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

Nicholas Tilipman†

November 12, 2021

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 overweighting the preferences of certain employees, specifically older workers and those in regions with less provider competition, over the preferences 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 (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 (ACA) marketplaces, with some states restricting choice sets to only a few plans and others 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 weighs 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 health care landscape, 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 whether a portion of these rising costs can be attributed to a mismatch between employer and employee preferences.

I focus my analysis on employer decisions over health plan provider networks and, in particular, 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 individual marketplaces, 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 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. Several features make it a notable setting for this study. First, 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 of narrow-network products and offers plans with considerable variation in both hospitals and physicians covered. This includes the addition of two

[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 ACA 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 fewer than 25% of the physicians in the market (Polsky and Weiner, 2015).

new narrow-network products midway through my sample. Second, the GIC designs plans both to satisfy its pool of state government workers, but also to attract new, municipal employees. As such, it behaves similarly to an employer, and it also sees variation in the set of employees entering the pool over time. Finally, in 2012, the GIC instituted a "premium holiday" in which it 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 a 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 plan inertia, 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 incorporates significant observed and unobserved heterogeneity to ensure precise 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.[4] 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.[5] I show that these inclusions are important: access to a broad network of *physicians* explains about 80% of the total network utility for broad-network plans, with only about 20% 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; 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 simply offer a counterfactual menu and compensate employees either through lower individual premium contributions ("co-premiums") or through wages for any lost utility.

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 is suggestive of a fundamental mismatch

[4] Together, these specialties comprise approximately 65% of all physician office visits. https://www.cdc.gov/nchs/fastats/physician-visits.htm

[5] 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.

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 at 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 \$52 per-household-per-month, implying that this mismatch results in about a \$620 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) frictions in altering co-premiums or non-insurance benefits; (b) employer mistakes or misperceptions; and (c) heterogeneity in the types of employees the employer values when designing benefits. While I am not able to causally separate these channels, I find substantial evidence for (c). In simulations where I shift the demographics of the employee pool, I find that the GIC placing a higher weight on the preferences of older workers may potentially account for up to 30% of the estimated employer-employee mismatch. More strikingly, favoring workers in less dense geographic regions—with less provider competition—may potentially account for as much as 80% of the mismatch.[6] Importantly, these households do *not* seem to be the ones with the highest health risk or those with the highest ex-ante probability of health care utilization.[7]

These patterns persist even in the network choices of a simulated sample of large, private employers, who theoretically have more flexibility in designing benefits and responding to labor market frictions. In fact, for private firms, I show that favoring older workers may be a more salient factor. All together, there is strong evidence that networks may be driven either out of regional equity concerns or employer beliefs that attracting the marginal older employee may yield productivity benefits for the firm that satisfying the inframarginal worker would not. The latter would also be consistent with employers favoring the preferences of managers, executives, or employees that otherwise have stronger labor market bargaining power.

While these results suggest a mismatch between employee preferences and employer incentives, the observed outcomes are exacerbated 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 these regions. 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 "overprovision" of broad networks cannot simply be resolved by offering more choice of plans without also substantially widening the premiums between those plans.[8] In other words, the mere presence of a narrow-network option, despite being lower

[6]These simulations should not be interpreted as causal decompositions, but rather potential upper bounds on plausible channels for how an employer-employee mismatch might arise.

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

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

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 moves from its current scheme of subsidizing 75% of total premiums to an "Enthoven" style managed-competition scheme.[9] In implementing this policy, I find that the price of the broad-network plans rises to such an extent that the employer ceases offering its flagship broad-network products altogether. While this approach would lead to large aggregate gains (about $60 per household per month)—driven by the substantial decrease in employer spending—it would also impose severe distributional consequences. In particular, older households and those in regions with a sparser set of providers would see considerable losses from the removal of these products. 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 more frequently adopt Enthoven-style pricing.

