners appears to drive this increase. For the market as a whole, Partners admission share barely budged – despite being dropped by the second largest plan. The same was not true for the other, less prestigious dropped hospitals. Their shares at other plans did not increase much, and their overall market share fell. These hospital use patterns again illustrate the unique status of Partners in this market. Figure 3 shows the implications of these hospital use changes for plans' costs per hospital admission. Network Health's costs fell sharply in 2012 by about 15%, with the drop *entirely* driven by less use and lower prices at Partners hospitals.[38]

[35] In addition to the switchers, the group exiting the market was high-cost. While the reasons are unclear, exiting enrollees appear to be high-cost in other years and plans as well, not just in Network Health in 2011-2012.

[36] One fact suggesting reversion does not fully explain the difference is that the Partners patients who left Network Health had a much smaller 13% cost decline, despite having even higher baseline costs in 2011. Of course, this difference could also reflect selection. To further separate mean reversion from treatment effects, I could attempt to decompose the changes into quantity of services (e.g., number of hospitalizations) vs. price per service.

[37] Notably, these declines, while substantial, were less than 100% because patients can still use out-of-network hospitals in an emergency or if given prior authorization by their plan.

[38] Per-admission costs also fell sharply at Partners, with a likely explanation being exchange rules for out-of-network reimbursements when patients use them in an emergency. In these cases, plans are allowed to pay the hospital at the state's low Medicaid payment rates. Because this rule does not apply to non-emergency admissions that the plan authorizes, costs (and prices) at Partners are still higher than for an average hospital.

Conversely, per-admission costs at the two plans still covering Partners (NHP and CeltiCare) began rising in 2012, as more of their patients used the expensive hospitals.[39]

4 Structural Model: Hospital and Insurance Plan Demand

The reduced form evidence suggests that consumers select into plans covering the star Partners hospitals based on their preference for using those hospitals. Understanding the competitive and welfare implications of this selection requires estimating a structural model that can capture this correlation. In this section, I present and estimate the hospital and insurance demand portion of this model. I follow a method introduced by Capps, Dranove, and Satterthwaite (2003) to capture how different consumers value plans' hospital networks. I first estimate a hospital demand model that captures how patients weigh different factors (e.g., distance, hospital characteristics) when selecting hospitals. This hospital demand model generates an expected "network utility" metric capturing the attractiveness of each plan's network to a specific consumer. I then estimate an insurance plan demand model, including network utility as a covariate. If patients choose plans based on their hospital networks, the coefficient on network utility should be positive. Because of the importance of past Partners users in the reduced form results, I allow preferences in both the hospital and plan demand models to vary based on which hospitals an individual has previously used. This section proceeds by estimating hospital demand (Section 4.1) and deriving network utility (Section 4.2). I then present and estimate plan demand (Section 4.3-4.4).

4.1 Hospital Demand

I use the micro-data on inpatient hospital use to estimate a multinomial logit model capturing how patients choose hospitals. My method and specification are similar to much past work (e.g., Town and Vistnes 2001, Gaynor and Vogt 2003, Ho 2006). The main covariates are distance and hospital characteristics, and I allow preferences to vary with patient observables. I do not include patient fees as a covariate, since CommCare's copays are constant across in-network hospitals and therefore drop out.[40] In addition, I do not include an outside option, since I am focusing on patients sick enough to need hospital care and Massachusetts is a relatively complete hospital market.[41]

[39] Partners also does not appear to have given these plans price discounts after being dropped by Network Health. Costs per admission at Partners actually rise, but this is driven by a compositional shift among Partners hospitals, with little change in hospital-specific prices.
[40] Recent work by Ho and Pakes (2014) also finds that hospital price matters for referral patterns in plans where doctors are paid by capitation. Unlike their California setting, CommCare insurers pay doctors almost exclusively fee-for-service, with capitation accounting for less than 5% of physician service fees.
[41] The only significant exception is spillover of patients in Southeastern Mass. to hospitals in Providence, RI.

My model differs from past work in two main ways. First, based on the reduced form results, I allow hospital preferences to vary with whether a patient has used the hospital in the past (either for inpatient or outpatient care).[42] Although its interpretation is not obvious – it captures both heterogeneity and state dependence – I include past hospital use because of its importance as a channel for adverse selection. Second, because I observe a non-trivial number of out-of-network hospitalizations covered by plans (e.g., see Figure 2), I include out-of-network hospitals in the choice set. This captures the fact that patients can sometimes get plan authorization to see an out-of-network provider. To capture the associated hassle costs, I estimate a plan-specific out-of-network cost in the hospital choice model.[43] This specification generalizes previous work that disallows out-of-network admissions, which is equivalent to assuming an infinite hassle cost.

Consider an admission at time t for individual i (in plan j) who has principal diagnosis d. I specify the following model for the latent utility for hospital h:

$$u_{i,d,t,j,h} = \underbrace{\delta\left(Z_{i,d,t}\right) Dist_{i,h}}_{\text{Distance}} + \underbrace{\gamma\left(Z_{i,d,t}\right) X_h + \eta_h}_{\text{Hospital Characteristics}} + \underbrace{\lambda \cdot PastUse_{i,h}}_{\text{Past Use Dummy}} - \underbrace{\kappa_j \cdot 1\left\{h \notin N_{j,t}\right\}}_{\text{Out-of-Network Hassle Cost}} + \varepsilon_{i,d,t,h} \tag{5}$$

where $Dist_{i,h}$ is patient travel distance (and distance-squared), X_h are observed hospital characteristics, η_h is an unobserved characteristic (captured by hospital dummies), $PastUse_{i,h}$ are past use indicators, and $1\left\{h \notin N_{j,t}\right\}$ is an out-of-network dummy (and κ_j the hassle cost). I allow coefficients on distance and characteristics to vary with patient observables $Z_{i,d,t}$ to allow for preference heterogeneity. Finally, $\varepsilon_{i,d,t,j,h}$ is an i.i.d. Type 1 extreme value error that generates the logit demand form.

Because all of the covariates are observed, I estimate the model by maximum likelihood. Table 5 shows the results. Consistent with previous papers' estimates, patients have a disutility of traveling to more distant hospitals, with the estimates implying that an extra 10 miles distance reduces a hospital's share by 31% on average. The model estimates a sizeable hassle cost for out-of-network hospitals that reduces their shares by 63% on average.[44] The table shows the largest hospital service x diagnosis interactions; the remaining coefficients are almost all significantly positive.

Two sets of coefficients have implications for adverse selection. First, teaching hospitals and particularly the largest academic medical centers (including the two star Partners hospitals) attract the sickest patients – where severity is based on an index of the costliness of a patient's diagnoses defined in

[42] To rule out immediate readmissions, I require that the past use occurred more than 60 days beforehand.

[43] Patients can also use any hospital in an emergency (without needing plan authorization) but may need to be transferred once stabilized, creating a different type of hassle. I allow for an interaction between emergency status and the out-of-network cost but find little evidence that the cost is lower in emergencies.

[44] A 63% reduction from being out of network may seem low. However, it is consistent with a basic statistic from the data: only 25% of hospital choices are out of network but 8% of admissions are at out-of-network facilities.

Section 5.1. A one standard deviation increase in severity (a change of 0.3) increases the likelihood of using an academic medical center by 47%. Second, the past use dummies are very strong predictors of future hospital choices. For instance, patients who have previously used a hospital for outpatient care choose the same hospital in future admissions about 40% of the time. The model implies that this 40% share is an almost 5-fold increase above what would be expected otherwise.

The model fit is quite good, particularly when past hospital use variables are included. Calculating hospital shares at the service area-plan-year level,[45] the model explains 74% of the variance in shares, despite the absence of any year-specific interactions in the model. This indicates that conditional on network, hospital use patterns are fairly stable in the market over time. The left half of Table 5 shows estimates and fit from a simpler model (with only distance, out-of-network cost, and hospital dummies) for comparison. This simple model can also pick up 64% of the variance in shares. Most of the increase in fit from moving to the more complex model comes by adding the previous use covariates.

One concern with the out-of-network costs is that they are based on the network of a patient's chosen plan. Plan selection on observables (such as distance and past use) is okay, but if there is selection on unobservable hospital tastes, the out-of-network cost will be biased upward. This problem could be addressed econometrically by estimating the plan and hospital demand models jointly, allowing for unobserved hospital tastes to enter into plan choices (see Crawford and Yurukoglu 2012; Lee 2013). I have not implemented this method because of its computational complexity. One suggestion that any bias may not be too severe is that the model credibly matches hospital use patterns around Network Health's 2012 change in network (see Section 5.4). Nonetheless, the absence of plan selection on unobserved hospital preferences is a limitation of the model.

4.2 Hospital Network Utility

To generate a measure of network utility for plan demand, I follow the method of Capps, Dranove, and Satterthwaite (2003) and Ho (2006, 2009). I define network utility based on the expected utility metric from the hospital demand system. Conditional on needing to be hospitalized, a consumer's utility of access to network $N_{j,t}$ in plan *j* is:

$$HospEU_{i,d,t,j}\left(N_{j,t}\right) \equiv E\max_{h}\left\{\hat{u}_{i,d,t,j,h}\left(N_{j,t}\right)+\varepsilon_{i,d,t,j,h}\right\} = \log\left(\sum_{h}\exp\left(\hat{u}_{i,d,t,j,h}\left(N_{j,t}\right)\right)\right) \tag{6}$$

[45] Service areas are subregions defined by the exchange as the areas at which plans can choose whether or not to offer coverage. The five regions are divided into 38 service areas.

where $\hat{u}_{i,d,t,j,h}\left(N_{j,t}\right) \equiv u_{i,d,t,j,h} - \varepsilon_{i,d,t,j,h}$. At the time of plan choice, however, consumers do not know their hospital needs. Instead, they have expectations of their hospital use frequency for each diagnosis *d* over the coming year, which I denote $freq_{i,d,t}$. Given this expectation, the *ex-ante* expected network utility is:

$$NetworkUtil_{i,j,t}\left(N_{j,t}\right) \equiv \sum_{d} freq_{i,d,t} \cdot HospEU_{i,d,t,j}\left(N_{j,t}\right) \tag{7}$$

This network utility in (7) is what I include in plan demand. To calculate it, I first use my data to estimate a Poisson regression of the annual number of hospitalizations for each diagnosis on individuals' age and demographics.[46] I use the predicted values from these regressions for $freq_{i,d,t}$. Next, I calculate the value of $HospEU_{i,d,t,j}\left(N_{j,t}\right)$ for each plan and diagnosis, using the individual's location and demographics at the time of plan choice.[47] Finally, I input these values into equation (7) to calculate network utility. Because network utility does not have natural units, I normalize it so that 1.0 is the average decrease in utility for Boston-region residents when Network Health dropped Partners in 2012.

4.3 Plan Demand Model

I next estimate plan demand to capture how plan premiums and hospital networks influence consumers' choices. These estimates are important for capturing the extent of both market power (which is based on the price elasticity of demand) and adverse selection (which is based on the correlation between demand and cost). The demand estimates also generate a revealed-preference welfare measure capturing how individuals trade off generous networks against lower prices when choosing plans.

I use the dataset described in Section 2 to estimate a multinomial logit plan choice model for both new and current enrollees (allowing inertia for the latter, as I discuss below). I treat individuals' timing of entry/exit from the exchange as exogenous and model just their choices among exchange plans.[48] For new/re-enrollee *i* making a choice at time *t*, the model for utility of plan *j* is:

$$U_{ijt} = \underbrace{\alpha\left(Z_i\right) \cdot Prem_{j,t,Reg_i,Inc_i}}_{\text{Plan Premium}} + \underbrace{Network_{ijt}}_{\text{Hospital Network Vars.}} + \underbrace{\xi_{ijt}}_{\text{Unobs. Quality}} + \underbrace{\varepsilon_{ijt}^{Plan}}_{\text{Logit Error}} \tag{8}$$

---

[46] I choose not to use diagnoses in this regression because past diagnoses are unavailable for new enrollees. I plan to explore a robustness check in which for current enrollees I use past diagnoses and for new enrollees, I use a separate model including chronic disease diagnoses observed in the subsequent plan year.

[47] The two hospitalization variables that remain to be filled in are severity and emergency status. For emergency status, I use the average emergency probability for each diagnosis to take an average of the values of *EU* for each possibility. For severity, I regress severity in the hospitalizations data on age-gender groups and emergency status and use the predicted value from this regression for each individual.

[48] Because exogenous factors like income and job status determine exchange eligibility and generous subsidies incentivize participation, this assumption seems reasonable. This assumption implies that in my model, changes in plan prices and networks do not induce people to substitute into/out of the market. Although it would be nice to weaken this assumption, I do not have sufficient data on people choosing the outside option (largely uninsurance) to estimate a model incorporating it as a choice.

where:

$$Network_{ijt} = \beta_1(Z_i) \cdot NetworkUtil_{ijt} + \beta_2(Z_i) \cdot CoverPastUsed_{ijt}$$
$$\xi_{ijt} = \xi_{j,Reg_i,Inc_i} + \xi_{j,t,Reg_i}$$

and ε_{ijt}^{Plan} is an i.i.d. Type 1 extreme value error that gives demand its logit form. Plan utility depends on three sets of plan attributes: premiums, networks, and unobserved quality. Premiums – which vary across plans and within-plan across years, regions, and income groups – are observed, and I include them directly. Hospital networks are more difficult because while observed, the value of a given network varies across individuals. To capture this heterogeneity, I include two terms: the consumer-specific network utility measure (see Section 4.2) and a direct variable for whether the plan covers a consumer's previously used hospitals (or the share covered if there are multiple). Of course, these two variables are related, since past use entered hospital demand and therefore influenced network utility. However, the direct variables may predict demand beyond their impact on hospital network utility for several reasons. First, they may capture loyalty to doctors, who in Massachusetts are often hospital-affiliated and covered/dropped along with the hospital.[49] Second, it may be picking up error in hospital demand or the sickness frequency prediction. Finally, it may matter simply because plan and hospital choices are driven by different things. People may choose plans based on whether it covers their regular provider but hospitals based on many other factors (e.g., which hospital is closest in an emergency).

The third set of covariates in plan demand (ξ_{ijt}) are plan dummies capturing unobserved quality – e.g., customer service and plan reputation.[50] To aid identification of the premium coefficient (see discussion below), I allow these to vary at a detailed region-year and region-income group level.

Preference heterogeneity enters this model in two ways. First, I allow observed heterogeneity by income, age, and gender groups for the premium coefficient and by income group for network utility. Second, the network variables also incorporate heterogeneity, since (for the same plan) they vary by consumer location, sickness, and past relationships with providers. This heterogeneity is useful for capturing substitution patterns and adverse selection.

Current Enrollees and Inertia: The model so far has applied to new/re-enrollees, who I can be sure are making active choices. A final issue is how to treat current enrollees, who can switch plans at annual open enrollment but are defaulted into their current plan if they take no action. There is growing evidence that defaults and inertia matter in health insurance (Ericson 2014, Handel 2013), and consistent with this, I find that fewer than 5% of enrollees switch plans each year. However, how to interpret this

[49] Though I do have information on physician networks and utilization, I have not yet modeled physician demand or network utility because of its complexity.
[50] Past work has found reputation to be an important driver of demand in the Medigap insurance market (Starc 2011), and based on my discussions with market participants, reputation is also important in CommCare.

low switching rate is less clear. It may reflect a combination of true inertia/switching costs (a form of state dependence) and preference heterogeneity causing optimal choices to be serially correlated.[51]

While I am not able to fully separate these factors, I want the model to capture switching behavior because of its implications for selection. To do so, I take a reduced form approach. In addition to the terms in equation (8), current enrollees' utility includes a dummy for their current plan. I allow the coefficients on this dummy to vary with observed demographics and (based on the evidence in Section 3) whether the plan has just dropped a previously used hospital. These inertia coefficients can be interpreted as either switching costs or reduced form coefficients capturing the likelihood of consumers being passive/inattentive in their switching choice, and I report statistics for both interpretations.[52]

Including current plan dummies ensures that the model will match average switching rates for each group with a separate coefficient. However, the coefficients themselves will pick up both true inertia and any unobserved heterogeneity driving choice persistence. For matching static adverse selection, it is not clear that it is critical to distinguish these factors. Where the two specifications will primarily differ is in their implications for dynamic competition, which I do not study in my counterfactuals. However, in interpreting the inertia estimates, readers should keep in mind that these coefficients are also picking up unobserved heterogeneity.[53]

Identification and Estimation: I estimate the model using a micro-data method of moments estimator similar to Berry, Levinsohn, and Pakes (2004). A key difference in my setting is that the main plan attributes – premium and network utility – vary across individuals even for the same product in the same market and year. As a result (under assumptions discussed below), I can estimate the premium and network coefficients consistently from the micro-data alone, without needing instruments.

To identify the premium coefficients, I use within-plan variation induced by CommCare's subsidies. The key variation is that higher price plans have higher premiums for above-poverty enrollees but the same premium (always $0) for fully subsidized below-poverty enrollees. This structure also creates differential premium changes across years, which I use for identification. Figure 4 shows how these differential changes work with an example from Network Health in the Boston region in 2010-2011. In 2010, Network Health was the cheapest plan for all groups. In 2011, its relative price increased but while above-poverty groups' premiums increased, below-poverty premiums were unchanged (still $0).

[51] This low switching rate does not appear to only reflect heterogeneity. Enrollees who enter the exchange just after prices have changed end up with very different market shares overall than enrollees who entered just before the price change. This group-level share difference is strongly suggestive the true state dependence is involved.
[52] In Appendix B, I show how this maps into a particular two-step model of inattention, where the first step models whether an enrollee is passive or active and a second step models plan choice conditional on being active.
[53] In a future revision, I plan to do a robustness check with a demand model that includes time-invariant unobserved heterogeneity through random coefficients on premiums and plan dummies. I will use the choice patterns of re-enrollees to separately identify the random coefficient variances from the switching costs.

