Employer Incentives and Distortions in Health Insurance Design: Implications for Welfare and Costs*

Nicholas Tilipman†

November 6, 2020

Abstract

This paper studies employer incentives in designing health insurance provider networks and whether observed offerings reflect preferences that are aligned with employees. I estimate a model of supply and demand where I endogenize employer health plan offerings with respect to hospital and physician networks. I find that employers "overprovide" broad networks by systematically overweighting the preferences of certain employees, specifically those in regions with less provider competition, over those of the average employee household. Shifting employers towards offering different provider networks in different geographic markets could yield substantial gains to surplus, with minimal distributional or selection effects.

Keywords: health insurance, narrow networks, switching costs
JEL Classification Codes: I11, I13, D83, G22

*This paper was previously circulated as *Narrow Physician Networks, Switching Costs, and Product Variety in Employer Markets.* I am deeply grateful to my dissertation committee, Sean Nicholson, Panle Jia Barwick, and Samuel Kleiner for their valuable advice and encouragement. I also thank Matthew Backus, Colleen Carey, Jean-Francois Houde, Pierre Leger, Anthony LoSasso, Claudio Lucarelli, Ross Milton, Elena Prager, Yuta Toyama, Wesley Yin, and numerous seminar participants for their comments and insight. This project was also made possible by a Cornell University Human Ecology Alumni Association grant and an Institute for Social Sciences grant. All mistakes are my own.

†Division of Health Policy and Administration, School of Public Health, University of Illinois at Chicago. Contact: tilipman@uic.edu.

1 Introduction

A central tension in insurance markets is how to optimally design product options that balance consumer preference for risk protection with minimizing moral hazard and health spending (Arrow, 1965; Pauly, 1968). In the market for health insurance, these tradeoffs are typically managed by intermediaries. For example, state and federal governments play an active role in determining plan menus on the Affordable Care Act exchanges, with some states restricting choice sets to only a few plans and other permitting a wide variety of options (Scheffler et al., 2016). Conversely, individuals who purchase coverage through their employer are typically exposed to fewer choices than in the individual market.[1] An important question is therefore whether these intermediaries serve as effective agents for their risk pools. An intermediary that weights each consumer equally will choose optimal coverage levels that differ to one that has private incentives or faces frictions in plan design.

In this paper, I study the determinants of health plan offerings among large employers and whether these plan choices reflect preferences that are aligned with that of employees. This is an important market to study this issue: employer-sponsored insurance (ESI) is a significant part of the healthcare landscape (ESI), representing approximately 30% of health expenditures. Moreover, costs in the employer market have been rising rapidly in recent years. Per-enrollee expenditures in the private ESI market have increased about 46% between 2008 and 2018, compared with an approximate 21% increase in Medicare per-enrollee spending over the same period.[2] This paper sheds light on the whether a portion of these rising costs can be attributed to mismatch between employer and employee preferences.

I focus my analysis on employer decisions over health plan provider networks and, in particular, the decision of whether to offer "narrow-network" benefit designs as part of their plan menus. Health insurers and employers have increasingly started offering these insurance plans as a means of containing spending and offering consumers low-cost options.[3] Despite the increasing popularity of narrow-network plans on the health insurance Exchanges, however, employers have been slower to adopt, design, and offer such products. In 2016, only 7% of employers nationally offered a narrow network as part of their plan menu (Hall and Fronstin, 2016).

To investigate employer incentives in network design, I estimate a model of supply and demand for health insurance plans for a large-group purchaser (employer) in Massachusetts: the Group Insurance Commission (GIC). The GIC offers coverage to public employees in the state, including approximately 300,000 active state government employees, as well as retirees and the employees of several municipalities.[4] It is an ideal setting for studying the welfare effects of narrow-network

[1] Indeed, most employers typically offer one or two plans to their employees, with coverage that tends to be quite comprehensive (Buchmueller et al., 2013; Dafny et al., 2013).

[2] https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.

[3] These plans achieve lower costs, and lower premiums, by significantly limiting the set of hospitals and physicians that an insurer will cover to only those with lower negotiated reimbursement rates. Approximately 70% of the plans available on the Affordable Care Act (ACA) health insurance Exchanges have been found to be "limited network" plans, covering fewer than 30% of the 20 largest hospitals in the market (McKinsey Center for U.S. Health System Reform, 2013) and about 40% of the plans cover less than 25% of the physicians in the market (Polsky and Weiner, 2015).

[4] In this way, it acts as a sort of employer Exchange, offering various products to all employees who participate in the group.

products for several reasons. It has, in the last several years, held most non-premium aspects of its plans fixed, with the exception of its networks. In fact, it has been active in encouraging the creation and adoption of narrow-network products and offers plans with considerable variation in both hospitals and physicians covered. This includes the addition of two new narrow-network products midway through my sample. Second, the GIC competes for business from municipal employers and, as such, sees variation in the set of employees in the pool over time. Finally, in 2012, the GIC instituted a "premium holiday" in which they forced all active state employee to re-enroll in a health plan, while simultaneously offering three months of free coverage if they switched from a broad-network to narrow-network product (Gruber and McKnight, 2016). This policy change is instrumental: not only does it allow me to more cleanly estimate price elasticities for broad-network vs. narrow-network products, but it also aids in the identification of health plan switching costs, a critical determinant of broad-network preferences in my setting. This identification is similar to prior work on switching costs, most notably Handel (2013).

The main part of the paper estimates a four-stage model of demand, pricing, and employer plan offer decisions. The model has several novel features. First, the demand side of the model incorporates significant observed and unobserved heterogeneity to ensure correct identification of network preferences. Notably, I estimate demand for not only hospitals, but also physician networks for three specialty groups: primary care, cardiology, and orthopedics.[5] To my knowledge, this is the first paper that incorporates willingness-to-pay for physician networks–in addition to hospital networks–into a model of insurance demand.[6] I show that these inclusions are important: access to a broad network of *physicians* explains about 85% of the preferences for broad networks, with only about 15% explained by hospital networks. Moreover, incorporating preferences for physician networks nearly triples premium elasticities, implying the average household may be highly sensitive to price. These results have important implications for optimal employer plan offerings.

Second, I endogenize employer plan menus (both in terms of the provider networks offered to employees and the *number* of different networks offered) by fully specifying an employer objective function that I estimate using moment inequalities. The employer function includes three key components: the value of the offered plan menu to the pool of employees; the net spending on premiums incurred by the employer given the offered products; and the fixed costs of offering multiple plans. I capture the extent to which employer and employee preferences differ through a single parameter (hereafter referred to as a "mismatch" parameter) that estimates the weight the employer places on enrollee surplus through the observed networks relative to its predicted net spending. Intuitively: if employer and employee preferences were fully aligned, then the employer ought to value consumer surplus at nearly an equivalent level to spending. Otherwise, the employer could

[5]Together, these specialties comprise approximately 65% of all physician office visits. $https : //www.cdc.gov/nchs/data/ahcd/namcs_{s}ummary/2013_{n}amcs_{w}eb_{t}ables.pdf$

[6]Much of the existing literature on networks has exclusively focused on hospitals (Ho, 2009; Shepard, 2016; Prager, 2016; Ho and Lee, 2019; Ghili, 2020; Liebman, 2018) and has ignored the role of physicians in determining consumer choice of insurance plans.This is likely due to three factors. First, until recently, physician markets were often thought to be less interesting than hospital markets, as physicians had very little bargaining power to leverage high prices from insurance plans. Second, estimation of physician demand is complicated by dimensionality: whereas there are typically a small number of hospitals in any given market, there are often thousands of physicians of various specialties, rendering the study of physician markets difficult in structural IO models. Finally, there is the lack of available data allowing researchers to both link individual physicians to their respective medical groups and construct physician networks of insurance plans.

simply offer a counterfactual menu and compensate employees either through lower co-premiums or through wages for any lost utility.[7]

My principal finding is that the persistence of broad networks does not fundamentally reflect the preferences of the average employee. Rather, employers place a significantly higher weight on the value of broad networks relative to how the average household values those networks. Specifically, I estimate that the employer values a dollar of consumer utility from a network by nearly four times as much as it values a dollar on premium spending. This results, in some sense, in an over-provision of broad networks: the employer maintains access to these plans for its employees even when removing them may lead to substantial savings in costs in excess of lost utility. This is suggestive of a fundamental mismatch between employer and employee incentives: the *average* employee in the pool would prefer only to have access to narrow-network products and to be compensated for the lack of access to their preferred providers in the form or lower co-premiums, or higher wages. However, the employer prefers to offer relatively more generous coverage and higher premiums. The implications of this are substantial: if the employer valued enrollee utility the same level as premium spending, my model predicts they would not only drop all broad-network plans in favor of more narrow-networks, but would reduce the overall *number* of plans offered as well. If employees could be compensated fully for the utility loss of their network, this change of plan menu would imply a social surplus gain of about \$40 per-household-per-month, implying that this mismatch results in about a \$480 per-household-per-year welfare loss.

This leaves the question of *why* employers exhibit this behavior in menu choice. I explore several candidate possibilities: (a) inability to alter co-premiums or benefits (by virtue of the employer I study being a public rather than private employer); (b) employer mistakes or misperceptions; and (c) heterogeneity in the types of employees the employer values most when designing benefits. While I am not able to fully separate these channels, I find substantial evidence for (c). Specifically, about 30% of the estimated employer-employee mismatch can be attributed to employers placing a higher weight on the network preferences of the oldest workers in the distribution, while as much as 80% of the mismatch can be explained by employers emphasizing the preferences of employees in certain geographic *regions*. These regions tend to be ones that are less dense, have fewer competition among health care providers, and are situated close to the state border. As such, households in these regions stand to lose the most utility from a loss of a notable provider. Importantly, households in these regions do *not* seem to be the ones with the highest health risk or those with the highest ex-ante probability of health care utilization.[8] These patterns persist even among large, private employers, who theoretically have more flexibility in benefit design and responding to labor market frictions. As such, there is strong evidence that observed network design may be driven either out of equity concerns for employees who live in regions with a sparser set of providers or beliefs that attracting older employees may yield productivity benefits for the firm. The latter would be consistent with, for instance, employers favoring the preferences of managers, executives, or employees that otherwise have stronger labor market bargaining power.

[7] Consider the case where the employer valued enrollee surplus *more* than what it spends on premiums, then there exists a feasible plan menu that such that, if offered, the dollarized utility change for employees relative to the current menu would be smaller than the decrease in premium spending. As such, the employer could offer this counterfactual menu and fully compensate employees for the lost utility in a way that would enhance social surplus.

[8] In fact, I find very little evidence that, apart from age, employers favor the *sickest* employees in the pool.

While these results are suggestive of a mismatch between employee preferences and employer incentives, the observed outcomes are buoyed by several market frictions and constraints. First, employers are prohibited by law from risk-rating or offering different benefits based on health status. Further, most large employers, even when operating in many geographic markets, offer identical plan designs across market. Due to this uniformity of benefits, employers cannot fully pass on the price of broad networks to consumers. As a result, consumers in many regions enroll in broad-network products even if they do not value those networks at their full cost. These households would lose relatively little utility by switching to narrow-network plans but would significantly reduce the group's costs if they did so, similar in spirit to the result in Bundorf et al. (2012). Second, the presence of strong health plan switching frictions implies that this phenomenon cannot simply be resolved by offering more choice of networks without substantially widening the premiums between those plans.[9] In other words, the mere presence of a narrow-network option, despite being lower cost, is not enough to induce those previously enrolled in broad-networks to switch, even though I estimate these households would see considerable savings from doing so.

Motivated by these findings, I next use my estimated coefficients to simulate employer equilibrium plan offerings, as well as total spending, utility, and social welfare changes from several alternate pricing and plan design policies. I first consider a uniform pricing policy in which the GIC is moved from its current scheme of subsidizing 75% of all premiums to an "Enthoven" style managed-competition scheme. Under this approach, the employer fully reimburses the premiums of the lowest-cost plan offered and employees bear the full increment of enrolling in any plan that has costs exceed the benchmark plan. I find that while such a policy would lead to large aggregate gains—approximately \$44 per household per month—the distributional consequences would be severe. In implementing this policy, the price of the broad-network plans rise to such an extent, that the employer ceases offering its flagship broad-network product altogether. Consumers in rating regions with weak preferences for broad-networks would be more than fully compensated by the increased subsidies from the employer to purchase narrow plans, leading to the large utility gains. However, households in regions with strong preferences for broad networks see considerable utility *declines* through the removal of this plan and through the increases prices of the remaining broad networks. As these are the very households that the employer tends to overweight in its benefit design, this may shed light on why employers do not often move to an Enthoven-style approach in equilibrium.

The story changes drastically, however, if I allow the employer to deviate from a uniform pricing and benefits structure. I next permit the employer to both rate premiums and set networks differentially by *region*. I find that such a policy would bring about substantial social surplus without the aforementioned distributional impacts. In this scenario, the employer is predicted to drop access to the costliest broad-network plans in three of the seven rating regions in the state, while preserving access to all broad networks in the remaining regions. Across *all* rating regions, the employer significantly reduces the overall number of plans it offers and, in some instances, drops certain insurers altogether. While this decrease in choice is predicted to lead to a \$7 per household per month decline in utility, it also leads to a \$32 per household per month decline in spending. If the employer could compensate the utility lost with these savings, this translates to

[9] This can unravel the market if there is substantial adverse selection (Marone and Sabety, 2020).

an approximately \$18 per household per month gain in social welfare.

While these gains are not as sizable as the uniform pricing approach in aggregate, the policy does see fewer adverse distributional consequences. Since there is little correlation between regions with the highest willingness-to-pay for broad networks and health risk, the utility losses in this scenario are primarily concentrated on employees with lower valuations for broad networks. In other words, permitting region-rating and benefit design essentially allows the employer to shift consumers with low value for broad-network plans into lower-cost products, while still preserving access to broad networks to the highest willingness-to-pay consumers. I find that this results in socials surplus gains across the age and location distribution.

This paper relates to several strands of literature. The first strand includes studies on behavioral and switching frictions in health insurance (Handel, 2013; Polyakova, 2016; Abaluck and Gruber, 2016; Ho et al., 2017). I also contribute to the literature on product entry, innovation, and variety that endogenizes firm product quality choices (Nosko, 2014; Eizenberg, 2014; Mohapatra and Chatterjee, 2015). A third strand focuses on network formation (Ho, 2006, 2009; Shepard, 2016; Lee, 2013; Ho and Lee, 2019; Liebman, 2018; Ghili, 2020; Prager, 2016) and valuation of narrow-network plans (Gruber and McKnight, 2016; LoSasso and Atwood, 2016; Dafny et al., 2015; Ericson and Starc, 2015a). Of particular importance is Shepard (2016), who uses a similar demand model to study whether adverse selection leads to the narrowing of networks on the individual market.

Most importantly, I contribute to the literature on the determinants and value of insurance plan choice, competition, and provision (Einav et al., 2013; Ericson and Starc, 2015b, 2016; Dafny, 2010; Dafny et al., 2012, 2013; Scheffler et al., 2016; Bundorf et al., 2012). In recent studies, Ho and Lee (2020) and Marone and Sabety (2020) explore analgous questions of employer menu design, focused on financial dimensions such as cost-sharing. My paper contributes to this literature not only by focusing on plan offerings as it pertains to provider networks, but also by endogenizing firm offers. As such, it explores the determinants of employer plan menus as well as the question of *why* employer menus may deviate from the choices we might expect to see made by a social planner.

The paper proceeds as follows: Section 2 outlines the data and setting for my study and presents some empirical patterns. Section 3 details the model, estimation, and parameter results. Section 4 discusses potential sources of employer-employee preference mismatches. Section 5 presents the results of counterfactual policy simulations. Section 6 concludes.

2 Data and Institutional Setting

2.1 Setting

Group Insurance Commission: The focus of this paper is the Group Insurance Commission (GIC) in Massachusetts, a large purchasing organization in Massachusetts that services the state's government employees—both employees of the state itself and local municipal governments. Though state employees constitute the bulk of GIC members, since 2007, municipalities have increasingly abandoned their existing insurance arrangements in favor of getting insurance through the GIC and, as such, there are a large number of municipal entrants in subsequent years. Therefore, the GIC has an interest in not only providing satisfactory health benefits for its existing members, but potentially competing for new members as well. In total, there are approximately 300,000 enrollees on the GIC, representing approximately 8% of the Massachusetts employer-sponsored-insurance market.

The GIC contracts with multiple health insurance carriers and provides multiple competing plans for enrollees. In particular, it contracted with six carriers throughout my sample period: Fallon Community Health Plan, Harvard Pilgrim Health Care, Health New England, Neighborhood Health Plan, Tufts Health Plan, and Unicare Health. Each carrier offers multiple plans at different total premiums. The GIC subsidizes 75% of these total premiums, leaving employees to pay a 25% "co-premium" to enroll in a plan. These premiums do not vary by consumer risk type or geography, but rather only vary with whether the household is a single-member ("individual") or multi-member ("family") household. Specifically, all GIC family plans are 2.4 times the individual rate. Apart from premiums, these plans are entirely standardized with the exception of two dimensions. The first is that the GIC employs tiered copay arrangements, which generates variation in copays across providers (discussed at length in Prager (2016).

The second is the actual network of included providers on each plan. In 2009 and 2010, four of the carriers offered narrow-network products with varying degrees of network breadth. In 2011, the GIC enacted a major change to the choice set by introducing narrow-network plans from both remaining insurers (Harvard Pilgrim and Tufts Health Plan, two dominant players in the state).[10] These plans are approximately 20% cheaper on average than their respective broad networks, though generally cover more providers than the narrow-network plans offered by the same insurers in other market segments.[11]

Premium Holiday: Though the GIC has promoted the adoption of narrow-network products, enrollment in these products was fairly limited in 2011 and health care spending among the group continued to rise. As a result, in 2012, the GIC offered a three-month "premium holiday" for all active state employees who chose to switch to a narrow-network plan. For households choosing to make the switch, the holiday entailed that they pay *no* premiums for three months of the fiscal year. Importantly, this holiday was not extended to municipal workers, but rather just active state employees. This served as the basis for prior work on evaluating the impact of narrow-network

[10] These plans are called "Harvard Primary Choice" and "Tufts Spirit", but will hereafter be referred to as "Harvard Narrow" and "Tufts Narrow."

[11] For example, Harvard offers a narrow-network plan in the small-group market known as "Harvard Focus," which is considerably narrower than the "Primary Choice" plan offered on the GIC.

product introduction (Gruber and McKnight, 2016). The holiday was fairly successful, inducing approximately 10% of enrollees to switch and resulting in approximately 20% savings in spending for those enrollees, largely due to the use of lower-cost providers.

Table 1 shows the market shares and premiums for all the plans offered on the GIC in 2012, the year *after* Harvard and Tufts both introduced narrow-network products. This also coincides with the first year of the premium holiday. The most expensive plans on the market are Unicare's Indemnity plan, as well as Harvard Independence (hereafter "Harvard Broad") and Tufts Navigator (hereafter referred to as "Tufts Broad"). The broad plans have the highest market shares, with Tufts and Harvard each making up about 25%-30% of the market. Their narrow plans, however, had much more limited enrollment in 2012, with about 5% for Harvard Primary Choice and 2% for Tufts Spirit. This is up from 2% and 1%, respectively, in 2011, due in large part to the premium holiday inducing members to switch to these narrow plans. Interestingly, despite having lower co-premiums, Tufts Spirit had a significantly lower market share than Harvard Primary Choice.[12] This is a point that I will return to below.

Table 1: GIC Summary Statistics, 2012

Insurer	Network Coverage	Market Share	Co-Premium ($PMPM)
Fallon Select	Broad	0.03	139.39
Fallon Direct	Narrow	0.02	112.97
Harvard Independence	Broad	0.21	163.98
Harvard Primary Choice	Narrow	0.05	131.50
Health New England	Narrow	0.06	110.34
Neighborhood Health Plan	Broad	0.02	113.02
Tufts Navigator	Broad	0.27	148.43
Tufts Spirit	Narrow	0.02	119.06
Unicare Indemnity	Broad	0.13	247.07
Unicare Plus	Broad	0.08	207.27
Unicare Community Choice	Narrow	0.10	111.61
Number of Enrollees in GIC	293,125		
Average Age	36.07		
Average Subscriber Age	48.04		

Notes: GIC plans for 2012. Co-Premiums refer to the enrollee share of the per-member-per-month premiums (25% of the overall premium).

2.2 Data Sources

I use two primary data sources to conduct the analyses in this paper: enrollment and claims data from the Massachusetts All-Payer Claims Database (APCD) and physician affiliation data from the SK&A database of physicians.

Enrollment and Claims Data: Enrollment and claims data for the GIC come from the Massachussetts APCD, a comprehensive database of medical claims from public and private payers in Massachusetts from 2009-2013. The claims data contain detailed information on both hospi-

[12]Though 5% versus 2% market share seems low, this represents a difference in almost 12,000 members.

tal and physician visits, with variables indicating the patient's primary and secondary diagnoses (through ICD9 codes), procedures performed (CPT codes), patient demographics (including patient and provider 5-digit zip codes, which allow me to estimate the effect of distance on provider demand), longitudinal patient identifiers, physician and facility identifiers, physician specialty, insurance and plan identifiers, and a wide variety of payment variables. Importantly, these payment variables contain not only hospital "charges amounts," but the amounts the insurers actually paid each provider for each claim, as well as the out-of-pocket amount received from the patient. In addition, the enrollment data from the APCD contain a record for each member enrolled in a health plan in Massachusetts. Enrollee characteristics include age, location, gender, and dates of enrollment in each plan. Plan characteristics include product type (e.g. HMO, PPO, etc.), specific plan identifier, market identifier (e.g. individual, group, GIC, etc.), and cost-sharing features.

I create samples for hospital admissions, physician visits, and insurance plan choice. I focus, in particular, on patient demand for primary care physicians (PCPs), cardiologists, and orthopedists. A detailed description of the different subsamples I create pertaining to different stages of my model are presented in Appendix A, along with summary statistics for the different specialties.

Physician Data: I focus my analysis of physicians on the *practice* level. In order to obtain physician characteristics as well as link physicians to their practices, I use proprietary data from the SK&A database for 2009 and 2013. The database includes information on each individual physician's name, location, specialty, NPI, affiliated medical group, affiliated hospital, and affiliated health system. It also contains characteristics for the site of the physician practice, including number of physicians on staff, the specialty of the practice, and the number of physicians on staff across all the locations of the particular medical group. The SK&A includes approximately 95% of all office-based physicians practicing in the United States, and the data is verified by the proprietors over the telephone.

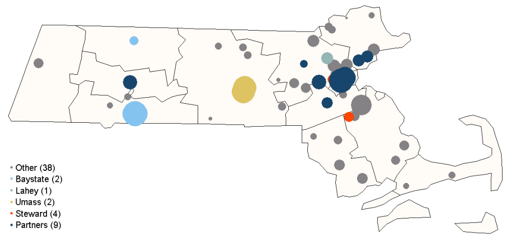
Premium and Network Data: I obtain premiums for GIC plans between 2009 and 2013, as well as detailed hospital network data for each plan, from publicly available GIC enrollee documents. I construct physician practice networks through a combination of linking each physician to his or her hospital or health system owner via the SK&A; verification of each practice's network status using each insurers' reported networks on their websites; and inferring network status using frequencies of in-network claims from the APCD. Details of the network construction are found in Appendix A.

2.3 Empirical Patterns

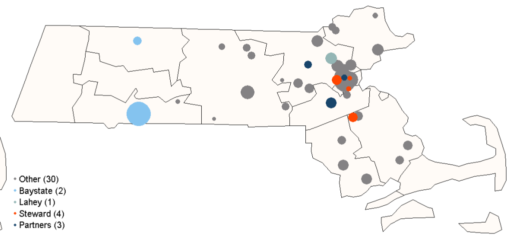
Variation in Physician Networks: Figure 1 and Figure 2 show the hospital and PCP networks for a select group of products available on the GIC in 2011: Harvard Broad, Harvard Narrow, and Tufts Narrow. The colors of the points on the maps refer to physician practices that are owned by the largest health systems in Massachusetts: Partners, Steward, Atrius, UMass, Lahey, and Baystate, with the gray points aggregating all other practices. The sizes of the points are in proportion to total market share of the practice for the particular physician specialty. Looking at primary care practices, it is clear that Partners (navy blue) and Atrius Health (red) dominate

Figure 1: Hospital Networks by Plan, 2011

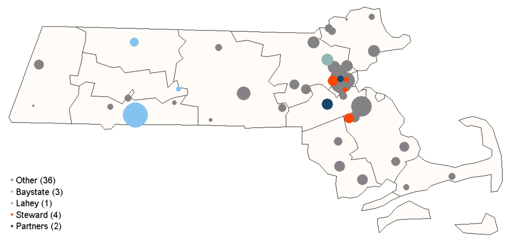
(a) Harvard / Tufts Broad



(b) Harvard Narrow



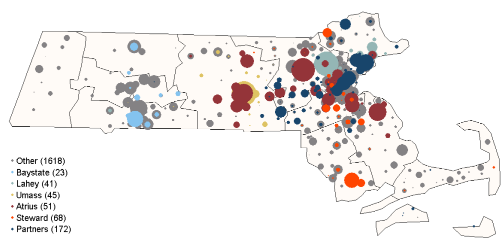
(c) Tufts Narrow



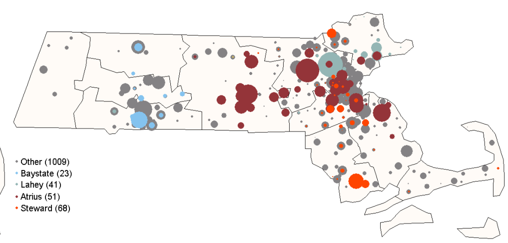
Notes: This figure plots the hospital networks of specified plans on the GIC in 2011. Sizes of the data points reflect relative market shares of the practices. Colors reflect ownership status (which health systems owns which practice).

Figure 2: Primary Care Practice Networks by Plan, 2011

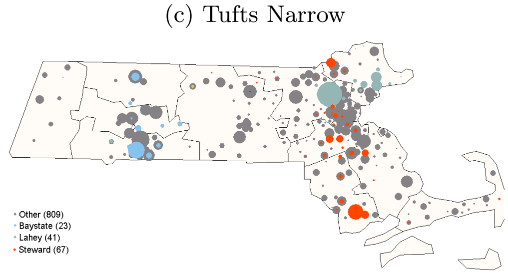
(a) Harvard / Tufts Broad



(b) Harvard Narrow



(c) Tufts Narrow



Notes: This figure plots the physician practice networks of specified plans on the GIC in 2011. Sizes of the data points reflect relative market shares of the practices. Colors reflect ownership status (which health systems owns which practice).

much of the primary care physicians in Massachusetts, with Partners owning 172 practices and Atrius owning approximately 51.[13] Panel (a) of Figure 2 shows that these practices are largely concentrated in eastern Massachusetts, particularly around Boston and the surrounding suburbs. However, Atrius Health also owns practices in central Massachusetts.[14].

Panels (b) and (c) of Figure 1 and Figure 2 reveal that the Harvard Narrow and Tufts Narrow still cover a large number of hospitals and physicians in Massachusetts. Interestingly, the hospital networks of both narrow plans are relatively similar. The only major difference between the broad and narrow hospital networks is that most Partners hospitals were dropped from each narrow plan. However, as noted in Table 1, Harvard's narrow plan has a significantly higher market share than the Tufts narrow network, with almost three times the number of enrollees in 2012. Given that Tufts covers a larger number of hospitals, it is therefore unlikely that hospital networks explain this discrepancy in market shares.

Turning to physician networks, however, provides more clues that help to explain employee plan choices. Figure 2 reveals that the Harvard's narrow physician network is considerably more comprehensive than the Tufts' narrow physician network. This is largely due to the fact that Harvard, but not Tufts, covers Atrius Health (noted by the red points in the map). This indicates that physician networks may be an important determinant of plan choice. Moreover, given that Partners physicians were primarily located in the Boston metro area, which also faces competition from Atrius, Lahey, Care Group, as well as many independent and solo practitioners, its removal from the network has minimal impact for choice of provider (a point I return to later).

Appendix B shows additional figures that depict the variation in hospital and physician networks across plans, over time, and across rating regions in which each plan operates.

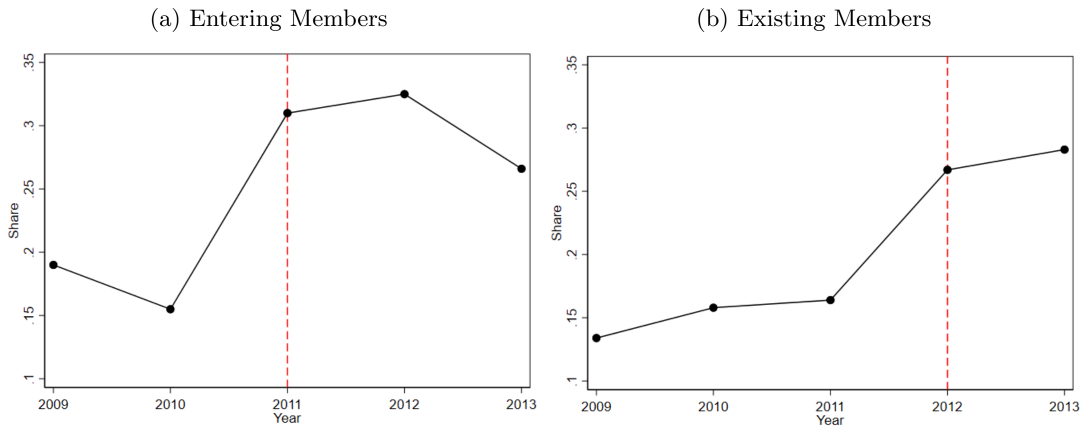
Variation in Narrow-Network Plan Enrollment: There is significant heterogeneity in terms of who is enrolling in narrow-network plans. However, one significant predictor of narrow-network enrollment is whether the household was previously employed by the group. Figure 3 depicts the share of GIC consumers enrolling in narrow-network plans by year and by whether they were new to the GIC that year (i.e. "entering members" or "active choosers") or whether they were existing GIC members who were automatically re-enrolled in their current plan unless they took action (i.e. "existing members" or "passive choosers"). Panel (a) depicts the share of new members enrolling in narrow-network plans, whereas panel (b) depicts the share of existing members. In 2009 and 2010, the share of enrollment of active choosers and passive choosers in narrow-network plans both hovered around 15%. However, there is a large spike in the share of new members enrolling in narrow-network plans in 2011 (to 30%), when the GIC introduced Harvard Narrow and Tufts Narrow. Conversely, once the GIC introduced the "premium holiday" in 2012, there is a significant spike in the share of *existing* members (for whom the policy applied, panel (b)) enrolling in narrow-network plans.