The story changes drastically, however, when I allow the employer to deviate from the uniform benefits structure by setting premiums and networks differentially by region. In this scenario, the employer is predicted to drop access to the costliest broad-network plans in only three of the seven rating regions in the state, while preserving access in the remaining regions. The gains from this move are more modest than the Enthoven approach in aggregate (about $25 per household per month), but the policy sees minimal adverse distributional consequences. Indeed, 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 for the highest willingness-to-pay consumers.

This paper contributes 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 analogous 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

[9] Under this approach, the employer maintains its reimbursement for the lowest-price plan, but fully passes through the additional cost for more expensive options.

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), 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 purchasing insurance through the GIC. Therefore, the GIC has an interest in not only providing satisfactory health benefits for its existing members, but competing for new members as well. In total, there are approximately 300,000 enrollees on the GIC per year, representing approximately 8% of the Massachusetts ESI market.

The GIC contracts with multiple health insurance carriers and provides multiple competing plans for enrollees. It contracted with six carriers throughout my sample period: Fallon Community Health Plan, Harvard Pilgrim Health Care, Health New England (HNE), Neighborhood Health Plan (NHP), Tufts Health Plan, and Unicare Health. While most of these carriers offer coverage in each region of the state, Fallon operates primarily in the central part of the state (e.g. Worcester) and Health New England primarily serves Western Massachusetts (see Appendix B). 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 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 cost 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 (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). 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.[10]

[10] 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.

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 evaluating the impact of narrow-network product introduction (Gruber and McKnight, 2016). The holiday was fairly successful, inducing approximately 10% of enrollees to switch, 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 individual co-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 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 "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 (hereafter "Harvard Narrow") and 2% for Tufts Spirit (hereafter "Tufts Narrow"). Interestingly, despite having lower co-premiums, Tufts Narrow had a significantly lower market share than Harvard Narrow.[11] 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	Narrow	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	293,125		
Average Member Age	36.07		
Average Subscriber Age	48.04		

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

[11] This represents a difference of almost 12,000 members.

2.2 Data Sources

I use two primary data sources 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 Massachusetts APCD, a comprehensive database of medical claims from public and private payers in Massachusetts from 2009-2013 (Massachusetts Center for Health Information and Analysis, 2013). The claims data contain detailed information on both hospital 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 five-digit zip codes), 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 amounts charged to the patient (e.g. copays). 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 identifiers, market identifiers (e.g. individual, group, GIC, etc.), and cost-sharing features.

I create samples for hospital admissions, physician visits, and insurance plan choice. Throughout the paper, I focus on patient demand for hospitals as well as three physician specialties: primary care physicians (PCPs), cardiologists, and orthopedists. A detailed description of the different subsamples is presented in Appendix A, along with summary statistics for each provider group.

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 2009-2013 (IMS Health and Quintiles, 2013). The SK&A includes information on approximately 95% of all office-based physicians practicing in the United States and the data are verified by the proprietors over the telephone. Data include 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 across all the locations of the medical group.

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 data on the GIC website (Group Insurance Commission, 2013). 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 presented in Appendix subsection A.2.

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 2013: Harvard Broad, Harvard Narrow, and Tufts Narrow.[12] The colors of the points on the maps refer to providers that are owned by the largest health systems in Massachusetts: Partners, Steward, Atrius, UMass, Lahey, and Baystate, with the gray points aggregating all others. The points are sized according to the market share with respect to the particular specialty. Looking at both figures, it is clear that Partners (navy blue) and Atrius (red) dominate much of the providers in Massachusetts, with Partners owning 133 PCP practices and Atrius owning approximately 46.[13] Panel (a) of both figures shows that there is a large density of providers in Eastern Massachusetts (particularly Boston and the surrounding suburbs) and that Partners and Atrius are largely concentrated in these areas. Atrius also owns several practices in Central Massachusetts, owing to its purchase of the Fallon Clinic in Worcester in 2011.