I use these differential premium changes for identification by absorbing all other premium variation with a detailed set of plan dummies. Recall that because of regulation, premiums vary only across plans, years, regions and income groups. The first set of dummies (ξ_{j,Reg_i,Inc_i}) absorb any persistent demand differences for plan *j* across income groups (within a region). The second set of dummies (ξ_{j,t,Reg_i}) absorb demand differences across regions and over time. The remaining variation is from within-region differential premium changes across income groups. Because I allow a separate premium coefficient for each above-poverty group, the main identification comes from comparing demand changes for each above-poverty income group to those of below-poverty enrollees.

This identification strategy is a nonlinear version of the standard difference-in-differences approach. Thus, the key assumption is that any changes in unobserved plan quality evolve in parallel for low- and high-income enrollees. This assumption seems reasonable because all groups have access to a plan under the same brand name, with the same provider network and member services. However, to test its validity, I employ the standard parallel trends test for difference-in-differences. This test compares trends for the outcome (market shares) around price changes for the treatment group (above-poverty) versus the control group (below-poverty). Figure 5 shows this test, plotting average market shares for new enrollees in each month around price changes, separately for price cuts (top graph) and price increases (bottom graph).[54] Consistent with the parallel trends assumption, market share trends are flat and parallel for both groups at all times except when prices change. At price changes, price-paying groups' shares jump sharply in the expected direction, while zero-price groups' shares are essentially unchanged.

The detailed plan dummies are also helpful for proper identification of the network utility coefficients. The potential identification threat is that plans with better networks also have better unobserved quality. However, with the plan dummies, the network utility coefficients are identified from *within-plan* variation across individuals in the same region and year. A key source of variation is enrollees' location relative to covered hospitals, since this strongly predicts hospital utility.

I estimate the model using moments similar to those used in Berry, Levinsohn, and Pakes (2004).[55] For plan dummies, I match observed market shares for the relevant plan and enrollee group. For plan characteristics (whose coefficients vary with observed demographics), I match the average interaction between the characteristic and the demographic among chosen plans in the data. Appendix B shows the formulas for these moments.

[54] The analysis is restricted to fiscal years 2008-2011. I drop 2007 because above-poverty enrollees did not start enrolling in the market until mid-way through 2007. I drop 2012+ because below-poverty new enrollees become subject to a limited choice policy that required them to choose lower-price plans.

[55] I use method of moments rather than maximum likelihood for two reasons. First, my network utility covariates are not observed, and I employ a standard error correction that is valid for method of moments. Second, in future revisions, I plan to include random coefficients, for which simulated method of moments is more appropriate.

4.4 Plan Demand Estimates

The demand estimates are shown in Table 6. Premiums (in dollars per month) enter negatively and significantly for all income groups. (I normalize the average premium coefficient to -1.0, so the remaining coefficients can be interpreted as dollar values for an average enrollee.) Premium sensitivity decreases monotonically with income, with the highest-income group's coefficient less than half as large as the lowest-income group's. Premium sensitivity also decreases with age, although much less sharply. Overall, these estimates imply that new enrollees are quite premium sensitive. A $1 increase in monthly premium decreases the average plan's market share among premium-paying enrollees by 3.0%.[56]

Enrollees place positive and significant value on both measures of hospital network quality. Recall that network utility was normalized so that 1.0 was the average utility change for Boston-area enrollees when Network Health dropped Partners in 2012. Thus, for an average Boston enrollee with no previous Partners use, the estimates indicate a modest $6-8 monthly value of Partners access. This positive but modest average value of broader networks is consistent with the findings of Ho (2006, 2009), who estimated a similar model for employer-sponsored insurance. However, this average masks significant heterogeneity both in the network utility of Partners and in the marginal utility of money. In addition, I estimate substantial coefficients on the direct measure of whether a plan covers an enrollee's previously-used hospitals. For non-Partners hospitals, I estimate an additional value of $5.41 per month and for Partners hospitals, the total effect is $17.04 per month.

As expected, I find substantial inertia in consumers' plan switching decisions. Rationalizing observed switching rates requires an average switching cost of $96.8 per month, or equivalently, an average 94.6% probability of passively choosing.[57] Though large, these estimates are actually a bit smaller than the average switching costs found in an employer insurance setting by Handel (2013) of $2,032 per year (or $169 per month). What is most interesting for selection on networks is that estimated inertia decreases when a plan drops an enrollee's past used hospital from network. For dropped non-Partners hospitals, enrollees are 19% points less likely to be passive and for Partners hospitals, they are 43% points less likely to be passive. A possible explanation is that when an enrollee's regular provider is being dropped, the provider contacts the patient and encourages them to switch plans. Whatever the reason, this inertia reduction exacerbates adverse selection, consistent with the findings of Handel (2013).

[56] Because prices are subsidized, there are two ways to convert this semi-elasticity into an elasticity. Relative to consumers' relatively low premiums (which average about $45 for premium payers), the elasticity averages a relatively modest -1.35. However, relative to insurers' full prices (about $400 on average) – the statistic relevant for insurers' markups – the demand elasticity is -11.9.

[57] While these estimates are also picking up unobserved heterogeneity, a simple calculation suggests that the passive probability would still be about 90% with a realistic degree of heterogeneity (based on the 55% rate at which re-enrollees choose the same plan as they had before). If there were 55% persistence among current enrollees who were active choosers, 91% of people must have been making passive choices to explain a 96% non-switching rate.

Here, the inertia reduction is particularly important because it occurs precisely among some of a plan's most expensive consumers, past patients of the Partners hospitals.

5 Structural Model: Insurer Cost and Profit Functions

The adverse selection implications of hospital networks depend on the interaction between demand and costs. In this section, I specify a model for insurer costs and (combining this with demand) derive the insurer profit function. The goal is to capture insurer incentives to cover or exclude high-price star hospitals like those in the Partners system. These incentives depend both on how covering Partners affects individual-level costs and how it affects plan selection by individuals of different costliness.

I start by modeling how individual-level costs would vary in plans with different hospital networks. Section 5.1 describes how I model insurer costs for hospital care, which uses my hospital demand model and a set of estimated hospital prices. Section 5.2 then presents my model for non-hospital costs. In Section 5.3, I aggregate this individual-level cost model up to the insurer level (using plan demand to predict plan choices) and derive the insurer profit function. Finally, Section 5.4 considers model fit and analyzes the 2012 change when Network Health dropped Partners.

5.1 Hospital Prices and Insurer Costs for Hospital Care

To model insurer costs for inpatient hospital care, I start from an individual-level model. I condition on each person's set of observed hospitalizations (and their diagnoses) and ask how hospital choices and costs would have changed if the patient had been in a different plan with a different hospital network. An advantage of this approach is that it lets me capture the correlation between hospital use and enrollee attributes (which determine plan selection) in a rich, nonparametric way.[58] Nonetheless, this approach assumes networks do not affect the *number* of hospitalizations, only the hospitals chosen when sick.[59]

I first estimate the prices insurers pay to hospitals for inpatient care using the payment data in the insurer claims. Because actual payment rules are unknown (and likely quite complicated), there is a need for simplification. I follow past work (Gowrisankaran, Town, Nevo 2013) in estimating *average* payment

[58] The potential danger is over-fitting. Because I have a large sample and consider only insurer actions that affect a large set of individuals (prices and coverage of Partners), over-fitting is less of a concern for my purposes.

[59] This assumption is likely a good first approximation but is not perfect. Recent evidence from Gruber and McKnight (2014) finds small reductions in the number of hospitalizations in limited network plans. If applicable in my setting, my model will somewhat understate the cost savings from plans' limiting their networks.

factors that capture proportional differences across hospital-insurer pairs.[60] I estimate a Poisson regression (also known as a generalized linear model with a log link) of the form:

$$E\left[Payment_{i,j,h,t,a} \mid Diag_{ita}, Z_{ita}\right] = \exp\left(\rho_{j,h,t} + Diag_{ita}\lambda + Z_{ita}\gamma\right) \tag{9}$$

where a indexes the admission, $Diag_{ita}$ is the principal diagnosis, and Z_{ita} is other patient covariates.[61] The key term is $\rho_{j,h,t}$, which is a coefficient that captures average payment differences across hospitals, insurers, and years.[62] This effect is assumed to be proportional across all types of admissions, which is surely not exactly right but should capture a valid average effect. Appendix B discusses additional details on the hospital price estimation.

I use the estimates of (9) to define hospital prices as $\hat{P}_{j,h,t} \equiv \exp\left(\hat{\rho}_{j,h,t}\right)$ and an admission-specific severity measure as $\hat{\omega}_{i,t,a} \equiv \exp\left(Diag_{ita}\hat{\lambda} + Z_{ita}\hat{\gamma}\right)$. I scale $\hat{\omega}_{i,t,a}$ so that its mean is 1.0 and divide $\hat{P}_{j,h,t}$ by the same factor, so it can be interpreted as the hospital price for a patient of average severity. The average prices and severities for the 10 most expensive hospitals are shown in Table 1.

I use these severities and prices to model how hospital costs would differ in counterfactual plans and networks. As discussed above, I condition on each individual's observed admissions (or lack thereof) and severities $(\hat{\omega}_{i,t,a})$ and use hospital demand to predict how these admissions shift across hospitals. The hospital costs for individual i in year t in plan j (with network N_{jt}) is:

$$c_{ijt}^{Hosp}\left(N_{jt}\right) = \sum_{a=1}^{NAdmits_{it}} \hat{\omega}_{i,t,a} \cdot \left(\sum_{h} \hat{P}_{j,h,t} \cdot s_{i,d,t,j,h}^{Hosp}\left(N_{jt}\right)\right) \tag{10}$$

For most hospitals, I use only the plans' observed networks so hold hospital prices fixed at the estimated values. However, for Partners hospitals, I also consider adding/dropping them and therefore need a counterfactual price model. For this, I use a simple average of prices paid by insurers that actually covered (excluded) the Partners hospital in a given year. The main limitation of this approach is that it does not capture insurer-hospital bargaining dynamics, something I have not yet modeled.[63]

---

[60] Following convention, I refer to these payment factors as "prices," although they are distinct from the actual negotiated prices. These payment factors capture both price differences and service quantity differences across hospitals (conditional on diagnosis) since both affect insurers' payment differences across hospitals.

[61] For the principal diagnosis, I use the Clinical Classification Software (CCS) dummies defined by the U.S. government's Agency for Healthcare Research and Quality. The additional covariates include age, gender, income, and Elixhauser comorbidity dummies for the secondary diagnoses.

[62] As discussed in Appendix B, I specify a restricted model for $\rho_{j,h,t}$ to avoid over-fitting for hospital-insurer-year cells with small samples. I allow for flexible hospital-insurer and insurer-year dummies, separately by in- and out-of-network status, plus a separate insurer-year factor for each of the six largest hospital systems.

[63] Two facts suggest this approach may be a reasonable approximation in this setting. First, within-year price variation across insurers for the main Partners hospitals is small in practice – standard deviations for Brigham and

5.2 Non-Hospital Costs

I complete the cost model by considering all costs other than inpatient hospital care. Unfortunately, I do not have a provider choice model for non-hospital care through which I could define costs analogously to my hospital cost model. Instead, I take a reduced form approach. I calculate monthly non-inpatient costs for each enrollee-year and use them to estimate the following Poisson regression model:

$$E\left(NonHospCost_{it} \mid Z_{it}\right) = \exp\left(\eta_{j(i),Reg(i),t} + Z_{it}\mu\right) \tag{11}$$

where Z_{it} are detailed enrollee diagnoses and demographics.[64] I use these estimates to define a region-year-specific plan effect $\hat{C}_{j,Reg,t} \equiv \exp\left(\hat{\eta}_{j,Reg,t}\right)$, an enrollee severity $\hat{\varsigma}_{it} \equiv \exp\left(Z_{it}\hat{\mu}\right)$, and an enrollee residual $\hat{\nu}_{it} \equiv NonHospCost_{it} / \left(\hat{C}_{j(i),Reg(i),t} \cdot \hat{\varsigma}_{it}\right)$. If an enrollee switches to plan *k*, I assume that his severity and residual are unchanged but that the plan effect switches to the counterfactual plan, so the enrollee's new cost is $\hat{C}_{k,Reg,t} \cdot \hat{\varsigma}_{it} \cdot \hat{\nu}_{it}$. This reduced form approach is clearly an approximation. However, the $\hat{C}_{j,Reg,t}$ estimates should capture a valid average plan effect on costs absent unobserved cost-based selection into plans. Given that I have documented unobserved selection based on the exchange's risk adjustment, this assumption is clearly imperfect.[65] If there is residual selection, I will understate costs for plans attracting residually healthier enrollees and overstate costs in the opposite case.[66] This will affect my estimates of the *level* of non-inpatient costs at observed networks but not the cost *difference* from network changes, which I specify separately next.

Networks may affect non-inpatient costs both through outpatient hospital care and through secondary effects on services like drugs and post-acute care. For the effect of adding/dropping Partners, I again specify a reduced form adjustment.[67] I first use the hospital cost model to calculate inpatient costs

---

Mass. General are just \$359 and \$846, respectively. Second, when Network Health drops Partners, I see little change over the next two years in Partners prices paid by the plans that still cover it.

[64] For diagnoses, I use the Hierarchical Condition Categories (HCC) defined by Medicare for its risk adjustment. I use HCCs observed in the *current* plan year so I can include diagnoses for new enrollees.

[65] The covariates in (11) will do somewhat better than the exchange risk adjustment because they include concurrently observed diagnoses, which allows for including diagnoses for new enrollees.

[66] To address this potential bias, I plan in a future revision to instrument for plan enrollment using the *timing* when an enrollee entered the exchange. Because of inertia, enrollees who enter just before a price change will have different plan shares at a later date *t* than enrollees who enter just after the price change. Assuming that entry timing does not independently affect costs and that attrition is independent of unobserved costs, then entry timing in the exchange is a valid instrument for current plan enrollment.

[67] Past structural work on hospital networks has generally either ignored non-inpatient costs or assumed that they did not change with the hospital network. My reduced form method, though imperfect, improves on the past literature.

with and without the network change. I then assume that a plan's non-inpatient costs change in proportion to the *regional average* change in inpatient costs.[68] The final non-inpatient cost model is:

$$c_{ijt}^{NonHosp}\left(N_{jt}\right)=\underbrace{\hat{C}_{j,Reg,t}}_{\text{Plan Effect}}\cdot\underbrace{\left(\hat{\varsigma}_{it}\cdot\hat{v}_{it}\right)}_{\text{Ind. Severity and Residual}}\cdot\underbrace{\left(1+\lambda\cdot\%\Delta HospCost_{j,Reg,t}\left(N_{jt}\right)\right)}_{\text{Network Cost Adjustment}} \tag{12}$$

Based on a risk-adjusted regression at the plan-region-year level, I find that each 10% increase in inpatient costs is typically associated with a 3.8% increase in non-inpatient costs and therefore set $\lambda=0.38$. However, I can do robustness checks with alternate values of λ.

5.3 Total Costs and Insurer Profits

With a model for both individual-level inpatient hospital and other costs, I sum them to define total costs, $c_{ijt}^{Total}\left(N_{jt}\right)$. I also include in total costs a measure of variable plan administrative costs (e.g., for claims processing) based on plan financial reports to the exchange.[69] The final model step is to aggregate costs and revenue up to the plan level using the demand function. The annual profit function for plan *j* is:

$$\pi_{jt}=\sum_i\left(\varphi_{it}P_{jt}-c_{ijt}^{Total}\left(N_{jt}\right)\right)\cdot D_{ijt}\left(\mathbf{Prem},\mathbf{N}\right) \tag{13}$$

where P_{jt} is the plan's price, φ_{it} is the exchange's risk adjustment score for enrollee *i*, and $D_{ijt}(.)$ is the enrollee's demand for plan *j*. Demand is in units of member-months and is the product of two terms:

$$D_{ij}\left(\mathbf{Prem},\mathbf{N}\right)=nMon_i\cdot S_{ij}\left(\mathbf{Prem},\mathbf{N}\right)$$

The first is the number of months an individual is enrolled in the exchange during the year. Many enrollees enter or leave in the middle of the year (e.g., because of a change in jobs that affects their eligibility), and I assume this enrollment churn is exogenous and hold $nMon_i$ fixed as observed. The second term is consumer *i*'s predicted share for plan *j* from the logit demand system.

5.4 Model Fit and Analysis of 2012 Network Health Change

Figure 6 shows the model fit for plans' average monthly medical costs per enrollee. The model averages are calculated using the model's cost and demand functions (as in (13)), creating two potential sources of errors versus the costs in the data. Nonetheless, the fit is quite good, with an R^2 at a plan-year level of 0.926. Importantly, the model captures very well the large fall in costs for Network Health in 2012 when it dropped Partners. The largest errors are predicting too high costs for CeltiCare in 2010 and 2011 (when

[68] A limitation with this method is that it does not capture *differential* percent changes for the people most likely to use Partners. In future revisions, I could consider using Network Health's 2012 dropping of Partners to estimate a more flexible model of non-hospital cost changes that varies with individual covariates.

[69] To do so, I estimate a regression of plan's reported administrative costs on their total enrollment. I find an almost perfect linear fit with a coefficient of about $30 per member-month, which I use for the model.

it was a new plan and had very low enrollment), although the model does capture its large cost increase in 2012 after Network Health dropped Partners.

I next consider in more detail how well the model matches the cost and demand patterns for Network Health in 2012. Appendix C shows a series of figures and tables similar to those analyzed in Section 3, with values predicted by the model added on. The model captures the variations in switching rates among Network Health's current enrollees quite well. Past Partners' patients switching rate is matched almost perfectly – since the model's interaction of switching costs with dropping Partners is largely identified from the 2012 change. It also captures the intermediate level of switching for patients of other dropped hospitals. The next table shows how the model matches the cost change from 2011 to 2012. For the average costs and cost changes, the model matches almost perfectly. Breaking it into enrollee subgroups, the model captures the basic pattern that enrollees who left the plan in 2011 were much more expensive and that the cost decrease for stayers was smaller than the overall decrease, though it slightly underestimates the former and overestimates the latter.