These differences in behavior between active and passive choosers suggest the presence of a

[13] In general, Partners is widely known to be one of the dominant players in the Massachusetts provider market, owning several large academic medical centers, including Mass. General Hospital and Brigham and Womens Hospital. Similarly, Atrius Health is one of the dominant players in the physician market, owning several key medical groups, including Harvard Vanguard.

[14] This is due to its purchase of the Fallon Clinic, later renamed "Reliant Medical Group" in Worcester in 2011

Figure 3: Share of People in Narrow Network Plans by Year and Whether New to GIC



Notes: This figure plots the share of members selecting narrow-network plans. Panel (a) plots the share of new members to the GIC enrolling in narrow-network plans. The dashed red line represents the year when the GIC introduced two new narrow-network plans from Harvard and Tufts. Panel (b) plots the share of existing members enrolling in narrow-network plans. The dashed red line represents the year of the "premium holiday."

high degree of consumer inertia in choosing health plans (Handel, 2013). One potential criticism of this conclusion is that active choosers might have different preferences for networks than passive choosers, or might otherwise be fundamentally different in ways that drive their choice of health plans. For instance, new employees of firms tend to be younger, and younger individuals tend to be more price sensitive in choosing health insurance plans than older families.

However, institutional details would suggest that these demographic differences between active and passive choosers is fairly minimal. First, most new members to the GIC are from new municipalities in Massachusetts contracting with the GIC for health plans, rather than new employees entering the firm. These municipalities tend to be geographically dispersed across the state, and the cohort of new workers entering from these municipalities tend to have similar observables to those already on the GIC. I present evidence for this in subsection B.2. In the appendix, I also show evidence that new consumers react to changes in plan emphpremium changes over time, whereas existing enrollees tend to remain inertial to plans even as co-premiums change significantly.

Taken together, these figures provide some suggestive evidence of two behaviors. The first is that consumers may exhibit a high degree of inertia after their initial choice of health plans, with existing members sticking to their choices, even as premiums grow relative to other similar plans on the market or as new options appear that are considerably cheaper. Second, once taking this inertia into account, consumers may actually be quite price sensitive in their choices of insurance plans, which is a characteristic often not attributed to purchasers of employer-sponsored insurance. These two stylized facts motivate my inclusion of inertia in the model I present in the next section.

3 Model

The model proceeds in four stages. A brief summary of these stages is as follows:

1. Employers select a number of products to offer to their enrollees and the network design of the plan. In selecting these plans, employers incur a fixed cost of adding each additional product.

2. Given the products selected, employers set premiums for self-insured products. Insurers set premiums for fully-insured products.

3. Consumers in each market select from the menu of insurance plans given their network breadth and composition, premiums, and various quality characteristics.

4. Consumers face some probability of contracting an illness, and based on that illness, along with individual and provider characteristics, patients select a hospital or doctor from one among their chosen insurance plan's network.

I now describe the model in detail from the latest stage through the earliest stage.

3.1 Patient Demand for Providers

The final stage of the model involves patient i enrolled in insurance plan j choosing a provider. The patient either has a condition that requires hospital care, l, in which case he or she chooses a hospital h from among the set of hospitals in insurance network N_{jt}^{H}, or the patient requires procedure r from specialist type s, in which case he or she chooses physician practice d among a set of practices within that specialty within the plan's network N_{jt}^{S}. Consumer utility for patient of type i, with either illness l or procedure r, from visiting a provider takes the following form:

$$u_{ilht} = \underbrace{T_{iht}\lambda_1 + T_{iht}v_{ilt}\lambda_2 + T_{iht}x_{ht}\lambda_3 + x_{ht}v_{ilt}\lambda_4 + \mathbb{1}(ih_t = ih_{t-1})\lambda_5 + \gamma_h}_{\phi_{ilht}\text{ (Hospitals)}} + \varepsilon_{ilht} \tag{1}$$

$$u_{irdt}^{s} = \underbrace{T_{idt}^{s}\lambda_1^{s} + T_{idt}^{s}v_{irt}\lambda_2^{s} + T_{idt}^{s}x_{dt}^{s}\lambda_3^{s} + x_{dt}^{s}v_{irt}\lambda_4^{s} + \mathbb{1}(id_t^{s} = id_{t-1}^{s})\lambda_5^{s} + \gamma_d^{s}}_{\phi_{irdt}^{s}\text{ (Physician Specialty } s)} + \varepsilon_{irdt}^{s} \tag{2}$$

where x_{ht} is a vector of observed hospital characteristics, x_{dt}^{s} is a vector of observed physician practice characteristics for specialty type s, v_{itl} and v_{irt} are observed characteristics of patient i with diagnosis l or requiring procedure r, T_{idt}^{s} and T_{iht} is the distance in miles from patient i's location to provider d or h's location, γ_d^{s} and γ_h are provider fixed effects, and ε are Type 1 Extreme Value error terms. Finally, $\mathbb{1}(ih_t = ih_{t-1})$ refers to whether patient i has used hospital h in any year prior to t, and $\mathbb{1}(id_t^{s} = id_{t-1}^{s})$ refers to whether individual i saw physician practice d for specialty care s in any year prior to t. The latter parameter represents inertia to previously used physicians. Details of the provider demand specification, as well as its estimation, identification, and parameter estimates are presented in subsection C.1.

3.2 Consumer Demand for Insurance Plans

I assume that choice of health plan is done at the household level. Therefore, the utility of household I for plan j at time t is given by the following:

$$u_{Ijt} = \underbrace{-r_{Ijt}\alpha_I + EU^H_{Ijt}\beta_1 + \sum_s EU^s_{Ijt}\beta^s_{2I} + \mathbb{1}(Ij_t = Ij_{t-1})\beta_3 + \eta_j}_{\delta_{Ijt}} + \omega_{Ijt} \tag{3}$$

Here, r_{Ijt} refers to the plan rate, or household's' co-premium, which varies only by whether the consumer has purchased individual coverage or family coverage. I allow the premium coefficient, α_I, to vary by age of the oldest member of the household as well as the household's geographic rating region. EU^H_{Ijt} is the expected utility from the plan's hospital network and EU^s_{Ijt} is the expected utility from the plan's network of physician specialty s. They measure household I's willingness-to-pay for a particular insurance plan's provider network, incorporating not just network size, but relative quality of the providers in the network as determined by the provider demand stage.[15] Details of their construction are in subsection C.5. η_j is the unobserved plan characteristics component, captured by a full set of plan fixed effects, reflecting the fact that plan demand may be driven by preferences for a particular plan unobserved by the econometrician, and ω_{Ijt} is the idiosyncratic, Type 1 Extreme Value error. Plan switching costs are captured by $\mathbb{1}(Ij_t = Ij_{t-1})$, which is an indicator function for whether household I was enrolled in plan j in year $t-1$.

The model allows preferences for physician networks, EU^s_{IJt}, to be a function of both observed and unobserved heterogeneity. The observed heterogeneity stems from demographic factors that determine an individual's preference for physicians from subsection 3.1 (i.e. age, diagnosis, location, etc.). In addition, I allow the coefficients, β^s_2, to vary by household geographic rating region. Unobserved heterogeneity comes from random coefficients on each of the utility terms. These capture heterogeneity for physician networks that may explain persistence in plan choice. For instance, certain households may be more risk-averse than others, conditional on age and location, and as such would prefer to remain on a broader network, even while an identical household with similar demographics might be more inclined to switch to a narrower plan with a lower premium. These random coefficients serve the dual purpose of (a) better predicting switching behavior in the face of a change in plan menus and (b) helping ensure that persistence in plan choice is not misattributed to plan switching costs.[16]

The coefficients for network utility are therefore specified as:

$$\beta^s_{2I} = \beta^s_2 + v^s_I \tag{4}$$

where $v^s_I\ N(0, \sigma^s)$. Here, β^s_2 represents the mean network valuation for specialty s and σ^s is the standard deviation of that network valuation across households.

The market share of households of type I for plan j in market t is derived as the familiar logit

[15] A network may, for instance, have fewer providers, and yet still yield a higher value of EU^s_{Ijt} for specialty group s is the physicians included are of higher demand than the larger network.

[16] Indeed, previous literature has shown that not accounting for unobserved preference heterogeneity results in considerable bias in estimates of switching costs. Similar approaches have been taken in Polyakova (2016) and Ho et al. (2017).

share, integrated over the distribution of β across households::

$$s_{Ijt} = \int \frac{exp(\delta_{Ijt})}{\sum_{k=1}^{J} exp(\delta_{Ikt})} F(\beta) d\beta \tag{5}$$

Estimation: Given that the share equation in Equation 5 is integrated over a distribution over the disturbances in β, the shares have no analytic, closed-form solution. Therefore, the model is estimated using maximum simulated likelihood, as in Train (2009), on the years 2009-2013. Additional details are presented in subsection C.5.

Identification: The mean expected utility coefficients, β_1 and β_2, are identified from within-plan variation in utility of provider networks across individuals. These differences in expected utility stem from differences in household ages, locations (i.e. households that live closer to more prestigious doctors and hospitals than others), and illness histories (i.e. individuals with a higher disease burden).

The premium coefficient is identified through within-plan variation in co-premiums generated by differences in family type. For households with only one member, individuals pay a base co-premium, and for households with more than one member, the household pays a total of 2.4 times the base co-premium, while expected utility from the provider network is linear in the number of household members.[17]

Identification of the inertia parameter, β_3, relies on two conditions to be true. The first is that choices made by "active choosers" need to be different from choices made from "passive choosers." The second condition is that the choice set changes over the sample period. In my setting, several features allow for clean identification of switching costs and separation of these costs from unobserved preference heterogeneity. First, throughout my sample period, I observe a panel of households making consecutive choices over time as the choice set changes. As is standard in the literature, observing these choices along with the inclusion of random coefficients is meant to capture household-specific persistence in preferences for broad networks that might inhibit switching, while the lagged plan choice variable is meant to capture the switching cost parameter.

Second, I observe a substantial number of enrollees making choices for the first time, driven by households from municipalities entering the GIC between 2009 and 2013. As such, the parameters describing unobserved persistent preferences can be estimated from the choices made from new enrollees alone. Further, as described in subsection 2.3, these municipal "active choosers," conditional on a rich set of observables, make extremely different choices in plans than members previously enrolled in a GIC. If inertia to previous plans was driven by preference heterogeneity, we would not expect such considerable differences between these two groups of enrollees.

[17] See Prager (2016) and Ho and Lee (2017) for a discussion of this identification strategy. Though there may be some concern that base co-premiums are set endogenously, which might bias my coefficient, premiums in Massachusetts adhere to medical loss ratio laws, which require that plan premiums be set no higher than prespecified amounts by the state government. The GIC is also quite active in negotiating lower premiums with insurers, and has traditionally upheld a medical loss ratio of approximately 90% on all plans (Prager, 2016). Therefore, I take the plan premiums as effectively exogenous conditional on utilization of health care services and expected plan costs, both of which are captured by EU_{Ijt}, and controlling for unobserved plan characteristics that might be correlated with ω_{Ijt}.

Third, the "premium holiday" in 2012 forced all active state employees to re-enroll in a plan at the same time the GIC both introduced new plans into the choice set and significantly decreased the co-premiums for a subset of those plans (the effect of this holiday on enrollment is shown in Figure 3). Similar in spirit to Handel (2013), the primary identifying assumption is that, controlling for detailed ex-ante health risk as well as observed and unobserved preferences for networks, β_3 should identify "true" inertia (switching costs) rather than preference heterogeneity. Polyakova (2016) relies on similar identification techniques.

Estimates: Table 2 reports the results for the insurance plan demand model. Due to the high dimensionality of the data,[18] I only run the model on a subset of 5,000 households across the five years of data. As I cannot observe Unicare products in the data, I run each model on the set of enrollees in all other GIC plans. Omitted from the table are plan fixed effects, as well as premium and network utility interactions with observables (e.g. age and location).

The first four columns present results with no random coefficients on network utility. The first two columns present estimates without accounting for plan switching costs, while columns 3-4 present estimates that account for switching costs by estimating a parameter on an indicator for whether a household was previously enrolled in a particular plan. Within each set, the first column (i.e. columns 1 and 3) present results only focusing on hospital utility (EU^H_{Ij}), while the second column present results with hospital and physician networks of all three specialties.

Panel A reports the estimated parameters. The monthly premium parameter, α_I, is negative and significant across all six specifications, suggesting that households are averse to paying higher co-premiums for health insurance. The coefficients on expected utility are also positive and significant across all of the models. Overall, the results indicate that households have a positive valuation of plan networks, consistent with prior literature.

There is, however, significant heterogeneity in network preferences across households and provider types. Looking at the models with no plan switching costs (models 1-2), it is clear that moving from a model with only hospital networks to a model that includes physician networks has a significant effect on the estimated parameters, particularly premium elasticities. In particular, including physician networks nearly triples the estimated premium disutility, while significantly reducing the estimated coefficient on hospital utility.

The effect of including physician networks can be more clearly seen in panel B, which translates the estimated parameters to dollarized "willingness-to-pay" (WTP) amounts for networks. These estimates report what single-member (individual) households on Harvard Broad would need to be paid to have their network reduced to that of Harvard Narrow. In a model with only hospital networks, individuals would need to be paid approximately \$247 per month to have their network reduced.[19] Moving to a model that includes physician networks, however, drastically reduces the implied WTP for hospital networks to merely \$21 per month, while yielding an implied WTP for physician networks of \$55 per month. Decomposing this valuation of physician networks, approximately \$21 per month comes from WTP for PCPs, \$20 per month from WTP for cardiologists,

[18]There are approximately 200,000 GIC members per year multiplied by about 70 hospitals, 18 potential diagnoses, 50 practices in seven rating regions, and three different specialties groups.

[19]Note that this figure is considerably higher than the actual premium differential between Harvard Broad and Harvard Narrow, which averaged approximately \$30 across the five-year period.

Table 2: Results of Plan Demand Models

Variable	(1)	(2)	(3)	(4)	(5)
	Panel A: Estimated Parameters				
	No Switching Costs		Switching Costs		Random Coef.
Prem (PM)	-0.0017**	-0.0048***	-0.0207***	-0.0235***	-0.0242***
	(0.0008)	(0.0008)	(0.0017)	(0.0017)	(0.0017)
EU^{H}_{Ijt}	7.7342***	1.9446***	5.8190***	1.9317**	2.1736**
	(0.6273)	(0.5097)	(0.8661)	(0.8618)	(0.8982)
EU^{PCP}_{Ijt}		0.1668***		0.0988***	0.2176***
		(0.0219)		(0.0253)	(0.057)
EU^{CAR}_{Ijt}		0.5073***		0.5190***	0.7776***
		(0.0843)		(0.1231)	(0.2139)
EU^{ORS}_{Ijt}		0.9910***		0.3348**	1.0688***
		(0.1236)		(0.1659)	(0.3204)
σ_{PCP}					0.1655***
					(0.0388)
σ_{CAR}					0.5188***
					(0.1698)
σ_{ORS}					0.9581***
					(0.2367)
Prior Plan			4.9352***	4.8906***	4.9242***
			(0.0883)	(0.0898)	(0.0936)
Plan FE	Yes	Yes	Yes	Yes	Yes
Obs.	41,673	41,673	41,673	41,673	41,673
Pseudo R^2	0.29	0.32	0.79	0.80	–
	Panel B: Willingness-to-Pay for Harvard Broad v. Harvard Narrow				
	No Switching Costs		Switching Costs		Random Coef.
WTP Hosp	\$247	\$21	\$15	\$4	\$5
WTP PCP		\$21		\$3	\$5
WTP CAR		\$20		\$4	\$6
WTP ORS		\$15		\$1	\$3
Switching Cost			\$238	\$208	\$204

Notes: Columns 1-2 are results for models without the plan inertia coefficient. Columns 3-4 include these coefficients. Column 5 reports results from a model with random coefficients. EU^{H}_{Ijt} refers to the household's expected utility for the hospital network, EU^{PCP}_{Ijt} refers to the utility of the primary care network, EU^{CAR}_{Ijt} refers to the utility of the cardiology network, and EU^{ORS}_{Ijt} refers to the utility of the orthopedic network. σ_s refers to the estimated standard deviation on network utility for specialty s. "WTP" refers to "willingness-to-pay" for Harvard Pilgrim's broad hospital and physician networks relative to its narrow network. "Switching cost" refers to the estimated dollarized plan switching cost. The premium variable is reported in monthly terms. Omitted from the table are PCP and specialist copays for the highest tier, premium interactions with region, age, and income, as well as physician utility interactions with region.

and $15 from WTP for orthopedists.

Turning to models with plan switching costs included (models 3-4) again significantly increases the magnitudes of the premium elasticities, with further increases seen when including physician networks in addition to hospital networks. When only hospital networks are included in the model, the estimates imply that individuals would need to be paid an average of $15 per month to move from Harvard Broad to Harvard Narrow. The estimated switching cost in this model is $238 per month. When physicians are included in the model, the WTP for hospital utility decreases to a mere $4 per month, while the estimated average WTP for the physician network differential is about $8 per month. Including physicians in the model decreases the estimated switching cost declines by about $30 to $208 per month.[20]

Finally, column 5 adds random coefficients to each of the physician network utility measures. To the extent that persistence in broad-network enrollment, even as narrow networks become available, is driven by unobserved preferences for large networks of physicians (rather than true switching costs), these coefficients should capture this behavior. Indeed, the standard deviations, σ, on all three specialists is large and significant, suggesting there is considerable variation in unobserved preferences. For instance, the standard deviation on utility for PCPs is estimated to be 0.17, relative to the estimated mean of 0.22. This has the predictable effect of increasing average WTP to move from a narrow to broad plan. In the model with random coefficients, individuals would need to be paid, on average, $5 per month more to move from Harvard Broad's PCP network to Harvard Narrow's physician network (compared with $3 per month in a model without random coefficients). WTP for cardiology and orthopedic networks also increased by $2 per month more relative to the coefficients in column 4. In aggregate, including unobserved heterogeneity in the model increased the WTP for Harvard Broad v. Harvard Narrow by about 60% (a $7 per month increase). The switching cost estimate, however, only declines by $4 per month to $204, suggesting that including unobserved preference heterogeneity only had a marginal effect on the estimated inertia parameter.[21]

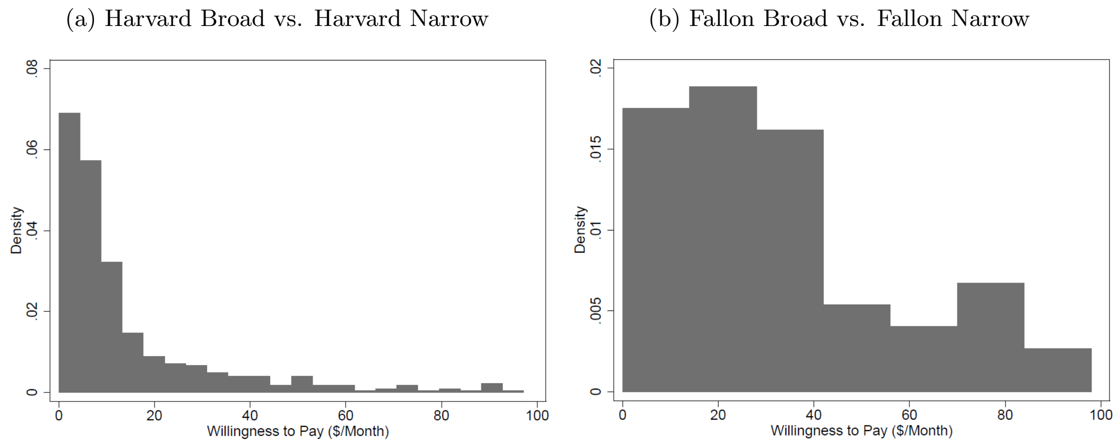
The WTP estimates in Table 2 are averages and only reported for the network differential of Harvard Broad and Harvard Narrow. To show the heterogeneity in WTP across consumers and different network types, I plot the distributions of WTP across households for Harvard Broad versus Harvard Narrow and for Fallon Broad versus Fallon Narrow derived from model 5 (the preferred specification). These distributions are reported in Figure 4. Two conclusions emerge from this figure. First, although the mean reported values for Harvard reported in Table 2 were around $19 per month, there is clearly significant heterogeneity, with certain households willing to pay nearly $100 per month for access to the broader network. Second, the overall WTP for Fallon's broad versus narrow network is larger than Harvard. This makes sense given that the difference in the networks is more substantial between Fallon plans.

These results suggest that heterogeneity for preferences in physician networks is an important determinant of plan choice. Ignoring physician networks in models of insurance demand yield

[20]Though high, these estimates are in range of prior work. In particular, Polyakova (2016) finds switching frictions in Medicare part D to be about twice to four times as large as premiums. A switching friction of $208 per month is approximately 1.58 times the individual premium for a broad network and approximately 63% of the family premium for the same network.

[21]This also suggests that the remainder of the new WTP were loaded onto plan fixed effects in the previous models

Figure 4: Willingness-to-Pay for Broad Versus Narrow Networks



Notes: This figure plots the distribution of willingness-to-pay across households for various networks. Panel (a) reports willingness-to-pay for Harvard Broad versus Harvard Narrow. Panel (b) reports willingness-to-pay for Fallon Broad versus Fallon Narrow. Estimates are in per-household-per-month dollars.

premium elasticities that are likely underestimated. This has the effect of making households on employer markets appear less price sensitive than they are. In fact, the results suggest that households tend to select into broad-network plans not necessarily because they are price-insensitive, but because of a combination of (a) fairly high valuations for physician networks, particularly those that include physicians with whom they have formed relationships, and (b) a high degree of plan choice inertia. This conforms with patterns shown in subsection 2.3.

3.3 Premium Setting

Consistent with prior literature, I assume that insurers' and employers' health care cost reimbursements to a particular provider can be decomposed into an insurer-provider-specific base negotiated rate, p_{jht} and p_{jdt}, scaled by a disease or procedure weight. Let the insurer/employer's marginal cost of *hospital* care to cover enrollees on plan j therefore be given by:

$$c^o_{jtH}(N^H_{jt}) = \sum_{i \in I} \sum_{l} f_{il} w_{lt} \sum_{h \in N^H_{jt}} \sigma_{ilht}(N^H_{jt}) p_{jht} \tag{6}$$

Here, f_{il} refers to the ex-ante probability that a type i individual contracts diagnosis l, while w_{lt} is the disease weight for that diagnosis and N^H_{jt} refers to the hospital network of plan j. Let the marginal cost of *physician* care to cover enrollees plan j be given by:

$$c^o_{jtS}(N^S_{jt}) = \sum_{i \in I} \sum_{s} \sum_{r} f_{ir} RVU_{rt} \sum_{d \in N^s_{jt}} \sigma^s_{irdt}(N^s_{jt}) p^s_{jdt} \tag{7}$$

where f_{ir} is the ex-ante probability that a type i individual needs procedure r, RVU_{rt} refers to the RVU weight assigned to a particular physician procedure, and N^s_{jt} is the physician network of plan j. Given these cost specifications, insurer m's profits are given by:

$$\pi_{mt} = \sum_{j \in J_m} \sum_{I} \left(s_{Ijt}(\delta_{Jt}, \theta) \left[R_{jt}(\delta_{Jt}, \theta)\theta_I^R - \underbrace{c_{jtH}^o(N_{jt}^H) - c_{jtS}^o(N_{jt}^S)}_{c_{jt}^o(N_{jt})} - c_{jt}^u(N_{jt})\theta_I^c \right] \right) \tag{8}$$

In the equation above, J_m refers to the set of products offered by MCO m and N_{jt} refers to the overall network of plan j in time t. R_{jt} refers to the *total* (employee+employer) "base" premium for each plan in each year. These base premiums scale by household type (individual vs. family) by a coefficient θ_I^R, which is 2.4 times the base premium if the household is a family (regardless of household size). Finally, c_{jt}^u refers to to a base "unobserved cost" of health care for plan j in time t.[22] I assume that these costs scale linearly across household type, i.e.:

$$c_{Ijt}^u = c_{jt}^u \theta_I^c$$

where θ_I^c is the parameter that scales these base unobserved costs across households.

In ordinary settings, one can take the first-order condition of Equation 8 and assume insurers set premiums according to a multi-product Nash-Bertrand function. However, for several reasons, such assumptions are fairly strong for this setting. First, two of the largest plans offered by the GIC (Harvard Broad and Tufts Broad) are on self-insured arrangements, and as such, the marginal health care costs given by Equation 6 and Equation 7 for enrollees on these plans are reimbursed by the employer directly.[23] Second, the GIC, as a large employer group that covers about 8-9% of the state's commercially-insured enrollees, has considerable bargaining leverage with insurers to reduce premiums, thereby inhibiting insurers on fully-insured contracts from setting markups that are too high.[24] Finally, plans in Massachusetts are bound by state medical-loss-ratio (MLR) regulation requiring that plans spend no less than 85% of premium dollars on medical care expenses. For these reasons, plans on the GIC are observed to set premiums, on average, at about 10% over their medical expenditures (Prager, 2016).

Therefore, as my primary pricing assumption, I assume that the employer/insurers set premiums for each plan at a fixed 10% markup over marginal health care costs. The pricing equation then becomes:

$$\sum_{I} s_{Ijt} R_{Ijt} \theta_I^R = 1.10(c_{jt}^o(N_{jt}) + c_{jt}^u(N_{jt})\theta_I^c) \tag{9}$$

Details of the construction of p_{jht} and p_{jdt}^s, as well as for estimating unobserved marginal costs, c_{Ijt}^u are presented in subsection C.6.

[22] These costs include physician specialties not modeled in this paper, pharmaceutical spending, etc.

[23] Industry experts note that, for these plans, insurers offer the GIC a "suggested" premium based on anticipated costs, but that GIC is ultimately free to set co-premiums for consumers at their discretion. As a result, the GIC has incentives to keep premiums low.

[24] An industry expert noted that insurance plans gain considerably from contracting with the GIC and, as such, are largely willing to capitulate to the GIC's requests for premiums and plan designs. See Ho and Lee (2017) for a model that incorporates employer-insurer bargaining over premiums using data from CalPers (an employer group similar to the GIC) in California.

3.4 Employer Objective Function and Network Design

I assume that the employer, in selecting product quality and setting prices, maximizes a weighted measure of consumer surplus from the chosen plans less the amount paid out in either medical expenditures (in the case of self-insured products) or premiums to insurers (in the case of fully-insured products). The consumer surplus measure is meant to capture the fact that employers care about satisfying the health care needs of their employees. A product menu that can more closely match the needs of its employees would allow the employer to retain employees for longer periods of time, as well as attract new enrollees from other firms. This implies that the more heterogeneous a firm's employees (or potential employees) are in terms of demographics, geography, and health preferences, the more employers should be willing to expand their product menu in order to accommodate the needs of the diverse employee preferences.

On the other hand, offering plans that are more generous (i.e. broader network) means that the firm pays out more in premiums, due to the presence of high-cost providers in the network. Moreover, offering multiple plans is costly for firms. I therefore assume that the employer's plan choices are subject to a fixed cost for each additional product chosen. These costs can reflect tangible, monetary expenses, such as the fact that offering multiple plans means that employers need to bear the additional expenses of designing the products, informing consumers, collecting and setting premiums, and negotiating with insurers (Bundorf, 2002; Moran et al., 2001).[25] However, they also include non-monetary opportunity or switching costs. For instance, fixed costs may reflect employers' belief that offering more plans might contribute to consumer confusion or suboptimal plan choice among employees.[26]

Formally the employer objective is:

$$W_t = \underbrace{\rho CS(\delta_{Jt}, \theta)}_{\text{Weighed Consumer Surplus}} - \underbrace{\sum_I \sum_j (1-\tau) s_{Ijt}(\delta_{Jt}, \theta) R_{Ijt}(\delta_{Jt}, \theta)}_{\text{Net Health Spending}} - \underbrace{\sum_j FC_j}_{\text{Fixed Costs}} \tag{10}$$

where:

$$CS(\delta_{Jt}, \theta) = \sum_I \frac{1}{\alpha_I} log\left(\sum_j^J exp(\delta_{Ijt})\right)$$

The term on the left-hand-side of the function, $CS(\delta_{Jt}, \theta)$ is the consumer surplus from the employer offering J products to its employees. This consumer surplus is a function of estimated demand parameters, θ, and the employer's chosen plan menu, δ_{Jt}. R_{Ijt} refers to the *total* premium (i.e. the enrollee plus the employer share). The term τ represents the mapping from the full premium to the employee co-premium, i.e. the percentage of premium that is to be paid by the enrollee.[27] The second term in the equation represents the payment in premiums to insurers the

[25] Bundorf (2002) notes that firms report that these costs inhibit them from offering more choice and variety to their consumers.

[26] In particular, recent research has shown that consumers facing a large number of choices often feel overwhelmed, resulting in the choice of "dominated" plans that are financially inferior to other options (Liu and Sydnor, 2018). In my setting, I abstract from distinguishing between these two types of costs.

[27] During the years of my sample period, the GIC set its enrollee share for employees hired prior to 2003 as 20%, while those hired after 2003 at 25%.

employer contracts with. Note that for self-insured plans, this term would reflect the cost of medical claims incurred by the employer plus administrative expenses. The third term, FC_j represents the fixed cost to the employer of offering plan j to its enrollees. It is a parameter to be estimated.

Finally, ρ (hereafter referred to as the "mismatch parameter") refers to the relative weight that the employer places on the sum of its employees' consumer surplus over total dollars spent on premiums (or medical claims) and fixed costs. This is the *key* parameter of the model and that one I use to determine the extent to which employer and employee preferences for network breadth are distorted. The intuition is as follows: if employers have the ability to pass-through shocks to health care prices through the reduction of other benefits, then the employer ought to value consumer surplus equally to premium spending, and therefore we would expect $\rho = 1$.

Consider, for example, a scenario in which, by narrowing the network of an existing product, the employer could induce savings in excess of the total utility loss from the change. The employer could then theoretically compensate the employees for the loss through other benefits or lower-co-premiums, while still achieving social welfare gains through the cost savings. If the employer valued surplus equally to premium spending, it would then narrow its network in such a way. If, however, the employer did *not* make such a move, it would be suggestive that the employer, perhaps, valued the dollar consumer surplus from *the network* by more than a dollar from lower co-premiums, higher wages, or other benefits. Such plan designs would, therefore, be indicative of a mismatch between employer and employee preferences to the degree that ρ deviated from 1.

To estimate ρ and FC_j, I assume that the employer chooses plan menu δ_{Jt} out of a feasible set of products to maximize its expected surplus, where the expectation reflects uncertainty over potential enrollees and demand preferences:

$$\max_{\delta_{Jt}} \left[E\left(\underbrace{\rho CS(\delta_{Jt}, \theta) - \sum_I (1-\tau) s_{Ijt}(\delta_{Jt}, \theta) R_{Ijt}(\delta_{Jt}, \theta)}_{S_t(\delta_{Jt}, \theta)} \right) - \sum_j FC_j \right] \tag{11}$$

where $S_t(\delta_{Jt}, \theta)$ refers to the *marginal* social surplus from having product menu J.