Panels (b) and (c) of Figure 1 reveal that the Harvard Narrow and Tufts Narrow still cover a large number of hospitals and physicians in Massachusetts, with one major difference between the broad and narrow networks is that most Partners hospitals were dropped from each narrow plan. As noted in Table 1, Harvard Narrow has a significantly higher market share than the Tufts Narrow, with almost three times the number of enrollees. However, as seen here, the hospital networks of both narrow plans are relatively similar. 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 this phenomenon. Figure 2 reveals that 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 (again, seen in the red dots). Moreover, relative to its impact in the hospital market, the removal of Partners *physicians* appear to play a smaller role, given the breadth of PCPs remaining in Harvard Narrow. This indicates that physician networks may be an important determinant of plan choice.

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

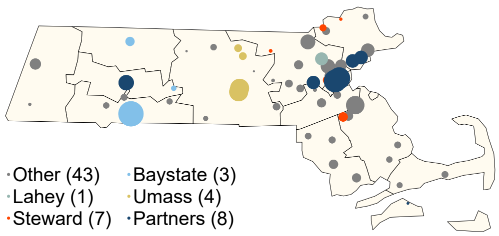
Variation in Narrow-Network Plan Enrollment: There is significant heterogeneity in terms of who enrolls in narrow-network plans. One significant predictor of narrow-network enrollment is whether the household is new to the GIC in any given year. Figure 3 depicts the share of consumers enrolling in narrow-network plans by year and by whether they were new to the GIC ("entering members") or whether they were existing GIC members ("existing members"). In 2009 and 2010, both types of members enrolled in narrow-network plans at rates between 10% and 17%. However, there is a large spike in the share of entering members enrolling in narrow-network plans in 2011 (to 30%), when the GIC introduced Harvard Narrow and Tufts Narrow (see Panel

[12] The maps use county shape file pulled from United States Census Bureau (2015).

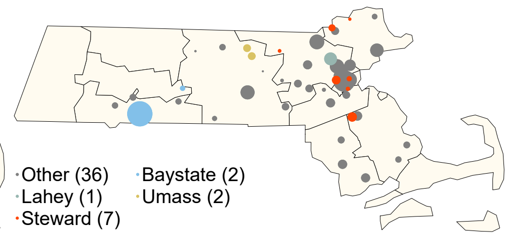
[13] In general, Partners is widely known to be one of the dominant players in the Massachusetts provider market, owning several large academic medical centers, including Mass. General Hospital and Brigham and Womens Hospital. Similarly, Atrius Health is one of the dominant players in the physician market,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

Figure 1: Hospital Networks by Plan, 2013

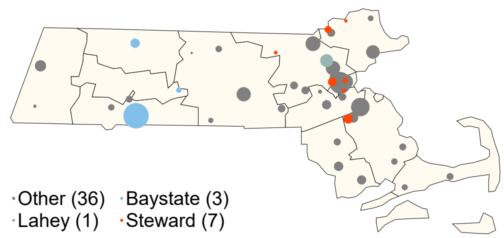
(a) Harvard / Tufts Broad



(b) Harvard Narrow



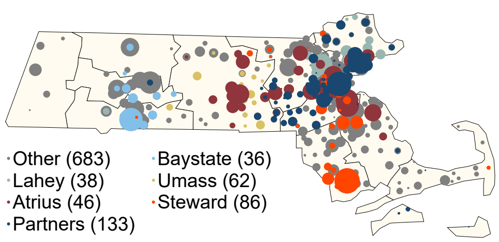
(c) Tufts Narrow



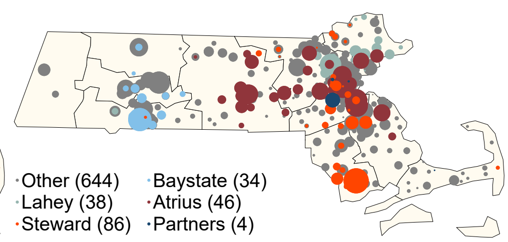
Notes: This figure plots the hospital networks of specified plans on the GIC in 2013. Sizes of the data points reflect relative market shares of the hospitals. Colors reflect ownership status (which health systems owns which hospital).