The final set of figures analyze how the hospital model captures changes in admission shares and costs at Partners and other dropped hospitals.[70] In all cases, the fit is quite good. In particular, the model matches the striking fact that Partners admissions fell for Network Health, rose at other plans, and barely changed overall. It also matches Network Health's and other plans' costs per hospital admission in levels and trends (including the 15% drop for Network Health in 2012).[71]

Finally, I use the model to decompose how much of the 15% decline in Network Health's risk-adjusted costs was due to selection versus "real" cost reductions. One indication that selection played a large role is that costs declined just 6% on a fixed population of stayers in the plan in both 2011 and 2012 (see Table 4). However, this statistic does not capture the full effect of real cost cuts, which would also have applied to the people who switched plans had they not left. Instead, I use the model to decompose how changes in plan selection versus changes in the cost function affected costs. Formally, I can decompose the 2011-2012 change in costs into:[72]

$$Cost_{2012} - Cost_{2011} = \underbrace{\sum_i \left(c_{ij,2012} - c_{ij,2011}\right) \cdot D_{ij,2012}}_{\text{Cost Function Change}} + \underbrace{\sum_i \left(D_{ij,2012} - D_{ij,2011}\right) \cdot c_{ij,2011}}_{\text{Selection}}$$

where $j = \text{Network Health}$. Based on this decomposition, I find that selection explains 50% of Network Health's reduction in costs, with the rest due to a lower cost function for a fixed population. Notice that

[70] To focus on the hospital demand and cost model's ability to fit patterns, these figures condition on people's actual plans, rather predicting plan shares using the plan demand model.

[71] I have found that including the past hospital use variables in the hospital demand system are important to matching these patterns so well. Without these covariates, for instance, the model cannot match the sharp rise in Partners admissions for plans other than Network Health in 2012.

[72] Because this decomposition requires observing individuals in both years, I restrict the sample accordingly.

this decomposition calculates the cost function effect with 2012 shares and the selection effect with 2011 costs. If instead, I calculate the cost function effect with 2011 shares, the cost reductions are larger, and selection explains 36% of the decline. This difference implies that many of the people whose costs would have declined the most selected out of Network Health in 2012. Selection attenuated the cost-reducing effects of a change in networks.[73] Either way, however, selection was important, explaining between 36-50% of Network Health's cost reduction.

6 Model Analysis: Heterogeneity in Value and Cost of Partners

Having estimated the model of insurance and hospital demand, I use the estimates to study heterogeneity in consumers' costs and value of Partners coverage. For simplicity, I focus on current enrollees in the exchange at the start of 2012, when Network Health dropped Partners. I define utility for Partners based on the difference in plan utility for Network Health, excluding switching costs (U_{ijt} in equation (8)), with and without Partners covered. I convert utilities into dollar values by dividing by each individual's marginal utility of money (the negative of their premium coefficient).[74] I calculate costs based on the cost function for Network Health with and without Partners.

Table 7 shows these estimates for all current enrollees in the exchange at the start of 2012.[75] The rows are sorted by the measure of Partners value. About 80% of enrollees have relatively little value for Partners coverage, with a monthly value of $4.30 or less – quite small compared to the typical variation in plan premiums of $20-60 per month. But value for Partners rises sharply in the top 10-20% of enrollees, with the top 5% valuing Partners at $46.80 per month. For these enrollees, almost all of whom are past Partners patients, Partners coverage plays a determinative role in their plan choices.

The remainder of Table 7 shows how these differences in value for Partners correlate with costs. I distinguish between two sources of adverse selection discussed in the theory in Section 1: selection on unobserved risk and selection driven by use of Partners' high-price care. Columns (2)-(3) suggest that unobserved risk is important. Even without Partners covered, people in the top decile of Partners value have risk-adjusted monthly costs of about $350, which is $50 (or 17%) higher than those who value Partners the least. Column (5) indicates that selection on use of Partners is also important. The ΔCost from covering Partners rises from $8.0 (2.7%) for the lowest-value group to $48.5 (10.0% of a larger base) for the highest-value group. Combining both types of selection, the people in the top decile of Partners values are $84 (or 27%) more expensive (after risk adjustment) in a plan covering Partners than

[73] This is consistent with one of the implications of the selection on moral hazard model in Einav, et al. (2013).
[74] I exclude below-poverty enrollees from this calculation because I cannot estimate their premium coefficient.
[75] I focus on current enrollees because their past hospital use (a key model covariate) is more likely to be observed.

people with below-median values. Of this $84 difference, about 60% is due to selection on unobserved risk and the remainder due to differential use of the Partners system.

A final insight from Table 7 is that for each group, the estimated consumer value from access to Partners falls short of the increase in insurer costs. Even aside from adverse selection, this fact gives insurers a strong incentive to drop Partners. However, this does not prove that the welfare effect of covering Partners is negative for all groups. Part of plans' higher costs represent markups to the Partners hospitals, which may be used for socially valuable purposes like teaching and research. To account for these markups, I draw on a Massachusetts government estimate of the per-admission costs of Partners (CHIA 2014).[76] Based on these estimates, the cost per admission at the two star Partners hospitals in 2012 were about $12,500 (Mass. General) and $13,800 (Brigham), implying margins of about 30-35% relative to my estimated prices. Column (7) shows the net cost increase, subtracting the change in Partners net revenue for inpatient care from insurer costs.[77] After doing so, the net cost increase for people in the top decile of Partners values is substantially lower. Their value for Partners coverage now exceeds the estimate of net costs. However, value still falls short of net costs for people in the bottom 90% of Partners valuations.

7 Equilibrium and Analysis of Policy Solutions

This section uses the demand and cost estimates to simulate equilibrium in a model of insurance competition. I use this to examine the impact of different policies used to address adverse selection in insurance exchanges. In general, insurer competition on prices and hospital networks is extremely complicated and subject to multiple equilibria.[78] To make progress, I focus on a static model where insurers compete only on price and coverage of the expensive Partners hospitals, holding hospital-insurer prices and other aspects of the network fixed. Although stylized, this model goes beyond most past empirical work on selection, which studies pricing holding fixed product characteristics.[79]

[76] The measure is of hospitals' "inpatient cost per case mix adjusted discharge". The calculation, which is based on hospitals' cost reports to the state, is intended to be a comprehensive measure of average hospital costs (including fixed costs of facilities), excluding physician compensation and graduate medical education costs.

[77] Note that this values each $1 of Partners net revenue as $1 of social value. This calculation is imperfect because it excludes Partners markups for non-inpatient care and reductions in net revenue for non-Partners hospitals. The latter, however, are likely to be small; non-star hospitals often have low or negative margins (Ho 2009).

[78] For an innovative model incorporating hospital-insurer bargaining and network formation in a dynamic equilibrium framework, see Lee and Fong (2013).

[79] For example, see Einav, Finkelstein and Cullen (2010), Starc (2014), and Handel, Hendel, and Whinston (2013). One recent paper by Einav, Jenkins and Levin (2012) does consider the effect of selection on product design in consumer credit markets but does so in a setting with a monopolist firm.

7.1 Equilibrium Simulations: Method and Results

Consider a model of insurance market equilibrium for a particular year (e.g., 2012) in the Massachusetts exchange. As in Massachusetts, I assume that each insurer offers a single plan with exchange-specified consumer cost sharing and covered service rules.[80] I condition on the plan's past history, including past network coverage and the set of current enrollees entering the year. I also hold fixed (at observed values) each plan's network and payment rates for all non-Partners hospitals. Before the year, the exchange announces policies (e.g., subsidy and risk adjustment rules). Insurers then compete in the following game:

Insurer Competition: 1. Insurers choose whether to cover Partners hospitals

2. Insurers set plan prices

Consumer Demand: 3. Consumers choose plans

4. Sick consumers choose providers (based on plan network)

I assume that insurers observe networks from stage 1 when setting prices and that they have full information on all demand and cost functions.

Insurers make choices to maximize profits, following the model for demand, costs and profits estimated in Sections 4-5. However, there is one additional simulation issue: how to incorporate dynamics arising because of enrollee inertia. When a plan lowers its price and attracts more enrollees today, it increases its future demand because some enrollees will passively stick with the plan in following years. This can lead to an invest-then-harvest equilibrium in which plans cycle between low and high prices. I choose not to specify a fully dynamic model, which would be both complicated and unrealistic unless it modeled uncertainty about policy changes (which occurred frequently in Massachusetts). Instead, I take a simple static approach that approximates a dynamic model. I assume that enrolling someone today increases future profits in proportion to the person's future duration in the market and the current profit margin on the individual. This "future profit effect" gives insurers an added incentive to keep prices low and helps offset the lower price elasticity of demand due to inertia. Appendix D shows the modified pricing first-order conditions and lays out additional details for the simulation method.

In full-information Nash equilibrium, each insurer sets prices in step 2 to satisfy its first-order conditions given all other insurers' prices and networks. In step 1, they choose Partners coverage knowing the pricing equilibrium that will prevail for each network possibility. For Partners coverage, I assume a binary choice: either sticking with their actual coverage of Partners or adding/dropping all of the Partners hospitals. I do not model the vertical relationship between Partners and Neighborhood Health Plan (NHP)

[80] In the ACA, insurers can offer multiple plans with varying networks across four tiers of cost-sharing generosity. Unfortunately, my data (in which all plans have the same cost sharing) do not make it possible to model cost-sharing differentiation. However, in future analysis, I could study equilibrium when insurers can offer two plans that vary in whether they cover Partners.

but allow it to flexibly cover/drop Partners.[81] Nash equilibrium occurs at a set of networks $\mathbf{N}$ if no insurer wishes to unilaterally deviate: $\pi_j\left(N_j,\mathbf{N}_{-j}\right)\geq\pi_j\left(\tilde{N}_j,\mathbf{N}_{-j}\right)\ \forall\tilde{N}_j,j$. While there is no guarantee of a unique equilibrium, I do not find multiplicity in my main results.

Table 8 shows equilibrium insurer choices for several simulations.[82] The top panel shows equilibrium under the actual Massachusetts subsidy and pricing policies in 2011, comparing these to the observed prices and networks.[83] The model's prices match extremely well. But this occurs largely because Massachusetts had a narrow allowed price range, and all plans bid at or near the range's min or max.[84] Nonetheless, the model captures well which insurers priced near the min versus the max. For networks, the model predicts just one plan (CeltiCare) willing to cover Partners, while in reality Network Health and NHP also covered Partners in 2011. However, Network Health did drop Partners in 2012, and Partners announced intentions to buy NHP in August 2011, a factor that I do not model. It is interesting that the model can rationalize CeltiCare's surprising decision (as the low-price plan) to cover Partners. In the model, CeltiCare is willing to do so because of the binding price floor. Without a price floor, CeltiCare instead cuts its price and drops Partners.

Because many of Massachusetts' distinct rules did not continue under the ACA, I perform the rest of the analysis using rules closer to those in the ACA. Specifically, I include only the 100-300% poverty population (those below poverty generally get Medicaid in the ACA), set subsidies as a flat amount for all plans, and do not impose minimum or maximum prices.[85] Panel B of Table 8 shows the simulation results. Under ACA-like policies, the model predicts that all plans drop Partners, and this result is robust across all the simulation years, 2011-2013. When an insurer deviates to cover Partners, its costs go up for all of its enrollees, and it particularly attracts the enrollees who most value Partners and whose risk-adjusted costs are high. But by raising its price to compensate, it reduces demand among a large number of lower-cost enrollees. As a result, total profits go down when a plan covers Partners.

[81] To simplify, I also hold fixed the observed choice of Fallon (which is not available in most of the Boston area) not to cover Partners.

[82] To speed computing time, all of the simulations I report here have been conducted on a 10% random sample of enrollees. I will perform the simulations on the full sample in future revisions.

[83] While I would like to perform a similar model fit test for other years, data limitations and policy complexities make this difficult. Prior to 2010, the pricing process was much more complicated and involved some negotiation with the exchange. In 2010, I am missing the risk adjustment scores. And in 2012-13, the exchange introduced a limited choice policy that creates auction-like dynamics that I have not yet modeled.

[84] Massachusetts used maximum prices to lower costs given that it fully subsidizes below-poverty enrollees for any plan they choose. Minimum prices were imposed by federal actuarial soundness rules, which are designed to prevent insurers from pricing so low that they are unable to pay for the required medical benefits.

[85] The main remaining differences with the ACA are the lack of higher-income unsubsidized enrollees (who represent about 20% of ACA enrollees) and the absence of multiple plans per insurer across four cost-sharing generosity tiers. There is not much I can do to incorporate these factors, since I do not have data on higher income people or a way or estimating preferences for different levels of cost sharing. Therefore, the simulations should be seen as illustrative of the economic forces involved, not a prediction of what will occur in the ACA.

7.2 Welfare Analysis

An important question is whether this unravelling is socially inefficient. Answering this requires a welfare function, which is not obvious to define in this setting. My starting point is a social surplus approach, in which welfare equals consumer plan value (plan utility divided by marginal utility of money[86]) minus insurer costs. But I make several adjustments. First, I choose to exclude the switching cost, treating them as pure inattention. Recall that I estimated that switching costs were much lower when a plan dropped a consumer's hospital, and I do not want the welfare analysis to be driven by this difference.[87] Once I exclude switching costs, however, the standard inclusive value formula for expected utility in a logit model does not apply. Instead, I define expected plan value for consumer *i* as:

$$ConsValue_i = \frac{1}{-\alpha_i} \sum_j \hat{s}_{ij}^{Plan} \cdot \hat{U}_{ij}$$

where α_i is the premium coefficient, $\hat{s}_{ij}^{Plan}$ is the model's predicted share for consumer *i* choosing plan *j*, and $\hat{U}_{ij}$ is plan utility excluding switching costs and the logit error.

A second adjustment to social welfare is that I allow for an excess cost of government subsidies, to reflect the distortionary cost of tax financing. As a baseline, I assume an excess cost of government funds (ECF) of 30%, but I also consider an ECF of zero as in a textbook social surplus calculation. Finally, I add to social welfare an estimate of the markup of Partners' hospital prices above cost, based on the Massachusetts government estimate (discussed above in Section 6). As a starting point, I value each dollar of markup as $1 of social welfare, although alternate assumptions are possible. How to value these markups depends on the social value of the hospital activities they fund, including teaching, research, and uncompensated care.

Table 9 compares welfare at the ACA-like equilibrium for 2012 (with no plans covering Partners) to a hypothetical in which all plans (except Fallon) cover Partners. The latter captures the hypothetical effect of ensuring full coverage of Partners (e.g., via a coverage mandate) if this could be done without Partners raising its prices. Plans, however, can adjust their prices, and consumers can re-sort across plans. Overall (in the last column), Partners access increases plan value by a relatively small $5.7 per member-month and increases costs (net of the Partners markup) by a slightly larger $11.2. Thus, on average, covering Partners lower net welfare, even without considering the increase in government costs. However, the averages mask important heterogeneity between people who strongly value Partners (again,

[86] The marginal utility of money is the negative of the premium coefficient in the plan demand system. I do not need to worry about the premium coefficient for the below-poverty group (which I could not estimate) because they are excluded from the ACA-like population.
[87] I have also done the welfare analysis with switching costs included. The results are qualitatively similar, but past Partners' patients value of coverage is higher because of the switching cost interaction. However, this difference is not enough to change the net result of the welfare calculation.

proxied by past Partners use) and others. Past Partners patients value access by \$30.2 per month, an order of magnitude more than all others. Although past Partners patients' net cost increases are also larger (\$26.6 per month), the value gains are large enough to produce a small net welfare increase of \$3.6 per month from giving them access. By contrast, all other consumers value Partners access relatively little but once Partners is covered, they use it and increase costs non-trivially. Because this broader group is 90% of the population, their effect dominates and social surplus falls.[88]

These welfare results highlight the fundamental issues involved with coverage of star hospitals. On the one hand, the group preferring Partners is high cost, so insurers have too great an incentive to avoid Partners because of adverse selection. On the other hand, covering Partners creates moral hazard among people who do not value it much but for whom expensive Partners hospitals are now available at no extra fee. Since Partners is very costly, the moral hazard costs are large, and reversing unravelling is not socially optimal. The welfare calculus, however, might be different if Partners were less expensive, since moral hazard would be smaller but adverse selection would still deter covering Partners.

7.3 Policy Counterfactuals

The welfare analysis above studied a hypothetical in which plans covered Partners despite it being profitable to drop them. In reality, reversing unravelling would require policy changes to address adverse selection. In this section, I examine two policies to offset adverse selection and encourage coverage of the Partners hospitals: modified risk adjustment and subsidies. I examine how plans' prices and Partners coverage decisions change under alternate policies, continuing to hold Partners' hospital prices fixed.

The first policy modifies risk adjustment by increasing how much it compensates for high-risk types and reducing it for low risks – a form of the "over-payment" that Glazer and McGuire (2000) find to be optimal for risk adjustment. The logic for over-payment is that plans covering Partners attract consumers who are both observably and unobservably high-cost. The modified risk adjustment over-pays based on observed risk to compensate for the high unobserved risk of enrollees in plans covering Partners. To implement this, I multiply all risk scores above the mean by a factor $(1+\phi)$, divide all below-mean risk scores by the same factor, and renormalize the distribution to be mean 1.0. The potential downside of this policy is that insurers have incentives to avoid covering people with low observed risk (e.g., young people). If low risks are more price sensitive (as I found for young people in the plan demand estimates), insurers will respond by raising prices and markups.