Estimation: To estimate the employer-employee mismatch parameter and fixed costs, I closely follow work by Ho (2009), Pakes et al. (2015), and Pakes (2010) in constructing moment inequalities to bound the estimates of ρ and FC_j, rather than imposing an equilibrium through distributional assumptions on the parameters.[28] Such moment inequalities approaches were subsequently used to estimate fixed and sunk costs of product introductions in markets such as computers, pharmaceuticals, and smartphones (Eizenberg, 2014; Nosko, 2014; Mohapatra and Chatterjee, 2015; Fan and Yang, 2020).The critical identifying assumption underlying the moment inequality approach is that the employer's expected surplus in offering a particular set of plans with particular networks is greater than any alternate set of plans and networks it could have chosen at a particular time.

[28] An alternate estimator would be to specify a multinomial logit model, similar to the provider and plan demand models, with a logit error shock. However, this is not well-suited for the context of GIC network decisions. In particular, the choice set of possible plans to offer, hospital networks of those plans, and physician networks of those plans is so large that assigning a logit shock to each *potential* choice is likely to produce unreliable, biased estimates.

Formally, let the expectation of the employer from offering a particular plan menu, δ_{Jt}, conditional on information set, $\mathcal{J}$ be given by:

$$E\left[W_t(\delta_{Jt},\theta)|\mathcal{J}\right] = E\left[S_t(\delta_{Jt},\theta|\mathcal{J}) - \sum_j FC_j\right] \tag{12}$$

where S_t is the marginal social surplus at time t as defined in Equation 11. Let δ^a_{Jt} be alternate plan menu offered in time t and $E\left[\Delta W_t(\delta_{Jt},\delta^a_{Jt},\theta)|\mathcal{J}\right]$ be the expected change in surplus of the employer from offering δ_{Jt} relative to δ^a_{Jt}. Then, to satisfy the identifying assumptions, it must follow that:

$$E\left[\Delta W_t(\delta_{Jt},\delta^a_{Jt},\theta)|\mathcal{J}\right] = E\left[W_t(\delta_{Jt},\theta)|\mathcal{J}\right] - E\left[W_t(\delta^a_{Jt},\theta)|\mathcal{J}\right] \geq 0 \tag{13}$$

Let $v_{1,\delta_{Jt}}$ be the difference between the employer's realized surplus and expected surplus such that:

$$v_{1,\delta_{Jt}} = W_t(\delta_{Jt},\theta) - E\left[W_t(\delta_{Jt},\theta)|\mathcal{J}\right] \tag{14}$$

It follows that:

$$E\left[\Delta W_t(\delta_{Jt},\delta^a_{Jt},\theta)|\mathcal{J}\right] = W_t(\delta_{Jt},\theta) - W_t(\delta^a_{Jt},\theta) \underbrace{-v_{1,\delta_{Jt}} + v_{1,\delta^a_{Jt}}}_{v_{1,\delta_{Jt},\delta^a_{Jt}}} \geq 0 \tag{15}$$

Assuming that $E[v_{1,\delta^k_{Jt}}|\mathcal{J}] = 0 \ \forall \ k$, considering an instrument set $z \in \mathcal{J}$, and taking sample averages, this becomes:

$$m(\delta_J,\delta^a_J,\theta,z) = \frac{1}{T}\sum_t \left[\left(W_t(\delta_{Jt},\theta) - W_t(\delta^a_{Jt},\theta)\right) \otimes g(z)\right] \geq 0 \tag{16}$$

where $g(z)$ is any positive function of instruments z.

For simplicity, the equations above omitted any unobserved heterogeneity in demand and, therefore, the moment above is for a fixed set of θ. To properly account for the presence of unobserved heterogeneity in estimating the employer objective function, I simulate 10 distributions of β^s_{2I} from Equation 4 (in effect letting $\theta = \theta_s$ for simulation s), construct a separate set of moments for each simulation, and take the mean across simulations. Formally, the unconditional moment then becomes:

$$m(\delta_J,\delta^a_J,\theta,z) = \sum_{s=1}^{10}\left(\frac{1}{T}\sum_t \left[\left(W_t(\delta_{Jt},\theta_s) - W_t(\delta^a_{Jt},\theta_s)\right) \otimes g(z)\right]\right) \geq 0 \tag{17}$$

I search for any values of ρ and FC_j that satisfy Equation 17. If no values satisfy all the inequalities, I find the values that minimize the squared deviations for all inequalities which were violated. More specifically, let:

$$\mathbf{Z} = -m(\delta_J, \delta_J^a, \theta, z) \qquad \text{if } m(\delta_J, \delta_J^a, \theta, z) < 0$$
$$\mathbf{Z} = 0 \qquad \text{if } m(\delta_J, \delta_J^a, \theta, z) \geq 0$$

I then estimate the equivalent of a GMM model where:

$$\underset{\rho, FC_j}{min}(\mathbf{Z'Z}) \tag{18}$$

As part of my estimation procedure, I assume the employer has expectations on the number of new employees to the firm each year as well as their demographic distributions. Additional details on this, as well as details on error assumptions and alternate estimators are in subsection C.7. One concern with the estimation procedure outlined above is that since I allow the employer to add and remove plans from the market, some of the surplus estimates, $CS(\delta_{Jt}, \theta)$, may be driven by the logit error shock, which may overestimates surplus gains or losses from product entry or exit. For robustness, I also report estimates in Appendix E in which I calculate $CS(\delta_{Jt}, \theta)$, ρ, and FC_j assuming that the logit shock is zero.

Restricting the Potential Choice Set: The combination of number of products offered and networks of those products, given the number of hospital and physicians in Massachusetts, are nearly infinite, thereby making estimation largely infeasible. I therefore make several restrictions on the choice set of the GIC for estimation. First, I assume that the GIC cannot cease contracting with any insurer, but can adjust the number of plans offered by any insurer.[29] Second, because the smaller insurers typically offer fixed, non-adjustable designs, I assume the GIC can alter only the *number* of plans offered by Fallon, but cannot alter the networks of its plans. Similarly, the GIC can only offer the sole plan by Health New England and Neighborhood Health Plan[30] Third, I assume that the GIC may freely adjust both the number of plans offered and the networks of Harvard Pilgrim and Tufts plans, but must limit the number of plans offered by either to four.[31] Finally, I assume that the GIC can offer Harvard and Tufts networks equivalent to the Fallon Direct network (hereafter "Very Narrow"), the Tufts Spirit network ("Narrow 1"), the Tufts Select network ("Narrow 2," sold primarily to small employers outside the GIC), the Harvard Primary Choice network ("Medium"), and the Harvard Independence network ("Broad").[32] In total, this leaves permutations of 14 potential products.

Identification: Identification of the employer-employee mismatch, ρ, comes from variation in

[29] The GIC engages in long-term, five-year contracts with insurers. During my sample period of 2009-2013, insurers were under their contract. Therefore, an assumption that the GIC could simply cease to offer any particular insurer would add a choice to the set that was not there in reality, thereby biasing my estimates of fixed costs and weight on consumer surplus.

[30] This is motivated by the fact that Fallon, Health New England, and Neighborhood Health Plan are all fully-insured products that also operate largely in the broader employer marketplace. Unlike Harvard Pilgrim and Tufts, which are self-insured products and offer GIC-specific network designs, the smaller insurers typically offer fixed, non-adjustable designs.

[31] Employers rarely offer more than two narrow-network designs from the same insurer.

[32] I restricted to these particular networks as (1) they were all observed to be offered in Massachusetts during my sample period and (2) they span a considerable range of network breadth, both in terms of hospitals and physicians.

the characteristics of the potential networks not chosen relative to the ones that were *conditional on the employer offering the same number of plans.* Intuitively, suppose the employer could broaden the network of one of the existing plans such that consumer surplus in Equation 11, $CS(\delta_{Jt}, \theta)$, increased. The fact that the employer did not choose to offer this potential network would imply that its weight on consumer surplus was low relative to the added expenditures broadening that network would bring, thus dampening ρ. Conversely, if narrowing an existing network reduced $CS(\delta_{Jt}, \theta)$ while lowering spending, the fact that the GIC did not do this would raise the value of ρ. Since this holds the number of plans constant, FC_j is not affected by these scenarios.

Identification of FC_j relies on the assumption that $FC_j = FC$; over-time variation in the number of products offered; and finally the variation in the potential surplus that could be achieved from offering additional products or reducing the number of products within a time period. Within period, if the employer could offer an additional plan, but did not, then the fixed cost of offering it must outweigh the surplus gained from its introduction. If the employer could have removed a product, but did not, it must be that the fixed costs are lower than the surplus gained from keeping that product. Over time, fixed costs are pinpointed by changes in the market driven by changes in the underlying provider costs, p_{jth} and p^s_{jtd}; changes in provider ownership structure (which ultimately change demand for providers); changes in the number of entering municipalities; and the risk profile of entering municipal households.

Estimates: Estimates for the employer objective function are presented in Table 3. Panel A reports the main specification: estimates of both ρ and FC_j from the full moment inequality approach. Panel B reports the approach that sets $\rho = 1$ and uses one-step deviations in plans to estimate bounds on FC_j.

Table 3: Results of Employer Objective Function Estimation

	ρ	Lower Bound FC_j	Upper Bound FC_j
Panel A: Estimating ρ and FC_j			
GIC/Employer ($Millions)	3.67	4.07	4.07
Percentage of Net Spending		0.42	0.42
Percentage of Net Surplus		0.91	0.91
Panel B: One-Step Deviations, GIC			
GIC/Employer ($Millions)	1	1.15	6.64
Percentage of Net Spending	1	0.12	0.70
Percentage of Net Surplus	1	0.26	1.50

Results from ρ and FC_j estimation for 2009-2013. Panel A reports the main specification, estimating both parameters using moment inequalities. ρ and FC_j are point estimates in this panel, therefore the "lower bound" and "upper bound" on FC_j are identical. Panel B reports the results only estimating FC_j using one-step deviations in plans. The corresponding percentages of fixed costs relative to net employer health spending and net employer marginal surplus (consumer utility minus health spending) are also reported. FC_j reported in millions of dollars.

The estimates from Panel A are presented as point estimates rather than bounds, as no set of parameters, ρ and FC_j, satisfied each of the inequalities presented in Equation 17.[33] As such,

[33] This is to be expected, given the large number of inequalities (the large number of potential plans and networks the GIC could offer in a given year).

the values reported in Panel A are the same in the "lower bound FC_j" column and the "upper bound FC_j" column. The estimate of the mismatch parameter, ρ, is 3.67, suggesting that the GIC places considerable weight on consumer surplus relative to net spending. This is indicative of a systematic mismatch between employer and employee preferences. Indeed, if the employer could flexibly adjust other components of benefits in response to health shocks, we ought to expect this parameter to be close to 1, as the employer could pass back savings from a move to narrow-network plans onto consumers. Instead, the employer appears to prefer offering a plan menu that is more generous in terms of its network configuration than the average employee would prefer. I explore the implications of this estimate as well as potential mechanisms leading to this distortion in section 4.

The point estimate of fixed costs is \$4.07 million for each plan. Though this estimate appears quite high, it is actually a very small fraction of the GIC's overall net spending.[34] To test the sensitivities of the fixed cost estimates, Panel B, using merely one-step deviations in plans and setting $\rho = 1$, suggests that the GIC spends between \$1.15 and \$6.64 million a year on fixed costs for each plan offered. Although these estimates, in theory, contain both tangible and non-tangible fixed costs, as described above, the magnitudes are actually in line with reported administrative costs estimates by insurers in Massachusetts.[35]

Three caveats should be noted regarding these estimates, particularly FC_j. The first is that although the lower bound has a fairly large sample due to the wide availability of various product networks in Massachusetts, fairly few of these networks were offered during my sample. Therefore, the upper bound estimates have a very small sample size. Second, and related, the low rate of offered products in the GIC may be driving up the estimates. Since there are only 8 products offered in a given year in the GIC, any particular product removed, if it has a large enough market share, would cause a large decrease in consumer utility, which when averaged over a small sample, may bias the estimates upward. I try to correct for this by omitting products with large shares from the upper bound, but the range may still be upwardly biased. Combined these two issues produce fairly wide bounds, as evidenced than Panel B. While the upper bound is estimated less precisely than the lower bound, its closeness in proximity to costs reported by insurers is cause to believe that these estimates are nonetheless reliable.

The third caveat is that some of the estimates of $CS(\delta_{Jt}, \theta)$ may be driven by the addition and removal of logit error shocks. In Appendix E, I report estimates of ρ and FC_j assuming the shock is zero. Indeed, doing so yields a mismatch estimate that is similar to the baseline estimate, while significantly reduces the estimates of FC_j to approximately \$1.6 million per plan. I report counterfactual estimates in the Appendix F using these alternate estimates of ρ and FC_j. Though this does change the equilibrium menus a bit, the qualitative results and welfare estimates remain similar.

[34] When "net spending" is defined as either premium revenue less medical spending (in the case of self-insured plans) or 75% of premiums paid out to insurers (in the case of fully-insured plans), the estimates of fixed costs represent about 0.42% of net spending. Similarly, these estimates represent about 0.91% of the net social surplus (again, defined as total consumer welfare less net spending). Therefore, relative to the overall budget that the GIC allocates towards health expenditures (nearly \$1 billion per year), fixed costs associated with managing multiple plans remains a small, but important component of its objective function.

[35] In a 2010 hearing held in Massachusetts with the state's major health insurers, at least one plan identified its expenditure of costs and resources associated with implementing new products as varying between \$1 and \$3 million in total costs, which is nearly identical to the range of estimates I am finding (Murray, 2010).

4 What Drives Employer-Employee Mismatch?

The high estimate of ρ in Table 3 indicates that that the employer weigh total consumer utility from a plan menu by nearly four times what its spends on health care and premiums. The implication of this estimate is that is that, when selecting plan menus for its employees, it is *overweighting* consumer WTP for broad-network plans, relative to the how the average employee values those networks. This is particularly surprising since, in theory, if employees value the insurance benefits by more than the cost to employers, then employers ought to be able to offer those benefits and either raise co-premiums or reduce other means of compensation (e.g. wages, other benefits) to adjust for the added cost (Summers, 1989; Gruber and Krueger, 1991; Bundorf, 2002).

There are several potential drivers of this mismatch between employer and employee preferences. The first is that there is considerable heterogeneity in both medical costs and preferences for broad networks among the employees. The ability to precisely target an insurance plan menu to match the preferences of a diverse set of employees is limited and, as such, employers might use heuristics or otherwise have a tendency towards satisfying the preferences of certain groups when designing benefits. For example, an employer may place increased weight on segments of employees who have high expected medical costs or otherwise value comprehensive insurance highly. Similarly, employers may place emphasis on employees who have high labor market bargaining power or those who they believe will be most productive at the firm. This includes, for example, older employees or those in managerial positions.

A second potential explanation is that the particular employer considered in this paper, the GIC, does not operate as a traditional private-sector employer and, while having the responsibility to determine plan benefits for all state employees of Massachusetts, may not have the ability to adjust other forms of public-sector employee compensation in tandem. For example, prior work has found that the premium-wage tradeoff among public-sector employees is considerably lower than the tradeoff for private sector employees (Qin and Chernew, 2014; Lubotsky and Olson, 2015). This may be particularly true in situations where public-sector employees are unionized, as is the case in Massachusetts (Clemens and Cutler, 2014).[36] Further, public-sector benefits are often subject to voter influence, as they are funded through taxpayers.[37] The full cost of fringe benefits, in particular, are often thought of as "shrouded" from the view of the local taxpayer (Glaeser and Ponzetto, 2014). As a result, health insurance benefits for public-sector employees may skew towards more comprehensive coverage (Lubotsky and Olson, 2015).

A third possibility may be that the employer is making mistakes in plan offerings. This could be the case if, for instance, the employer found it difficult to obtain accurate information about expected medical costs of its employees or obtain information about employee preferences (Dafny

[36] As union membership tends to traditionally scale towards older employees (https://www.bls.gov/news.release/union2.nr0.htm)–and since older employees tend to be more active in union functioning–bargaining with municipalities and the state over health benefits might therefore skew towards the preferences of those employees over the average employee, who is likely younger and healthier. This dovetails nicely with the explanation above

[37] There is indeed some evidence that this is the case with the GIC. In early 2018, the GIC decided to eliminate plans by Harvard Pilgrim, Tufts Health Plan, and Fallon, leaving only a streamlined set of health plans at considerably lower costs. After considerable pushback, they relented on the decision. https://www.wbur.org/commonhealth/2018/01/25/open-meeting-investigation-is-latest-twist-in-gic-health-plan-controversy

et al., 2010). It might also be the case that employers *misattribute* employee switching costs as genuine preferences for broad-network plans. This would be consistent with the GIC's decision not to, for instance, reduce the network breadth of some of its flagship plans (e.g. Harvard Pilgrim) when the cost savings from this move would be substantial.

4.1 Heterogeneity in Employer Preferences

To address the possibility that the employer's preferences may be aligned with certain segments of the employee pool, but misaligned with the *average* employee in the population, I re-estimate ρ and FC_j after reweighting the employee population to skew more heavily towards older employees. If it were the case that plan offerings reflected the preferences of older employees, then we should expect the estimate of ρ to attenuate towards 1 as the average employee shifts older.

To test this, I first reweight the distribution of employees such that that 90% of the employee pool is greater than 55 years old.[38] At baseline, the share of employees 55 and older represent only about 30-35% of the population, depending on the year. The results of this simulation is depicted in Panels A of Table 4. Column 1 reproduces the baseline estimates from Table 3. Column 2 reports the estimates from the shifted distribution. Here, the estimate of the employer-employee mismatch noticeably decline: ρ decreases to 2.89, while FC_j decreases to \$3.69 million. This conforms to expectation: the GIC indeed places an outsizes value on older employees in its plan design. That is, as the population gets older, the employer's observed plan offerings become more aligned with that of the average employee and, as a result, the mismatch parameter declines. The results are similar when reweighting the population to those with "unstable" chronic conditions (Column 3).[39]

I next consider additional simulations, in which I reweight the population such that 90% of employees are 55 and older *and* live in select rating regions in the state. Columns 4 and 5 report the estimates of ρ and FC_j for these simulations. I find that the strongest effects on the employer-employee mismatch occur by shifting the distribution towards employees residing in three of the seven geographic regions: region 1, region 4, and region 6 (Column 4).[40] In aggregate, enrollees in these areas make up 17% of the GIC population. The weight on consumer surplus falls again, this time to 2.41, a 34% decline from the baseline estimate. The effect is the most prominent in Region 4, as reported in Column 5. Here, the implied employer-employee mismatch falls to dramatically, to just 1.54, a nearly 60% decline. This is noticeably closer to 1, lending credence to the theory that the employer is, at least in part, choosing plans that disproportionately benefit a smaller share of the population. The fixed cost estimates for these simulations also declines, as the marginal benefit of offering *additional* narrow-network products falls for these groups.[41]

[38] As depicted in Table A.2, the average age of employees who saw a primary care doctor in the GIC is 47 year olds, with a standard deviation of 15 years old.

[39] I define this as the member having a diagnosed chronic condition, according to the AHRQ definition, and the member having previously experienced at least one inpatient hospital stay.

[40] Rating Area 1 encompasses employees who live in Western Massachusetts, primarily around Amherst, Springfield, and the Berkshires. Rating Area 4 encompasses employees who live towards the North Shore in Massachusetts. Rating Area 6 encompasses people who live in Bristol and Plymouth counties, including near Rehoboth, Massachusetts. I tested *every* rating region for these calculations, but omitted the results for rating regions 2,3, and 5 for brevity. The implied employer-employee mismatch in those regions were *higher* than the baseline estimate.

[41] As a robustness test, Panel B reports estimates for the employer-employee mismatch only for moments that hold fixed the number of plans the employer could offer. This, in effect, isolates estimation of ρ by negating the effect of the fixed costs and, thus, ensuring that the mismatch term is not potentially driven by error in the estimates of

Table 4: Employer Objective Function Parameters For Different Populations

	Baseline	Older	Chronic	Older, Regions 1,4,6	Older, Region 4
	Panel A: Estimation with All Moments				
ρ	3.67	2.89	2.81	2.41	1.54
FC_j	4.07	3.69	3.69	2.64	2.39
	Panel B: Estimation with Fixed Number of Plans				
ρ	3.67	2.89	2.77	2.41	1.51
FC_j	–	–	–	–	–
	Panel C: Estimation on Private Employers				
ρ	2.59	1.17	2.64	1.05	1.65
FC_j	–	–	–	–	–
	Panel D: Estimation On New, Municipal Entrants				
ρ	1.96	1.60	1.58	1.56	1.10
FC_j	–	–	–	–	–
	Panel E: Estimation Restricting Harvard Replacement				
ρ	1.79	1.29	1.62	1.22	0.97
FC_j	–	–	–	–	–

Results from ρ and FC_j estimation for 2009-2013. Column 1 presents estimates for the current population of GIC enrollees. Column 2 presents estimates that reweights the population such that 90% of the population is comprised of adults 55 and older. Column 3 presents estimates that reweights the population such that 90% of the population have unstable chronic conditions. Column 4 reweights the population such that 90% of the population are older *and* residing in rating regions 1, 4, or 6. Column 5 reweights the population such that 90% of the population are older and residing in region 4. Panel A reports the estimates of the employer weight on consumer surplus, ρ, and fixed cost, FC_j, using all moment inequalities (the main specification). Panel B estimates these parameters restricting the employer to keeping its existing *number* of plans (but allowing networks to change). Panel C presents estimates conducted on a simulated sample of private employers. Panel D reports estimates on only new, municipal entrants. Panel E presents estimates restricting the GIC from moving Harvard's broad-network to the Harvard Primary Choice. FC_j reported in millions of dollars.

These areas are united by three common factors. First, each region is served by a hospital that is both a member of one of the state's flagship health systems and absent from narrow-network plans.[43] Second, each region is less dense than the major metropolitan areas of the state (e.g. Boston, Worcester, etc.), sees less competition among health care providers, and requires more travel to providers. As such, employees stand to lose substantially more utility by losing access to a dominant provider in those regions. Finally, these rating areas are close to the state border and may therefore cater to households who desire access to providers in neighboring states.[44] Taken together, the characteristics of these regions imply a natural preference towards broad-network plans, relative to areas in the state with a high volume of provider options in the immediate area. Indeed, the high variation in market dynamics across the regions within the state highlights the difficult task the GIC faces in designing a single menu of options to satisfy all its members.[45]

Overall, the results indicate that the observed health plan offerings appear to be reflective of the preferences of a small subset of employees with high WTP for networks. This is highly suggestive of two plausible phenomena. First, employers may be weighing workers with higher labor market bargaining power, such as older workers more likely to be employed in managerial roles or those with strong union protections (see Column 2). However, even more strongly, employers appear to be largely driven out of equity concerns for workers living in less competitive provider markets (Columns 4 and 5). In other words, employers appear to want to ensure that *each* employee has access to a wide array of providers. As shown in section 5, in absence of offering different benefits to different employees, this tendency leads employers to both offer *more* plans and more *comprehensive* plans in equilibrium than would be offered under a social planner weighing each employee equally. The reason is that, while indeed ensuring access to all employees, it also provides a mechanism for employees in higher-cost markets who have lower WTP for networks to also enroll in broad-network products.

4.2 Benefit Adjustment Frictions for Public Employers

I test the extent to which the estimate of employer-employee mismatch in Table 3 is driven by differences in incentives between private and public sector employers by re-estimating Equation 17 on a sample of large, self-insured private employers in Massachusetts. If the employer-employee mismatch is driven by the fact that the employer considered in this study is a public employer—perhaps with limited ability to defray increases in health costs through other benefit adjustments—then we ought to expect the estimate of ρ to be driven down to 1 when re-estimated on a different set of employers. To do this, I use the APCD to construct a sample of private employers and simulate narrow-network offer distributions by supplementing the data with microdata from the Kaiser/HRET annual employer survey. Overall, the sample produced 86 large firms, approximately 6% of which offered a narrow-network plan in 2013. Details of the construction of the private employer sample are given in Appendix D.

FC_j.[42] The results are encouragingly quite similar to the results in Panel A.

[43] For instance, Rating Region 1 has the UMass health system and Region 4 has the Partners system.

[44] Rating Area 1 borders New York and Connecticut, Rating Area 4 borders New Hampshire, and Rating Area 6 borders Providence, Rhode Island.

[45] One can imagine similar dynamics existing in designing plan menus for companies whose employees live in many different states and markets.

Panel C in Table 4 reports the results from this exercise for each of the subpopulation groups considered above. To isolate the specific effects on the mismatch parameter, I only consider counterfactual plan menus that contain the same number of products as the firm is observed to offer, rendering the fixed costs irrelevant. For the baseline population, the implied employer-employee mismatch for these firms is 2.6, suggesting that private employers still substantially "overweight" consumer preferences for health insurance relative to health spending, though by a smaller amount than the GIC. This does lend some credence to the theory that public employers may be more constrained in their ability to adjust other margins of employee compensation in response to health cost shocks.

However, this is still a sizable mismatch between employer and employee preferences. Moreover, most of the patterns across subgroups still persistent, albeit with different relative importance. The implied employer-employee mismatch among the older population is nearly eliminated, dropping to just 1.17, a 55% decline, compared with just a 20% decline for the GIC. Interestingly, the mismatch for the chronic-condition population barely budges. The mismatch term does decline further when considering the same regions that drove the result in the GIC sample, but by a lesser extent.[46] Taken together, this suggests that the majority of the mismatch for private employers appears to be driven by firms' weighting the preferences of older—though not sicker —employees. Compared to the GIC, which appears more driven by regional equity concerns, this could reflect the increased influence that managers or executives at private companies have over benefit design.

4.3 Employer Misperceptions or Mistakes

The Role of Switching Costs: A possible explanation for the persistence in broad-network plans is that employers misperceive the true loss in employee utility from a loss of providers. This would most commonly be the case if they mistook enrollee inertia for "true' network utility. This is a fairly difficult phenomenon to test for. Indeed, if the entirety of the switching cost parameter were shifted to network utility, then the mismatch parameter would mechanically shift downward as the utility gap between broad and narrow networks would widen. To get a sense of the precise magnitude, one possibility would be to re-estimate the plan demand model but simply omit the plan switching cost term. However, as seen in Table 2, this results in implausibly low premium sensitivity estimates.[47] Another approach would be to shift some portion of the switching cost estimate towards the network utility. However, this approach is difficult to implement empirically as it requires making assumptions as to how switching costs–a flat per-plan cost–maps to network utility, which scales by plan.

I instead take an alternate approach: I re-estimate the plan demand model for a specific subsegment of the population for whom it is likely the employer believes have strong preferences for broad networks. I then apply these estimates to the entire population. Specifically, I focus on new entrants to the GIC coming specifically from municipalities entering for the first time. This solves the premium elasticity issue mentioned above, as these are all new entrants. However, unlike other new employees, these municipal entrants have previously lived and worked in the state, and have

[46] The weighting for Boston among private employers was 2.63, suggesting that, like the GIC, the populations in dense, urban areas do *not* appear to be the ones employer are emphasizing in their plan designs

[47] Some older households under this specification are predicted to have *positive* utility from higher premiums.

also been previously enrolled in private health insurance. Moreover, those plans were uniformly generous, broad-network plans.[48]

Panel D of Table 4 reports these estimates, again focusing only on moments that include the same number of products to isolate the effects on ρ. The employer-employee mismatch term drops to 1.96, about a 46% decline from the baseline estimate (Column 1). Indeed, this does suggest a role for employer misperceptions. However, even in this case, the employer continues to overweight the average household's preferences by about 2-to-1. Moreover, the patterns across household types persist. Assuming that this entire discrepancy is driven by misperceptions, about 60% of the employer-employee mismtach can be attributed to employer misperceptions of switching costs, while the remaining 40% would be attributed to unequal weighting in household preferences.

Which Plans Drive the Result? I run a series of robustness checks in which I restrict the estimation to certain moments to ascertain which plans drive the result. The estimates are robust to most alternate specifications. The one glaring exception is when I restrict the GIC's ability to reduce the network breadth of Harvard Broad to either its medium-sized narrow-network that it introduced in 2011 ("Primary Choice") and a similarly-sized network that it offers on the small group market ("Harvard Focus"). Notably, these networks are broader than all the other narrow-network options: they contains all hospitals and providers in Massachusetts, with exception of those owned by Partners. The networks each include physicians who are part of Atrius Health System, which owns the Harvard Vanguard medical group, a very prominent group in the Boston region.

Panel E of Table 4 reports these estimates, while also holding fixed the number of plans (as in Panel B). The results are quite large: the estimate of ρ for the baseline population falls from 3.67 to 1.79. The implication is that the mismatch appears to be mostly driven by employers' unwillingness to make small, yet impactful network changes. In this instance, the moments responsible for the result are the ones in which the employer could remove from its network some flagship and costly hospital systems (e.g. Partners) but preserve a wide network of *physicians* (e.g. Atrius). If the employer made this move, the utility differences for most employees would be small, given consumers' strong preferences for physician networks seen in Table 2, even as the cost implications would be substantial. This result is consistent with the insights from Shepard (2016).

Interestingly, even in simulations restricting the GIC's ability to narrow Harvard's broad-network plan, the same heterogeneity implications from above continue to persist. Reweighting the population to the older sample described above, the estimate of ρ falls further to 1.29, while it falls to 1.22 when rescaling towards the employees residing in regions 1,4, and 6. Therefore, these result lends support to either the interpretation that the employer is making mistakes in its plan offerings[49] or that its offerings overweight specific segments of the population.[50]

[48] About 50% of municipal entrants were previously insured by Blue Cross Blue Shield, which at the time had no narrow-network products. About 90% of entrants were enrolled in a plan with 0 deductible.

[49] Since the time of this sample period, the GIC froze enrollment in the Harvard broad-network plan and it has remained this way since 2016. This was in direct response to concern that this plan's premiums were growing considerably more than other products offered on the menu, despite having similar benefits. https://www.lowellsun.com/2016/03/03/gic-bars-harvard-pilgrim-plan/. Existing members were allowed to remain on the product, but members could not switch to it and new enrollees could not select it. While not dispositive, this does lend support to the idea that the employer may be attempting to issue a corrective for a prior mistake in plan design.

[50] In particular, it may be the case that these moves are made in deference to populations that I am not capturing

5 Welfare Implications and Policy Simulations

5.1 Welfare and Cost Implications of the Mismatch

As demonstrated in section 4, the prevalence of broad-network plans among large employer groups appears to be driven not by the underlying network preferences of the average employee, but rather by diverging *employer* incentives from that of the average employee. I now proceed to discuss the welfare and cost implications of these differing incentives.