[88] One might expect these results to be different if just one or two plans covered Partners, since consumers could sort based on Partners preference. However, I find results to be qualitatively similar. Consumers choose plans based on many factors, so there are still consumers who value Partners little but end up in a plan covering it.

The top of Table 10 shows the simulation results for modified risk adjustment. A ϕ of 50% is sufficient to reverse the unravelling, with NHP choosing to cover Partners. The change increases consumer surplus (by \$5.4 per member-month), insurer profits (by \$6.9), and Partners net revenue (by \$1.1, about one-fourth of the increase as when all insurers cover Partners in Table 9). However, the largest change is in government subsidy costs, which increase by \$14.4 per member-month (or 4.4%). Government costs increase because subsidies are set based on the lowest plan's price, which rises from \$365 to \$381.[89] The low-price plan (CeltiCare) tends to selects low-risk people, and the modified risk adjustment penalizes it more for doing so. In addition, it has less incentive to keep markups low to attract the healthy, as discussed above. Therefore, CeltiCare raises its price. The cost of higher subsidies depends on whether there is an excess marginal cost of government funds (ECF). If there is no excess cost (ECF = 0), this is a pure transfer, and social surplus changes only slightly. With a more typical ECF of 30% (the final column in Table 10), social surplus falls more substantially.

I consider a second policy to address adverse selection: differentially subsidizing high-price plans. Rather than a fixed subsidy S_0 for all plans, a plan's subsidy equals $S_0 + \sigma \cdot \left(P_j - \min_k P_k\right)$, which is linked to its price P_j. I call this policy "marginal subsidies" because a plan's subsidy increases on the margin as it raises its price. Marginal subsidies decrease plans' incentive to compete on prices and therefore increase the incentive to raise quality (here, Partners coverage) – as shown by the classic analysis of Dorfman and Steiner (1954). Marginal subsidies also decrease the inefficiently high premiums a plan covering Partners charges because of selection, which I highlighted in the theory in Section 1. The downside is that plans have greater incentive to markup prices, regardless of whether they cover Partners.

The bottom part of Table 10 shows the simulations for marginal subsidies. Although conceptually different, subsidies have similar qualitative effects as risk adjustment. Marginal subsidy rates exceeding 25% induce BMC plan and at 50%, also NHP to cover Partners. Consumer surplus, insurer profit, and Partners net revenue increase at the expense of higher government spending. Relative to risk adjustment, however, consumer surplus increases less and insurer profits increase more. A key difference is the pattern of price increases across plans. With risk adjustment, the low-price plan raises its price, while all higher-price plans raise prices relatively little, since they benefit from the greater compensation for their relatively sick consumers. However, with subsidies, all plans increase their prices in tandem. As a result, insurer profits increase a bit more, and social surplus falls a bit more.

[89] These simulations follow the Massachusetts rule of setting subsidies so that the cheapest plan's premium equals a pre-specified affordable amount. The ACA sets subsidies based on the second-cheapest silver-tier plan, which I do not follow because I do not have plans across multiple generosity tiers. Note that the increase in the lowest plan price is slightly larger than the increase in subsidies because the risk adjustment is not perfectly budget neutral.

This analysis points to a more general tradeoff involved with mitigating adverse selection in settings with imperfect competition, as shown in recent work by Starc (2014) and Mahoney and Weyl (2014). When sicker people differentially choose higher-price plans, insurers have an incentive to keep prices low to avoid the sick. If risk adjustment or subsidies offset this effect, insurers raise price markups. In insurance exchanges, higher markups may be a greater public policy concern than in typical markets, since government subsidies are linked to prices. Higher markups raise government costs, which create a direct efficiency cost because of the excess cost of tax-financed public funds.

An important limitation of this analysis is that I have throughout held fixed the prices of Partners hospitals. This may be sensible for the relatively small CommCare exchange (covering about 3% of Massachusetts' population), and indeed, I found that Partners did not change its prices much after Network Health dropped it in 2012. However, if plans in a broader array of markets dropped Partners, Partners would be forced to respond. Analyzing this response would require modeling hospital-insurer bargaining, something I have not yet done because of its complexity. However, part of the logic in such a model seems clear. Adverse selection that discourages plans from covering Partners should pressure Partners to lower its prices – while policies that offset selection should reduce this pressure. These effects are qualitatively similar to the effects on insurer prices discussed above.

With Partners, however, the welfare effects of higher prices are more complicated. Higher prices at star academic hospitals partly fund activities like teaching and medical research. Whether the government should subsidize plans to cover Partners depends on the social value of these activities. The above analysis has valued these at cost, but the true social value may be higher or lower. How to assess high star hospital prices is beyond the scope of this paper but an important topic for future research.

8 Conclusion

As health insurance programs like the ACA increasingly use exchanges to provide coverage, an important question is how well insurance competition will work. A key part of that question is whether adverse selection is still a concern, despite exchange regulations and risk adjustment used to combat it. This paper has shown evidence from the Massachusetts exchange that there is meaningful residual selection against plans covering expensive star hospitals. Studying a 2012 case where a large plan dropped the star Partners hospitals, I find that selection explains between 35-50% of the plan's cost reductions. The selection is driven by people who strongly prefer the star hospitals and are willing to switch plans to maintain access to them. I find that this group has high risk-adjusted costs both because of greater unobserved risk and because conditional on medical risk, they are more likely to use the high-price hospitals. Improved risk adjustment can mitigate the selection on unobserved risk, but existing risk adjustment methods are not designed to address selection on use of high-price providers.

In many ways, the implications of this adverse selection are standard. Plans have disincentive to cover star hospitals. And when they do, their costs (and therefore prices) are increased in a way that sub-optimally allocates consumers across plans. For example, some people who would like to use star providers only for a severe disease like cancer must pay higher premiums that reflect the costs of people who use high-price providers for all their health care. This inability of a single premium to efficiently sort people with heterogeneous costs across plans is related to a point made in a different context by Bundorf, Levin, and Mahoney (2011). I show that this problem is also related to adverse selection, which gives plans an incentive to exclude star hospitals from network.

This inefficiency is fundamentally related to a sorting challenge: which patients should get access to the expensive services star academic hospitals provide? In standard markets, prices at the point of use create the sorting mechanism – only those willing and able to pay get access. In health insurance, plans cover all or most of hospitals' prices. Instead, people choose their hospital access when they choose plans. This system can lead to a type of moral hazard – when a plan covers star hospitals, its enrollees switch to using these high-price facilities rather than lower-price alternatives. Policies that reduce this moral hazard may also mitigate the adverse selection I find (see Einav et al. 2013). Examples include tiered patient copays (higher fees for more expensive providers) and supply-side incentives for doctors to steer patients to lower-price facilities (e.g., partial capitation; see Song, et al. 2011; Ho and Pakes 2014). How best to sort patients across hospitals of varying costs is an important question for future research.

A key driver of the selection I find is the high prices of star hospitals. Researchers are increasingly recognizing the importance of provider prices in driving both cost increases and variations across providers (HCCI 2014; Newhouse et al. 2013). This study contributes an additional finding: providers with high prices create adverse selection against plans covering them.

This selection has implications for the health insurance exchanges in the ACA. It calls into question the efficiency of the sharp rise in limited network plans in the ACA's first year (McKinsey 2014). Narrow network plans (covering less than 70% of area hospitals) represented almost half of exchange plans and about 70% of the lowest-price plans. These plans, which are particularly likely to exclude academic hospitals, may grow because of favorable selection at the expense of broad network plans. This pressure on insurers may lead star providers to respond by cutting their prices and costs. It may also add to incentives for these providers to merge with or create an insurer – as Partners did with NHP in Massachusetts and as hospitals elsewhere have done or are considering (Frakt 2014).

The policy implications of my adverse selection findings, however, are less clear. On the one hand, selection against plans covering star hospitals suggests a benefit to subsidizing these plans, through modifications to risk adjustment or subsidies. However, as I showed in simulations, these policies reduce incentives for both insurers and the star providers to lower prices, worsening pre-existing market power. A key question for assessing this tradeoff is what high prices at star academic hospitals fund. If high prices fund valuable teaching, medical research, and uncompensated care for the poor, then pressure to reduce prices may be a public policy concern. If high prices fund higher physician salaries and fancier medical facilities, the policy calculus of subsidizing them would be different. Optimal policy also depends on whether there are more efficient means of subsidizing these activities than through the insurance system. These issues are important questions for future research.

References

Akerlof, G. A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *The quarterly journal of economics*, 488-500.

Association of American Medical Colleges. Center for Workforce Studies. (2013) 2013 State Physician Workforce Data Book. Retrieved from https://www.aamc.org/download/362168/data/2013statephysicianworkforcedatabook.pdf

Berry, S., Levinsohn, J., & Pakes, A. (2004). Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market.*Journal of Political Economy*, *112*(1), 68-105.

Brown, J., Duggan, M., Kuziemko, I., & Woolston, W. (2011). How does risk selection respond to risk adjustment? Evidence from the Medicare Advantage program (Working Paper No. w16977). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/ w16977

Bundorf, M. K., Levin, J., & Mahoney, N. (2012). Pricing and Welfare in Health Plan Choice. *American Economic Review*, *102*(7), 3214-48.

Cardon, J. H., & Hendel, I. (2001). Asymmetric information in health insurance: evidence from the National Medical Expenditure Survey. *RAND Journal of Economics*, 408-427.

Capps, C., Dranove, D., & Satterthwaite, M. (2003). Competition and market power in option demand markets. *RAND Journal of Economics*, 737-763.

Chandra, A., Gruber, J., & McKnight, R. (2010). Patient cost-sharing and hospitalization offsets in the elderly. *The American economic review*, *100*(1), 193.

Chandra, A., Gruber, J., & McKnight, R. (2011). The importance of the individual mandate—evidence from Massachusetts. *New England Journal of Medicine*, *364*(4), 293-295.

Chandra, A., Gruber, J., & McKnight, R. (2014). The impact of patient cost-sharing on low-income populations: Evidence from Massachusetts. *Journal of health economics*, *33*, 57-66.

Chiappori, P. A., & Salanie, B. (2000). Testing for asymmetric information in insurance markets. *Journal of Political Economy*, *108*(1), 56-78.

Commonwealth of Massachusetts, Center for Health Information and Analysis (CHIA) (2013). Annual Report on the Massachusetts Health Care Market. Retrieved from http://www.mass.gov/chia/docs/r/pubs/13/ar-ma-health-care-market-2013.pdf

Commonwealth of Massachusetts, Center for Health Information and Analysis (CHIA) (2014). Massachusetts Hospital Profiles, Acute Hospitals Databook. Retrieved from http://www.mass.gov/chia/docs/r/hospital-profiles/2012/massachusetts-hospital-profiles-acute-databook-2.xlsx

Crawford, G. S., & Yurukoglu, A. (2012). The Welfare Effects of Bundling in Multichannel Television Markets. *American Economic Review*, *102*(2), 643-85.

Cutler, D. M., & Reber, S. (1998). Paying for Health Insurance: The Trade-off between Competition and Adverse Selection. *Quarterly Journal of Economics*,*113*(2).

Dorfman, R., & Steiner, P. O. (1954). Optimal advertising and optimal quality.*The American Economic Review*, 826-836.

Einav, L., Finkelstein, A., & Cullen, M. R. (2010). Estimating Welfare in Insurance Markets Using Variation in Prices. *The Quarterly Journal of Economics*, *125*(3), 877-921.

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P., & Cullen, M. R. (2013). Selection on Moral Hazard in Health Insurance. *American Economic Review*,*103*(1), 178-219.

Einav, L., Jenkins, M., & Levin, J. (2012). Contract pricing in consumer credit markets. *Econometrica*, *80*(4), 1387-1432.

Ericson, K. M.M., & Starc, A. (2012). Heuristics and heterogeneity in health insurance exchanges: Evidence from the Massachusetts connector. *The American Economic Review*, *102*(3), 493-497.

Ericson, K. M. M., & Starc, A. (2013). How Product Standardization Affects Choice: Evidence from the Massachusetts Health Insurance Exchange (Working Paper No. w19527). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w19527

Ericson, K. M. M. (2014). Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange. *American Economic Journal: Economic Policy*, *6*(1), 38-64.

Finkelstein, A., & Poterba, J. (2013). Testing for Asymmetric Information Using “Unused Observables” in Insurance Markets: Evidence from the UK Annuity Market. *Journal of Risk and Insurance*.

Frakt, A. (2014, May 26). When Hospital Systems Buy Health Insurers. *The New York Times*. Retrieved from http://www.nytimes.com/2014/05/26/upshot/when-hospital-systems-buy-health-insurers.html?abt=0002&abg=1

Gaynor, M., & Vogt, W. B. (2003). Competition among Hospitals. *RAND Journal of Economics*, 764-785.

Gowrisankaran, G., Nevo, A., & Town, R. (2013). Mergers when prices are negotiated: Evidence from the hospital industry (Working Paper No. w18875). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w18875

Gruber, J., & McKnight, R. (2014). Controlling Health Care Costs Through Limited Network Insurance Plans: Evidence from Massachusetts State Employees (Working Paper No. w20462). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w20462

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review*, *103*(7), 2643-2682

Handel, B. R., Hendel, I., & Whinston, M. D. (2013). Equilibria in Health Exchanges: Adverse Selection vs. Reclassification Risk (Working Paper No. w19399). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w19399

Health Cost Institute Inc.. (2014). 2013 Health Care Cost and Utilization Report. Retrieved from http://www.healthcostinstitute.org/files/2013%20HCCUR%2010-28-14.pdf

Health Policy Commission. Health Care Cost Trends in the Commonwealth, Hearing, September 16, 2013 (#13632). Retrieved from http://www.mass.gov/anf/docs/hpc/celticare.pdf

Ho, K. (2006). The welfare effects of restricted hospital choice in the US medical care market. *Journal of Applied Econometrics*, *21*(7), 1039-1079.

Ho, K. (2009). Insurer-Provider Networks in the Medical Care Market. *The American Economic Review*, *99*(1), 393-430.

Ho, K., & Pakes, A. (2014). Physician Payment Reform and Hospital Referrals.*The American Economic Review*, *104*(5), 200-205.

Lee, R. S. (2013). Vertical integration and exclusivity in platform and two-sided markets. *The American Economic Review*, *103*(7), 2960-3000.

Mahoney, N., & Weyl, E. G. (2013). Imperfect Competition in Selection Markets. *Unpublished Manuscript*.

McKinsey for U.S. Health System Reform. (2014). Hospital networks: Updated national view of configurations on the exchanges. Retrieved from http://healthcare.mckinsey.com/hospital-networks-updated-national-view-configurations-exchanges

McWilliams, J. M., Hsu, J., & Newhouse, J. P. (2012). New risk-adjustment system was associated with reduced favorable selection in Medicare Advantage. *Health Affairs*, *31*(12), 2630-2640.

Miller, R. H., & Luft, H. S. (1997). Does managed care lead to better or worse quality of care?. *Health affairs*, *16*(5), 7-25

Newhouse, J. P., Garber, A. M., Graham, R. P., McCoy, M. A., Mancher, M., & Kibria, A. (Eds.). (2013). *Variation in health care spending: target decision making, not geography*. National Academies Press.

Office of Attorney General Martha Coakley. (2013). Examination of Health Care Cost Trends and Cost Drivers. Retrieved from http://www.mass.gov/ago/docs/healthcare/2013-hcctd.pdf

Richards, Tori. (October 30, 2013) "Top Hospitals Opt Out of Obamacare." *U.S. News & World Report*. Retrieved from: http://health.usnews.com/health-news/hospital-of-tomorrow/articles/2013/10/30/top-hospitals-opt-out-of-obamacare.

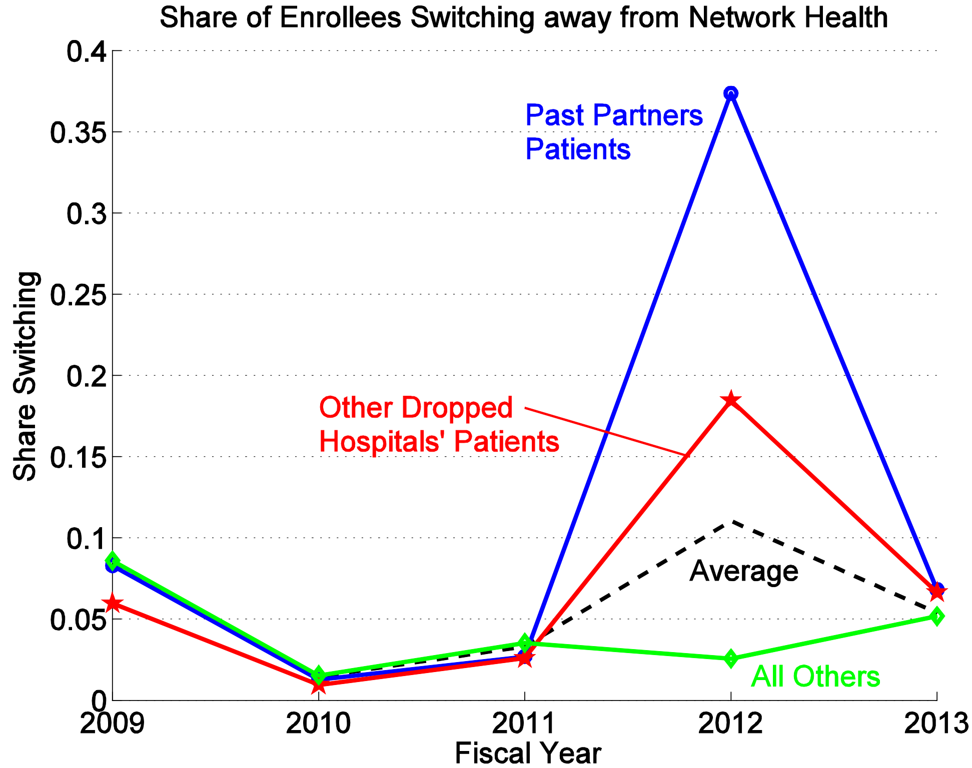
Rothschild, M., & Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics*, *90*(4), 629-649.