I first simulate the employer's choices of plan menus and premiums in 2011 (the year that the GIC introduced Harvard Narrow and Tufts Narrow), assuming that the employer weighed each employee equally (or that the employer's incentives were fully aligned with the average employee). To do so, I simply set the estimated parameter, ρ, to 1. To conduct the simulations, I fix the potential product space to a set of 14 potential plans offered by 5 insurers. Among the two largest insurers, Harvard and Tufts, I allow four different counterfactual networks:

- B: A "broad" network equivalent to the two insurers' currently-offered widest network
- M: A "medium" network where Partners hospitals and doctors are removed[51]
- N2: A "narrow" network where Partners *and* Atrius are removed[52]
- N1: A "narrow" network where Partners *and* Atrius are removed, along with other hospital and physician groups[53]
- VN: A "very narrow" network where many hospital and physician groups are removed[54]

I allow the employer to completely sever a relationship with a carrier. However, I make a "no uninsurance assumption." That is, between all the plans offered on the GIC, every single member must have access to at least one plan with adequate network coverage. To do so, I leverage data on the counties in which each network are currently offered and impose that any counterfactual network of similar size must be offered in those same counties. For example, if the GIC chooses to reduce Harvard's narrow-network plan to much smaller network equivalent to Fallon's, then that new network can only be offered in counties in which Fallon's plan is offered. Across all the offered plans, individuals in all counties must be offered insurance. The simulation procedure used to evaluate these counterfactuals is described in more detail in Appendix F.

I report measures of consumer surplus, health spending, fixed cost, and total surplus changes. Consumer surplus is defined as:

$$\Delta CS(\delta_{Jt}, \delta^a_{Jt}, \theta) = \sum_I \frac{1}{\alpha_I} ln\left(\sum_j^J exp(\delta^{1,a}_{Ijt})\right) - \sum_I \frac{1}{\alpha_I} ln\left(\sum_j^J exp(\delta^{1}_{Ijt})\right) \tag{19}$$

where $\delta^1_{Ijt} = \delta_{Ijt}$ from Equation 3.4, $\delta^{1,a}_{Ijt}$ is the counterfactual plan menu.

through age, region, or the presence of a chronic condition.

[51] This is equivalent to the Harvard Primary Choice network.

[52] This is equivalent to the Tufts Select network available in the small-group market.

[53] This is equivalent to Tufts Spirit.

[54] This is equivalent to Fallon Direct.

Table 5 reports the equilibrium predicted products/networks offered from the simulations. I also report the observed and the predicted plan menus at baseline. Encouragingly, the predicted networks match the observed plan menus very well. The only difference is that the model does not predict Tufts to introduce its narrow-network product, and instead assigns this product to Harvard.[55] Column 5 reports the simulations from the assumption that $\rho = 1$. Under this scenario, the employer is predicted to drop *all* broad-network plans in favor of narrower products across the board. Overall, the number of plans falls from 8 to 7. Under this scenario, consumer surplus significantly declines, by about \$70 per household per month. However, this loss is more than compensated for by a significant decrease in health spending of approximately \$113 per household per month and an additional \$4 decrease in fixed costs. Total surplus therefore is about \$40 per household per month higher than at baseline, implying that the employer-employee mismatch generates a surplus shortfall of about \$480 per household per year.

5.2 Uniform Pricing and Plan Menus

One of the major drivers of the persistence in broad-network enrollment is that employers offer uniform insurance plans to all employees in the risk pool (with exception of plans that do not have adequate network coverage in a region). If these plan offerings are driven primarily by the preferences of a relatively small share of high WTP employees, as demonstrated in section 4, then this implies such plans are also available to low WTP households. Under the GIC's current pricing scheme—where it subsidies 25% of enrollee premiums regardless of which plans they take up—households who may not value broad networks at their full cost still do enroll in those plans, which further drives up costs for the group. One natural solution would be to offer multiple plans at different network levels. However, at current pricing, the premium differential between the different network levels is not wide enough to generate enrollment shifts among employees with high switching costs (as shown in Table 2).

Enthoven Approach with Fixed Menus: A potential solution to this, while maintaining uniform pricing and plan menus for all employees, is to widen the premium differential between narrow and broad plans such that only the high WTP consumers enroll in broader coverage.[56] I conduct a simulation of one such scenario: moving the GIC to an "Enthoven"-style managed competition approach, wherein flat premium subsidies are pegged to the lowest-cost plan offered and households are asked to pay the full premium differential for any plan that is more expensive.[57]

Results of this simulation are presented in Column 6 of Table 5. Under this new pricing structure, if plan menus are held fixed at their 2011 observed values, then employer costs increase by \$17 per household per month (Panel B). Yet this is offset by an increase in consumer surplus by \$56 per household per month, leading to total surplus gains of about \$38, approximately the

[55] This is driven by the fact that Tufts Narrow has consistently received little enrollment in my sample period and, therefore, my model has difficulty rationalizing its inclusion.

[56] If WTP for networks is highly correlated with ex-ante health risk, then this approach might result in the unraveling of broad-network plans due to selection, as in Shepard (2016). However, if WTP is not significantly correlated with risk, then this approach should result in more efficient sorting across plans without such unraveling.

[57] In effect, this scenario simulates the same effect as the GIC premium holiday: the employer offers a subsidy to employees to switch to lower-cost plans. This exercise is similar to Bundorf et al. (2012).

Table 5: Counterfactuals: Equilibrium Networks Chosen Under Uniform Pricing

Insurer	Network	Observed	Pred.	$\rho = 1$	Enthoven
Panel A: Equilibrium Plan Menus/Networks					
Fallon	VN	x	x	x	
Fallon	B	x	x		x
HPHC	VN			x	
HPHC	N1				x
HPHC	N2		x		x
HPHC	M	x	x	x	x
HPHC	B	x	x		
HNE	N	x	x	x	x
NHP	N	x	x	x	
Tufts	VN			x	
Tufts	N1	x		x	
Tufts	N2				
Tufts	M				
Tufts	B	x	x		x
Total Plans		8	8	7	6
Panel B: Welfare and Spending Holding Plan Menu Fixed					
ΔCS (Fixed)				–	\$55.85
$\Delta Costs$ (Fixed)				–	\$17.74
ΔFC (Fixed)				–	–
$\Delta Surplus$ (Fixed)				–	\$38.11
Panel C: Welfare and Spending Allowing Plan Menu to Change					
ΔCS (Change)				-\$68.99	\$121.21
$\Delta Costs$ (Change)				-\$113.61	\$85.31
ΔFC (Change)				-\$3.81	-\$7.63
$\Delta Surplus$ (Change)				\$40.81	\$43.53

Notes: GIC observed and predicted products offered under various counterfactual assumptions. "$\rho = 1$" refers to predicted networks when the estimated employer-employee mismatch is eliminated. "Enthoven" refers to predicted networks when employers are moved to a uniform pricing mechanism that pegs plan subsidies to the lowest-price plan offered in each market. Panel B reports the welfare and cost changes assuming plan menus remain fixed at observed 2011 levels. Panel C reports these changes allowing endogenous employer changes to menus. "ΔCS" refers to change in consumer surplus per-household-per-month. "$\Delta Costs$" refer to the change in total GIC costs per-household-per-month. "ΔFC" refer to changes in fixed costs from the new menus.

same aggregate gains that would be achieved if the entirety of the employer-employee mismatch were eliminated.

The gains in consumer surplus from the new pricing scheme are achieved through substantial increases in subsidies to purchase lower-cost plans. Indeed, under this scenario, the lowest-cost plan offered in any market are free for the households to purchase by definition. This is also the source of increased employer health spending. This nets out to aggregate surplus gains due to the fact that the premium differential between narrow and broad plans becomes wide enough such that a large share of consumers are induced to switch away from broad-network plans. These results are displayed in Table 6. Under the Enthoven counterfactual, individual co-premiums for broad-network plans increase, while co-premiums for narrow-network plans substantially decrease. Harvard Broad, in particular, increases from its observed value of \$152 per month to \$201 per month. Meanwhile, Fallon's "very narrow" network declines from \$105 to just \$7 per month. Not coincidentally, these plans also see substantial shifts in enrollment. The share of enrollees in Harvard Broad declines from 34% to just 12%, with many of those households shifting to Fallon, HNE, and NHP. As a result of these shifts, the overall spending burden for the GIC rises by *less* than the rise in consumer surplus.

Table 6: Counterfactuals: Shares and Premiums for Enthoven Approach, 2011

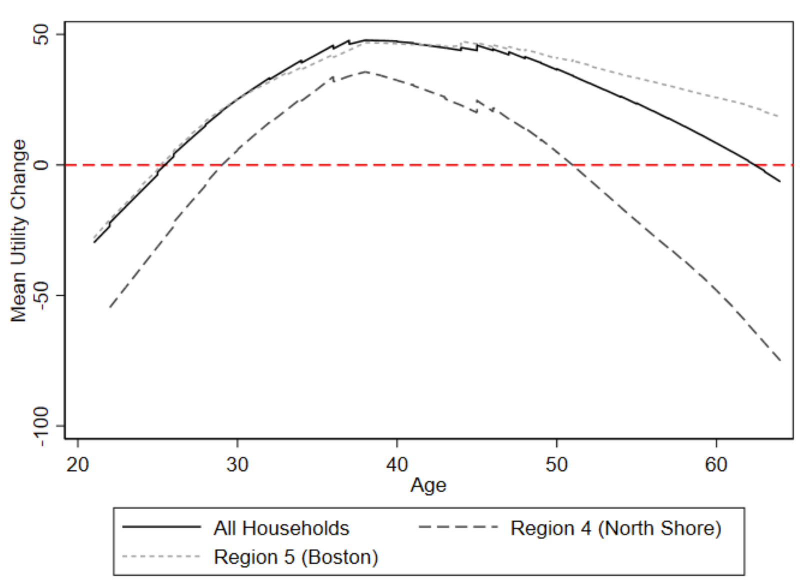
Insurer	Network	Baseline	Counterfactual	Baseline	Counterfactual
		Market Shares		Co-Premiums	
Fallon	VN	0.02	0.08	\$105	\$7
Fallon	Broad	0.05	0.03	\$126	\$92
HPHC	Med	0.04	0.05	\$122	\$75
HPHC	Broad	0.34	0.12	\$152	\$201
HNE	N	0.11	0.22	\$105	\$0
NHP	N	0.03	0.28	\$106	\$5
Tufts	N	0.01	0.02	\$117	\$57
Tufts	Broad	0.41	0.20	\$147	\$167

Notes: Market shares and individual monthly premiums for baseline and counterfactual predictions, holding the GIC's product menu fixed. Individual premiums for counterfactual plans computed only in regions where Health New England was offered.

Enthoven Approach with Endogenous Menus: If the GIC is allowed to endogenize its menu in response to this policy change, then it is predicted switch Harvard Broad to a narrow plan, dropping Fallon's currently-offered narrow plan, and drop NHP, leaving a total of 6 plans (as seen in Table 5). Note that this varies by year: in simulations done on 2013, for instance, the GIC drops Tufts' broad network as well. This may seem surprising, given the large enrollment in Fallon and NHP seen in Table 6. The reason for this change is that doing so allows the GIC to actually peg its premium subsidies towards a higher-cost plan. Due to the higher subsidies, GIC costs soar by \$85 per household per month (Panel C), which is once again offset by increases to consumer surplus of \$121. Ultimately, total surplus increase of \$43 per household per month.

Distributional Consequences: While the Enthoven approach yields surplus gains *on average*, employers may find it undesirable to impose a policy that results in the elimination of a flagship broad-network product, particularly if the distributional consequences are severe. In Figure 5, I plot the predicted surplus changes across households by age and region that result from such an approach.[58]

Figure 5: Total Surplus Changes by Age, Enthoven Approach



Notes: This figure plots the average utility change across households by age from implementing an Enthoven-style pricing approach, while allowing the GIC to alter its plan menus. Curves are plotted for all households, for households in Rating Region 4 (the North Shore of Massachusetts), and for Rating Region 5 (which includes the Boston metro area). Surplus is presented in dollarized terms, assuming that the employer fully passes back spending changes to consumers equally across households.

I predict that such a policy change would result in net surplus increases for most households where the eldest member is between the ages of 30 and 50. Indeed, this represents *most* households on the GIC. However, if the GIC were to shift the entirety of its spending increases onto consumers uniformly, households at the low and high ends of the age distribution would incur substantial losses. For younger households, this is due to the fact that such households are comprised primarily of single adults who rarely interact with the health care system. As such, these households would prefer to take their benefits, for example, through wages than as generous subsidies to purchase insurance. Conversely, for the oldest households, the loss of access to Harvard Broad, even with the presence of large subsidies, significantly decreases their utility. Indeed, these two sets of households represent a tiny share of the overall pool. Most households in between those age groups see substantial surplus increases—even net of any potential employer passback of spending increases—that bring up the average. These households care about purchasing health insurance, though are not the households

[58]Note that assessing welfare changes is going to critically depend on the extent that the employer imposes the additional premium subsidy spending incurred from Table 5 onto employees through slower wage increases or reductions to other amenities. For the purposes of this exercise, I assume that the employer fully passes this increase onto employees evenly across households. Though perhaps unrealistic, any alternate assumption would simply rescale the magnitudes of the results. The qualitative implications would remain similar.

with the highest WTP for broad networks.

There is, however, substantial regional variation in surplus changes as well. Notably, averaging across all households implies that only employees at the very upper end of the age distribution (right around 65) start seeing surplus losses. While *no* elderly households in Region 5 (which includes Boston) suffer net utility losses from this policy, households residing in Region 4 (the North Shore area) start seeing surplus declines at around age 50. This coincides precisely with the results in section 4. Namely, these are the households who value broad networks the most. As such, their removal, particularly for older employees, represents a significant loss.

Overall, then, while moving to an Enthoven-style approach would yield surplus gains for most households, it would likely adversely affect households in regions that value broad-networks the most, where competition among providers is more scarce. This is particularly true for older households in these regions. Given that these are precisely the households that the employers "overweight" relative to the average when designing their benefits, it is then clear why employers do not more uniformly adopt this approach in practice.

5.3 Group Rating and Plan Menus

Region-Based Rating: I now consider an approach where I permit the GIC to set benefits and prices differently for employees in each of the rating regions in the state.[59] The results are reported in Table 7. When the employer holds plan menus fixed at their observed level, household utility and spending from the insurance plans remain virtually unchanged. The reason is that the employer is merely shifting some costs of enrolling in broad-network plans onto certain *regions*, while reducing those costs for other regions, leading to a virtual wash. The pricing differences are, in other words, simply not large enough to induce significant enrollment shifts without actually altering the plan menus for different groups.

However, when I allow the GIC to endogenously alter its plan menus, there are significant effects to plan design, welfare, and spending. In Table 7, I report the plan offer choices for three select rating regions for illustration: Region 1 (Western Massachusetts), Region 4 (the North Shore), and Region 5 (Boston). Here, plan menus change considerably. In Region 4, the employer preserves access to most of the existing plans plans. In Region 1, however, the employer drops all but 4 plans and *only* retains broad-network products. Finally, for the Boston region, the employer significantly *narrows* the networks in its menu. Specifically, the GIC drops 3 plans from the menu, including both Harvard Broad and Fallon Broad. In fact, it drops *all* plans by Fallon HNE, and instead adds a narrow-network offered by Harvard. These changes result in spending declines of about \$32 per household per month, which is partially offset by utility declines of about \$7 per household per month due to the loss of choice. In addition, since the employer now offers more *total* products across all regions, fixed costs increase by about \$7 per household per month. As a result, total surplus increases by \$18 per household per month.

[59] Region-based rating is the norm in the individual marketplaces under the ACA exchanges. However, in practice, most large employers offer identical benefits, premiums, and plan designs to employees irrespective of their residence location. This may be due to discrimination concerns, putting employers at increased risk of a lawsuit, as described in https://www.safegardgroup.com/blog/2017/12/19/health-plan-treating-employees/.

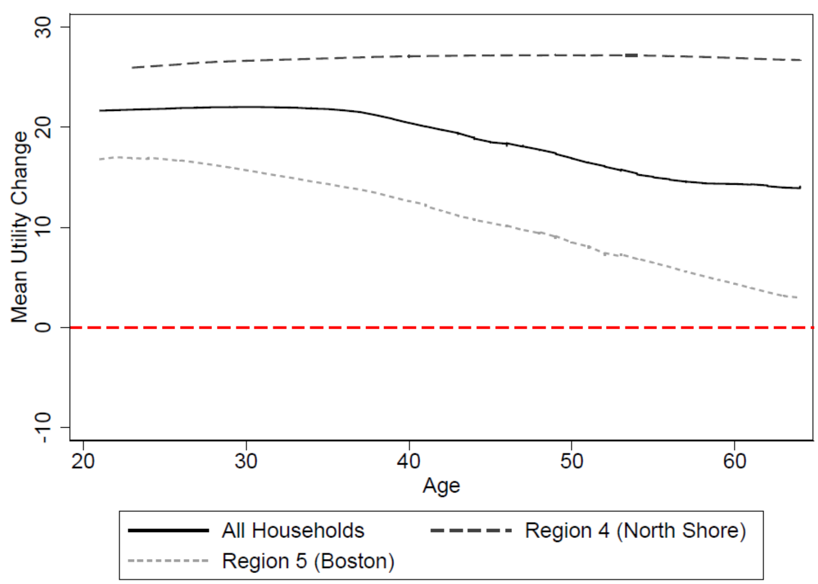
Table 7: Counterfactuals: Equilibrium Networks Chosen Under Region-Based Pricing

Insurer	Network	Observed	Region		
			R1	R4	R5
Panel A: Equilibrium Plan Menus/Networks					
Fallon	VN	x		x	
Fallon	B	x	x	x	
HPHC	VN				
HPHC	N1				x
HPHC	N2			x	x
HPHC	M	x		x	x
HPHC	B	x	x	x	
HNE	N	x	x		
NHP	N	x		x	x
Tufts	VN				
Tufts	N1	x			
Tufts	N2				
Tufts	M				
Tufts	B	x	x	x	x
Total Plans		8	4	7	5
Panel B: Welfare and Spending Holding Plan Menu Fixed					
ΔCS (Fixed)				-\$0.47	
$\Delta Costs$ (Fixed)				\$0.07	
ΔFC (Fixed)				–	
$\Delta Surplus$ (Fixed)				-\$0.54	
Panel C: Welfare and Spending Allowing Plan Menu to Change					
ΔCS (Change)				-\$7.12	
$\Delta Costs$ (Change)				-\$32.27	
ΔFC (Change)				\$7.62	
$\Delta Surplus$ (Change)				\$17.53	

Notes: GIC observed and predicted products offered under region-based rating. "R1" refers to plan networks for region 1, etc. Panel B reports the welfare and cost changes assuming plan menus remain fixed at observed 2011 levels. Panel C reports these changes allowing endogenous employer changes to menus. "ΔCS" refers to change in consumer surplus per-household-per-month. "$\Delta Costs$" refer to the change in total GIC costs per-household-per-month. "ΔFC" refer to changes in fixed costs.

Distributional Consequences: While the aggregate surplus gains from a region-based rating approach are only half of that the size of moving towards an Enthoven-style pricing approach, the distributional consequences are considerably less severe. In Figure 6, I reproduce the same plot of utility changes from Figure 5, but for the region-rating approach. Indeed, averaged across all households, utility from the menu change declines as consumers age. This particularly pronounced for households residing in the Boston region, where the slope of the utility decline is sharper beginning at around age 40. However, unlike the Enthoven approach, *each* household's net utility change remains above 0, regardless of location or age. While households in Boston do stand to lose the most from the change—both due to loss of access to a large number of plans and to the loss of Harvard's broad network—they can be more than compensated by the employer due to the spending savings. Meanwhile, households residing in the North Shore—the ones the GIC appears to value the most—sees the largest utility gains from the new plan menus. This is due to the fact that in equilibrium, the existing menus are almost identical to those observed in the data. Given that the employer now sees spending savings due to plan changes in *other* regions–and is assumed to equally distribute these savings to all households–these consumers now see large welfare gains, in spite of seeing no benefits changes.

Figure 6: Total Surplus Changes by Age, Region-Rating Approach



Notes: This figure plots the average utility change across households by age from implementing an region-based benefits and rating approach, while allowing the GIC to alter its plan menus. Curves are plotted for all households, for households in rating region 4 (the North Shore of Massachusetts), and for rating region 5 (which includes the Boston metro area). Surplus is presented in dollarized terms, net of the predicted increased spending to the GIC.

6 Conclusion

The rollout of the Affordable Care Act has brought a renewed focus on managed competition in health insurance markets. Particularly as new types of insurance innovations emerge (including narrow networks, tiered networks, health savings account, and high-deductible health plans), policymakers and employers have struggled with balancing offering consumers choices that provide risk protection, while keeping premiums and spending low, keeping consumers well-informed, and preventing confusion. With regards to the exchanges, states vary dramatically in the plan choices available and the levels of plan standardization. As a result, states have very different experiences in terms of consumer enrollment, premiums, and spending.

Employers similarly struggle to strike this balance. As companies grow and cater to employees with much more heterogeneous preferences, firms have increasingly turned to offering not only more choices of plans, but also offering different types of products. So far, most of this choice has been among financial dimensions of health plans: copayments, coinsurance, and deductibles. As in Brot-Goldberg et al. (2017), these types of products, particularly high-deductible health plans, are often difficult to navigate for consumers. This is partly because the burden is on the consumer to investigate underlying health care prices and make informed decisions on which providers to utilize. Narrow-network plans, conversely, put the onus on the insurers and employers to form the networks that consumers may choose.

In this paper, I show that moving towards offering employees narrow-network coverage does have the potential to significantly decrease costs and increase surplus. In particular, I show that a majority of employees at large firms would be better off under a scenario in which they had less choice of physicians and hospitals, but were compensated for that loss. This begs the question of why, then, employers do not yet offer these plans in large numbers. I provide evidence that this is driven by a combination of factors. First, as demonstrated in previous literature (Handel, 2013; Polyakova, 2016; Liu and Sydnor, 2018), consumers often select into products because of inertia or the presence of other behavioral frictions. I show that in the context of narrow networks, switching frictions drive a substantial portion of enrollment into broad-network products. To the extent that employers misperceive these frictions as true preference for broad networks, it may inhibit them from offering such products.

Second, I show that employer plan choices—even conditional on possible misperceptions and frictions—are driven by placing high weight on the preferences of older consumers and those in select geographic markets. Importantly, the geographic markets in which consumers see the most value from access to broad networks are *not* those in which households have the highest risk or the highest ex-ante probability of health care utilization. They are, however, regions that have less competition among providers and are less dense. This is highly suggestive that employers may be driven by equity concerns in their plan design, rather than maximizing the total surplus of its employee risk pool. Indeed, I show that while moving away from broad networks entirely would produce the largest *aggregate* surplus gains, the distributional consequences—particularly for older employees—would be severe. Conversely, for large employers who operate in multiple geographic markets, switching to a system where they may offer different provider networks in different regions has the potential both to improve aggregate surplus, while having minimal adverse distributional effects.

This analysis has some limitations. First, I only consider valuations for a limited set of physician specialties. This leaves out some heterogeneity that may drive choice into broad networks that I am not picking up. For instance, consumers may have extremely high valuations of certain high-cost provider types, such as oncologists, that may drive their preferences for health plans, and subsequently employer offerings. Second, while I am able to separate health *plan* switching costs from unobserved preference heterogeneity, I am not able to fully separate *physician* switching costs from unobserved preference heterogeneity. Finally, the model does not consider bargaining effects of altering plan menus. Indeed, if employers decide to place additional emphasis on narrow-network plan designs, this may impact the negotiations between insurance carriers, hospitals, and physicians.

Overall, this paper contributes to our understanding of what consumers in employer markets value in their choice of plan, and how employers aggregate those preferences to design insurance choices for those employees. Given that employers seek to maximize equity in addition to efficiency, a first-best approach may be to offer employers more flexibility in how they design their plans across different segments of consumers. This is especially true if this does not create sorting on risk.

References

Abaluck, J. and Gruber, J. (2016). Evolving choice inconsistencies in choice of prescription drug insurance. *American Economic Review*, 106(8).

Arrow, K. J. (1965). Uncertainty and the welfare economics of medical care: Reply (the implications of transaction costs and adjustment lags). *American Economic Review*, 55.

Berry, S. T. and Pakes, A. (2007). The pure characteristics demand model. *International Economic Review*, 48(4):1993–1225.

Brot-Goldberg, Z. C., Chandra, A., Handel, B. R., and Kolstad, J. T. (2017). What does a deductible do? the impact of cost-sharing on health care prices, quantities, and spending dynamics. *Quarterly Journal of Economics*, 132(3).

Buchmueller, T., Carey, C., and Levy, H. G. (2013). Will employers drop health insurance coverage because of the affordable care act? *Health Affairs*, 32(9).

Bundorf, K. (2002). Employee demand for health insurance and employer plan choices. *Journal of Health Economics*, 21.

Bundorf, M. K., Levin, J., and Mahoney, N. (2012). Pricing and welfare in health plan choice. *American Economic Review*, 102(7).

Center for Health Information and Analytics (2013). Annual report on the massachusetts health care market. Technical report, Center for Health Information and Analytics.

Clemens, J. and Cutler, D. M. (2014). Who pays for public employee health costs? *Journal of Health Economics*, 38.

Dafny, L. (2010). Are health insurance markets competitive? *American Economic Review*, 100(4):1399–1431.

Dafny, L., Duggan, M., and Ramanarayanan, S. (2012). Paying a premium on your premium? consolidation in the us health insurance industry. *American Economic Review*, 102(2).

Dafny, L., Hendel, I., and Wilson, N. (2015). Narrow networks on the health insurance exchanges: What do they look like and how do they affect pricing? a case study of texas. *American Economic Review: Papers and Proceedings*, 105(5).

Dafny, L., Ho, K., and Varela, M. (2010). An individual health plan exchange: Which employees would benefit and why? *American Economic Review: Papers and Proceedings*, 100.

Dafny, L., Ho, K., and Varela, M. (2013). Let them have choice: Gains from shifting from employer-sponsored health insurance and toward an individual exchange. *American Economic Journal: Economic Policy*, 5(1).

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P., and Cullen, M. R. (2013). Selection on moral hazard in health insurance. *American Economic Review*, 103(1).

Eizenberg, A. (2014). Upstream innovation and product variety in the u.s. home pc market. *Review of Economic Studies*, 81.

Ericson, K. M. and Starc, A. (2015a). Measuring consumer valuation of limited provider networks. *American Economic Review: Papers and Proceedings*, 105(5).

Ericson, K. M. and Starc, A. (2015b). Pricing regulation and imperfect competition on the massachusetts health insurance exchange. *Review of Economics and Statistics*, 97(3).

Ericson, K. M. and Starc, A. (2016). How product standardization affects choice: Evidence from the massachusetts health insurance exchange. *Journal of Health Economics*, 50.

Fan, Y. and Yang, C. (2020). Competition, product proliferation and welfare: A study of the u.s. smartphone market. *American Economic Journal: Microeconomics*, 12(2).

Ghili, S. (2020). Network formation and bargaining in vertical markets: The case of narrow networks in health insurance. *Working Paper*.

Glaeser, E. L. and Ponzetto, G. A. M. (2014). Shrouded costs of government: The political economy of state and local public pensions. *Journal of Public Economics*, 116.

Gowrisankaran, G., Nevo, A., and Town, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review*, 105(1):172–203.

Gruber, J. and Krueger, A. B. (1991). The incidence of mandated employer-provided insurance: Lessons from workers' compensation insurance. In Bradford, D., editor, *Tax Policy and the Economy 5*, pages 111–143. The MIT Press, Cambridge, MA.

Gruber, J. and McKnight, R. (2016). Controlling health care costs through limited network insurance plans: Evidence from massachusetts state employees. *American Economic Journal: Economic Policy*, 8(2).

Hall, M. A. and Fronstin, P. (2016). Narrow provider networks for employer plans. Technical report, Employee Benefit Research Institute.

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

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

Ho, K. (2009). Insurer-provider networks in the medical care market. *American Economic Review*, 99(1):393–430.

Ho, K., Hogan, J., and Morton, F. S. (2017). The impact of consumer inattention on insurer pricing in the medicare part d program. *RAND Journal of Economics*, 48(4).

Ho, K. and Lee, R. (2017). Insurer competition in health care markets. *Econometrica*, 85(2):379–417.

Ho, K. and Lee, R. (2019). Equilibrium provider networks: Bargaining and exclusion in health care markets. *American Economic Review*, 109(2).

Ho, K. and Lee, R. (2020). Health insurance menu design for large employers. *Working Paper*.

Kleiner, S., White, W. D., and Lyons, S. (2015). Market power and provider consolidation in physician markets. *International Journal of Health Economics and Management*, 15(1):99–126.

Lee, R. (2013). Markov-perfect network formation: An applied framework for bilateral oligopoy and bargaining in buyer-seller networks. *Working Paper*.

Liebman, E. (2018). Bargaining in markets with exclusion: An analysis of health insurance networks. *Working Paper*.

Liu, C. and Sydnor, J. R. (2018). Dominated options in health-insurance plans. *NBER Working Paper*.

LoSasso, A. T. and Atwood, A. (2016). The effect of narrow preovider networks on health care use. *Journal of Health Economics*, 50.

Lubotsky, D. and Olson, C. A. (2015). Premium copayments and the trade-off between wages and employer-provided health insurance. *Journal of Health Economics*, 44.

Marone, V. and Sabety, A. (2020). Should there be vertical choice in health insurance markets? *Working Paper*.

McKinsey Center for U.S. Health System Reform (2013). Hospital networks: Configuration on the exchanges and their impact on premiums.

Mohapatra, D. P. and Chatterjee, C. (2015). Price control and access to drugs: The case of india's malarial market. *Working Paper*.

Moran, J. R., Chernew, M. E., and Hirth, R. A. (2001). Preference diversity and the breadth of employee health insurance options. *Health Services Research*, 36(5).

Murray, J. G. (2010). Small group health premiums in massachusetts. Technical report, Commissioner of Insurance.

Nosko, C. (2014). Competition and quality choice in the cpu market. *Working Paper*, Forthcoming.

Office of the Attorney General Martha Coakley (2013). Examination of health care cost trends and cost drivers. Technical report, Office of the Attorney General Martha Coakley.

Pakes, A. (2010). Alternative models for moment inequalities. *Econometrica*, 78(6).

Pakes, A., Porter, J., Ho, K., and Ishii, J. (2015). Moment inequalities and their applications. *Econometrica*, 83(1).

Pauly, M. V. (1968). The economics of moral hazard: Comment. *American Economic Review*, 58.

Polsky, D. and Weiner, J. (2015). The skinny on narrow networks in health insurance marketplace plans. Technical report, Leonard Davis Institute of Health Economics and Robert Wood Johnson Foundation.

Polyakova, M. (2016). Regulation of insurance with adverse selection and switching costs: Evidence from medicare part d. *American Economic Journal: Applied Economics*, 21(3).

Prager, E. (2016). Tiered hospital networks, health care demand, and prices. *Working Paper*.