Song, Z., Safran, D. G., Landon, B. E., He, Y., Ellis, R. P., Mechanic, R. E., Day, M.P. & Chernew, M. E. (2011). Health care spending and quality in year 1 of the alternative quality contract. *New England Journal of Medicine*, 365(10), 909-918.

Starc, A. (2014). Insurer pricing and consumer welfare: Evidence from medigap. *The RAND Journal of Economics*, *45*(1), 198-220.

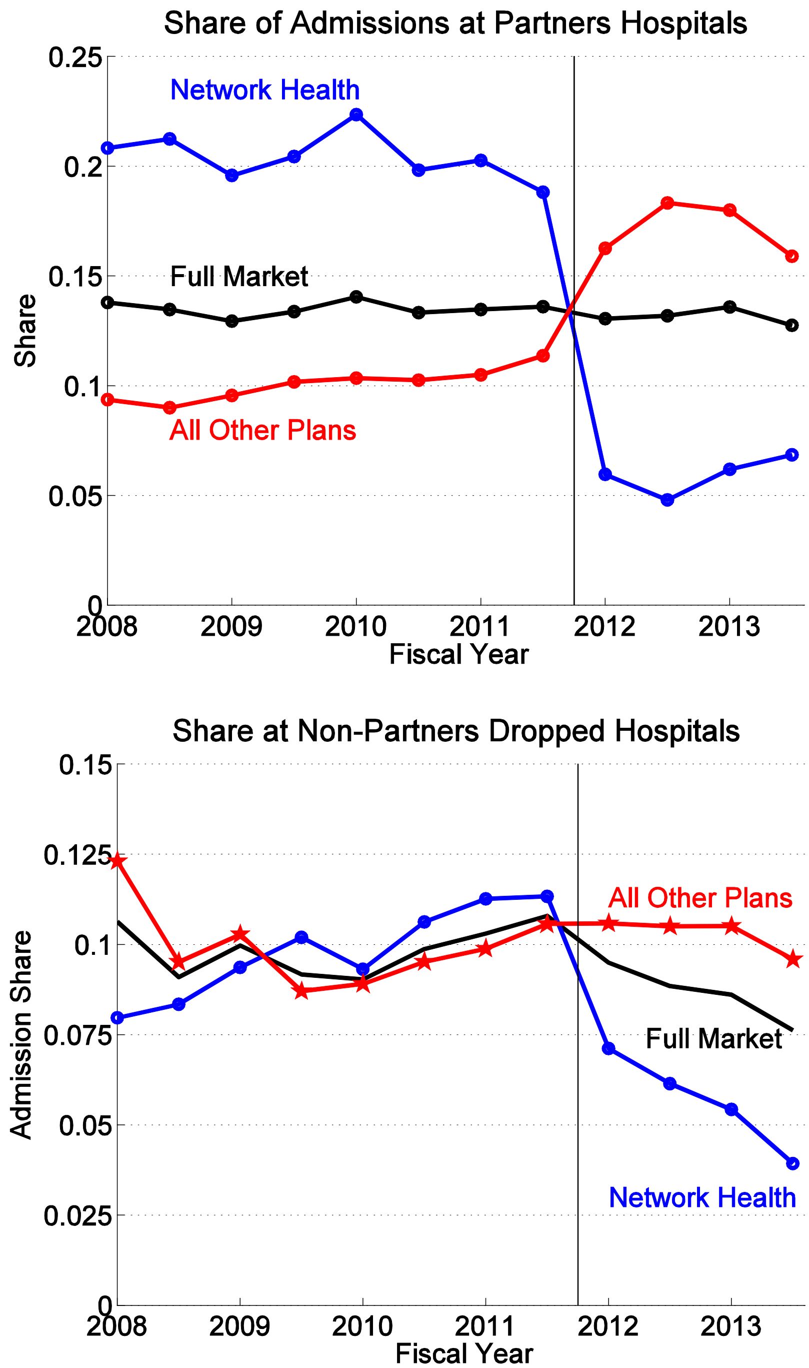
Town, R., & Vistnes, G. (2001). Hospital competition in HMO networks. *Journal of Health Economics*, *20*(5), 733-753.

Figure 1. Plan Switching after Network Health Dropped Partners in 2012



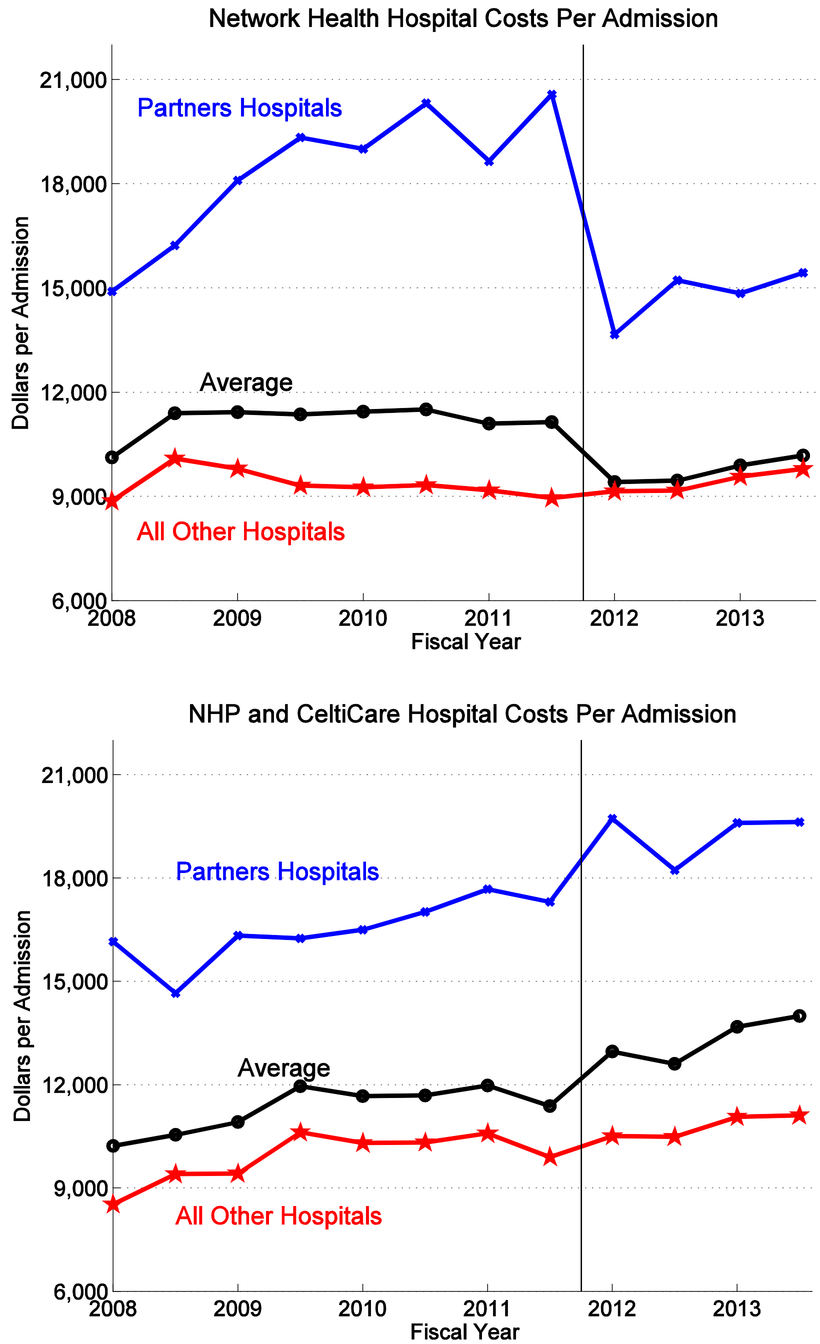
NOTE: This figure shows the share of enrollees in Network Health plan who switch to another plan at the start of each fiscal year (when all exchange enrollees are given an opportunity to switch plans). The black dashed lines show the average switching rate for all enrollees; the colored solid lines decompose this average into subgroups. In most years, switching rates are quite low, but in 2012, switching spiked after Network Health dropped the star Partners hospitals and eight other less prestigious hospitals. The graph shows a large switching spike among past patients of Partners (in blue) and a smaller spike among patients of the other dropped hospitals (in red). There was little change in switching rates among all other enrollees (in green). These results suggest that many patients are willing to switch plans to keep access to their regular hospital provider. As I show elsewhere, the past Partners patients were a particularly high-cost group, so these switching patterns contributed to favorable selection for Network Health when it dropped Partners.

Figure 2. Admission Shares at Hospitals Dropped by Network Health in 2012



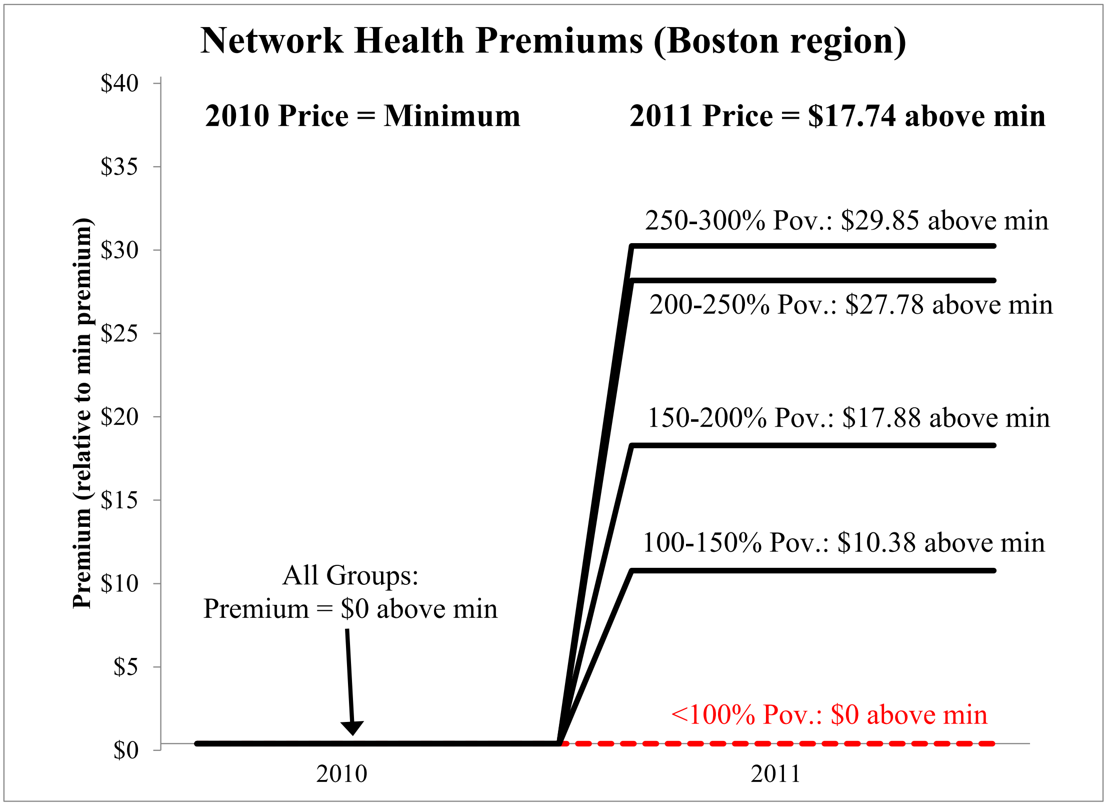
NOTE: These figures show the share of hospital admissions at hospitals that Network Health plan dropped from its network in 2012. The top figure shows shares for the star Partners hospitals; the bottom figure shows the eight other dropped hospitals. In both cases, shares fell sharply among Network Health enrollees in 2012. However, Partners' shares rose sharply at other plans (even though their coverage did not change), as the enrollees most likely to use Partners switched to plans that covered them. The same was not true for the less prestigious dropped hospitals, whose shares from other plans barely changed.

Figure 3. Changes in Cost per Hospital Admission around 2012Changes



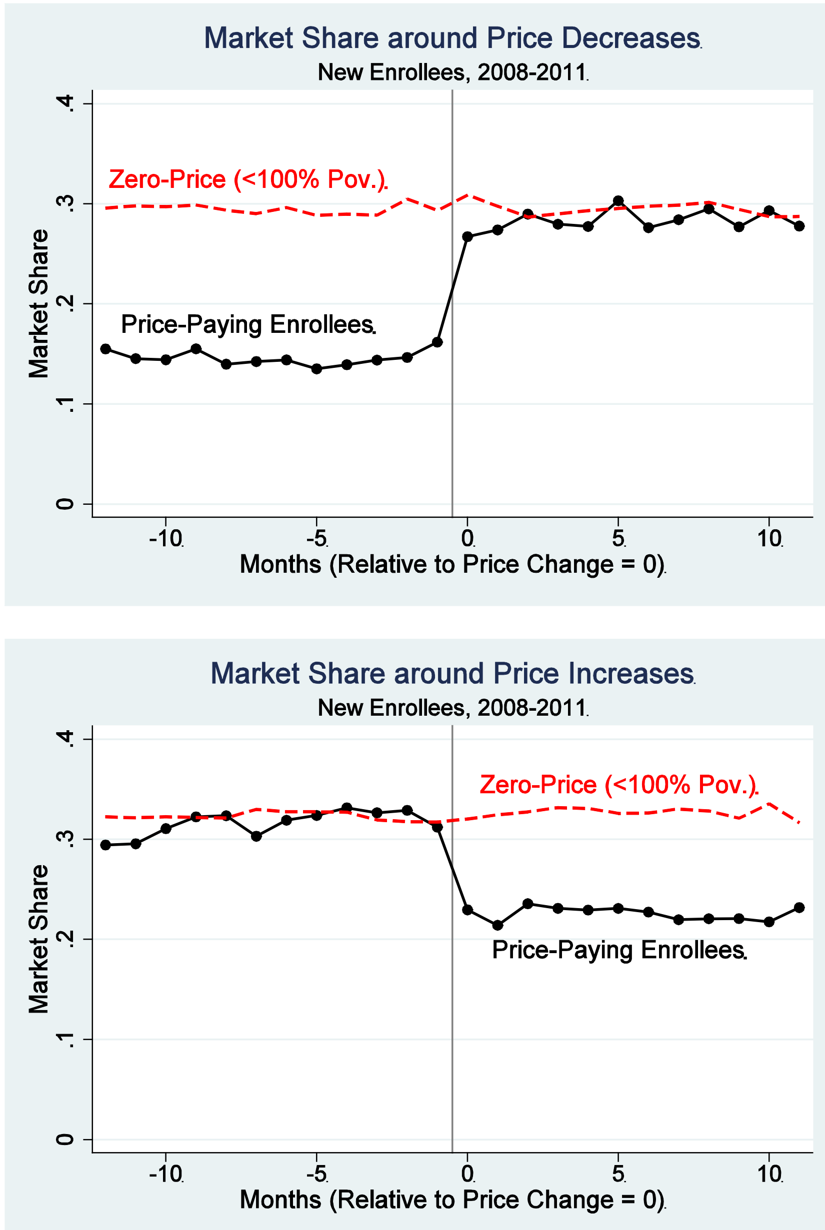
NOTE: These figures show average costs per hospital admission for enrollees of two sets of plans: Network Health (top figure), which dropped the star Partners hospitals in 2012, and NHP and CeltiCare (bottom figure), which continued to cover them. Each figure shows the plans' overall average (in black) and also separate averages for Partners (blue) and all other hospitals (red). Network Health's per-admission costs fell 15% in 2012, with all of the reductions explained by less use and lower prices at the Partners hospitals. Costs for NHP and CeltiCare increased in 2012 as more enrollees used the high-price Partners facilities. All averages were calculated after winsorizing per-admission costs at $150,000 (above the 99.9th percentile) to reduce the influence of outliers.