Qin, P. and Chernew, M. (2014). Compensating wage differentials and the impact of health insurance in the public sector on wages and hours. *Journal of Health Economics*, 38.

Raval, D., Rosenbaum, T., and Wilson, N. E. (2019). Using disaster induced closures to evaluate discrete choice models of hospital demand. *Working Paper*.

Scheffler, R. M., Arnold, D. R., Fulton, B. D., and Glied, S. A. (2016). Differing impacts of market concentration on affordable care act marketplace premiums. *Health Affairs*, 35(5).

Shepard, M. (2016). Hospital network competition and adverse selection. *NBER Working Paper Series, 22600*.

Song, M. (2007). Measuring consumer welfare in the cpu market: An application of the pure characteristics demand model. *Rand Journal of Economics*, 38(2):429–446.

Summers, L. H. (1989). Some simple economics of mandated benefits. *American Economic Review*, 79(2):177–183.

Train, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

For Online Publication

A Data Descriptions

A.1 APCD Sample Creation

Hospital Admissions: The first sample is the sample of hospital admissions, which I use to estimate the patient demand for hospitals, described in more detail in subsection C.1. To construct this data, I limit the APCD to any facility claim flagged as an inpatient admission between the five-year sample period and to any hospital that is located within the state of Massachusetts. I therefore exclude any admission of patients receiving hospital care outside the state (regardless of whether the patient resides in Massachusetts or not). For each hospital, I used the organization's National Provider Identification (NPI) number to match the hospital to a set of hospital characteristics from the American Hospital Association (AHA) Annual Survey. These characteristics include the type of hospital (teaching, critical-access, academic medical center, specialty, etc.) and hospital amenities (including number of beds and types of services offered). The data is aggregated to the hospital admission level, and the "allowed amounts" are summed over all service-lines for that particular admission, in order to construct a price-per-visit. For each admission, I link the primary diagnosis (ICD-9 code) to a set of Chronic Conditions Indicators (CCI) and Clinical Classifications Software (CCS) categories. These are indicators provided by the Agency for Healthcare Research and Quality (AHRQ) that allow me to aggregate diagnosis codes into a set of 18 distinct groups, and also to flag which patients suffer from chronic conditions.[60]

Table A.1: Hospital Sample Summary Statistics

	Mean	Std Dev
Patient Characteristics		
Age	52.14	25.98
Female	0.58	0.49
Chronic	0.53	0.49
Neurological	0.02	0.15
Cardiac	0.16	0.37
Obstetrics	0.22	0.42
Imaging	0.27	0.44
Hospital Characteristics		
Distance	9.95	12.06
NICU	0.87	0.33
Neuro	0.96	0.19
MRI	0.90	0.30
Critcal Access	0.01	0.08
Teaching	0.74	0.44
Specialty	0.02	0.14
Academic Medical Center	0.25	0.43
Would Recommend	0.74	0.12

Notes: Hospital sample summary statistics 2009-2013. Diagnosis characteristics (e.g. "Neurological," "Cardiac," etc.) are derived from AHRQ's Clinical Classification Software categoreis and Chronic conditions Indicators indicators.

Table A.1 contains the hospital sample summary statistics for hospital admissions from 2009-2013. On average, patients admitted to Massachusetts hospitals are 52 years old, and about half of the patients suffer from a chronic condition. Approximately 16% of patients are admitted with a primary cardiac condition,

[60] https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccsfactsheet.jsp

while about 22% are admitted with an obstetrics-related diagnosis. Patients are, on average, willing to travel approximately 10 miles to visit a hospital, and visit teaching hospitals approximately 74% of the time, while visiting academic medical centers approximately one-quarter of the time.

Physician Visits: The second constructed sample from the APCD is used to estimate the physician demand portion of the model. I construct it by limiting the data to professional claims only. These capture reimbursements specifically to medical providers that are separate from reimbursements for facilities, even though the particular service may have been performed in a facility. This includes patient visits to independent offices, larger medical groups, or non-inpatient visits to hospitals, outpatient centers, or clinics within hospitals such that a separate claim is generated to pay individual physicians. The data is then merged with SK&A data on physician affiliations (described in more detail below), and each individual practitioner is assigned to their primary medical group. After constructing these practice groups, I then stratify the data into three different specialty groups: primary care physicians (PCPs), cardiologists, and orthopedists. Primary care practices are defined as any medical group that contains at least one physician that is either an internist, general practitioner, family practice doctor, or geriatric doctor. Similarly, cardiology practices and orthopedic practices are defined as any practice that employs at least one physician of the relevant specialty. I consider these three specialties in order to capture three different component of medical care: primary care, which is the most common type of visit to a health care provider (at about 55% of all office visits), medical specialty care (exemplified by cardiology), and surgical care (exemplified by orthopedics).

For each service-line, I merge in Medicare Part B physician fee schedules from Center for Medicare and Medicaid Services (CMS).[61] These data contain annual federal updates to each procedure (CPT) code's "Relative-Value-Unit" (RVU) weight, which are constructed in order to assign each service an approximate measure capturing its relative intensity to other procedures. These weights are then used to determine Medicare payment rates. Specifically, each year CMS releases updates to its Medicare "conversion' factor" and to its RVUs. The "conversion factor" reflects the base Medicare payments per RVU that it pays to physicians in a given year. This factor is then scaled by the RVU for a particular procedure to determine the physician reimbursement.[62] I aggregate the data to the patient-visit level, summing over all the RVU weights of each service provided during a visit and summing over all the "allowed amounts" for each service to determine a total payment per visit and total RVUs performed per visit. I also use these RVUs in construction of insurers' negotiated rates with physician practices, described further in subsection C.6.

Table A.2 shows summary statistics for the physician samples. On average patients going to see PCPs are younger and have a higher likelihood of being female than those going to cardiologists, though patients seeing orthopedists tend to be the youngest on average. Average RVUs for orthopedic services are higher than for PCPs and cardiologists, with significantly higher standard deviations. This reflects the fact that while orthopedists often perform routine office-based procedures, they also perform surgeries which are more resource intensive and thus are assigned higher RVUs. About 57% of primary care patients saw a doctor between 2009 and 2013 that they also have seen previously, while this number was about 63% for cardiologists and about 60% for orthopedists. Distance traveled to any of the specialty groups are all about 10 miles, with the distance to see PCPs somewhat shorter. When seeing a PCP, patients on average visit practices with 26 doctors on site, whereas this number is significantly higher for orthopedic practices and, especially, for cardiology practices. Moreover, patients tend to visit cardiology practices with a greater number of locations and that disproportionately tend to be part of medical groups, owned by hospitals, or owned by health systems.

[61] https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/PFS-Relative-Value-Files.html.

[62] As an example, if the Medicare Conversion Factor for a given year is \$36, a procedure performed in that year with an RVU of 2 will receive a total reimbursement of \$72.

Table A.2: Physician Sample Summary Statistics

	PCPs	Cardiologists	Orthopedists
Age	47.64	54.09	44.34
	(15.59)	(13.89)	(18.51)
Female	0.56	0.43	0.52
	(0.50)	(0.49)	(0.50)
RVU	2.64	2.95	5.55
	(1.79)	(4.89)	(12.54)
Used Doc Previously	0.57	0.63	0.60
	(0.49)	(0.48)	(0.49)
Used Med Group Previously	0.58	0.70	0.64
	(0.49)	(0.46)	(0.48)
Used System Previously	0.59	0.73	0.66
	(0.49)	(0.44)	(0.47)
Distance	7.21	9.67	9.83
	(9.18)	(11.13)	(10.59)
Doctors on Site	26.44	97.51	49.66
	(75.80)	(159.11)	(109.73)
Number of Locations	6.93	7.79	4.18
	(7.56)	(7.94)	(6.27)
Part of Medical Group	0.60	0.67	0.59
	(0.49)	(0.47)	(0.49)
Owned by Hospital	0.22	0.32	0.16
	(0.22)	(0.46)	(0.37)
Owned by System	0.40	0.45	0.24
	(0.49)	(0.50)	(0.43)

Notes: Physician sample summary statistics for select variables for primary care physicians, cardiologists, and orthopedic surgeons 2009-2013. For practice characteristics (e.g. "doctors on site", "number of locations," etc.), these estimates reflect means and standard deviations weighted by patient visits. In other words, "Doctors on Site" reflects the number of doctors at a particular practice location weighted by patient visits to that practice.

GIC Member Data: The final subsample constructed is a sample of GIC members by year, which is used to estimate the insurance demand portion of the model. In addition to claims data, the APCD contains an enrollment file, where each insurer provides a list of each of its enrollees by market, plan, and year. These files also come with a rich set of enrollee demographics, including 5-digit zip code, age, gender, employer industry code, employer zip code, monthly plan premium, annual plan individual and family deductible, enrollment start date, and enrollment end date. I limit this file to all enrollees who are part of the GIC between 2009 and 2013. The file also allows me to link individual enrollees to their family members when estimating insurance demand. Finally, I merge this list of GIC members to external data on GIC annual plan premiums and hospital networks. An advantage of studying this particular market is that plan premiums are the same for each member across the state, and only vary by family type ("Individual" versus "Family"). Each year, the GIC publishes these premium rates for each family type. It also publishes an annual list of the hospitals included in each plan's network for each of the commission's narrow-network plans. I merge this public information onto the enrollee dataset in order to obtain a full set of plan characteristics for each enrollee. For the year 2012, the year of the premium holiday, I assume that each active employee under the age of 65 pays only 9 of the 12 months of the annual premium if they switch to a narrow-network plan in that year.

A.2 SK&A Sample Creation

Matching Physicians to Practices: Given the breadth of the data as well as the inconsistencies in reporting between the APCD and SK&A, linking the two datasets involved several steps. First, I matched every available physician in the SK&A to the APCD via the NPI variable and provider zip-code variables in each dataset. This ensures that all the matches were not only to the correct physician, but also to the correct practice location for each physician. In cases where this did not match, I then matched only by the NPI and assumed that the closest location in the SK&A to that where the service was rendered in the APCD was the correct practice.

However, not all insurers in the APCD report physician NPIs, opting instead to bill using the organizational NPI. For instance, Health New England only reports the NPI for the hospital or medical group when processing claims. Given that the SK&A only contains individual doctors' NPIs, in instances where this occurs, I conduct an iterative string-matching algorithm to match the medical practice data. I use the first and last name fields in the APCD and match the provider's names and zip codes to the names and zip codes from the SK&A. For all records that did not match, I then match only by first and last name. Then I repeat this just for last name and zip code. These set of steps allowed me to match approximately 80% of the claims from the APCD to an appropriate physician from the SK&A.

After completing this procedure, I define two different variables. The first is a "practice" variable, which is the unit used in the demand analysis. This variable refers to any particular physician-practice-location triple in the data that billed more than 50 claims in any particular year. If a physician was not reported as being employed by a medical group in the SK&A, I consider the physician-hospital-location triple as the practice definition. These are physicians who are employed by hospitals but may be billed for physician services separately (for example if they take outpatient or office visits in the hospital clinic). If there is no medical group or hospital reported, I consider this variable to be just the physician-location double, and assume the physician is a solo-practitioner. I assume that when selecting a physician, individuals choose at this "practice" level.

The second variable I define is an "ownership" variable, which is used in defining networks. This refers to the highest level of vertical integration for the physician. If a particular physician's highest reported ownership in the SK&A is a medical group, then this variable is coded as the group. If the highest level of ownership is a particular hospital (i.e. a hospital-owned physician practice), then this variable is coded as the hospital. Finally, if the highest level of ownership is reported as a health system (e.g. Partners Health Care, Steward Health System), then this variable is coded as the system. In considering counterfactual networks that the employer could offer, I make the assumption that the insurers contract at the "ownership" level. Therefore, if the employer chooses to eliminate a Partners physician, it must eliminate all physicians employed by the Partners health system.

I then assign each physician a specialty according to the specialty reported in either the APCD or the SK&A. For example, if a particular physician is reported as a cardiologist in either dataset, I flag that physician as a cardiologist. I consider any practice a cardiology practice if it employs at least one physician flagged as a cardiologist, or if the SK&A reports that the practice is a cardiology practice.

Constructing Physician Practice Networks: The final task involves determining which physician practices are in a particular insurance plan's network. While some GIC insurers actually report the medical groups that they cover in their narrow networks (i.e. Fallon), others only report the list of hospitals. I therefore assume for simplicity that if a particular hospital is excluded from a particular plan's network, then any physician, physician practice, or medical group that is owned by that particular hospital is also excluded from the network. Similarly, as bargaining between insurers and providers is typically done as the *system* level, I assume that if any particular system is excluded from a plan's network in its entirety (e.g. if a particular plan excluded all Partners hospitals), I assume that any physicians or groups that are owned by Partners (even though they may not be affiliated with any particular hospital) are also excluded.[63] For

[63]In practice, this is a close approximation of contracts observed on the GIC. Harvard Primary Choice and Tufts Spirit, for instance, cease contracting with all Partners-owned medical groups as well as Partners hospitals. The exception appears to be for Fallon Direct, which does contract with certain Partners-affiliated medical groups (e.g. Charles River Medical Associates) and certain Atrius-affiliated groups (e.g. Reliant Medical Group). Fortunately, Fallon reports these covered groups on its website and, as such, I was able to incorporate them into the network

any large medical group that is not affiliated with a particular hospital or system, I conduct manual checks on the insurers' websites to see whether these groups are covered by the plans. For all remaining practices, if they are not owned by any hospital or system, I use the claims to infer whether the practices are in a particular plan's network. In particular, I assume that any practice that has more than 10 in-network claims from a particular plan is considered in-network. For robustness, I also construct networks that default each each of these small practices as being in-network unless a majority of claims that are processed for these practices by a particular plan is flagged as being "out of network."

---

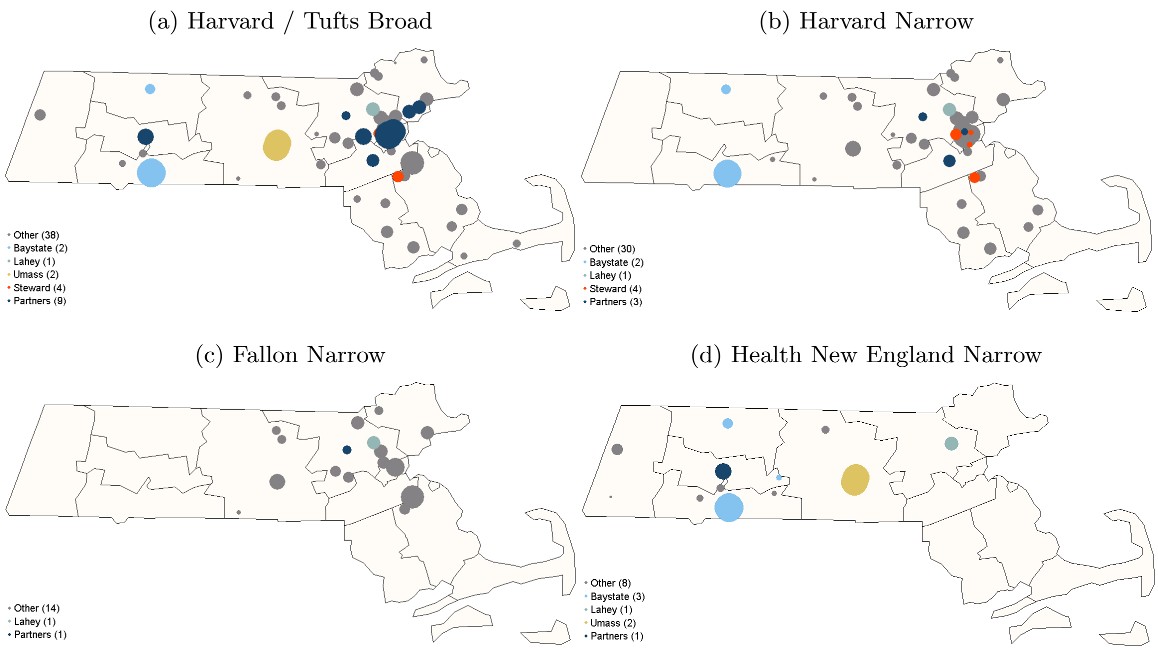
structure.

B Additional Descriptives

B.1 Additional Network Figures

Below, I present figures depicting the hospital, PCP, cardiology, and orthpedic practice network coverage across Massachusetts of Harvard and Tufts Broad, Harvard Narrow, Fallon Narrow, and HNE Narrow.

Figure B.1: Hospital Networks by Plan, 2011



(a) Harvard / Tufts Broad

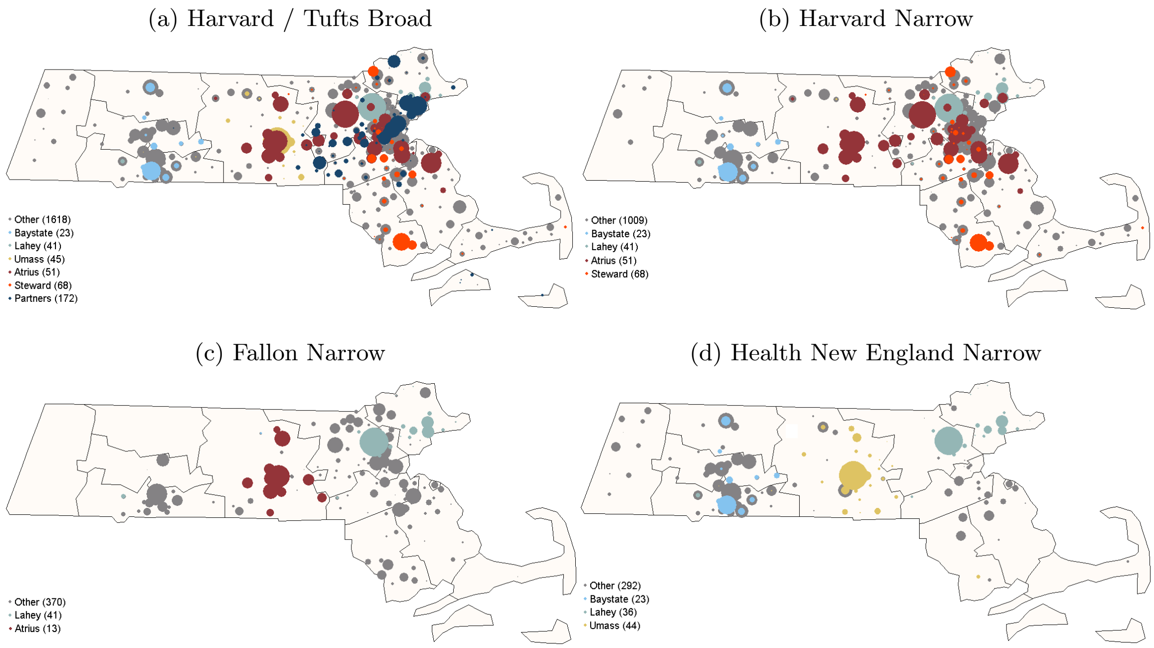
(b) Harvard Narrow

(c) Fallon Narrow

(d) Health New England Narrow

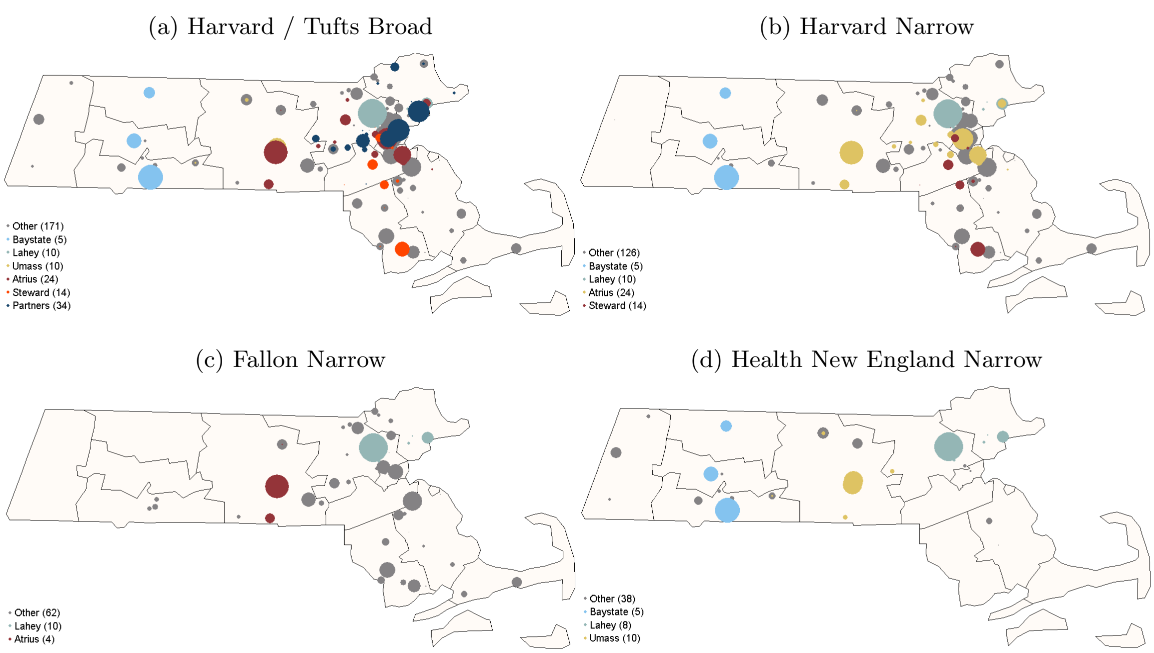
Notes: This figure plots the hospital networks of specified plans on the GIC in 2011. Sizes of the data points reflect relative market shares of the practices. Colors reflect ownership status (which health systems owns which practice).

Figure B.2: Primary Care Practice Networks by Plan, 2011



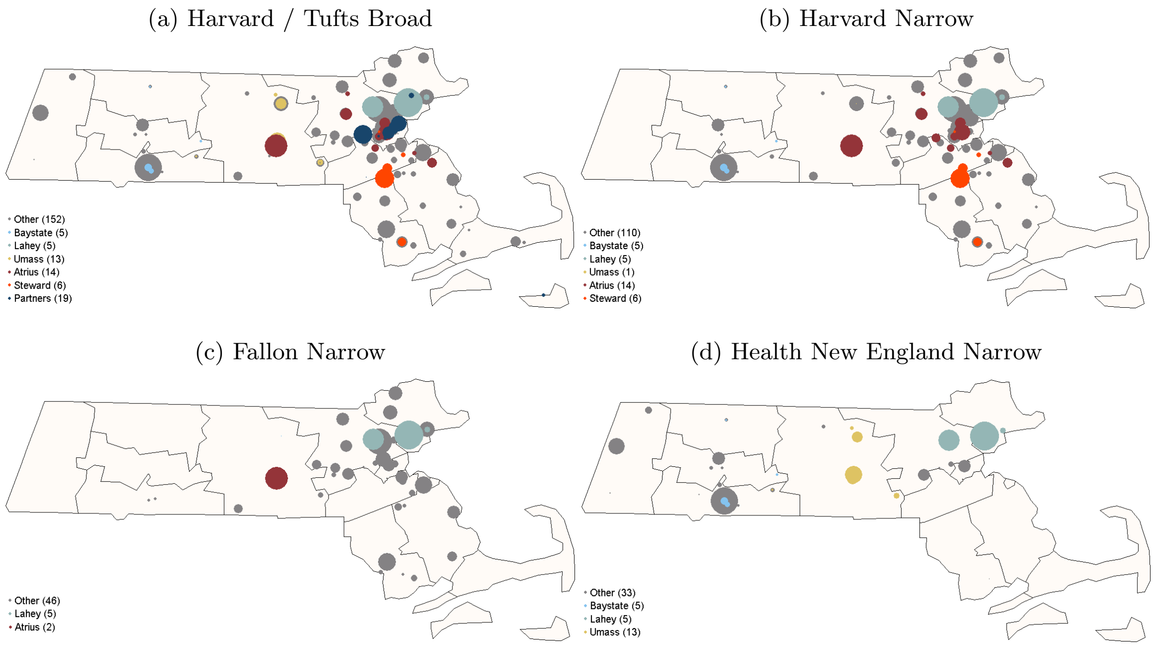
Notes: This figure plots the primary care practice networks of specified plans on the GIC in 2011. Sizes of the data points reflect relative market shares of the practices. Colors reflect ownership status (which health systems owns which practice).

Figure B.3: Cardiology Networks by Plan, 2011



Notes: This figure plots the cardiology practice networks of specified plans on the GIC in 2011. Sizes of the data points reflect relative market shares of the practices. Colors reflect ownership status (which health systems owns which practice).

Figure B.4: Orthopedic Networks by Plan, 2011



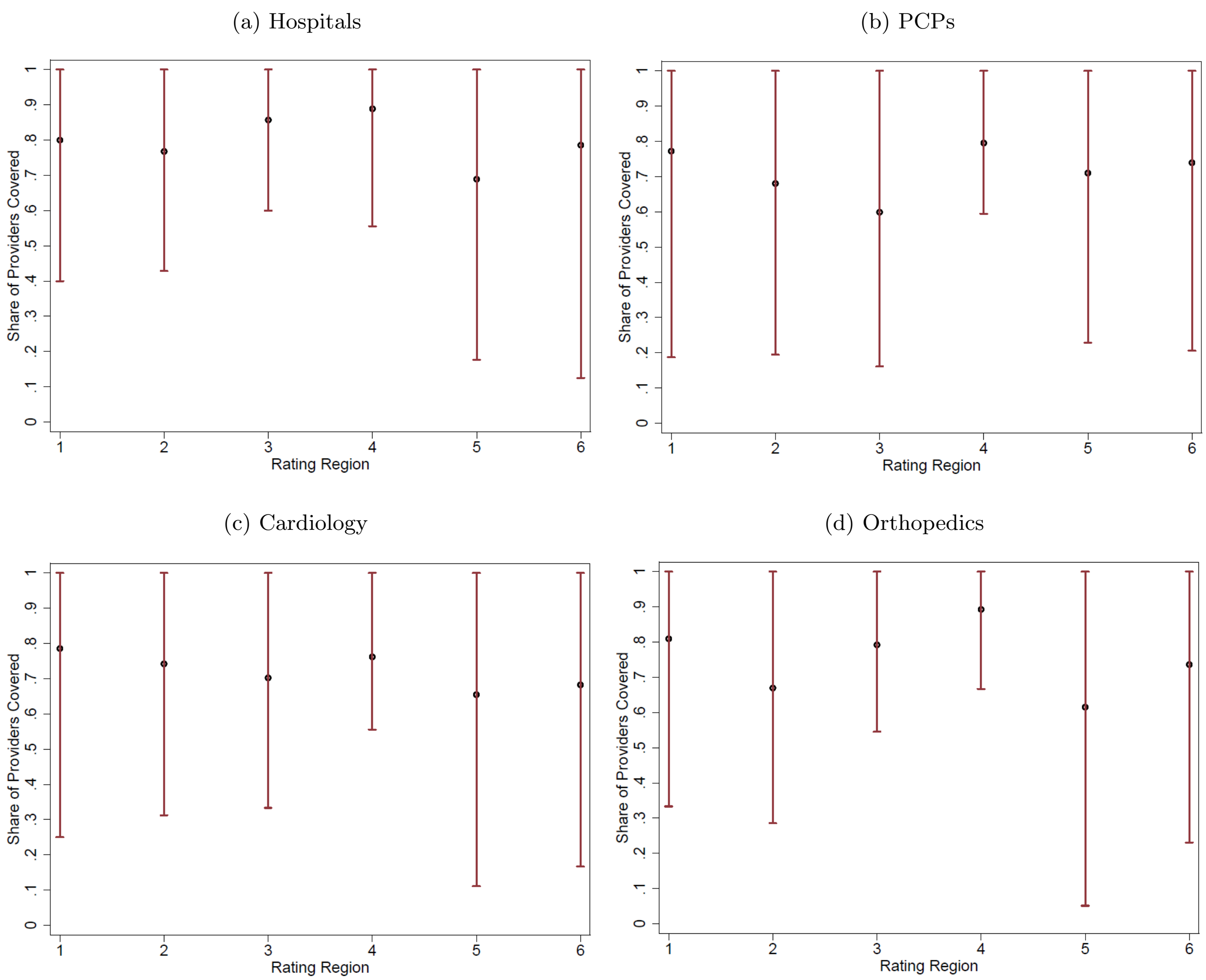
Notes: This figure plots the orthopedic practice networks of specified plans on the GIC in 2011. Sizes of the data points reflect relative market shares of the practices. Colors reflect ownership status (which health systems owns which practice).

To dive deeper into the variation in networks by geographic market across plans, Figure B.5 plots the variation in hospital and physician networks across plan and rating region in Massachusetts in 2011.[64] The y-axis represents the share of providers operating in the rating region that each plan covers in-network (hereafter referred to as "network breadth"). The dots represent the average network breadth across the plans on the GIC that operate in the respective rating regions, and the bars represent the range of network breadth in that region. For physicians, providers were limited to just the top 50 practices (by number of claims) in each rating region, to avoid measurement error. For each specialty, there is considerable variation in network breadth, both across and within rating region. Across rating regions, *average* network breadth for PCPs, for instance, ranges from about 60% to about 80% depending on the region. Within rating region, the broadest plans cover virtually all the top 50 practices and hospitals in the region, while the narrowest plans cover only about 20% of the providers. In Rating Region 5—the region including Boston—average network breadth for hospitals, cardiologists, and orthopedists is quite low, reflecting the fact that many of the narrow-network plans exclude providers in the Boston region. Noticeably, the narrowest plan operating in the region only covers about 10% of the top 50 orthopedic practices in the region.

The network shares also vary over time in addition to across plan and region. Figure B.6 displays the average network breadth over time. The y-axis here represents the share of the state's hospitals and physicians practices covered, averaged across all plans operating statewide. This again limits the data to only the top 50 physician practices for each specialty in each rating region. While hospital networks remain fairly stable over time, with the exception of a small uptick in 2013, the network breadth for the three physician specialty groups considered seem to be increasing over time, ranging from about 50% of physicians covered in 2009 to between 65 and 70% coverage in 2013. This change is driven primarily by three factors. First, during this time period there were some physician exits, as well as mergers between physician groups that resulted in a change in network status. Second, during this period there were significant hospital acquisitions

[64] Rating regions are defined according to CMS definitions: https://www.cms.gov/CCIIO/Programs-and-Initiatives/Health-Insurance-Market-Reforms/ma-gra. Rating Region 7 (Cape Cod) is omitted from analysis due to the low number of households on the GIC residing in this region.

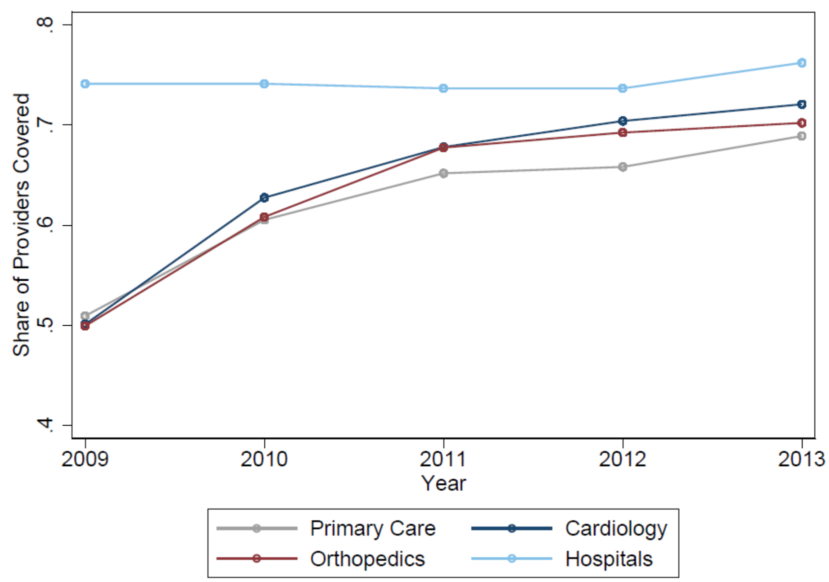
Figure B.5: Share of Providers Covered by Rating Region and Specialty, 2011



Notes: This figure plots the share of all hospital and physician practices covered by each plan on the GIC by rating region. Each dot represents the average share of providers in the respective rating covered across all plans operating in those regions. Red bars represent the range of coverage across plans in that region. For PCP, cardiology, and orthopedic networks, data is limited to the top 50 practices (by number of claims) in each rating region.

of physician practices. Third, certain narrow plans over time began covering more groups.[65]

Figure B.6: Share of Providers Covered by Year and Specialty



Notes: This figure plots, by year, the unweighted average share of all hospital and physician practices covered across plans operating statewide on the GIC. For PCP, cardiology, and orthopedic networks, data is limited to the top 50 practices (by number of claims) in each rating region.

B.2 Additional Evidence of Inertia

In Table B.1 I present a regression of enrollment in narrow-network plans on a set of household observables, as well as an indicator for whether the household was new to the GIC that year. Indeed, older households are less likely to enroll in a narrow-network plan, as are households with at least one member with a chronic illness. Larger households are also less likely to enroll in a narrow-network plan. However, even controlling for these, as well as year and county fixed effects, existing members of the GIC are, on average, 11% less likely to be enrolled in a narrow-network plan than new members.

To see more evidence that new cohorts behave differently than older cohorts, one need not look only at enrollment in broad versus narrow-network plans, but also at the stickiness of enrollment in broad-network plans as the characteristics of those plans change. To that end, I note that in 2010, the premiums for Harvard and Tufts were fairly similar, while beginning in 2011, the premium difference between the two plans began to rise thereafter, with Harvard Broad growing significantly more expensive than Tufts.

Figure B.7 shows the ratio of enrollment in Harvard Broad versus Tufts Broad over time, along with the change in the ratio of Harvard Broad premiums to Tufts Broad premiums. The black line represents the Harvard-to-Tufts enrollment ratio for new members to the GIC, while the light grey line represents the Harvard-to-Tufts enrollment ratio for existing GIC members. First, it is notable that as Harvard's premiums rise relative to Tufts', enrollment in Tufts relative to Harvard rises dramatically among new members to the GIC. By 2012, Tufts' premiums were about 10% less than Harvard's (representing about $30 per month for families). Enrollment in Harvard among new members, meanwhile, declined from more than three times

[65] As an example of the first phenomenon, the Atrius Health system gradually purchased several prominent medical groups (including Harvard Vanguard and the Fallon Clinic) which were previously separate entities. As a result, plans that may have covered some, but not all, of these practices, began covering all of them under the Atrius umbrella. As an example of the third phenomenon, Fallon Health Plan did not cover the Partners system (including its physicians) until 2013.

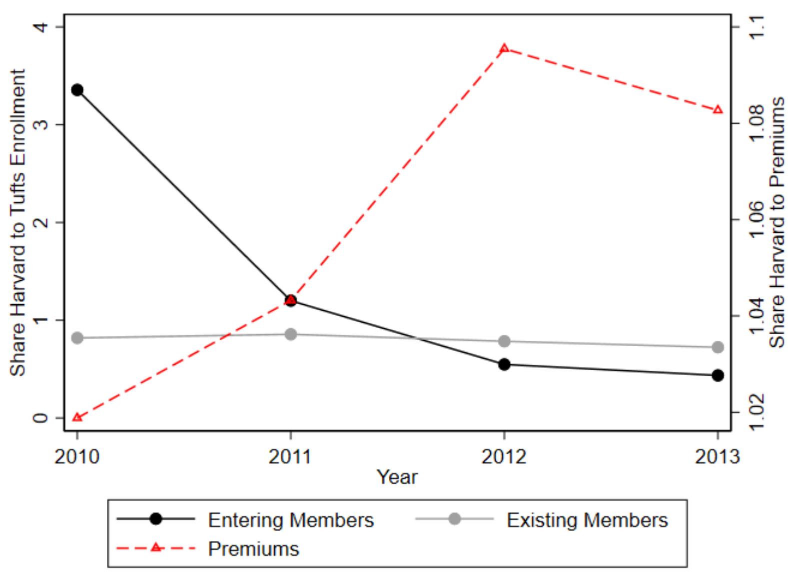
Table B.1: Probability of Enrolling in a Narrow Plan

Variable	Coefficient	Standard Error
Existing GIC Member	-0.113***	0.003
Age	-0.003***	0.000
Female	0.009***	0.002
Chronic Condition	-0.021***	0.003
Members in HH	-0.009***	0.003
Constant	0.670***	0.006
Year FE	Yes	
County FE	Yes	
Obs.	151,331	
Adj R2	0.432	

Notes: Results from regression of enrollment in a narrow network plan on household characteristics. GIC sample 2009-2013.

that of Tufts in 2009 to about 90% that of Tufts in 2012. Second, existing members exhibit no such changes in enrollment patterns. Between 2010 and 2013, enrollment among existing members in Harvard relative to Tufts barely budged, even as the premium difference widened considerably.

Figure B.7: Share of Members Enrolling in Tufts Broad Network Plan by Whether New to GIC



Notes: This figure plots the share of ratio of members selecting Harvard's broad-network plan over Tufts' broad-network plan as well as the ratio of the premium difference between Harvard and Tufts. The dark line plots the ratio of entering (new) members to the GIC that year. The light grey line plots the ratio of existing members on the GIC. The dashed red line plots the premium ratios.

C Model Details

C.1 Provider Demand Estimation

Market Shares: The probability that patient i and diagnosis l will choose hospital h in time t is given by:

$$\sigma_{ilht} = \frac{exp(\phi_{ilht})}{\sum_{k=1}^{N_{ijt}^{H}} exp(\phi_{ilkt})} \tag{20}$$

where N_{it}^{H} refers to the number of hospitals in individual i's network in time t. Similarly, the probability that patient i needing a procedure with RVU r from specialist group s will chose physician practice d is:

$$\sigma_{ird}^{s} = \frac{exp(\phi_{irdt}^{s})}{\sum_{k=1}^{N_{ijt}^{S}} exp(\phi_{irkt}^{s})} \tag{21}$$

where N_{ijt}^{S} is the network of practices of type s in individual i's network.

Estimation: The patient choice of providers is estimated using maximum likelihood. Estimation of hospital demand follows techniques standard in the literature (Ho, 2006). For estimating the physician models, I make additional assumptions in order to reduce the dimensionality of the estimation, described below. Further, I estimate the models separately by the seven Massachusetts health rating regions[66] and by specialty group (PCP, cardiology, and orthopedics).

The model includes patient characteristics interacted with provider characteristics, travel time interacted with both patient and provider characteristics, and a full set of provider fixed effects (interacted with diagnosis/procedure intensity weights) in order to account for unobserved heterogeneity across the providers in the data. The patient characteristics include five-digit zip code, age, an indicator for female, patient diagnosis (in the case of hospital care), patient procedure required (in case of physician care), and whether the patient has ever been treated for a chronic condition.

For hospital care, patient diagnoses, l, are grouped into 18 Clinical Classification Software (CCS) categories. Chronic conditions are grouped according to HCUP indicators mapping chronic conditions from ICD-9 diagnosis codes. Given that my data span 2009-2013, I define patient i in time t as having a chronic condition if that patient has gone to see any provider at any time prior to t for a diagnosis that is considered to be "chronic." Each of the 18 diagnosis categories are further assigned numerical weights that proxy for the intensity of the particular diagnosis (the construction of these weights follow closely to work by Shepard (2016); a discussion of their construction can be found in subsection C.6). Hospital characteristics include location, number of beds, whether the hospital had a NICU, whether the hospital provided imaging services (including an MRI), and whether the hospital included a catheterization lab. I include indicators for whether the hospital is a critical access hospital, a teaching hospital, a specialty hospital (such as cancer center or children's hospital), or whether the hospital is an academic medical center. I further interact these hospital characteristics with each of the 18 disease categories. In addition, I include a full set of hospital fixed effects in the model to account for any unobserved quality components of hospitals not captured by the model. In order to capture additional heterogeneity, I interact these fixed effects with the numerical weights for the patient diagnoses, in effect allowing patients with different disease severities to prefer seeking care from different hospitals.

For patients requiring care from physicians, I match procedures performed (CPT codes) to Medicare

[66] The rating regions are detailed in https://www.cms.gov/CCIIO/Programs-and-Initiatives/Health-Insurance-Market-Reforms/ma-gra.html.

RVU weights, r, which serves as a proxy for procedure intensity. For physician practice characteristics, I include a number of variables from the SK&A including the number of doctors at the particular practice's location, the number doctors across *all* the practice's locations, the share of the doctors at the practice who are specialists (relative to PCPs), whether the practice is part of a medical group, whether the practice is owned by a hospital or health system, and the number of total locations of the medical group. I interact each of these with patient characteristics, including the patient's RVU weight, as well as with the 18 CCS diagnosis categories. I also include a full set of practice fixed effects within each specialty group, and interact those fixed effects with RVU weights.

To capture physician inertia, I include three separate indicators: whether a patient had sought care from this particular physician practice previously; whether a patient had sought care from any of the practice's locations previously; and whether a patient had previously sought care from any provider employed by the hospital or health system that owns the particular practice. I interact each inertia variable with a proxy measure for the length of a particular patient-provider relationship. To construct this, I infer from the claims the earliest visit a particular patient had with a particular provider, and calculate the number of years to the present day.[67]

I run the model separately for hospitals, PCPs, cardiologists, and orthopedists. I assume these all can be thought of as separate markets that do not compete with one another. For instance, patients who require a procedure for knee surgery would be unlikely to select a cardiology practice for that procedure. One limitation of this approach is that it abstracts away from referral networks across specialties and between physician groups and hospitals.[68]

Dimensionality Reduction Perhaps the most salient issue in estimation of the physician models is the presence of tens of thousands of physicians within each specialty group in Massachusetts, making estimation of parameters through a multinomial logit framework difficult. I take three primary approaches to reduce the dimensionality problem. The first is that, as previously described, I estimate the provider demand model at the physician *practice*-zip-code level rather than the individual physician level. This reduces the patient choice set considerably. Second, I estimate the model separately by the seven rating regions in Massachusetts, as defined by CMS.[69] As individual practices are location-specific, this allows me to include a larger span of the full Massachusetts physician practice space in my estimation. In addition, it allows for estimation of flexible parameters that vary by region.[70]

Finally, I assume that only the top 50 practices (by market share) within each region and specialty group have an individual mean utility. All practices outside the top 50 are assumed to have identical mean utilities and only be differentiated on distance to the patient. In order to further narrow the choice set, I assume that practices outside the top 50 in a region can be grouped into a set of 7 discrete distance bands, b, where $b = 0$ to 5 miles, 5 to 10 miles, 10 to 15 miles, 15 to 30 miles, 30 to 50 miles, 50 to 100 miles, and over 100 miles. I assume that the distance between any given patient and physician practice, T_{id}, is constant within each of these bands and takes the value of the midpoint of the distance band, i.e. $\{T_{id} \in b\} = b^{mid}$.[71] Given these assumptions, and dropping the region and time subscripts for convenience, the model in Equation 1

[67] Indeed, this results in some measurement error, as the data is censored at the start of the sample period in 2009. Nevertheless, it provides an approximation as to how length of relationship influences stickiness.

[68] Indeed, patients often seek care initially from their PCPs, who may subsequently refer them to a cardiologist or orthopedist. My model, by treating these specialty groups as independent, does not capture these behaviors. This may bias the parameter estimates, particularly in the hospital and specialist models (unlikely, however, in the primary care model) as choice may be driven not by, say, distance, but by the recommendation of a previously used provider. Future work aims to quantify these physician referral networks, and to see how these drive demand for different specialties.

[69] https://www.cms.gov/CCIIO/Programs-and-Initiatives/Health-Insurance-Market-Reforms/ma-gra.html

[70] For instance, it is likely that individuals in Boston would be more averse to traveling for physician care than individual in Worcester, due to the density of patients and providers in the former relative to the latter.

[71] As an example, $b^{mid} = 2.5$ for distance band b =0 to 5 miles.

becomes:

$$u^s_{ird} = \underbrace{\phi^s_{ird} + \varepsilon^s_{ird}}_{\text{Utility for Top 50 Practices}} \tag{22}$$

$$u^s_{ird} = \underbrace{\sum_b \mathbb{1}\{T^s_{id} \in b\}(T^s_{ib}\lambda^s_1 + T^s_{ib}v_{ir}\lambda^s_2 + N^s_{ib}\gamma^s_b) + \varepsilon^s_{irb}}_{\text{Utility for Practices Outside Top 50}} \tag{23}$$

where N^s_{ib} is the number of physicians of specialty s in individual i's network in distance band b. This specification can be thought of as adding a single option to the choice set for each distance band b, rather than an individual option for each physician practice in those distance bands. γ^s_b, then, rather than estimating a fixed effect for each individual practice $d \in b$, simply estimates a fixed effect for each distance band b and scales it by the number of physicians in that band. This allows patient valuations of these options to vary by the number of doctors in those groups. As an example, if patient i's physician network removed a physician practice in distance band b, patient i's utility would decrease by $(N^s_{ib} - 1)\gamma^s_b$.

The assumption that practices outside the top 50 have the same mean utility conditional on distance bands may seem like a strong one. However, it makes sense given two empirical facts. First, the top 50 practices by market share in a given region account for most of patient claims.[72] Second, most practices outside the top 50 are included in all plans' networks, even narrow-network products. As a result, most of the variation in networks across plans comes from network choice among these top practices. Therefore, treating these smaller practices as essentially undifferentiated in quality (but for distance) not only has the benefit of making the model more easily estimable, but also likely to hold true given observed networks.

Outside Option: For the hospital choice model, I define the outside option to be any hospital outside the state of Massachusetts. For the physician models, I assign any physician practice in distance band $b = 7$ (i.e. outside of 100 miles from the patient's location) to be the outside option. I normalize these goods to be 0 in the utility models.

Identification: Each of the coefficients are identified through within-provider variation in patient characteristics. The parameter on distance, for example, is identified by differences in choice of a particular provider across patients who live in different zip-codes throughout Massachusetts. The identifying assumption is that patient choice of where to live is orthogonal to their preferences for providers.

Identification of the inertia coefficient, λ^s_5, relies on differences in choices made between patients who have never sought care from *any* physician within a particular specialty group and patients who previously sought care from a physician, conditional on other observables included int he model. I abstract away from decomposing the extent to which λ^s_5 is driven by true switching costs as opposed to unobserved preference heterogeneity. In particular, persistence in physician choice may be driven by three factors: physician-patient capital accumulated through repeated interactions (i.e. the patient *develops* utility for a particular physician ex-post); unobserved physician quality (i.e. the patient stays with the physician for factors unobserved to the econometrician); and true switching frictions or hassle costs irrespective of physician quality. In my setting, I choose to focus on the most conservative interpretation of physician inertia possible: that λ^s_5 entirely reflects physician-patient capital. In counterfactual exercises, when patients lose access to their previously used physicians or practices, I therefore treat this as a "welfare-relevant" utility loss.[73] However, to test the robustness of this, I also present estimates of the employer objective function in Appendix E that treat the

[72]In Boston, for instance, where there is the highest density of physicians, the top 50 PCP practices account for approximately 70% of all claims, while the top 50 cardiology and orthopedic practices account for nearly 90% of all claims.

[73]Shepard (2016) discusses this issue in detail in his context of hospital inertia.

inertia term as coming from the other two sources.

C.2 Hospital Demand Estimates

Table C.1 reports the results for the hospital demand model. Column 1 displays the main results, which are run on the full sample of hospital admissions in Massachusetts for consumers on the GIC between 2009 and 2013. Consistent with prior literature on hospital demand, the distance coefficient is negative and significant, implying that patients prefer to go to hospitals that are close to where they live. While this coefficient is difficult to interpret (the measure is in utils instead of a dollarized amount), comparing this coefficient with other parameter estimates shed some light on its practical magnitude. For instance, the estimates imply that hospital patients are on average willing to travel approximately 20 extra miles to reach the hospital with the highest unobserved quality parameter (i.e. the largest fixed effect estimate). This is indicative of the fact that patients are "willing-to-pay" in terms of extra miles traveled to access prestigious, academic medical centers, such as Mass. General and Brigham and Women's (both owned by Partners), Beth Israel, Lahey Medical Center, and others.

In addition, these models incorporates more flexible distance coefficients interacted with county identifiers in Massachusetts. This is done in order to allow patients to react differently to distance traveled to a particular hospital depending on where in Massachusetts they reside. Coefficients are for Barnstable county (the omitted variable), Worcester (Central Massachusetts), Hampden (Western Massachusetts), and Suffolk (Eastern Massachusetts). The distance coefficients are negative and significant in all reported counties. Notably, patients are far less reactive to distance in Barnstable, Hampden, and Worcester (where they are more likely to drive by car to find a hospital) than they are in Suffolk (which contains metropolitan Boston).

A second important finding concerns the large positive and significant coefficient on individuals who have used the hospital in the previous period. This "willingness-to-travel" to a hospital the patient has previously used varies by county, conditional on age, disease, and hospital characteristics. The estimates imply that consumers in Barnstable, for instance, are willing to travel an additional 13 miles on average in order to access a hospital they have used before. In Suffolk, however, they would only be willing to travel an additional 8 miles to access a previously used hospital.

Women are less likely to travel far to reach a hospital, and older individuals (conditional on diagnosis) also receive significant disutility from traveling. Conditional on age, however, patients with histories of chronic conditions (i.e. sicker patients) are willing to travel *more* to access a hospital of their choice. People are also on average more likely to travel to a hospital that has more beds, a specialty hospital (such as a children's hospital or a cancer center), or to travel for an academic medical center. This reinforces the point that prestigious academic medical centers in Massachusetts are able to generate high demand for their facilities.

Finally, I report the coefficients on a series of variables interacting patient diagnosis with hospital amenities. Each of these are, unsurprisingly, positive and significant. Patients with a cardiac CCS diagnosis significantly prefer hospitals with a catheterization laboratory, patients with obstetrics conditions significantly prefer hospitals with a neo-natal intensive care unit, and patients with a diagnosis requiring imaging (defined to be either a neurological, cardiac, or musculoskeletal diagnosis) prefer hospitals equipped with magnetic-resonance-imaging machines.

It is worth mentioning that this model omits copayments that plans charge to visit different hospitals. On the GIC, plans are differentiated in their premiums, their networks, and the copays that patients pay for a hospital admission across *plans*, across *hospitals*, and over time (Prager, 2016). In column 1, I exclude all observations where patients are either admitted through the hospital's emergency room or admissions resulting from a hospital transfer. This is done for two reasons. The first is that ER and transfer admissions may not necessarily reflect patient *choice* of a hospital. Faced with an emergency, a patient may be taken to the closest hospital rather than the hospital of his or her choice. The second reason is that the copays are

Table C.1: Results of Hospital Demand Model

Variable	(1)	(2)
Distance	-0.2171***	-0.2379***
	(0.0122)	(0.0079)
DistancexWorcester	-0.0334***	-0.0287***
	(0.0054)	(0.0041)
DistancexHampden	0.0135***	0.0091**
	(0.0048)	(0.0037)
DistancexSuffolk	-0.1346***	-0.1612***
	(0.0146)	(0.0109)
Used Hospital	2.8474***	2.8324***
	(0.0438)	(0.0299)
Copay	-0.0001*	-0.0000
	(0.0000)	(0.0001)
DistxFemale	-0.0048***	-0.0021
	(0.0017)	(0.0013)
DistxAge	-0.0003***	-0.0004***
	(0.0001)	(0.0000)
DistxChronic	0.0234***	0.0247***
	(0.0026)	(0.018)
DistxSpecialty	0.0326***	0.0454***
	(0.0026)	(0.0023)
DistxAcademic	0.0186***	0.0259***
	(0.0023)	(0.0018)
CardiacxCathLab	0.6072***	0.2523***
	(0.1180)	(0.0603)
ObstetricsxNICU	3.9403***	3.6289***
	(0.2797)	(0.2200)
ImagingxMRI	0.0832	0.1268
	(0.1242)	(0.0790)
Hospital FE	Yes	Yes
ER & Transfers	No	Yes
Obs.	1,021,481	1,949,285
Pseudo R2	0.52	0.54

Notes: Results from hospital demand model from years 2009-2013. Omitted distance category is for the Barnstable county. “Copay” refers to the plan-specific copayment amount in dollars for a particular hospital visit. “Chronic” refers to having a chronic condition, “Specialty” refers to being a specialty hospital. Omitted from the table are distance terms interacted with each of 18 CCS diagnosis categories, a full set of hospital fixed effects, hospital fixed effects interacted with disease weights, as well as other patientxhospital interaction variables.

typically different for hospital admissions through the ER and transfers rather than voluntary admissions. Therefore, observations that pick up transfers might register a copay amount that is not reflective of the full amount. Indeed, column 1 shows that the coefficient on copay is negative and somewhat significant. The result is similar in magnitude to Prager (2016). In column 2, where I include the full sample of admissions (including ER and transfers), the coefficient on copay reduces effectively to zero and becomes insignificant.

C.3 Physician Demand Estimates

Table C.2 reports the results of the physician demand models for PCP practices, cardiology practices, and orthopedic practices for the Boston rating region. Due to the large number of physician visits during my time frame, I run the model on a random sample of 50,000 visits across four years for each different specialty group.[74] As the model was estimated separately for each of the seven Massachusetts health rating regions, I only report here select coefficients for the Boston rating region. Table C.3 shows analagous parameter estimates for the Worcester region, for comparison. Following previous literature, I also assume there is no selection on unobservables in this model (that is, providers are not horizontally differentiated in ways unobserved to the econometrician). subsection C.4 addresses potential selection concerns in more detail.

Consistent with the results of the hospital demand model, distance plays an extremely important role in choosing physician practices. Across the three specialist groups, distance has a negative and significant effect on utility. While the magnitude of the coefficient is quite large for primary care physicians, it is about half the size for cardiology and orthopedic practices.[75] Across all three specialty groups, patients, on average, prefer visiting practices owned by hospitals or health systems, though the effect is considerably stronger for cardiology practices.[76]

Somewhat surprisingly, distance interacted with female and distance interacted with age are small and insignificant across most of the models, in contrast to the results in the hospital demand model. The only exceptions are a significant negative coefficient for distance interacted with female in the orthopedic model, and a significant negative coefficient for distance interacted with age in the PCP model. The former may be driven by the large number of sports injuries that orthopedists treat, which tend to be among patients who are disproportionately male. The latter is consistent with the result from hospital demand, namely that conditional on risk, older individuals prefer to travel smaller distances to seek care, particularly for routine primary care treatment.[77] For cardiologists and orthopedic practices, the presence of a chronic condition is associated with increased travel time, though this coefficient is insignificant in the PCP demand model. This is suggestive that sicker patients tend to have stronger preferences for specialists.

Patients seeking primary care are willing to travel further to access practices with more physicians on site. In addition, they are willing to travel further for practices with more locations and practices that are affiliated with medical groups. This result makes sense, particularly in the Boston rating area, as many physician practices are owned by larger groups, such as Partners and Atrius. However, this result is reversed for cardiologists and orthopedists. Patients are less willing to travel for larger practices, practices with multiple locations, and practices that are part of larger medical groups. While somewhat surprising, this is tempered by the age interactions, which show that older individuals significantly prefer visiting physicians from larger practice sites, physicians who are part of medical groups, and groups with multiple locations.[78] This is

[74] I omit year 2009, the earliest year of data in the claims, as I cannot observe patients' prior-use of physicians in that year.

[75] However, these coefficients should be interpreted with caution on their own. As these models are estimated separately, these coefficients are not directly comparable, as their magnitudes are driven in part by relation to practice fixed effects as well as well as scaling of the logit error.

[76] This is consistent with descriptive statistics showing that patient-weighted visits to cardiologists tend to be among larger practices. See Appendix A.

[77] In addition, the model includes distance interacted with RVU weight (omitted), which likely proxies for age.

[78] The exception is PCPs, which shows older individuals preferring smaller practice locations.

Table C.2: Results of Physician Demand Models (Boston)

Variable	PCP Practices	Cardiology Practices	Orthopedic Practices
Distance	-0.4187***	-0.3002***	-0.2330***
	(0.0186)	(0.0147)	(0.0160)
Owned by Hosp. or System	0.2555***	1.4405***	0.6391***
	(0.0982)	(0.0901)	(0.0867)
Used Prac Previously	3.6339***	1.0391***	2.3883***
	(0.0438)	(0.0354)	(0.0408)
x Length of Relationship	0.3832***	0.1996***	-0.1896***
	(0.0098)	(0.0105)	(0.0120)
x RVU	-0.0640***	0.0650***	0.0090***
	(0.0094)	(0.0053)	(0.0019)
Used Med Grp Previously	1.4471***	1.7003***	1.5240***
	(0.0401)	(0.0346)	(0.0431)
Used System Previously	0.5689***	0.8513***	1.1029***
	(0.0363)	(0.0303)	(0.0357)
		Interactions with Patient Characteristics	
DistxFemale	0.0005	-0.0046	-0.0060*
	(0.0015)	(0.0029)	(0.0034)
DistxAge	-0.0008***	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)
DistxChronic	-0.0018	0.0287***	0.0257***
	(0.0019)	(0.0079)	(0.0055)
		Interactions with Provider Characteristics	
DistxNumDocs	0.0004***	-0.0004***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)
DistxNumLocs	0.0036***	-0.0034***	-0.0052***
	(0.0009)	(0.0006)	(0.0007)
DistxMedGrp	0.0270***	-0.0378***	-0.0476***
	(0.0075)	(0.0080)	(0.0075)
AgexNumDocs (00s)	-0.0024***	0.0055***	0.0000
	(0.0000)	(0.0000)	(0.0000)
AgexNumLocs (00s)	0.0178	0.1391***	0.0804***
	(0.0165)	(0.0133)	(0.0149)
AgexMedGrp	0.0024*	0.0108***	0.0246***
	(0.0014)	(0.0015)	(0.0014)
Practice FE	Yes	Yes	Yes
Obs.	3,289,932	1,853,631	1,634,164
Pseudo R2	0.64	0.59	0.57

Notes: Results of physician demand models are for years 2010-2013 for Boston rating region only. Excluded from the table are distance, RVU weights, 18 CCS diagnosis categories, and gender interacted with additional practice characteristics: number of unique services as the practice, share of physicians at the practice who are specialists, number of doctors across the entire system, and number of practices owned by the system, Model contains a full set of practice fixed effects. Note that AgexNumDocs and AgexNumLocs are reported in hundreds (00*s*). Length of relationship is measured in years.

particularly pronounced for cardiologists, where the age effect on visiting larger practices is considerably larger than the other specialty groups.

All three of the physician inertia indicators are highly predictive of physician choice across all specialty groups, with having used the particular physician in the past being the biggest predictor and having used a provider owned by the same health system being the smallest. The estimates imply that a 35-year-old individual in average health would be on average willing to travel an additional 11.3 miles to access the same PCP practice, 14.4 miles to access the same cardiology practices, and 20.4 miles to access the same orthopedic practices. The magnitudes are quite similar to the magnitudes in the hospital demand model. The stickiness to previously used providers also varies significantly with patient health and the length of the patient-provider relationship. For PCPs and cardiologists, the longer a patient has been seeing a physician, the more likely they are to use the physician again next time. For orthopedic practices, this is reversed: the longer time has elapsed since the first time seeing the provider, the *less* likely a patient is to see that orthopedist again. This may be driven by the short-term nature of orthopedic care, which tends to more often than PCPs or cardiologists treat specific injuries on a one-off basis. For cardiologists and orthopedists, patients needing more intensive procedures (i.e. those who have higher RVU weights) are more likely to use physicians they have used in the past. However, this is not the case for PCPs, where those who have more intensive needs are likely to see a new PCP. Altogether, these results imply that inertia to previously used physicians play a significant role in provider choice.

For comparison, Table C.3 reports the results of the physician demand model for the Worcester rating region. The results are qualitatively similar to the results from the Boston rating region, however there are some notable exceptions. First, physician inertia, particularly to PCPs, plays a much larger role in Worcester than in Boston in terms of distance traveled. While in Boston, patients were on average willing to travel an additional 11.3 miles to access the same PCP practice, this figure is approximately 42 miles in Worcester. This may be, in part, due to high volume of PCPs in Boston relative to Worcester, or may be due to the fact that Worcester is an area that requires driving more so than walking.[79] Moreover, seeking care from a physician owned by a hospital or health system seems to have less of an effect in Worcester and is, in fact, *negative* for orthopedic practices. This may be reflective of the fact that, unlike Boston, Worcester has fewer prestigious academic medical centers.[80] Much like in Boston, older patients therefore significantly prefer seeking care from doctors that are part of medical groups and that work for practices which have multiple locations.

[79] The average distance traveled for PCPs in Boston is about half that of Worcester.

[80] Worcester does, however, contain a prominent medical group: the Fallon Clinic (later renamed Reliant Medical Group).