Figure 4. Premium Coefficient Identification Strategy



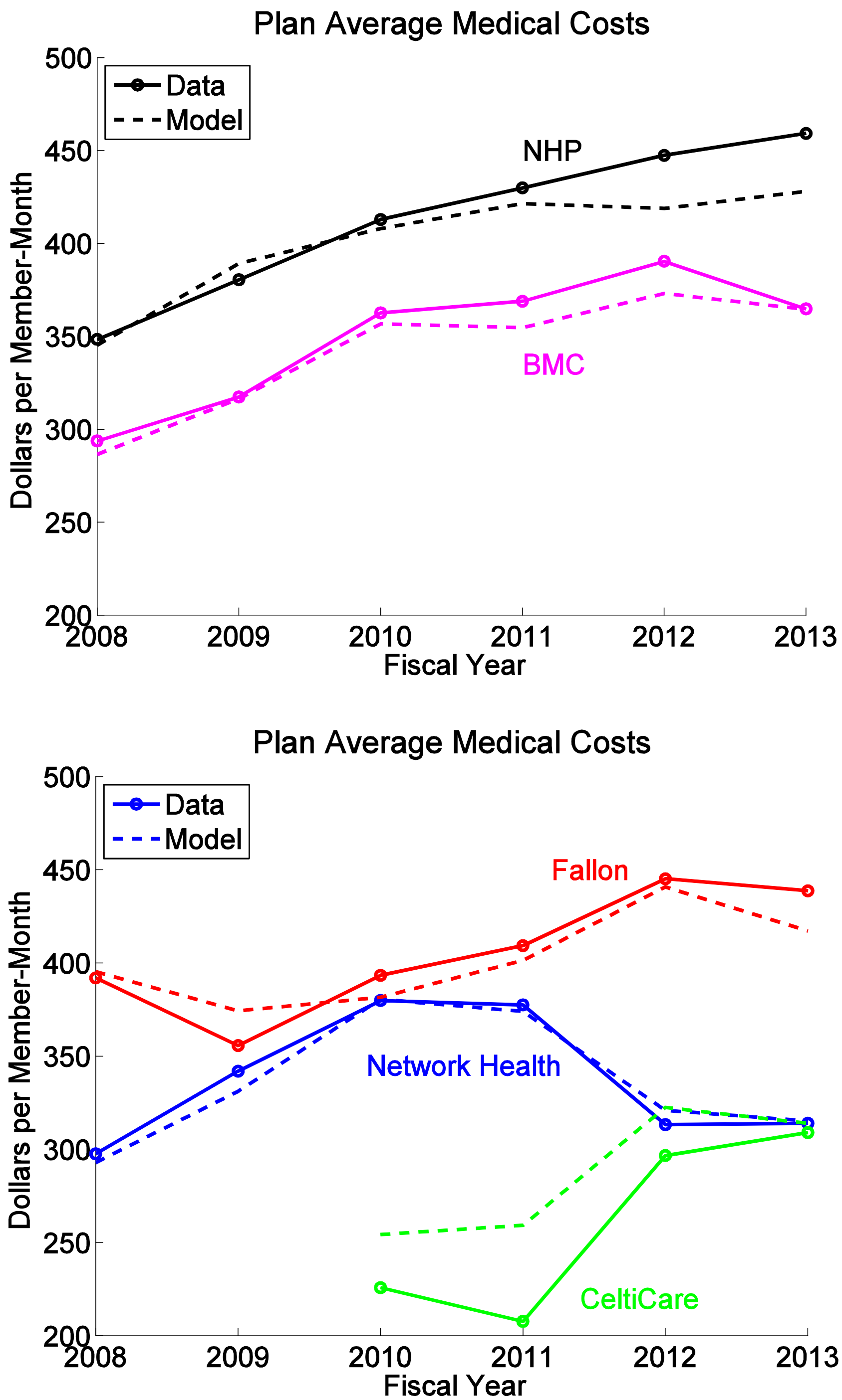
NOTE: This graph shows an example of the subsidy-driven premium variation for a single plan used to identify the premium coefficients in the plan demand model. The example is for Network Health plan in the Boston pricing region in 2011 and 2012. Network Health's premium in Boston was the cheapest in 2010, but in 2011 its relative premium increased. Because of subsidies, different income groups faced different premium increases. In particular, the below-poverty group faced *no increase* (its plans are always fully subsidized), while higher income groups faced different premium increases depending on the subsidy rules for each group. I use these within-region differential premium changes across income groups to identify the premium coefficients in plan demand. This identification approach is similar to difference-in-differences.

Figure 5. Premium Identification and Test of Parallel Trends Assumption



NOTE: These graphs show the source of identification for the premium coefficients in plan demand (see also Figure 4) and test the key parallel trends assumption for the difference-in-differences approach. Each graph shows average monthly plan market shares among new enrollees for plans that at time 0 decreased their prices (top figure) or increased their prices (bottom figure). Each point represents the shares for an independent set of new enrollees. The identification comes from comparing demand changes for above-poverty price-paying enrollees (for whom premium changes at time 0) versus below-poverty zero-price enrollees (for whom premiums are unchanged at $0). Consistent with the parallel trends assumption, trends in shares are flat and parallel for both groups at times other than the premium change but change sharply for price-payers only at the price change. The sample is limited to fiscal years 2008-2011. I drop 2007 because above-poverty enrollees did become eligible for the market until mid-way through the year and 2012+ because below-poverty new enrollees became subject to a limited choice policy that required them to choose lower-price plans.

Figure 6. Model Fit for Plan Average Medical Costs



NOTE: These figures show each plan's average medical costs (per member-month) in each fiscal year in the data (solid lines) and in the model (dashed lines). The cost model, described in Section 5, is an individual-level model of costs, which is aggregated up to the plan level based on predicted shares from the plan demand model.

Table 1. Massachusetts Hospital and Plan Network Statistics

Panel A: Most Expensive Hospitals

	Hospital	System	Hospital Type	Per-Admit Insurer Cost Stats: Total Payment	Risk-Adj. Price	Severity Index
1	Brigham & Women's	Partners	Acad. Med Ctr	$23,278	$20,474	1.12
2	Mass. General	Partners	Acad. Med Ctr	$21,428	$19,550	1.09
3	Boston Med. Ctr.	BMC	Acad. Med Ctr	$16,850	$15,919	1.05
4	Tufts Med. Ctr.	Tufts	Acad. Med Ctr	$15,328	$14,038	1.10
5	UMass Med. Ctr.	UMass	Acad. Med Ctr	$14,941	$14,111	1.07
6	Charlton Memorial	Southcoast	Community	$14,411	$14,210	1.03
7	Baystate Med. Ctr.	Baystate	Teaching	$13,715	$12,223	1.11
8	Lahey Clinic	Lahey	Teaching	$13,430	$11,742	1.13
9	Beth Israel Deaconess	CareGroup	Acad. Med Ctr	$12,971	$11,787	1.08
10	St. Vincent	Vanguard	Teaching	$11,881	$11,455	1.03
	All Other Hospitals	---	---	*$8,232*	*$8,585*	*0.95*

Panel B: Most Used Hospital Systems

	Hospital/System	Share of Admissions	Num. Hospitals	Hosp. Beds	Payment per Admission
1	Partners	13.6%	7	2,488	$18,016
2	Steward	12.2%	10	1,627	$7,695
3	UMass	8.5%	5	947	$13,140
4	CareGroup	7.7%	5	1,088	$11,481
5	Boston Med. Ctr.	6.6%	1	474	$16,850
6	Southcoast	5.7%	3	815	$12,642
7	Baystate	4.9%	3	774	$12,361
	All Others	*40.7%*	*30*	*5,391*	*$8,649*

NOTE: These tables show the most expensive hospitals and the most used hospital systems in the Massachusetts exchange (CommCare) data. The most expensive hospitals in Panel A are ranked by the average insurer payment per in-network hospital admission. I also show averages for risk-adjusted prices and an index of patient severity, both of whose estimation is described in Section 5.1. Hospital type is a classification defined by the Massachusetts state government. Panel B shows the most used hospital systems (by share of admissions in the Massachusetts data), including some other statistics on the hospitals. In this table, payment per admission is again the average insurer payment per in-network hospital admission.

Table 2. Hospital Network Coverage by Exchange Plans

Hospital Coverage by Year and Plan

Hospital Group	2009	2010	2011	2012	2013
		Boston Medical Center Plan (BMC)			
Partners: Academic Hospitals	No	No	No	No	No
Other Partners Hospitals	2/5	1/5	1/5	1/5	1/5
Share of Non-Partners Hospitals	*88%*	*92%*	*92%*	*92%*	*92%*
		Network Health			
Partners: Academic Hospitals	Yes	Yes	Yes	No	No
Other Partners Hospitals	5/5	5/5	5/5	2/5	2/5
Share of Non-Partners Hospitals	*74%*	*75%*	*81%*	*78%*	*85%*
		Neighborhood Health Plan (NHP)			
Partners: Academic Hospitals	Yes	Yes	Yes	Yes	Yes
Other Partners Hospitals	2/5	4/5	4/5	4/5	5/5
Share of Non-Partners Hospitals	*72%*	*84%*	*87%*	*87%*	*87%*
		CeltiCare *(new in 2010)*			
Partners: Academic Hospitals	---	Yes	Yes	Yes	Yes
Other Partners Hospitals		3/5	3/5	3/5	3/5
Share of Non-Partners Hospitals	---	*25%*	*33%*	*40%*	*40%*
		Fallon Health Plan *(central Mass only)*			
Partners: Academic Hospitals	No	No	No	No	No
Other Partners Hospitals	0/5	0/5	0/5	1/5	0/5
Share of Non-Partners Hospitals	*20%*	*21%*	*12%*	*10%*	*10%*

NOTE: This table shows statistics on the hospital network coverage of the five Massachusetts exchange plans in each plan year. For each plan, I list statistics separately for the Partners Healthcare System (the star hospital system this paper focuses on) and all other Massachusetts general acute care hospitals. For Partners, I list whether the plan covers its two star academic medical centers (Mass. General and Brigham & Women's hospitals) and the number of other Partners hospitals covered. For all other hospitals, I list the share of hospitals covered, weighted by number of hospital beds. The largest network change in the exchange history is for Network Health in 2012 when they dropped almost all of the Partners system, as well as eight other hospitals. I focus on this change in my reduced form analysis in Section 3.

Table 3. Positive Correlation Test Regressions

Panel A: Plan Selection and Costs

VARIABLES	Dep. Var.: Choose Plan Covering Partners		Dep. Var.: Costs (per month)		
	All Enrollees	Re-Enrollees	All Plans	Plans Not Covering Partners	Plans Covering Partners
	(1)	(2)	(3)	(4)	(5)
Past Partners Use	0.328***	0.268***	56.22***	36.45***	62.31***
	(0.002)	(0.004)	(5.54)	(8.04)	(7.67)
Control Variables					
Risk Score	-0.006***	-0.015***	405.03***	392.71***	419.18***
	(0.001)	(0.002)	(7.94)	(10.08)	(12.58)
Year Dummies	X	X	---	---	---
Plan-Year Dummies	---	---	X	X	X
Dependent Var. Mean	0.409	0.420	$387.0	$372.8	$407.5
Observations	843,779	131,120	843,779	498,773	345,006
R-Squared	0.126	0.201	0.10	0.11	0.10

NOTE: Panel A shows results from the positive correlation test for asymmetric information discussed in Section 3. It tests whether a factor not used in plan pricing (here, whether an enrollee has used Partners in the past) correlates with both probability of choosing a plan covering Partners and with risk-adjusted costs per month. The tests are performed for 2011-2013, the years for which I have full risk adjustment data. The cost regressions are performed separately for enrollees in all plans and in plans that cover and do not cover Partners.

Panel B: Other Dependent Variables and Tests

VARIABLES	Sample: Plans Covering Partners			
	Hospitalization Rate	Prob. Choose Partners Hosp.	Cost per Hospitalization	Costs: Diagn. Risk Adj Only
	(1)	(2)	(3)	(4)
Past Partners Use	0.0005	0.463***	1,953***	76.64***
	(0.0016)	(0.010)	(454.9)	(8.97)
Risk Score	0.0574***	-0.003	750.6***	379.85***
	(0.0020)	(0.002)	(169.6)	(13.70)
Plan-Year Dummies	X	X	X	X
Dependent Var. Mean	0.0439	0.244	$12,543	$394.50
Observations	345,006	11,759	11,759	172,731
R-Squared	0.042	0.208	0.016	0.13

*** p<0.001, ** p<0.01, * p<0.05

NOTE: Panel B explores why past Partners users may be higher cost by using a similar regression structure as Panel A with other dependent variables. The final column shows a robustness check on the cost results in Panel A by limiting the sample to individuals for whom the exchange used diagnosis-based risk adjustment.

Table 4. Analysis of Network Health's Cost Changes from 2011-2012

Network Health Cost Changes, 2011-2012

Enrollee Group	Avg. Costs			Risk Adj. Costs			Group
	2011	2012	%Δ	2011	2012	%Δ	Size*
All Enrollees	**$386**	**$306**	**-21%**	**$370**	**$313**	**-15%**	**---**
Stayers	$323	$303	-6%	$317	$300	-6%	36,768
Left Plan in 2012							
Switched Plans	$670	*[$616]*	---	$509	*[$425]*	---	4,640
Exited Market	$470	---	---	$459	---	---	22,617
Joined Plan in 2012							
Switched Plans	*[$283]*	$288	---	*[$303]*	$309	---	15,062
Entered Market	---	$315	---	---	$334	---	51,109

NOTE: This table shows the changes in medical costs per member-month for Network Health from 2011 (when it covered the star Partners hospitals) to 2012 (when it dropped them). The first set of columns show unadjusted costs. The next columns show risk-adjusted costs, defined as a group's average cost divided by average risk score (a measure defined by the exchange for risk adjustment). Group size is the number of enrollees in the relevant group during the year(s) they were enrolled in Network Health. Overall, Network Health's costs fell by 21%, or 15% after risk adjustment. The next rows break costs into enrollee subgroups: a fixed group of "stayers" (people in the plan in both years) and enrollees who left or newly joined the plan in 2012. These show that costs for stayers fell by just 6%. Selection out of Network Health by high-cost enrollees explains a large portion of the fall in costs.

Panel B: Breakdown by Partners Patient Status

Enrollee Group	Avg. Costs			Risk Adj. Costs			Group
	2011	2012	%Δ	2011	2012	%Δ	Size*
Stayers							
Partners Patients	$533	$397	-26%	$472	$332	-30%	5,308
All Others	$284	$289	2%	$285	$294	3%	31,460
Switched from Network Health in 2012							
Partners Patients	$778	*[$674]*	---	$571	*[$438]*	---	3,169
All Others	$403	*[$490]*	---	$335	*[$391]*	---	1,471
Switched to Network Health in 2012							
Partners Patients	*[$547]*	$435	---	*[$502]*	$364	---	1,303
All Others	*[$261]*	$275	---	*[$284]*	$302	---	13,759

* Number of enrollees during the relevant year they were enrolled in Network Health.

NOTE: Panel B breaks down cost changes for stayers and switchers (see note above) in Network Health into people who had ever been a patient at a Partners hospital (as of the start of 2012) and all other enrollees. Exit by high-cost Partners patients fully explains high costs among people who switched out of Network Health in 2012. Among stayers, Partners patients' costs fell sharply, while all other enrollees' costs rose slightly.

Table 5. Hospital Demand Estimates

	Simple Model		Full Model	
VARIABLE	Coeff.	Std. Error	Coeff.	Std. Error
Distance to Hospital:				
Distance in Miles (avg. coeff.)	-0.189***	(0.001)	-0.144***	(0.001)
Distance^2 (avg. coeff.)	0.0013***	(0.0000)	0.0009***	(1e-5)
Distance Interactions:				
x Income > Poverty			-0.006***	(0.0006)
x Age / 10			-0.003***	(0.0002)
x Severity Weight			-0.002	(0.0011)
x Emergency			-0.015***	(0.0006)
Out-of-Network Disutility				
Out-of-Network x BMC	-1.327***	(0.016)	-1.117***	(0.034)
Out-of-Network x CeltiCare	*(same for all plans)*		-1.464***	(0.058)
Out-of-Network x Fallon			-1.583***	(0.059)
Out-of-Network x NHP			-0.543***	(0.049)
Out-of-Network x Network			-1.011***	(0.036)
Out-of-Network x Emergency			0.010	(0.034)
Past Use of this Hospital *(>60 days before)*				
Inpatient Care			1.417***	(0.020)
Outpatient Care			2.202***	(0.013)
Hospital Characteristics				
Hospital Dummies	Yes		Yes	
Severity x Academic Med. Ctr. (avg).			2.076***	(0.044)
Severity x Teaching Hosp			1.026***	(0.045)
Diagnoses x Hospital Services (largest coeffs.):				
Mental: Psych. Services			1.844***	(0.040)
Pregnancy: Obstetrics Services			1.122***	(0.076)
Injury: Level 1 Trauma Center			0.805***	(0.037)
Cancer: Oncology Services			0.704***	(0.084)
Model Statistics:				
Pseudo-R^2 (McFadden's)	0.463		0.569	
R^2 in Shares (Area-Plan-Yr Level)	0.643		0.742	
Num. Choice Instances	74,383		74,383	

Std. Errors in parentheses. * = 5% sign., ** = 1% sign., *** = 0.1% sign.

NOTE: The table shows estimates for the multinomial logit hospital choice model described in Section 4.1. The left columns show a simple model, while the right columns show the full model used for all further analyses. The logit coefficients shown are interpretable as entering the latent utility function describing hospital choice. Past use variables are dummies for whether a patient has used each specific hospital at least 60 days before the current admission. Severity is an estimated summary measure of costs described in Section 5.1. In addition to the variables shown, the model includes: distance interacted with exchange region, detailed income group (by 50% of poverty), and gender; severity interacted with separate dummies for each academic medical center; and five additional diagnosis x hospital service interactions (circulatory diagnosis interacted with cath lab, interventional cardiology, and heart surgery services; pregnancy diagnosis x NICU; and musculoskeletal diagnosis x arthritis services).

Table 6. Insurance Plan Demand Estimates

VARIABLE	Coeff.	Std. Error	
Premium: Avg. Coeff. (*normalized*)	**-1.000*****	**(0.025)**	
x 0-100% Poverty -- *Omitted (no prems.)*	---		
x 100-150% Poverty	-1.340***	(0.038)	
x 150-200% Poverty	-0.935***	(0.024)	
x 200-250% Poverty	-0.712***	(0.015)	
x 250-300% Poverty	-0.656***	(0.016)	
x Age/5 (average effect)	0.029***	(0.002)	
Hospital Network			
Network Utility x <100% Poverty	6.355***	(0.885)	
Network Utility x 100-150% Poverty	7.371***	(0.939)	
Network Utility x 150-200% Poverty	7.453***	(0.962)	
Network Utility x 200-250% Poverty	7.736***	(1.270)	
Network Utility x 250-300% Poverty	8.541***	(1.878)	
Past-Used Hospitals Covered (share)	5.411***	(0.836)	
x Past-Used Partners Hospitals	11.631***	(0.773)	
Switching and Inertia			Passive Prob.
Average Inertia Coefficient	96.810***	(0.230)	94.6%
x Drops Past-Used Hospital (Non-Partners)	-29.905***	(1.142)	75.2%
x Drops Past-Used Hospital (Partners)	-51.048***	(0.962)	51.8%
Plan Brand Effects (average)			
BMC HealthNet (normalized)	0.000	---	
CeltiCare	-23.088***	(0.890)	
Fallon	14.021***	(1.023)	
Neighborhood Health Plan	-2.199***	(0.251)	
Network Health	-3.822***	(0.337)	
Model Statistics			
R^2 in Share (Area-Income-Year)	0.963		
Model w/ Only Avg. Plan Dummies	0.866		
No. Choice Instances	1,588,889		
No. Unique Individuals	611,455		

* = 5% sign., ** = 1% sign., *** = 0.1% sign.