Table C.3: Results of Physician Demand Models (Worcester)

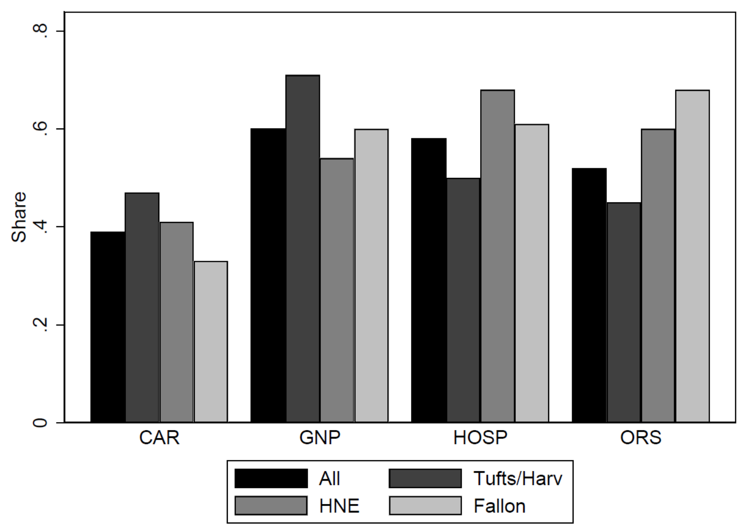
Variable	PCP Practices	Cardiology Practices	Orthopedic Practices
Distance	-0.1526***	-0.1597***	-0.2225***
	(0.0118)	(0.0104)	(0.0108)
Owned by Hosp. or System	-0.1552**	0.1103	-0.2953***
	(0.0690)	(0.0975)	(0.0905)
Used Prac Previously	4.6755***	1.3605***	3.1643***
	(0.0446)	(0.0525)	(0.0659)
x Length of Relationship	0.2552***	0.0986***	-0.3298***
	(0.0118)	(0.0186)	(0.0249)
x RVU	-0.1332***	0.0190***	0.0107***
	(0.0114)	(0.0052)	(0.0031)
Used Med Grp Previously	0.6694***	1.2402***	1.0790***
	(0.0416)	(0.0534)	(0.0667)
Used System previously	0.7643***	0.8361***	0.9833***
	(0.0411)	(0.0513)	(0.0630)
		Interactions with Patient Characteristics	
DistxFemale	0.0021*	-0.0003	-0.0030
	(0.0011)	(0.0025)	(0.0028)
DistxAge	-0.0004***	-0.0002*	0.0002**
	(0.0000)	(0.0001)	(0.0000)
DistxChronic	0.0071***	0.0180***	0.0510***
	(0.0017)	(0.0056)	(0.0047)
		Interactions with Provider Characteristics	
DistxNumDocs	-0.0001*	0.0000	-0.0002***
	(0.0000)	(0.0000)	(0.0000)
DistxNumLocs	-0.0100***	-0.0017***	0.0006
	(0.0013)	(0.0006)	(0.0007)
DistxMedGrp	-0.339***	0.0031	0.0032
	(0.0069)	(0.0059)	(0.0056)
AgexNumDocs (00s)	-0.0038**	-0.0024	0.0154***
	(0.0018)	(0.0024)	(0.0026)
AgexNumLocs (00s)	0.3938***	0.1861***	0.0081
	(0.0373)	(0.0405)	(0.0424)
AgexMedGrp	-0.0024	0.0107***	0.0098***
	(0.0023)	(0.0032)	(0.0023)
Practice FE	Yes	Yes	Yes
Obs.	2,662,897	686,687	560,253
Pseudo R2	0.60	0.62	0.62

Notes: Results of physician demand models are for years 2010-2013 for Worcester rating region only. Excluded from the table are distance, RVU weights, 18 CCS diagnosis categories, and gender interacted with additional practice characteristics: number of unique services as the practice, share of physicians at the practice who are specialists, number of doctors across the entire system, and number of practices owned by the system, Model contains a full set of practice fixed effects. Note that AgexNumDocs and AgexNumLocs are reported in hundreds (00*s*). Length of relationship is measured in years.

C.4 Selection on Unobservables in Provider Demand

A concern with two-part multinomial logit demand models of the type presented in section 3 is that they may suffer from a problem with selection on unobservables as a consequence of being estimated separately. Due to the fact that the models condition on the hospital and physician networks of each patient i at time t, N^H_{ijt} and N^S_{ijt}, the expected utility of a particular hospital and physician network, EU^H_{Ijt} and EU^S_{ijt} (defined below), is calculated assuming that there is no selection in the plan choice stage. This assumption may be violated, however, if individuals select into narrow-network plans differentially from broad-network plans for reasons unobserved by the econometrician (such as an unobserved aversion to high-cost providers, including Partners hospitals and Atrius physicians). If such selection were a major concern, this would bias EU_{Ijt}, and therefore subsequently bias the parameter estimates from the plan demand stage. Indeed, there is literature that such discrete choice models are prone to incorrect predictions when hospitals are exogenously removed from a patient's choice set (Raval et al., 2019).

Figure C.1: Share of Actual Choices Accurately Predicted, by Specialty



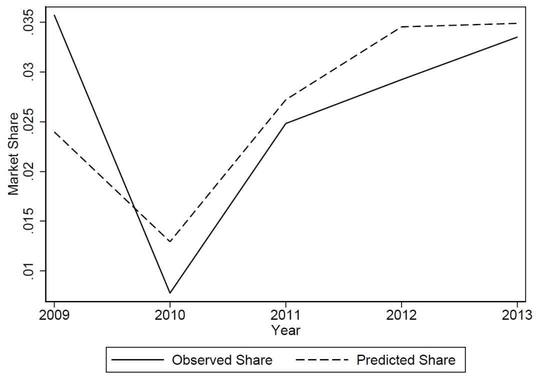
Notes: This figure plots the share of choices of providers made by individuals in narrow-network plans that are accurately predicted. Parameters used for prediction were estimated from a demand model among only individuals in broad-network plans.

I present here some reduced form evidence suggesting that such selection is not a major concern in my setting. Figure C.1 displays the share of individual choices of hospitals and physicians for individuals only in *narrow-network* plans that are accurately predicted by a model of provider demand run only on individuals in *broad-network* plans. The logic is that if unobserved selection into narrow-network plans were a big concern, we would expect a model of choice only run on patients in broad-network plans to significantly misrepresent the choices of patients with reduced choice sets. According to the figure, however, the logit model predicts the choices of narrow-network patients quite well. For PCPs (GNPs), the model accurately predicts about 60% of individual choices, and over 70% of the choices in the Tufts and Harvard narrow networks, in particular. The model also predicts hospital choices quite well, with a particularly good fit for patients in Health New England. The model does slightly worse for orthopedic surgeons, predicting about 55% of choices overall, and does worse still for cardiologists, with about 40% of choices predicted.

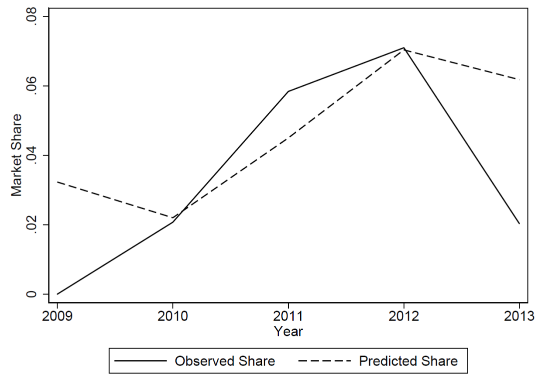
In addition, Figure C.2 plots the actual market share of selected medical centers versus the predicted market share among only narrow-network patients. For the most part, the model predicts these market shares very well. For the hospitals in the metropolitan Boston area (Tufts, Beth Israel, and Boston Medical Center), the model seems to have some trouble predicting accurate market shares in 2009, but then converges

Figure C.2: Observed versus Predicted Hospital Shares for Narrow Network Patients

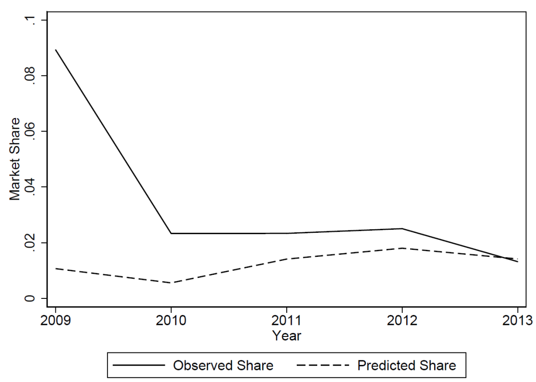
(a) Tufts Medical Center Share



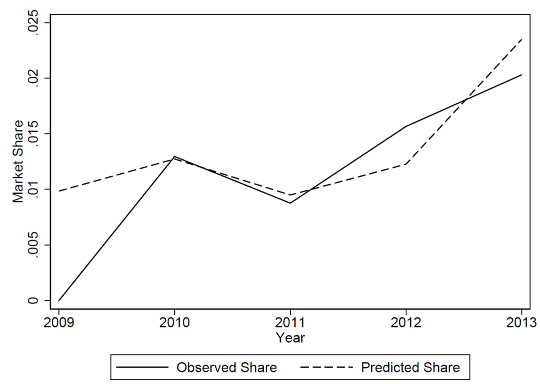
(b) Beth Israel Hospital Share



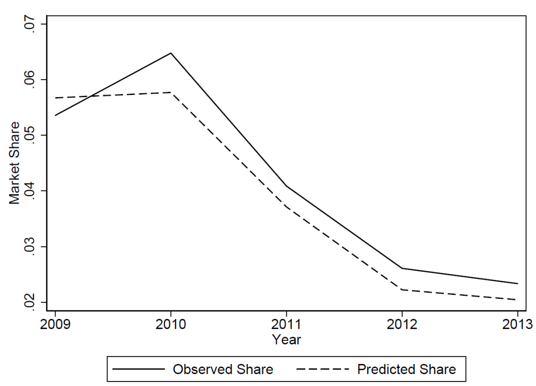
(c) Boston Medical Center Share



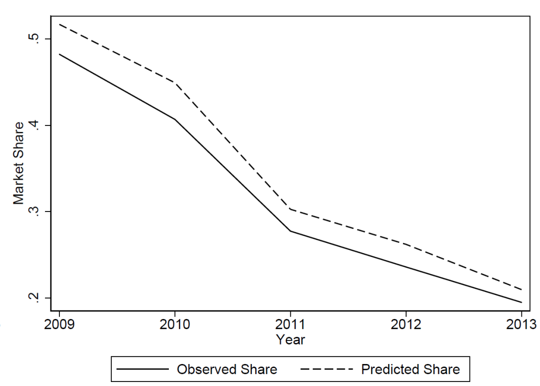
(d) Lahey Hospital Share



(e) Berkshire Medical Center Share



(f) Baystate Medical Center Share



Notes: This figure plots actual market shares of select medical centers against the predicted market shares of those medical centers among consumers in narrow-network plans. Parameters used for prediction were estimated from a demand model among only individuals in broad-network plans.

for every year after 2010.[81] Despite this, the model seems to predict the market share patterns across time very well, although it predicts a less steep decline in 2013 for Beth Israel (panel b) than the observed share. Finally, the model does extremely well in predicting the market shares of the Berkshire and Baystate medical centers, both of which are located in Eastern Massachusetts.

Taken together, these figures imply that selection is likely not a major concern in my model. Indeed, the predicted market shares for hospitals in the Boston area (which contains the highest number of academic medical centers and high-cost physicians excluded in narrow-network plans) for the most part track nicely with the observed shares, despite some difficulty in 2009. The hospitals in Eastern Massachusetts are predicted with much better accuracy.

C.5 Plan Demand

Construction of EU_{Ijt}: I define the expected utility for hospitals and physicians, respectively, as:

$$EU^H_{Ijt} = \sum_{i \in I} \sum_{l} f_{il} log \left(\sum_{h \in N^H_{jt}} exp(\phi_{ilht}) \right)$$

$$EU^s_{Ijt} = \sum_{r} f_{ir} log \left(\sum_{d \in N^S_{jt}} exp(\phi^s_{irdt}) \right)$$

where, f_{il} and f_{ir} are the ex-ante probabilities that individual i contracts diagnosis l (requiring hospital care) or requires procedure r (requiring physician care). Note that, as demand for insurance plans is at the *household* level, the expected utility variables are also aggregated to the household level by summing over each individual i's willingness-to-pay for the provider networks. The assumption is that a household's total utility for a particular hospital and physician network is a linear combination of all its individual household members. Both expected utility terms vary over time and across households.

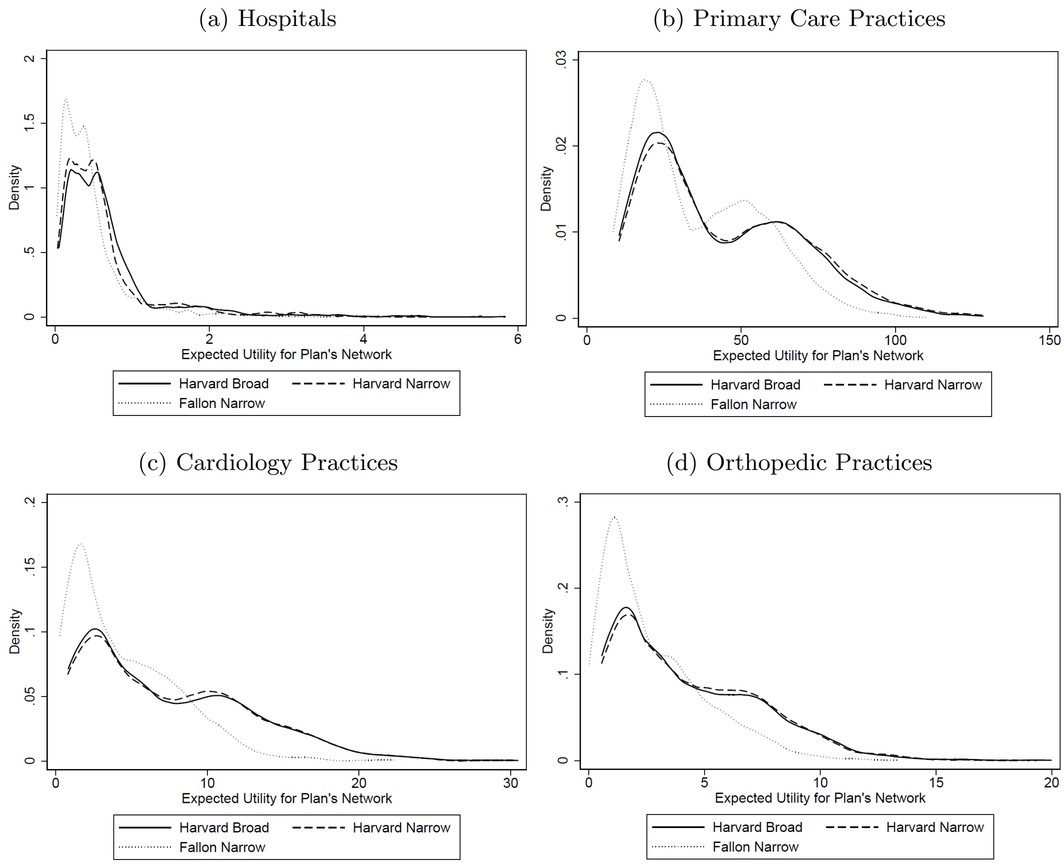
For the ex-ante illness probabilities, f_{il} and f_{ir}, individuals are grouped into distinct age-sex-chronic condition categories, with the following age bins: 0-19, 20-29, 30-39, 40-49, 50-64, 65+. f_{il} and f_{ir} are estimated directly from the claims data by averaging over the share of all GIC members of type i who sought medical treatment for diagnosis l or procedure r. For hospitalizations, diagnoses were grouped into the 18 CCS categories used in the demand estimation. For those seeking physician care, diagnoses were grouped first into the probability of requiring care from a cardiology, orthopedist, and PCP, and were subsequently grouped into bins of RVU weights: 0-1; 1-2; 2-5; 5-10; 10-20; 20-40; 40+. This reflects the fact that individuals of different ages, genders, and medical histories have differing probabilities not only of needing to see certain specialists, but also of requiring treatment of varying levels of complexities.[82]

Figure C.3 plot the density of each household's expected utility for hospitals and physician specialties for three plans' networks in the Boston rating region: Harvard Broad, Harvard Narrow, and Fallon Narrow. It is immediately clear from this series of charts that Harvard's narrow plans yield lower utility than its broad plans, and that Fallon's narrow plan yields even lower utility. This pattern is consistent across provider types. This makes sense given that Harvard's narrow network covers a fairly large number of providers—almost all excluding those owned by Partners—whereas, Fallon covers significantly fewer providers in Boston.

[81] This is likely due to small sample sizes of hospital admissions among narrow-network patients, which is particularly true in 2009 (prior to the introduction of the Tufts and Harvard narrow plans).

[82] A perhaps more robust model would specify the probability of requiring more specific procedures, rather than the probability of requiring a certain RVU-weight. Indeed, the probability of requiring knee surgery may be different than the probability of requiring shoulder surgery. However, given the number of procedures that any given specialists treats, this would present a significant computational burden. Grouping procedures into specialty-RVU categories is therefore a simplification towards computing ex-ante probabilities of valuing an insurer's provider network

Figure C.3: Expected Utility for Various Networks, Boston Rating Region



Notes: This figure plots the distribution of EU_{Ijt}^{H} and EU_{Ijt}^{s} for each physician specialty. Figures are plotted for households in the Boston rating region. Each figure plots the density of expected utility for three plans: Harvard Broad, Harvard Narrow, and Fallon Narrow.

However, the differences across provider types tells a more illuminating story. Panel (a) shows the distribution of total utility for hospitals, EU_{Ijt}^{H}. While the plot for the Harvard Broad network does skew slightly to the right to that of both narrow networks, the three network utilities virtually overlap one another for a significant portion of the density plot. Looking at panel (b), which shows the utility distribution for PCPs, EU_{Ijt}^{PCP}, consumers appear to view both Harvard plans quite similarly, whereas the Fallon Narrow plan noticeably skews left, suggesting that there is considerably more variation in the *physician* utilities across these networks than the hospital utilities. This becomes even more pronounced in panel (c) and panel (d), where the utility for cardiologists and orthopedists in Fallon's plans skews even further to the left.

Taken together, these figures show that accounting for physician services is an important part of consumer valuation of networks. While hospital networks do play a role in consumer choice, preferences diverge more strongly when considering the variation in availability of physicians between narrow and broad network plans.

Estimation Details: I leverage the panel structure of my data—the fact that I observe a sequence of household I making plan choices of plans J over time periods T—to estimate the plan demand model using maximum simulated likelihood, following the procedure outlined by Train (2009). Specifically, the probability that I observe household I making any particular sequence of choices over time is given by:

$$s_I = \int \sum_{t=1}^{T} \sum_{j=1}^{J} \left[\frac{exp(\delta_{Ijt}(\beta))}{\sum_{k=1}^{J} exp(\delta_{Ikt}(\beta))} \right]^{y_{Ijt}} F(\beta)d\beta \tag{24}$$

where y_{Ijt} is equal to 1 if household I chose plan j at time t and 0 otherwise. To construct a simulated likelihood function, I take r draws for household I from the distribution of β as outlined in Equation 4. For each draw, the likelihood function becomes:

$$\mathcal{L} = \sum_{I} ln \left\{ \frac{1}{R} \sum_{r=1}^{R} \sum_{t=1}^{T} \sum_{j=1}^{J} \left[\frac{exp(\delta_{Ijt}(\beta^r))}{\sum_{k=1}^{J} exp(\delta_{Ikt}(\beta^r))} \right]^{y_{Ijt}} \right\} \tag{25}$$

where β^r is draw r from the distribution of β. I search over 500 independent draws.

I do not observe Unicare products in my data, as the insurer does not contribute to the APCD. I therefore run the insurance demand model on the set of GIC enrollees who do not purchase Unicare products.

A full set of plan fixed effects are included. As with the provider demand model, I include an indicator variable for whether a particular plan matches an enrollee's plan choice from the previous year. This follows prior literature on plan inertia (Handel, 2013; Polyakova, 2016; Shepard, 2016) and is designed to capture enrollee switching costs from moving to a different plan. This variable is extremely important towards matching observed choice behavior in the GIC. Without it, the model would attribute what is really plan inertia to a low value of α_I (premium sensitivity parameter) or a high value for β_1 and β_2 (the network of the plan itself).

For the year 2012 (the year in which the GIC began offering its premium holiday), I adjust premiums to reflect the fact that members choosing a narrow-network plan would only pay for nine of the twelve months of the year. One caveat is that I cannot observe which members are active state employees and which members are municipal employees from years prior to 2012. Therefore, as a first-approximation, I match enrollee zip codes to public data on municipalities entering the GIC by year and do not extend the premium holiday to members with zip codes in the corresponding municipalities who joined during the corresponding years.[83]

C.6 Premium Setting Stage

Construction of of p_{jht} and p^s_{jdt}: In order to complete Equation 8 and construct the employer objective function, I construct a measure for the base reimbursement price between insurers and providers. I leverage the fact that insurers and providers do not typically negotiate over a full menu of prices for different services, but rather negotiate over a base price and then use a series of weights to scale the base price in order to arrive at a payment for each diagnosis and procedure. I use observed "allowed amounts" to to specify a base rate for each insurer-provider combination.[84]

For physicians, who are typically reimbursed on a fee-for-service basis for each procedure, r, I rely on observed RVU weights in addition to observed allowed amounts, as in Kleiner et al. (2015). I assume that price takes the following form:

[83] This is likely to produce some amount of measurement error, but sensitivity checks on the specific zip codes used revealed very minor fluctuations of the core coefficients. Moreover, running the model only on the set of new enrollees each year (i.e. those making an active choice) yields a similar premium coefficient and expected utility coefficient, indicating that any bias is likely small.

[84] Similar approaches have been taken by Gowrisankaran et al. (2015), Ho and Lee (2017), and others.

$$A^s_{irjdt} = p^s_{jdt} * RVU_{rt} \quad (26)$$
$$ln(A^s_{irjd}) = ln(p^s_{jdt}) + ln(RVU_{rt}) \quad (27)$$

A^s_{irjdt} refers to the allowed amount between plan j and physician practice d of specialty s for a patient i getting procedure r. Here, the allowed amount is a function of the base negotiated price, p^s_{jdt} between plan j and practice d, multiplied by the RVU weight for the procedure, RVU_{rt}. The model I estimate is:

$$ln(A^s_{irjdt}) = ln(RVU_{rt})\rho + \gamma^s_{jdt} + \epsilon^s_{irjdt} \quad (28)$$

where γ^s_{jdt} refers to plan-practice-time fixed effects. After estimating this model, I fix the RVU to 1 (i.e. $ln(RVU_{rt})$=0). The resulting predicted payments yield a price for each insurer-practice-specialty combination for a *standardized* procedure, and these are used as p^s_{jdt}.

In the case of hospitals, I assume that the negotiated amount is multiplied by a weight related to the "Diagnosis-Related Group (DRG)" of the particular illness that is being treated, as hospitals are reimbursed by diagnosis. These weights are typically assigned annually by CMS. Unfortunately, the APCD does not have a variable organizing the ICD-9 diagnosis codes into DRGs. Therefore, I follow Shepard (2016) and take a reduced-form approach towards estimating the hospital base price, by running the following model:

$$ln(A_{iljht}) = \gamma_{jht} + \psi_{lt} + x_{ilt} + \epsilon_{iljh} \quad (29)$$

Here, A_{iljht} refers to the observed allowed amount for patient i with diagnosis l on plan j seeking care from hospital h. γ_{jht} are fixed effects for every plan-hospital-year combination. Rather than incorporating a numerical weight with an estimated linear parameter, as done in the physician model, I proxy for diagnoses by including ψ_{lt}. These are a set of fixed effects for the 18 CCS diagnosis categories used in the demand model for hospitals. The model is therefore similar to the physician price construction model, except that by including these fixed effects, I estimate weights for each diagnosis rather than using observed weights. The model also includes Elixhauser comorbidity indexes for each of 12 secondary diagnoses, x_{ilt}. This is meant to capture nuances within diagnoses that may require heavier use of hospital resources than in generic cases (such as comas, hypertension, etc.). I use the model to predict prices for each insurer-hospital-year combination, $p_{jht} = exp(\gamma_{jht})$, and to predict the weights for each diagnosis group, $w_{lt} = exp(\psi_{lt})$. For each year, I then take the average predicted weight across admissions and consider this to be the "standardized diagnosis" for which base prices are negotiated between insurers and hospitals. I scale the predicted price by this factor in order to achieve the predicted base price for hospitals, p_{jht}.

Table C.4 reports the average negotiated base prices for hospitals and physicians and average weights by type of provider and facility type in 2011.[85] The table suggests that negotiated prices do not vary considerably across medical specialties in Massachusetts, on average. Within specialty, however, there is considerable variation. Facility-based cardiology practices, for instance, receive an average price-per-RVU of \$56, but with a standard deviation of \$20. Certain practices, therefore, receive more than \$80 per RVU. In the hospital market, the maximum base price in 2011 was \$17,306 while the minimum was \$3,545. Additionally, there are some notable differences in the average weights per procedure for physicians. Office-based PCPs, for instance, submit an average of 2.19 RVUs per visit, yielding an average of \$122 per visit. Orthopedists, however, perform an average of 4 RVUs per visit, implying an average payment of \$220 per visit.

[85] I define practices that are "office-based" are defined as practices in which more than 70% of the claims are conducted in an office-based setting. Any setting in which less than 70% of the claims are performed in an office is considered a "facility-based" setting. These include group practices in which services are primarily performed in outpatient settings of hospitals, or physicians performing services within hospital settings, but billing for professional services separately from inpatient admissions.

Table C.4: Estimated Price and Weight Measures, 2011

Variable	PCPs	Cardiologists	Orthopedists	Hospitals
		Office-Based		
Average Base Price	56.55	56.29	55.37	–
	(12.43)	(14.79)	(16.94)	–
Average Weight	2.19	2.74	3.99	–
	(0.60)	(1.25)	(2.45)	–
		Facility-Based		
Average Base Price	56.71	56.59	52.51	10,303.73
	(14.48)	(19.61)	(16.43)	(3,177.89)
Average Weight	2.35	2.07	5.38	1.00
	(1.13)	(1.95)	(5.22)	(0.34)

Notes: Standard deviations in parentheses. "Average base price" refers to the negotiated price for a standardized unit of health care. In the case of physician practices, this refers to a case where $RVU_l = 1$. In the case of hospitals, this refers to the case where $w_l = 1$. Hospital weights are scaled so that the yearly average is one, meaning that hospital base prices refer to the price fo a procedure of average weight. "Office-based" settings are defined are practices where more than 70% of claims are flagged as in an office-based setting.

I next examine whether the preference for broad-network plans translates into higher negotiated rates for those providers. Figure C.4 depicts the relationship between demand and negotiated provider price for one of the insurers on the GIC in the Boston rating region. Due to confidentiality concerns, I omit both the identity of the insurer and the actual negotiated rate. Instead, I report the negotiated rate relative to the insurer-specific average. The y-axis depicts this standardized rate, where the x-axis depicts the predicted market share from the provider demand models.

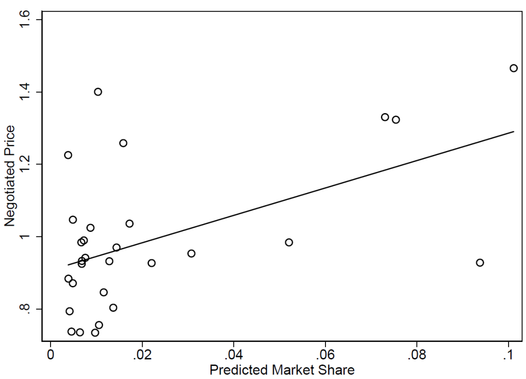
It is clear from the graphs that there is a distinct positive relationship between provider price and consumer valuation of a provider within the insurer's network. The relationship appears strongest for hospitals and, surprisingly, primary care providers, though there is still a positive relationship for cardiologists and orthopedists as well.[86] These results suggest that within specialty groups, including high-demand providers indeed tends to translate into higher prices for medical care. These prices then, in turn, translate into higher premiums for consumers. The inherent tradeoff for insurers and employer in offering plan choice thus becomes clear: to offer a broad-network plan to consumers would yield greater consumer surplus through the inclusion of high-valuation hospitals and doctors, but would also reduce surplus through higher premiums. This tradeoff is explored more in the next sections.

Estimating Unobserved Marginal Costs: To estimate $c^u_{Ijt}(N_{jt})$, I rely on standard inversion of the first-order condition specified in Equation 9. In traditional product markets, there are JT equations and JT unknowns, allowing for recovery of all necessary cost parameters. In health insurance markets, however, marginal costs do not merely vary by product, but also by consumer risk type. As a result, in my context, there are only JT equations but JTI unknowns, where I is household type. While the marginal costs for care from hospitals, PCPs, cardiologists, and orthopedists are observed in the my claims data, to recover unobserved marginal costs, I parameterize costs as $c^u_{Ijt}(N_{jt}) = c^u_{jt}(N_{jt})\theta^c_I$, where θ^c_I reflects a parameter that scales base plan-specific unobserved costs, $c_{jt}(N_{jt})$, across household type I. I assume that unobserved marginal costs only vary by age and whether the household is an individual or family. I infer θ^c_I directly from the data by aggregating all claims from providers that are not hospitals, PCPs, cardiologists, and

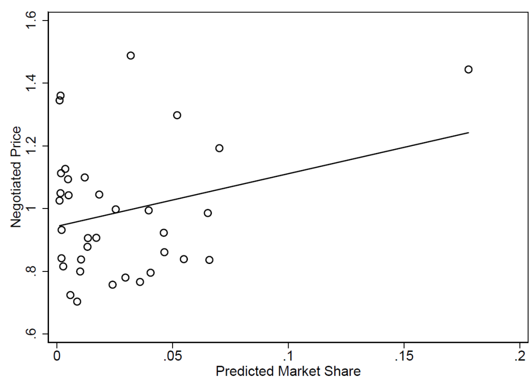
[86] This may be explained by the presence of Harvard Vanguard in the Boston rating region, which has considerable bargaining power. Other large primary care practices in the area likely hold similar bargaining power. Though modeling the full bargaining game between physician practices and insurers is outside the scope of this paper, it is an interesting subject for future work.

Figure C.4: Insurer Negotiated Price by Market Share, Boston Rating Region 2011

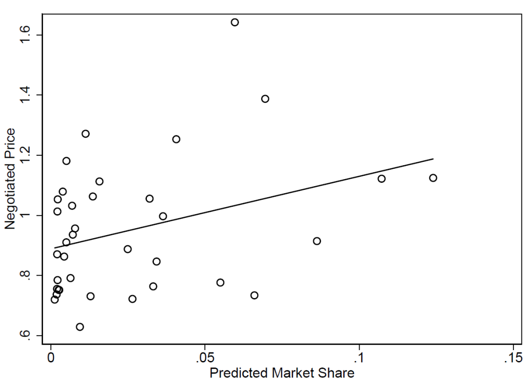
(a) Primary Carel Practices



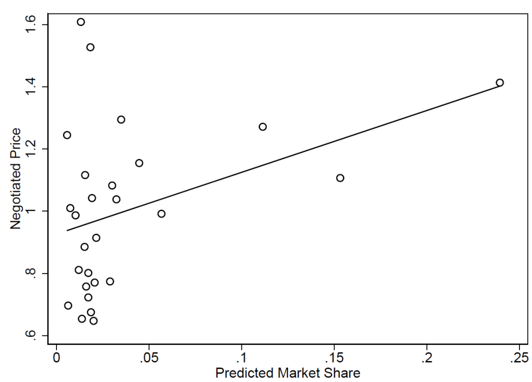
(b) Cardiology Practices



(c) Orthopedic Practices



(d) Hospitals



Notes: This figure plots the the negotiated price for hospitals, p_{jht}, and for physician practices, p^s_{jtd}, against predicted market share from the provider demand models. Prices are reported for a single insurer and relative to the insurer-specific mean. Data is for year 2011.

orthopedists, and regressing the observed allowed amounts for these claims on age and household type.[87] This reduces the number of unknowns to JT, allowing for full recovery of the base marginal costs, $c^u_{jt}(N_{jt})$.