NOTE: This table shows estimates for the multinomial logit plan choice model described in Section 4.3. Premium is the monthly plan price, which typically varies by $20-60 across plans. (In addition to the interactions shown, the full model contains interactions with 5-year age groups and gender.) I normalize the average consumer's premium coefficient to -1.0, so all other coefficients are interpretable as dollar values. Network utility is the consumer-specific expected utility measure for a plan's hospital network, derived in Section 4.2. Past-used hospitals coverage is the share of an enrollee's previously used hospitals that a plan covers, with a separate interaction for the star Partners hospitals. Switching and inertia are coefficients on a dummy variable for the current plan. The coefficients are interpretable as "switching costs" in dollars per month; the passive probabilities are the implied share of enrollees who passively stick with their current plan. The plan brand effects are coefficients on dummies for each plan. I show average values; the full model contains region-year- and region-income group-specific plan dummies.

Table 7. Model Estimates: Relationship between Value and Cost of Partners Coverage

Consumer Value of Partners Covg.		**Costs to Insurer (per month)**					
		Not Covering Partners		**Covering Partners**			
Percentiles	Avg. Value ($/month) (1)	Unadjusted Cost (2)	Risk Adj. Cost (3)	Risk Adj. Cost (4)	ΔCost (5)	%Δ (6)	ΔCost - Partners Hospital Mkup. (7)
0-50%	$0.5	$300.0	$301.2	$309.2	$8.0	2.7%	$7.0
50-70%	$2.2	$269.6	$294.5	$308.6	$14.0	5.2%	$10.6
70-79%	$4.3	$264.3	$292.7	$310.8	$18.1	6.8%	$12.4
80-89%	$8.8	$300.1	$311.8	$335.3	$23.5	7.8%	$14.0
90-95%	$23.6	$455.7	$360.4	$398.3	$37.9	8.3%	$21.1
96-100%	$46.8	$482.3	$340.1	$388.6	$48.5	10.0%	$23.3
Average	***$5.7***	***$308.8***	***$305.6***	***$321.2***	***$15.6***	***5.0%***	***$10.6***

NOTE: This table shows the estimated model's implication for the relationship between enrollees' costs and their value coverage of the star Partners hospitals – the key relationship driving adverse selection. Consumers are sorted into percentiles of Partners value, and each row shows average values and costs for people in the relevant percentiles. All values and costs are calculated for current enrollees in 2012 based on the value and cost if Partners were added to Network Health plan's network. Value of Partners is defined as the extra plan utility (excluding switching costs) if Partners is covered, divided the marginal utility of money – based on plan utility estimates shown in Table 6. Because I cannot estimate marginal utilities for below-poverty enrollees, they are excluded. Costs are defined using the estimated cost function without Partners covered (columns 2-3) and with it covered (columns 4-7), both based on the plan cost model in Section 5.3. Column 7 subtracts from the increase in cost an estimate of how much of these higher costs are funding higher markups for Partners. The table shows that most enrollees value Partners coverage little, but the top 10-20% values Partners substantially. The table also decomposes two different reasons people with high values for Partners are high-cost. First, they have higher risk adjusted costs even if Partners is not covered, which suggests they are unobservably sicker. Second, they have a larger increase in costs when Partners is covered (column 5) because they use Partners hospitals more often.

Table 8. Simulation Results

Equilibrium Simulation Results

Panel A: Mass. Exchange Population & Policies (2011)

Source	Year	Variable	Insurance Plan: BMC	Fallon	Network Hlth	NHP	CeltiCare
Observed	2011	Partners Covg.	No	No	Yes	Yes	Yes
		Price*	$424.6	$425.7	$422.6	$425.7	$404.9
Simulated	2011	Partners Covg.	No	No	No	No	Yes
		Price*	$425.7	$425.7	$425.7	$425.7	$404.9

** Exchange imposed maximum price of $425.7 and minimum price of $404.9*

Panel B: ACA-Like Population & Policies

Source	Year	Variable	Insurance Plan: BMC	Fallon	Network Hlth	NHP	CeltiCare
Simulated	2011	Partners Covg.	No	No	No	No	No
		Price	$407.2	$409.3	$389.4	$402.5	$318.8
Simulated	2012	Partners Covg.	No	No	No	No	No
		Price	$427.5	$464.5	$371.0	$417.6	$365.0
Simulated	2013	Partners Covg.	No	No	No	No	No
		Price	$437.2	$476.8	$432.9	$461.8	$419.4

NOTE: These tables show equilibrium results for the insurance market simulations described in Section 7.1. In the game, insurers first simultaneously choose whether or not to cover the Partners hospitals (holding fixed other hospital coverage) and then simultaneously choose their plan's price. The tables show their equilibrium choices of Partners coverage and price. Panel A shows simulations using the Massachusetts exchange's actual enrollee population and policies for 2011 – including required minimum and maximum prices – and compares simulated coverage and prices to the observed values. I do this comparison only for 2011 because of complications with analyzing other years. As discussed in Section 7.1, the model matches prices well but predicts even more dropping of Partners than actually occurred (although Network Health dropped Partners the following year).

Panel B conducts simulations with a population and policies closer to those in the ACA exchanges. Specifically, I exclude enrollees below poverty (who get Medicaid in the ACA), set subsidies as a flat amount for all plans (versus Massachusetts' higher subsidies for higher-price plans), and do not impose minimum and maximum prices. In these simulations, no insurer chooses to cover Partners partly because doing so attracts enrollees with high risk-adjusted costs and therefore lowers profits.

Table 9. Welfare Analysis of Partners Coverage

Welfare Analysis of Partners Coverage

ACA-Like Policies & Population (Sim. Yr. 2012), Units = $/member-month

Statistic		Enrollee Group: Past Partners Patients	Enrollee Group: All Others	Overall
Share of Enrollee Months		9.7%	90.3%	100.0%
Consumer ΔPlan Value		***$30.2***	***$3.0***	***$5.7***
Insurer Costs	No Partners Covg.	$521.8	$342.6	$360.0
	All Cover Partners*	$563.7	$355.6	$375.9
	Difference	***$41.9***	***$13.1***	***$15.9***
Partners Net Inpatient Revenue	No Partners Covg.	$3.5	$0.3	$0.6
	All Cover Partners*	$18.9	$3.8	$5.2
	Difference	***$15.3***	***$3.5***	***$4.6***
	Net Cost Difference	***$26.6***	***$9.6***	***$11.2***
Net Value - Cost of Partners Covg.		***$3.6***	***-$6.5***	***-$5.6***
Govt. Subsidy Cost Increase		*$26.2*	*$19.8*	*$20.4*
Additional Cost of Govt. Funds (ECF = 0.3)		***-$7.8***	***-$5.9***	***-$6.1***

** Except Fallon, which does not operate in most of the Boston area.*

NOTE: This table shows a welfare analysis of plans' coverage of Partners hospitals described in Section 7.2. It compares the consumer plan value and costs in two scenarios: no coverage of Partners and full coverage of Partners by all plans except Fallon (which does not operate in most of Boston). In each case, plan prices adjust into Nash equilibrium and consumers re-sort across plans based on the demand system. I show all values for the overall population and separately for past Partners patients and all other enrollees. Consumer plan value is defined in Section 7.2 and equals expected plan utility (excluding switching costs), divided by the marginal utility of money. Insurer costs are defined using my estimated cost function, and Partners net inpatient revenue is defined based on the hospital demand system, its estimated prices, and its per-admission cost estimates from the Massachusetts state government (CHIA 2014). The net cost difference equals the change in insurer costs, minus the change in Partners' net revenue, and the net value minus cost difference is consumer value minus this net cost. Overall, the results show that covering Partners increases value net of costs for past Partners patients, but raises costs for all others without increasing value much.

Table 10. Counterfactual Policy Simulations

Risk Adjustment

Over-Adjustment Factor	Plan Statistics			Welfare Analysis (per member-month)					
	Covering Partners	Minimum Price	Avg. Price Other Plans	ΔCons. Surplus	Insurer Profit	Partners Net Rev.	Govt. Costs	ΔSocial Surplus ECF = 0	ΔSocial Surplus ECF=0.3
None	None	\$365.0	\$420.1	\$0.0	\$26.5	\$0.6	\$322.7	\$0.0	\$0.0
25%	None	\$374.5	\$420.9	\$4.1	\$30.0	\$0.6	\$330.7	-\$0.4	-\$2.8
50%	NHP Only	\$381.3	\$426.4	\$5.4	\$33.4	\$1.7	\$337.1	-\$1.0	-\$5.3

Marginal Subsidies

Marginal Subsidy Rate	Plan Statistics			Welfare Analysis (per member-month)					
	Covering Partners	Minimum Price	Avg. Price Other Plans	ΔCons. Surplus	Insurer Profit	Partners Net Rev.	Govt. Costs	ΔSocial Surplus ECF = 0	ΔSocial Surplus ECF=0.3
None	None	\$365.0	\$420.1	\$0.0	\$26.5	\$0.6	\$322.7	\$0.0	\$0.0
15%	None	\$368.8	\$427.6	\$0.7	\$33.4	\$0.6	\$331.1	-\$0.8	-\$3.3
25%	BMC Only	\$372.1	\$435.9	\$0.7	\$39.5	\$1.0	\$338.8	-\$1.9	-\$6.8
50%	BMC + NHP	\$384.5	\$469.4	\$2.5	\$65.5	\$2.4	\$370.2	-\$4.1	-\$18.4

NOTE: This table shows results of simulations of counterfactual policies to address the adverse selection, as discussed in Section 7.3. The top table shows simulations that modify risk adjustment by over-paying by the listed “over-adjustment factor” for people with above-average risk scores (and under-paying by the same factor for below-average risks). The bottom table shows simulations with “marginal subsidies” that narrow price differences across plans by the listed marginal subsidy rate. All simulations are for the ACA-like population and policies in 2012, so the baseline results (in the top row of each table) are the same as the 2012 equilibrium in Table 8. Each table lists which plans cover Partners, the minimum plan price, and average price of all other plans. They also list welfare statistics in units of dollars per member-month: the change in consumer surplus (with the baseline normalized to \$0), insurer profit, Partners’ net inpatient hospital revenue, and government subsidy costs. The final columns show the change in social surplus, with an excess government cost of funds (ECF) of either 0 or 0.3. The latter values each \$1 of government subsidies as incurring a social cost of \$1.3 because of the excess burden of tax financing.

Appendix A. Sample Summary Statistics

Hospital Choice Sample

Patient Characteristics

Variable		Mean
No. of Hospitalizations		74,383
Age		44.6
Male		49%
Emergency Department		56%
Diagnoses	Mental Illness	16.7%
	Digestive	13.5%
	Circulatory	12.6%
	Injury / Poisoning	7.1%
	Respiratory	7.0%
	Cancer	6.4%
	Endocrine / Metabolic	6.0%
	Musculoskeletal	5.6%
	Genitourinary	5.1%
	Pregnancy / Childbirth	5.0%
	All Others	14.9%

Chosen Hospital Statistics

Variable	Mean	Std. Dev.
Distance: Chosen Hosp. (miles)	14.1	16.3
All Hospitals (miles)	48.4	25.9
Hospital Category		
Academic Med. Ctr.	29%	---
Teaching Hospital	19%	---
All Others	52%	---
Partners Hospital	14%	---
Out-of-Network	8%	---
Past Used Hospital (>60 days before)		
Any Use	54%	---
Inpatient Use	19%	---
Outpatient Use	51%	---
Total Cost to Insurer	$11,369	$15,711
Price (estimated)	$10,981	$4,112
Patient Severity (estimated)	1.000	0.310

Plan Choice Sample

Enrollee Characteristics

Variable	Mean	Std. Dev.
No. of Enrollees	611,455	---
Age	39.6	13.8
Male	46.5%	---
Income: <100% Poverty	47.1%	---
100-200% Poverty	39.6%	---
200-300% Poverty	13.3%	---
Past Hospital User	44.3%	---
Partners Hospitals	7.4%	---
Other Hospitals	40.3%	---
Risk Adjustment Score	0.99	0.90
Choice Type: New Enrollee	29.5%	---
Re-Enrollee	13.5%	---
Current Enrollee	57.1%	---

Plan Choice Statistics

Variable	Mean	Std. Dev.
No. of Choice Instances	1,588,889	---
Insurer Price	$380.7	$69.5
Cons. Premium: Below Poverty	$0.0	$0.0
Above Poverty	$47.3	$45.7
Costs per Month: Total	$371.5	$1,480
Hospital Inpatient	$81.5	$1,048
Non-Inpatient	$290.0	$873
Current Enr: Non-Switching	95.8%	---
Market Shares: BMC	35.5%	---
Network Health	34.7%	---
NHP	19.2%	---
CeltiCare	6.9%	---
Fallon	3.8%	---

Appendix B. Demand and Cost Model and Estimation Details

B.1. Insurance Plan Demand Estimation Details

I estimate the plan demand model parameters by matching moments that fall into two categories. First, for plan dummies, I match market shares for the appropriate region/year/income group *g*. These market shares uniquely identify plan mean utilities, which in my case are equivalent to the plan dummies.[90] The formula for these market share moments is:

$$G^{(1)}_{j,g}(\theta) = \tfrac{1}{N}\sum_{i,j,t} 1\{i,t \in g\} \cdot \left[1\{y_{it} = j\} - Pr(y_{it} = j \mid \theta)\right]$$

where θ is the parameter vector, $1\{y_{it} = j\}$ is an indicator for whether individual *i* chose plan *j* at time *t*, and $Pr(y_{it} = j \mid \theta)$ is the predicted choice share from the logit model.

Second, for the coefficients for premium, network utility, and other observed characteristics (which are interacted with observed enrollee attributes), I match the average values for chosen plans in the data to those in the model. Specifically, the moments for characteristic $X^{(k)}$ (e.g., premium) interacted with enrollee attribute $Z^{(r)}$ (e.g., income) are:

$$G^{(2)}_{k,r}(\theta) = \tfrac{1}{N}\sum_{i,j,t} X^{(k)}_{ijt} Z^{(r)}_{it} \cdot \left[1\{y_{it} = j\} - Pr(y_{it} = j \mid \theta)\right]$$

Another way of interpreting these is as matching the *covariance* between plan characteristics and household attributes. In the case of observing the full market, these moments are equivalent to the micro BLP covariance moments. These moments are also equivalent to first-order conditions from the associated maximum likelihood problem.

Stacking all of the moments into a vector $G(\theta)$, the MSM estimator searches for the parameter θ that minimizes the weighted sum of squared moments, $G(\theta)' \cdot W \cdot G(\theta)$. Because the system is just-identified (equal number of parameters and moments), I am able to match the moments exactly, making the solution invariant to the choice of *W*. I calculate standard errors using the standard GMM sandwich formula. To account for the fact that network utility variable is derived from the hospital demand estimates, I am planning to implement an adjustment following the lecture notes of Pakes (2013). However, I have not yet implemented this adjustment in the current draft.

B.2. Inattention Interpretation of Plan Inertia Coefficients

[90] A difference in my setting from the standard BLP approach is that I treat the plan dummies as parameters, with associated standard errors, since both they and the characteristics coefficients are estimated from a dataset of the same size (the full market data). In previous applications including Berry, Levinsohn, and Pakes (2004), the micro data came from a sample, while the market shares came from aggregate data on the whole market.

For current enrollees, I included in the logit demand model a dummy variable for their current plan, so their full demand utility was:

$$U_{ijt}^{Curr} = \hat{U}_{ijt} + \underbrace{\chi(Z_i)\cdot 1\{j = CurrPlan\}}_{\text{Switching Cost / Inertia}} + \varepsilon_{ijt}^{Plan} \tag{14}$$

where $\hat{U}_{ijt}$ is the plan utility for new enrollees (defined in Section 4.3), excluding the ε_{ijt}^{Plan}. In this equation, $\chi(Z_i)$ is interpreted as a switching cost – an extra utility for the current plan needed to rationalize the low level of plan switching. The plan demand estimates in Table 6 reports these switching costs but also an alternate interpretation based on an inattention model. I show here how I derive the inattention/passive probability reported in Table 6.

Consider a two-step model in which the first step models whether enrollees make an active choice, and the second step models plan choice conditional on being active. The second step is standard and follows the logit model for new enrollees (or current enrollees excluding switching cost):

$$Pr(y_{it} = j \mid Active) = \frac{\exp(\hat{U}_{ijt})}{\sum_k \exp(\hat{U}_{ikt})}$$

The first step is a reduced form model of being passive:

$$Pr_{it}(Passive) = \frac{\exp(\hat{U}_{i,j_{curr},t} + \tilde{\chi}_i)}{\exp(\hat{U}_{i,j_{curr},t} + \tilde{\chi}_i) + \exp(I_{i,Active,t})} \quad \text{where } I_{i,Active,t} = \log\left(\sum_k \exp(\hat{U}_{ikt})\right)$$

Notice that it is the choice probability from a two-choice logit model, where the utility of being passive is the current plan utility plus a reduced-form inertia coefficient $\tilde{\chi}_i$ (which is different from the switching cost χ). The utility of being active is $I_{i,Active,t}$, which is the inclusive value (or expected utility) from the second-stage active choice model.