To predict counterfactual $c^u_{jt}(N_{jt})$ with different networks of hospitals and physicians, I regress the recovered costs on a series of cost-shifters (and adding insurer subscript m back) such that:

$$c^u_{mjt}(N_{mjt}) = \kappa x_{mjt} + \gamma_m + \gamma_t + \varepsilon_{mjt} \tag{30}$$

In my estimation, these shifters include insurer fixed effects, year fixed effects, and an indicator, x_{mjt}, for whether or not the plan is a narrow-network plan.

Cost Estimates: Table C.5 reports the results Equation 30, regressing the log of unobserved marginal costs of health care on insurer fixed effects, year fixed effects, and an indicator for whether the plan is narrow or not. Year 2012 is omitted due to potential bias in estimates from it being the year of the premium holiday.

Table C.5: Unobserved Marginal Cost Estimates

Variable	Coefficient	Standard Error
Narrow Network	-0.174***	0.022
Harvard Pilgrim	0.056**	0.025
Health New England	-0.053*	0.030
Neighborhood Health Plan	-0.046	0.030
Tufts Health Plan	0.054**	0.025
2010	-0.006	0.026
2011	0.037	0.025
2013	0.082***	0.025
Constant	5.904***	0.024
Obs.	28	
Adjusted R2	0.87	

Notes: Results from marginal cost estimation. Dependent variable is the log of unobserved marginal costs. Omitted insurer is Fallon Health Plan. Omitted year is 2009. Year 2012 is left out of analysis due to concern about bias in estimates from it being the year of the premium holiday.

The results indicate that being a narrow-network plan reduces unobserved marginal costs of health care by approximately 17%. Among insurers, Harvard and Tufts each have higher relative unobserved costs, compared with Health New England, Neighborhood Health Plan, and Fallon. This indicates that Harvard and Tufts may have non-hospital, PCP, cardiology, and orthopedic expenditures that are higher, potentially due to contracting with larger set of providers unaccounted for by the chosen specialties.[88] Unobserved costs increase steadily over time, likely reflecting increases in negotiated prices with providers over time as well as general medical inflation. In particular, unobserved costs in 2013 are estimated to be approximately 8% higher than in 2009.

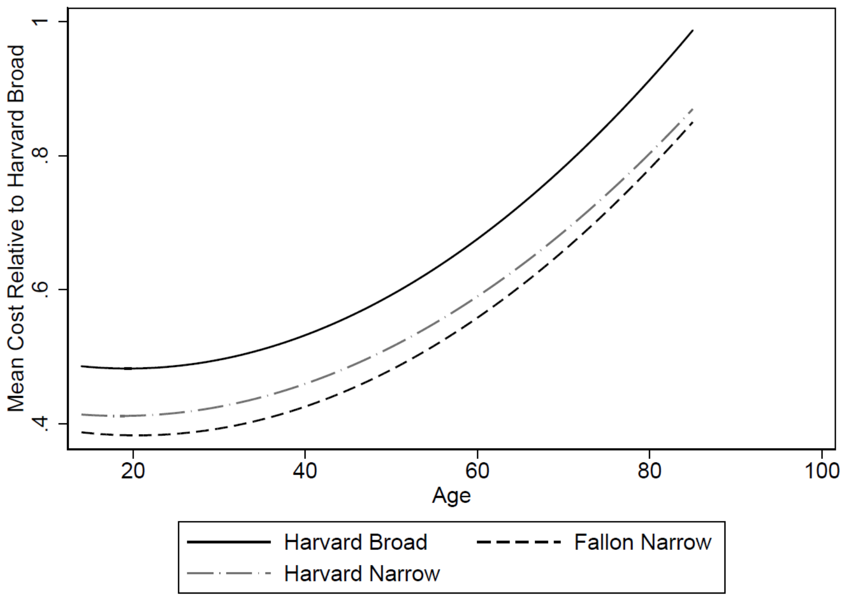
Figure C.5 plots the total estimated marginal costs of health care (hospital + PCP + cardiology + orthopedics + unobserved) against age for single-member households. I report estimated cost-curves for Harvard Broad, Harvard Narrow, and Fallon Narrow. As expected, predicted insurer costs rise rapidly with age. Moreover, the broad-network plan has consistently higher predicted costs than the narrow-network plans at all age levels. Further, the cost-curves do slope upward as similar rates, although Harvard Broad does have a slight uptick in the rate at which it rises after age 60 relative to the narrow products. This suggests the potential for selection on expensive providers, particularly among older individuals, conforming

[87] The critical assumption here is that all marginal costs that vary by more granular risk types are captured through *observed* hospital and physician costs, whereas *unobserved* costs only vary by age and family type. While strong, this seems reasonable as a first-order approximation.

[88] An alternate explanation is that these costs reflect higher administrative costs or more generous drug formularies.

to the results of the hospital and physician demand models.

Figure C.5: Estimated Insurer Marginal Costs



Notes: This figure plots estimated marginal cost curves for select plans in 2013. Note that the y axis reflects costs relative to the average cost of Harvard Broad.

C.7 Additional Estimation Details of Employer Objective Function

Error Assumptions: I make several assumptions to proceed with the estimation of ρ and FC_j. First, I assume that the only disturbances to the expected surplus, $v_{1,\delta_{Jt}}$, are composed of two sources: $v^a_{1,\delta_{Jt}}$ and $v^b_{1,\delta_{Jt}}$. The former refers specifically to uncertainty about which municipalities will enter the GIC in the coming year. The latter refers to all other uncertainty in demand, including measurement error. Both disturbances are unknown to the employer and the econometrician. I assume that $E[v^b_{1,\delta_{Jt}}] = 0$.

Rather than relying on instruments within the employer's information set, I instead use observed data on municipal entrants by year to specify a distribution of household entrants over which the employer has an expectation. I make a timing assumption that the GIC knows the number of municipalities that entered in the previous year and assumes the same number of municipalities enter the subsequent year, but does not known *which* municipalities, and therefore does not know the underlying risk and preferences (or location) of the households entering in any given year.[89,90] More formally:

$$E[v^a_{1,\delta_{Jt}}] = v^a_{1,\delta_{Jt-1}} + \omega_t$$

where $v^a_{1,\delta_{Jt-1}}$ is the realized disturbance from period $t-1$ and ω_t is a shock to the risk profile and location of entrants in year t. I assume $E[\omega_t] = 0$, or that the shocks to household risk in a given year, conditional on observing entrants in the prior year, are zero.

Translating to sample means, this implies:

89 Indeed, between 2009 and 2013, municipalities chose to enter the GIC during many different time-periods within a given year, leaving the GIC little room to incorporate those entrances into its menu decisions. As an example, if a municipality enters in April, it would be unreasonable to assume that the GIC could then reoptimize its product offerings to begin the following fiscal year in July.

90 It would be more sophisticated to fully specify a model in which the GIC competes for municipal business as a function of the networks and products offered. This model is outside the scope of this paper. However, future work will consider this issue more explicitly.

$$v^a_{1,\delta_{Jt-1}} + \underset{K\to\infty}{plim}\frac{1}{K}\sum_{j}^{K}\omega_k = 0$$

In the estimation of Equation 17, I take the average of 100 disturbances of ω_k. That is, I estimate the moment inequalities assuming 100 different potential random sample of entrants in each year given the number of municipalities who entered the previous year.

The second assumption is that there is no presence of a structural error component that the employer knows when making decisions, but the econometrician does not. Such structural errors would normally appear in the fixed cost term, FC_j, appearing as a potential disturbance such that, for instance,: $FC_j = FC + v_{2,j}$, where v_2 represents the structural shock to fixed costs. Eizenberg (2014) and Mohapatra and Chatterjee (2015) describe in detail a potential selection problem that would arise out of this formation if the error term varied by the type of product offered. In this setting, the GIC might choose to contract with certain insurers, offer certain products, or offer certain networks for which the fixed costs of doing so are lower. Without addition assumptions, this structural error would bias my estimates of both ρ and FC_j.

I circumvent this selection problem by assuming there is no structural error term and, namely, that the fixed costs do not vary by where the plan is in the quality space, i.e. $FC_j = FC$. While this may be a strong assumption in other settings that have wide variation in fixed or sunk costs of product introduction, it is a more reasonable approximation for this environment. This is a single-agent problem, and I am estimating the fixed costs associated with introducing additional plans under the umbrella of one large employer group. While such costs may differ across employers, the differences in fixed costs *within* employer group are likely smaller.[91]

Alternate Estimator for FC_j: I construct an additional estimator for comparison, setting $\rho = 1$ and constructing bounds on FC_j using merely exclusively one-step deviations in the *number* of products offered.

For this, I construct two counterfactual quality vectors. I define $\delta_{J+j,t}$ as the total product quality that would result from offering an additional product j that is not currently offered. I define $\delta_{J-j,t}$ as the total product quality that would result in the employer removing one of its currently offered products, j.

The estimation follows from a similar revealed preference assumption as the previous estimator, namely that the products I observe in the data are chosen in equilibrium. This establishes the necessary conditions that the employer would not choose to add a product ($\delta_{J+j,t}$) or remove a product ($\delta_{J-j,t}$) unless these deviations increased its objective function, W_t. These necessary conditions allow me to estimate bounds on the fixed cost parameter.

One side of the bound comes from the assumption that any product the employer chooses to offer must necessarily increase variable social surplus, $S(\delta_{Jt}, \theta)$. Therefore, by removing a product currently offered and computing counterfactual surplus, I can infer that the fixed costs for offering an additional product must be less than the surplus gained by offering the product. Formally this upper bound on fixed costs is given by:

$$FC_j \leq E\left[S(\delta_{Jt}, \theta) - S(\delta_{J-j,t}, \theta)\right] \equiv \overline{FC}_j \tag{31}$$

where $\overline{FC}_j$ refers to the upper bound on fixed costs. Similarly, I can obtain the lower bound as follows:

[91] Similar assumptions were made by Nosko (2014). This assumption may be violated if, for instance, offering a product that was broader in network size than another product also meant an increase the cost of the negotiation process. However, this is unlikely to apply to the GIC for two reasons. First, I do not allow the GIC to offer any plans for which the network is larger than the largest currently offered by the particular insurer anywhere in Massachusetts. In other words, insurers can only design plans that are narrower than what they currently offer, but not broader. This implies that there would be no additional contracting fixed costs for providers with whom any particular insurer does not currently negotiate with. Second, while employer groups negotiate premiums with different plans, they rarely ever negotiate base prices with providers. This task falls largely onto the insurers, and it is therefore unlikely that the added negotiation cost of offering broader network plans would result in additional fixed costs for the GIC itself.

$$FC_j \geq E\left[S(\delta_{J+j,t},\theta) - S(\delta_{Jt},\theta)\right] \equiv \underline{FC}_j \tag{32}$$

where $\underline{FC}_j$ is the lower bound on fixed costs. This side of the bound implies that if the employer can offer a potential product, but is not observed to, then it must be the case that fixed costs are larger than the change in marginal social surplus from introducing it.

Assume that the employer's expectation of its total surplus from adding or removing products follows the following form, where $v_{3,\delta_{Jt}}$ is a disturbance such that $E[v_{3,\delta_{Jt}}] = 0$:

$$E[S(\delta_{Jt},\theta)] = S(\delta_{Jt},\theta) + v_{3,\delta_{Jt}} \tag{33}$$

As long as the GIC has correct expectations on average, the estimation equation becomes:

$$\underset{K\to\infty}{plim}\frac{1}{K}\sum_{j}^{K}\left(S(\delta_{Jt},\theta) - S(\delta_{J-j,t},\theta)\right) \geq FC_j \geq \underset{K\to\infty}{plim}\frac{1}{K}\sum_{j}^{K}\left(S(\delta_{J+j,t};\theta) - S(\delta_{Jt};\theta)\right) \tag{34}$$

D Details on Private Employer Sample

Sample Construction: To construct the sample of large, private employers used in section 4, I limit the claims data to members employed by non-government firms with more than 50 employees and those who have at least one commercial insurance product that is "self-insured." Restricting the sample to large, self-insured firm makes the estimation of the moment inequalities considerably simpler than if I also included small employers, as it allows me to construct similar premium pricing rules as for the GIC and abstract from incorporating insurer profit functions. As the APCD does not contain firm identifiers, I instead create a sample of firms using the employer zip code field (herafter referred to as "employerzip"), Standard Industrial Classification (SIC) code, and product (plan) identifiers (IDs), the latter typically being unique within firm. As employers can offer multiple different plans—and therefore have multiple plan IDs—I use employee flows across plan IDS to determine the likelihood that any two IDs belong to the same firm. Specifically, if I observe two different plan IDs within an employerzip-SIC grouping and also observe that a non-trivial share of employees switch from one ID to the another (and vice versa), I assume that both IDs are part of the same firm.

I then simulate a distribution of firms offering narrow-network insurance plans using external micro data from Kaiser Family Foundation and Health Research & Education Trust (HRET). The Kaiser/HRET annual survey of employer-sponsored health benefits contains questions about employers' general characteristics, plan offerings, enrollment, health risk appraisals, and other topics. Beginning in 2014, the survey asked whether firms offered narrow-network plans. Since my APCD sample ranges from 2009 - 2013, I limit the APCD to 2013 and match simulate firm offer distributions from the 2014 Kaiser/HRET survey. Unfortunately, the survey only contains geographic information up to broad Census region categories. I am therefore not able to match the distribution of firms and plan offers to Massachusetts firms directly. Instead, I limit the Kaiser/HRET sample to only firms in the Northeast United States, and match narrow-network offer rates offer rates using data on firm size, industry, and number of plans offered.

I make a few additional simplifying assumptions in creation of the private firm sample. First, I limit the APCD to only members covered by the same insurers as in the GIC. This primarily has the effect of removing Blue Cross Blue Shield (BCBS) members from the data. While this represents a non-trivial share of commercial enrollment in Massachusetts,[92] it is nonetheless a sensible restriction to make. During my sample period, BCBS was the only carrier that did not offer *any* narrow-network products on the market (Office of the Attorney General Martha Coakley, 2013). Moreover, focusing exclusively on GIC carriers reduces computation burden significantly, as it enables me to use already-estimated demand parameters and negotiated prices (see subsection C.6) rather than re-estimating demand for the set of BCBS members.[93]

The second simplifying assumption I make is in the network offerings of employers. For products outside the GIC (where there is no publicly-available data on provider networks of plans), I do not observe the network breadth of each plan. One option would be to infer networks based on observed claims.[94] However, this approach is prone to significant measurement error, particularly for firms with fewer employees. Instead, I leverage institutional features of the Massachusetts insurance market. In particular, outside of the GIC, Harvard Pilgrim and Tufts Health Plan each only marketed one narrow-network insurance product to employers as of 2013.[95] I therefore assume that any firm offering Harvard Pilgrim products simulated to also offer a narrow-network plan gets assigned the respective Harvard narrow-network product available in Massachusetts at the time. Similarly, for firms offering Tufts products, I assign the network breadth

[92]BCBS represented about 45% of commercial payer market share in Massachusetts in 2012 (Center for Health Information and Analytics, 2013).

[93]In particular, I would be unable to use existing demand estimates on BCBS enrollees as I have no estimated "brand effect" for BCBS.

[94]See Gruber and McKnight (2016) for such an approach.

[95]These products are the Harvard Focus network and Tufts Select network, respectively (Office of the Attorney General Martha Coakley, 2013).

of the respective Tufts narrow plan. I impute the networks of these products using publicly-available network brochures for each of these plans (in a similar way to the construction of GIC networks, detailed in Appendix A).[96]

Table D.1: Summary Statistics for Simulated Private Employer Sample

Variable	Private Emp.	GIC	Private Emp.	GIC
	Firm-Level		Employee-Level	
Offer Narrow Network	0.0561	–	–	
	(0.2311)		–	–
Enrolled in Narrow Network	–	–	0.0237	0.1182
	–	–	(0.1521)	(0.3228)
Employees Over 55	–	–	0.1602	0.2010
	–	–	(0.3668)	(0.4008)
Employees Female	–	–	0.5352	0.5183
	–	–	(0.4988)	(0.4997)
Rating Area				
1	0.1667		0.0654	0.1585
	(0.3744)		(0.2474)	(0.3652)
2	0.0556		0.0871	0.1273
	(0.2301)		(0.2820)	(0.3333)
3	0.0648		0.1720	0.1004
	(0.2473)		(0.3334)	(0.3005)
4	0.1296		0.1698	0.2181
	(0.3375)		(0.3754)	(0.4129)
5	0.5370		0.3700	0.2402
	(0.5100)		(0.4828)	(0.4272)
6	0.0463		0.1272	0.1354
	(0.2111)		(0.3332)	(0.3422)
7	–		0.0086	0.0202
	–		(0.0921)	(0.1408)
Industry				
Health Care	0.3604		0.3882	
	(0.4823)		(0.4873)	
Service	0.3423		0.3993	
	(0.4767)		(0.4897)	
Wholesale	0.0180		0.0222	
	(0.1336)		(0.1474)	
Transportation, Communications, Utilities	0.1441		0.0729	
	(0.3528)		(0.2600)	
Manufacturing	0.1351		0.1174	
	(0.3434)		(0.3219)	
Number of Employers	123			

Notes: Summary statistics for simulated sample of private employers in Massachusetts in 2013 (Columns 1 and 3) and employees of the Group Insurance Commission (Columns 2 and 4). First two columns present firm-level characteristics, while last two columns presents employee-level characteristics.

Table D.1 reports summary statistics for the simulated private employer sample and compares them to the GIC. Overall, the sample contains 123 simulated large private employers in the state. Though many of the characteristics of the simulated sample look similar to the GIC, there are some notable differences. Approximately 6% of those employers offer narrow-network plans (Column 1), consistent with the share seen in the Kaiser/HRET survey. However, only 2% of *employees* across the state actually enrolled in narrow-network plans in 2013, compared with about 12% in the GIC (Columns 3 and 4).[97] The GIC sample is slightly older, with about 20% of employees being over age 55, compared about about 16% in the private employer sample. Together, the health care and service industry comprised 70% of the sample. In terms geographic distribution, most large private employers are headquarted in Boston (Rating Region 5). This

[96] In theory, as these firms are self-insured, they may have offered custom narrow-networks in a similar way that the Harvard Primary Choice and Tufts Spirit plans were designed for the GIC. However, this approach is likely to be fairly accurate for a first-order approximation.

[97] Recall that the recent premium holiday implemented in 2012 was somewhat responsible for this high share of enrollment.

translates to about 37% of all private employees working for firms in Boston, with the next largest share (16%) working for firms in Rating Region 4 (the North Shore). On the GIC, conversely, employees were more evenly distributed across regions. For instance, 24% of employees lived in Rating Region 5 and 21% of employees lived in Rating Region 4. Overall, then, private employers skew more heavily towards dense, urban areas than employees on the GIC.

Estimation of Employer Objective Function: Estimation of the employer objective function for private employers followed a very similar procedure outlined in subsection 3.4. However, I make several assumptions to accommodate features of the simulated sample. First, I use the same demand parameters as estimated in Table 2, essentially assuming that employee of large, self-insured, private firms, conditional on observables, have similar demand for health insurance as employees of the GIC. Second, in order to circumvent issues with estimating fixed costs across different employers (noted in subsection C.7), I restrict the moments for each employer to have the same number of plans they currently offer. For example, if an employer currently offers 2 plans, then for that employer, I only consider alternate plan menus/networks in which that employer offers 2 plans. This allows me to isolate the effect on the employer-employee mismatch term, ρ. Finally, for each alternate plan menu, I now construct moments by taking sample averages across employers. In other words, the moment equation from Equation 17 becomes:

$$m(\delta_J, \delta_J^a, \theta, z) = \sum_{s=1}^{10} \left(\frac{1}{F} \sum_f \left[(W(\delta_J, \theta_s) - W(\delta_J^a, \theta_s)) \otimes g(z) \right] \right) \geq 0 \tag{35}$$

where f is the subscript for employer f and F is the total number of private firms sampled.

E Robustness on Employer Objective Function

Alternate Assumptions on Physician Inertia: In my model of provider demand, my baseline estimates treat persistence in provider choice as a welfare-relevant utility component, such that if a physician a patient had previously used were removed from the network, the patient would suffer a utility loss from this network change in excess of what a patient who did *not* previously use the physician might experience. In other words, at baseline I treat physician inertia as patient-physician-specific capital. Should the relationship be severed, the patient would suffer a genuine loss, though the provider the model predicts the patient chooses *next* would not necessarily be a function of the characteristics of the provider that was just removed from the network.

There are two alternate interpretations of physician inertia in the model. The first is that persistence in choice of providers is driven by unobserved physician quality and not necessarily the patient-provider match. Here, the loss of a physician from the network would also imply a welfare-relevant loss. However, the main distinction from the baseline assumption is that the utility change from the loss of a provider will vary by (a) the patient's characteristics and preferences; (b) the characteristics of the provider and; (c) the characteristics of the remaining providers in the choice set. For example, if a high-quality physician were removed from the network with no close substitute in the resulting smaller network, the patient would incur a substantially higher utility loss than the baseline estimate. Conversely, if a physician were removed and the resulting network had many physicians remaining of similar quality, the utility loss—and hence welfare implications —would be *smaller* than baseline.

Finally, the inertia term may reflect switching or hassle costs irrespective of physician quality or match. Here, if a physician were removed from a network, the model ought to predict a similar second choice as with the baseline assumption. However, if persistence were driven by hassle costs, then it is possible the employer would not view such costs as welfare-relevant in its decision-making about networks.[98]

Each of these interpretations, through their impact on consumer utility of a network change, can have significant effects on estimation of the employer objective function—particularly on the estimate of the employer-employee mismatch. In Table E.1, I report results on ρ and FC_j assuming that the entirety of inertia term were driven by these various forces.[99] To test the impact of treating physician inertia as a switching/hassle cost, I re-estimate the employer objective function assuming that the utility change from losing a provider were "welfare irrelevant" from the eyes of the employer. In doing so, the estimate of the employer-employee mismatch increases significantly, from a baseline of 3.67 to 5.89 (Column 2). This makes sense: in this scenario, any narrowing of a network results in a smaller utility loss, but a similar decline in employer health spending. As such, the fact that the employer does *not* narrow the network implies a much larger mismatch between employer incentives and employee preferences.

To test the impact of treating physician inertia as unobserved provider quality, I re-estimate the provider demand model only on patients who had never seen *any* provider prior to their current visit. In effect, the assumption here is that if persistence in physician choice were driven mainly by unobserved physician quality (irrespective of physician-provider-specific match), this ought to be reflected the first-time choices made by brand new patients.[100]. Under this interpretation, the employer-employee mismatches again rises from baseline, though not as dramatically as when treating inertia as a hassle cost (Column 3). This suggests

[98] Two anecdotes support this point. First, the GIC actively encourages employees to switch to narrow-network plans, going so far as to implement a premium holiday in 2012. This is highly suggestive that they, at least in part, view persistence in provider choice as being driven by hassle costs. Second, if employers were forward-looking, then once employees choose a new provider, long-run utility ought to be fairly stable if persistence were driven by hassle costs. One can imagine a scenario, for instance, in which patients were simply "defaulted" to a new physician and would incur no such costs.

[99] Indeed, the inertia term might be driven by a combination of these forces. Treating the entirety of the term as being driven by one force or another is meant to show bounds on the relevant parameters for the employer.

[100] This is similar to the exercise presented in subsection 4.3

Table E.1: Employer Objective Function Parameters Under Alternate Assumptions

	Baseline	Inertia = Switching Costs	Inertia = Unobserved Quality	No Logit Error
ρ	3.67	5.89	4.60	3.99
FC_j	4.07	7.78	5.23	1.57

Results from ρ and FC_j estimation for 2009-2013. Column 1 presents estimates for the current population of GIC enrollees. It assumes that physician inertia is interpreted as provider-patient-specific capital. Column 2 presents estimates under the assuming that physician inertia is interpreted as a "welfare-irrelevant" switching cost. Column 3 presents estimates under the assumption that all of the physician inertia is interpreted as unobserved provider quality. Column 4 presents estimates assuming that the logit error in plan demand is set to 0. FC_j reported in millions of dollars.

two things. First, removing any physician yields a smaller utility loss for patients than at baseline, implying that patients are typically able to find close substitutes. Second, the baseline model does reasonably well as estimating unobserved provider quality.[101] Taken together, the fact that the baseline model yields the smallest mismatch parameter implies that it is most conservative interpretation of physician inertia. The "true" mismatch parameter, then, lies somewhere between 3.67 and 5.89, but it always considerably greater than 1.

Estimates with No Logit Error: The baseline results presented in Table 3 assume that consumer surplus, $CS(\delta_{Ijt}, \theta)$ is calculated in the traditional "logsum" way, which implicitly assigns consumers a positive valuation of any counterfactual product added to a choice set regardless of where that product lies in the quality space. This valuation is assigned through the idiosyncratic logit error. As a result, even if plans of "low quality" are introduced to the market, consumers may be made better off in a way that may bias the number of equilibrium products upward.[102] To remove this potential bias, I re-estimate Equation 17, allowing consumers to select plans as they would in a logit world, but setting the logit error to zero for the purposes of computing consumer surplus.[103]

The results on the mismatch term are, unsurprisingly, similar to the baseline estimate as identification of ρ relies primarily on utility changes holding the number of products fixed (Column 4 of Table E.1). However, the estimates of fixed costs decreased significantly from \$4.07 million to approximately \$1.6 million. Though a large decline, this is quite intuitive: each additional product that could have been offered but was not brings substantially less utility to consumers without the presence of the logit error. The employer is therefore sacrificing less in terms of utility by *not* offering additional choice, thereby reducing the estimate of FC_j. Appendix F shows how these logit assumptions affect the policy simulations and welfare implications.

[101] If the estimate revealed that the mismatch parameter substantially *declined*, this would imply the baseline model was not accurately capturing the utility loss from the removal of a flagship or high-quality provider from the network.

[102] This is particularly true in my setting, where the employer explicitly has $CS(\delta_{Ijt}, \theta)$ in its objective function.

[103] A more sophisticated approach would be to estimate a pure characteristics model of demand for insurance, assuming away the logit error, as in Berry and Pakes (2007), Nosko (2014), and Song (2007). However, due to the complexity of this task, I take a more simplified approach and simply preserve the estimated parameter estimated from a logit model with shocks, but remove those shocks for the purposes of computing surplus. Though this approach is a simplification, it is meant to be an approximation as to what reasonable bounds on product offerings might be due to the removal of switching costs.

F Additional Counterfactual Details

F.1 Simulation Procedure

I now describe the procedure used to implement the policy simulations in section 5. In order to reduce the dimensionality of the computation, as with the employer objective function estimation, I restrict the offset set to that outlined in subsection 3.4. This leaves a possible set of 14 products for the employer to offer. I proceed computing the equilibrium networks offered in a series of steps:

1. Construct a vector of $2^{14} = 16,384$ possible equilibria combinations of products offers
2. For each vector, compute the expected utility of the hospital and physician networks for each member, EU^H_{ijt} and EU^s_{ijt}, for each offered product's network using the estimates from the provider demand model
3. Compute the predicted marginal costs of health care to the employer, c^H_{Ijt} and c^S_{Ijt} for each household if they enrolled in any of the offered products, using the negotiated price construction
4. Compute the base "unobserved" marginal costs of health care, c^u_{Ijt}, using the parameters estimated from Equation 30.
5. Compute the expected market shares and premiums, $s_{Ijt}(\delta_{Jt}, \theta)$ and $R_{Ijt}(\delta_{Jt}, \theta)$, for each household in each offered product, using the results from the insurance plan demand model and the pricing equation in Equation 9.
6. Compute the estimated consumer surplus, $CS(\delta_{Jt}, \theta)$, and total outlays for the employer under the current product offered
7. Compute the employer's objective function using estimated $CS(\delta_{Jt}, \theta)$, total expenditures, and estimated mismatch parameter, ρ, and fixed costs, FC_j.
8. Repeat this procedure for each vector of possible equilibria, and take the max of all the computed welfare functions.

F.2 Policy Simulations Assuming No Logit Error

Counterfactual Product Offerings: I use these new estimates to re-estimate the counterfactuals presented in section 5. Table F.1 reports these results. The results remain largely consistent with those reported in Table 7. In particular, the employer still offers predominantly broad-network plans in Rating Region 4 and predominantly narrow-network products in Rating Region 5. The most notable change is the the *number* of product offered drops somewhat, with Rating Region 1 and Rating Region 4 each offering 5 plans, while Rating Region 5 only offers 4 plans. Despite these changes, the welfare implications remain similar, but for a slight increase in total surplus (Panel C) relative to the estimates in Table 7. This is driven by the fact that fixed cost estimates are smaller with the logit error is removed (as seen in Appendix E) and, as a result, social surplus is somewhat higher relative to the baseline scenario.

Table F.1: Counterfactuals: Equilibrium Networks Chosen Under Region-Based Pricing

Insurer	Network	Observed	Region		
			R1	R4	R5
Panel A: Equilibrium Plan Menus/Networks					
Fallon	VN	x		x	
Fallon	B	x	x	x	
HPHC	VN		x		
HPHC	N1				x
HPHC	N2				
HPHC	M	x	x		x
HPHC	B	x		x	
HNE	N	x	x		
NHP	N	x		x	x
Tufts	VN				
Tufts	N1	x			
Tufts	N2				
Tufts	M				
Tufts	B	x	x	x	x
Total Plans		8	5	5	4
Panel B: Welfare and Spending Holding Plan Menu Fixed					
ΔCS (Fixed)				-\$0.42	
$\Delta Costs$ (Fixed)				-\$1.12	
ΔFC (Fixed)				–	
$\Delta Surplus$ (Fixed)				-\$0.54	
Panel C: Welfare and Spending Allowing Plan Menu to Change					
ΔCS (Change)				-\$7.30	
$\Delta Costs$ (Change)				-\$34.38	
ΔFC (Change)				\$1.96	
$\Delta Surplus$ (Change)				\$25.12	

Notes: GIC observed and predicted products offered under region-based rating with no logit error. "R1" refers to plan networks for region 1, etc. Panel B reports the welfare and cost changes assuming plan menus remain fixed at observed 2011 levels. Panel C reports these changes allowing endogenous employer changes to menus. "ΔCS" refers to change in consumer surplus per-household-per-month. "$\Delta Costs$" refer to the change in total GIC costs per-household-per-month. "ΔFC" refer to changes in fixed costs.