I claim that if $\tilde{\chi}_i = \log(\exp(\chi(Z_i)) - 1)$, the switching cost and inattention models have identical predictions for choice probabilities. For the current plan, the inattention model predicts a probability that it is chosen of $Pr_{it}(Passive) + (1 - Pr_{it}(Passive)) \cdot Pr(y_{it} = j_{curr} \mid Active)$, which simplifies to:

$$Pr(y_{it} = j_{curr}) = \frac{\exp(\hat{U}_{i,j_{curr},t} + \chi(Z_i))}{\exp(\hat{U}_{i,j_{curr},t} + \chi(Z_i)) + \sum_{k \neq j_{curr}} \exp(\hat{U}_{ikt})}$$

This equals the current plan's choice probability in the switching cost model in (14). Further, the inattention model's probability of switching to another plan *j* is $(1 - Pr_{it}(Passive)) \cdot Pr(y_{it} = j \mid Active)$, which simplifies to:

$$Pr\left(y_{it}=j\right)=\frac{\exp\left(\hat{U}_{i,j,t}\right)}{\exp\left(\hat{U}_{i,j_{curr},t}+\chi\left(Z_i\right)\right)+\sum_{k\neq j_{curr}}\exp\left(\hat{U}_{ikt}\right)}$$

which is again equivalent to the choice probability from the switching cost model in (14).

Hence, these two models have equivalent predictions for choice probabilities. The plan demand results in Table 6 report both the average switching costs $\chi\left(Z_i\right)$ and the passive probability $Pr_{it}\left(Passive\right)$, as defined by the equation above.

B.3. Details of Hospital Price Model

As discussed in Section 5.1, I estimate a risk-adjusted hospital price model. Recall that I estimate a Poisson regression (also known as a generalized linear model with a log link) of the form:

$$E\left[Payment_{i,j,h,t,a}\mid Diag_{ita},Z_{ita}\right]=\exp\left(\rho_{j,h,t}+Diag_{ita}\lambda+Z_{ita}\gamma\right)$$

For the principal diagnosis ($Diag_{ita}$), I use the Clinical Classification Software (CCS) dummies defined by the U.S. government's Agency for Healthcare Research and Quality. The additional covariates (Z_{ita}) include age, gender, income, and Elixhauser comorbidity dummies for the secondary diagnoses.

I specify a restricted model for $\rho_{j,h,t}$ to avoid over-fitting for hospital-insurer-year cells with small samples. Specifically, I start from the model:

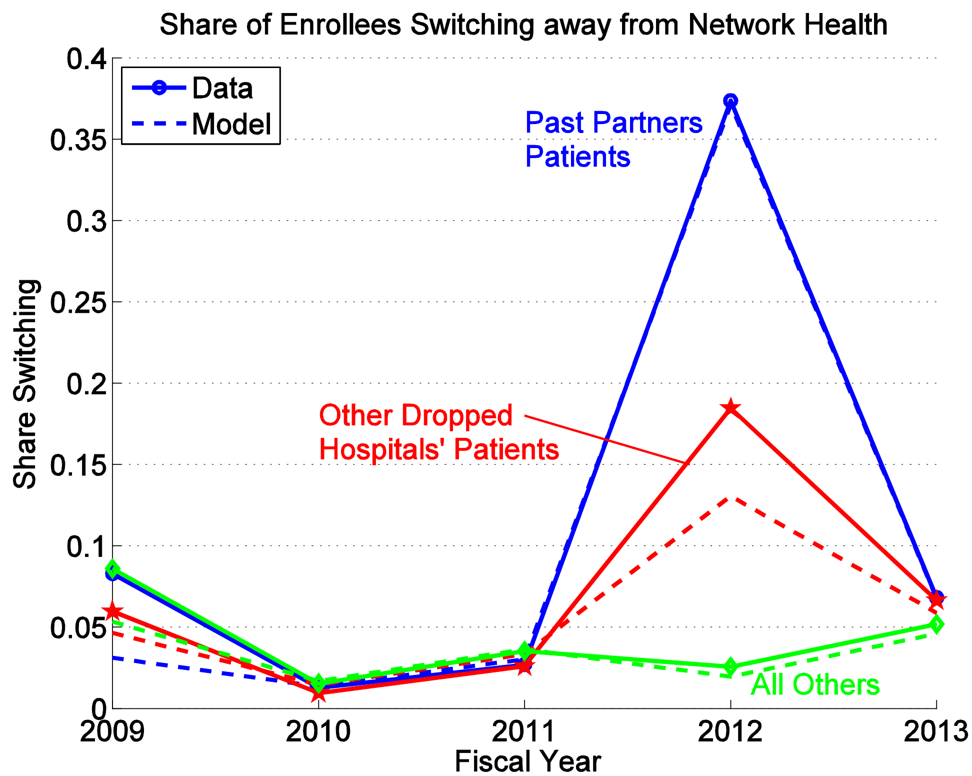
$$\rho_{j,h,t}=\rho_{j,h,NetwStat(h,t)}+\rho_{j,Sys(h),t}+\rho_{j,t,NetwStat(h,t)}$$

The first term, $\rho_{j,h,NetwStat(h,t)}$, is a coefficient on hospital-insurer-network status (i.e., in or out of network) dummies that is constant across years. I include this term for all cells with at least 50 observations; otherwise, I set it to zero. The second term, $\rho_{j,Sys(h),t}$, is a coefficient on insurer-hospital system-year dummies for the top six hospital systems. This allows for a separate hospital price paths over time for each of the largest systems (including Partners). I do not include this term for hospitals in smaller systems or when the large system is out-of-network, with the exception that I always include these dummies for Partners regardless of whether it is in-network. The final term, $\rho_{j,t,NetwStat(h,t)}$, is a residual that allows for a separate effect for each plan, year, and network status. This captures the average insurer-specific price path for all smaller hospitals not included in one of the six largest systems.

Appendix C. Model Fit Tables and Figures

This appendix shows tables and figures that display the model's ability to match the reduced form patterns around Network Health's dropping of the Partners hospitals in 2012, as discussed in Section 5.4.

Appendix Table C.1. Plan Switching Patterns

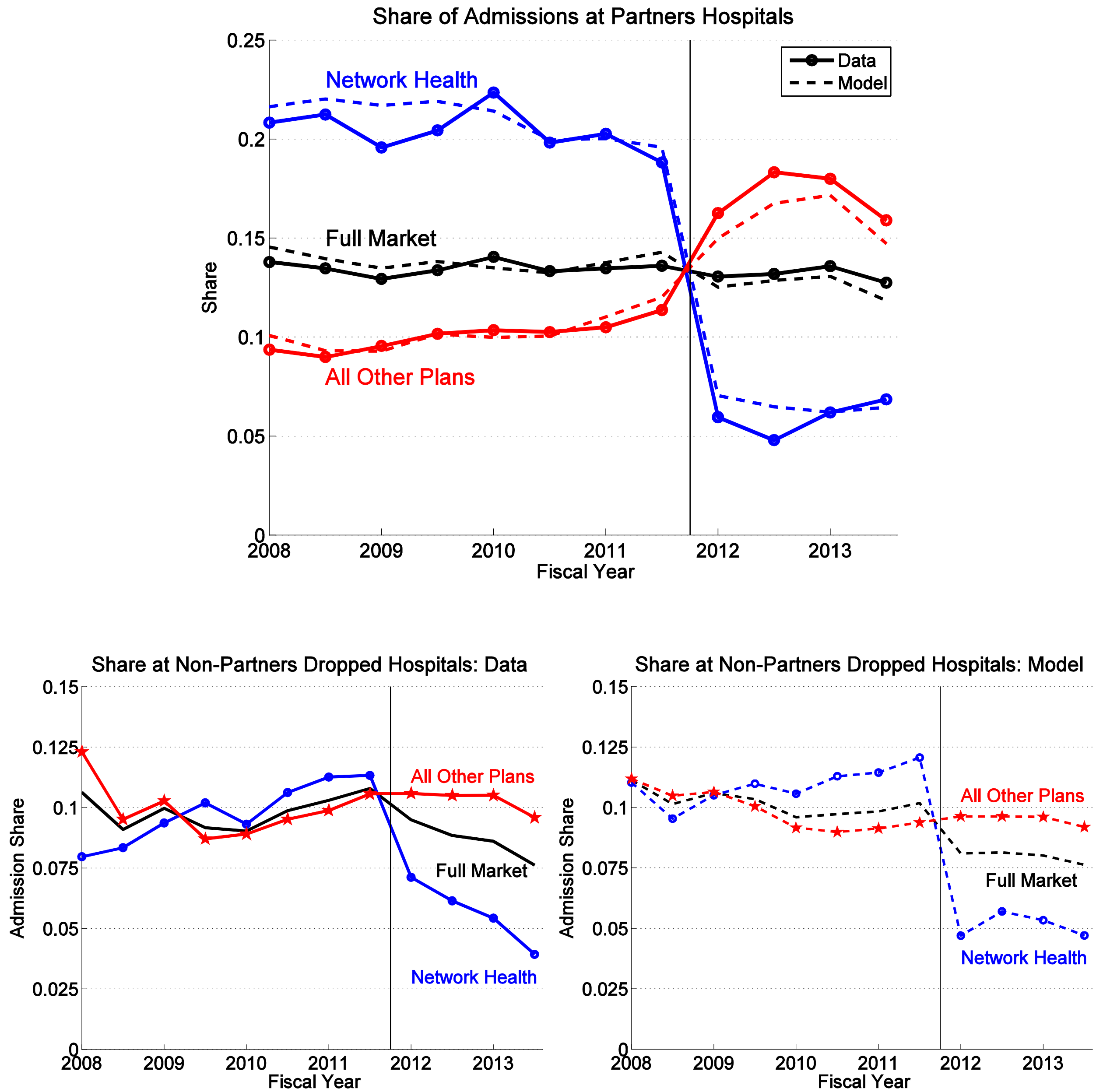


Appendix Table C.2. Cost Changes for Network Health

Network Health: Average Costs 2011-12

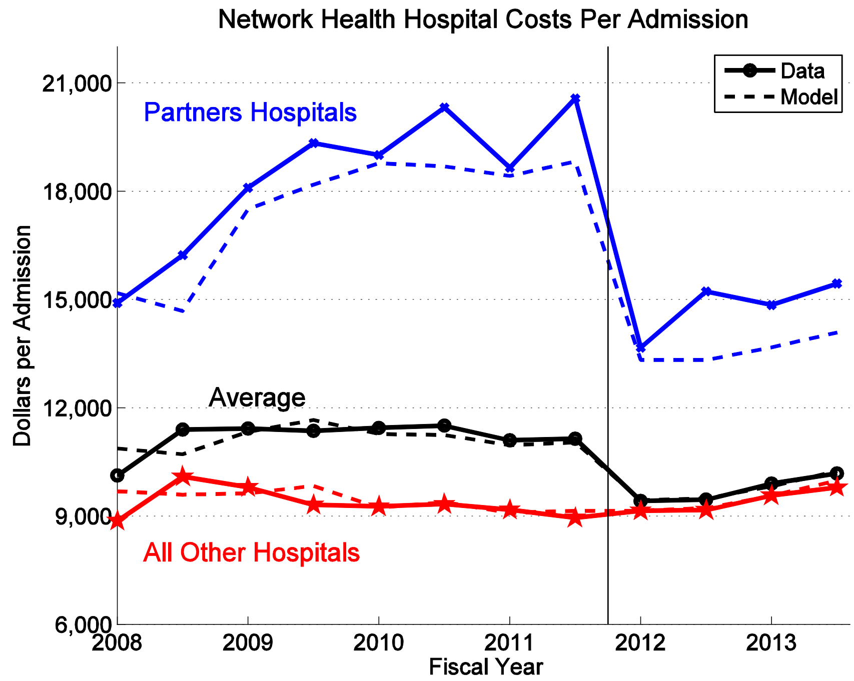
Enrollee Group	Data 2011	Data 2012	Data %Δ	Data Risk Adj. %Δ	Model 2011	Model 2012	Model %Δ	Model Risk Adj. %Δ
All Enrollees	**$378**	**$313**	**-17%**	**-15%**	**$374**	**$310**	**-17%**	**-16%**
Stayers (in plan both years)	$317	$305	-4%	-5%	$334	$312	-7%	-9%
2011 Only Enrollees	$476	---	---		$435	---	---	
2012 Only Enrollees	---	$310	---		---	$302	---	

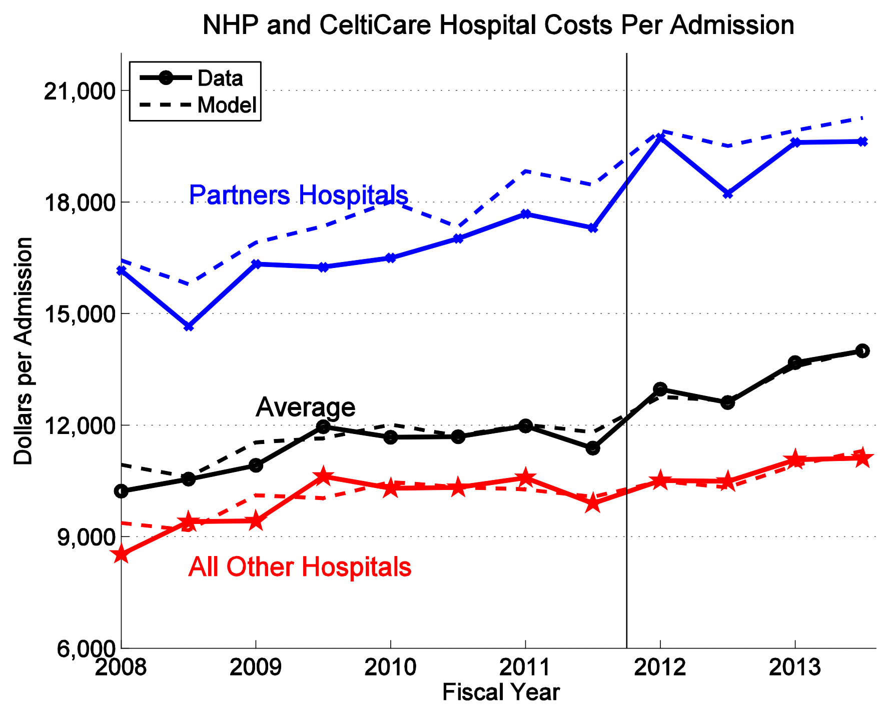
Appendix Table C.3. Admission Shares at Hospitals Dropped by Network Health in 2012



NOTE: These figures show the share of hospital admissions at hospitals that Network Health plan dropped from its network in 2012. See Figure 2 for a more detailed description of the values in the data. The dashed lines show the model's prediction for the same statistics. These are calculated holding fixed each individual's observed plan, not reassigning plan choices using the plan demand model.

Appendix Table C.4. Changes in Cost per Hospital Admission around 2012 Network Changes





NOTE: These figures show average costs per hospital admission for two sets of plans: Network Health (top figure), which dropped the star Partners hospitals in 2012, and NHP and CeltiCare (bottom figure), which continued to cover them. See Figure 3 for a more detailed description of the values in the data. The dashed lines show the model's prediction for the same statistics. These are calculated holding fixed each individual's observed plan, not reassigning plan choices using the plan demand model.

Appendix D. Simulation Method Details

This appendix details the simple approach I use to incorporate a future profit effect in a static pricing model for my simulations in Section 7. Note that in a dynamic model, an insurer's pricing FOC includes a term capturing the effect of changing today's price on future profits on consumer *i*. I model this "future profit effect" as the product of the change in future demand $(\partial D_{ij}^{Fut} / \partial P_j)$ times an expected profit margin M_{ij}^{Fut}, which is unaffected by today's price. For the change in future demand, a lower price increases demand today and therefore increases the number of inertial enrollees in the future. To simplify, I take future market enrollment $(nMon_{i,t+k})$ as given and assume an exogenous, constant inertia probability η at each year's switching choice, which I set at 89%.[91] Given these assumptions:

$$\frac{\partial D_{ij}^{Fut}}{\partial P_j} = \frac{\partial S_{ij}}{\partial P_j} \cdot \left(\sum_{k \geq 1} \eta^k \cdot nMon_{i,t+k} \right) \tag{15}$$

where $\partial S_{ij} / \partial P_j$ is the effect of price on current year's choice share.

Finally, I need to specify insurers' future profit margins. Although imperfect, I simply assume that insurers expect M_{ij}^{Fut} to equal current margins at the enrollee level – which assumes that prices and costs grow in parallel for each enrollee. Notice that I still treat M_{ij}^{Fut} as a constant in the pricing FOC but plug in the equilibrium margin ($= \varphi_i P_j^* - c_{ij}\left(N_j\right)$) for it at the end.

Combining these assumptions and defining the term in parentheses in (15) as $nFutMon_i$, the pricing FOC for insurer *j* is:

$$\begin{aligned} 0 &= \frac{\partial \pi_j}{\partial P_j} + \sum_i M_{ij}^{Fut} \cdot \frac{\partial D_{ij}^{Fut}}{\partial P_j} \\ &= \sum_i \varphi_i \cdot nMon_i \cdot S_{ij}\left(.\right) + \sum_i \left(\varphi_i P_j - c_{ij}\right)\left(nMon_i + nFutMon_i\right) \cdot \frac{\partial S_{ij}}{\partial P_j} \end{aligned} \tag{16}$$

Accounting for future profits adds the $nFutMon_i$ term to the FOC, which increases the incentive to lower prices (just like a steeper demand curve). This effect is likely to have a significant impact. Months in the current year $(nMon_i)$ average 6.2, and future months $(nFutMon_i)$ average 6.8. So the future profit effect works like a more than doubling of the demand slope.

[91] I use 89% rather than the 95% inertia probability reported in the plan demand estimates based on a rough correction for unobserved heterogeneity. Looking at re-enrollees (people who leave the market and return later), people tend to actively choose the same plan as during their prior spell about 55% of the time. For an inertia probability of ρ the overall non-switching probability is $\rho + (1-\rho) \cdot Pr_j^{Active}$. Plugging in $Pr_i^{Active} = 55\%$, $\rho = 89\%$ is required to rationalize a 95% overall non-switching probability.