Figure 2: Primary Care Practice Networks by Plan, 2013

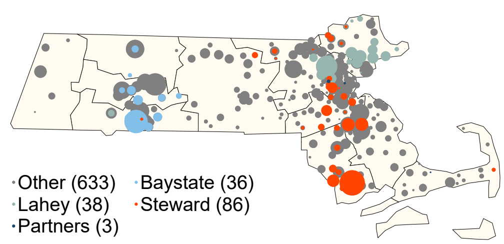
(a) Harvard / Tufts Broad



(b) Harvard Narrow



(c) Tufts Narrow

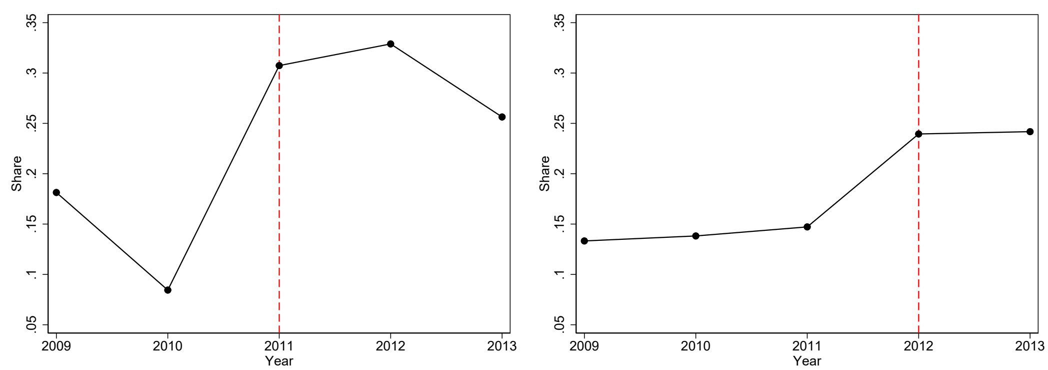


Notes: This figure plots the PCP practice networks of specified plans on the GIC in 2013. Sizes of the data points reflect relative market shares of the practices. Colors reflect ownership status (which health systems owns which practice).

(a)). This share remained high thereafter. Notably, the existing members saw no such spike in enrollment in 2011 with the arrival of these new products. Conversely, once the GIC introduced the "premium holiday" in 2012, there was a significant spike in the share of *existing* members (for whom the policy applied, Panel (b)) enrolling in narrow-network plans.

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

(a) Entering Members (b) Existing Members



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."

These differences in behavior between new members (or those making active choices) and existing members (those potentially making passive choices) suggest the presence of a high degree of inertia in choosing health plans (Handel, 2013). One potential criticism of this conclusion is that active choosers might have systematically different preferences for networks than passive choosers. For instance, new employees of large firms in general tend to be younger, which may make them more price-sensitive and more willing to choose narrow-network plans.

However, institutional details of the GIC would suggest that these demographic differences between new and existing members is fairly minimal. First, most new members to the GIC are entering from new municipalities contracting with the GIC for health plans, rather than new active state employees. 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 Appendix subsection B.2. I also show evidence that new consumers react to co-premium changes over time, whereas existing enrollees tend to remain inertial to plans even as co-premiums change significantly, further suggesting the role of inertia.

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 ESI. 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 each. In selecting these, employers incur a fixed cost of adding each additional product.
2. Given the products selected, employers set premiums for self-insured products, while 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 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 a set of in-network hospitals, 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 of that specialty, 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:

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

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

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_{ilt} and v_{irt} are observed characteristics of patient i with diagnosis l or requiring procedure r, T_{idt}^{s} and T_{iht} are 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.[14] 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. Details of the provider demand specification, as well as its estimation, identification, and parameter estimates, are presented in Appendix subsection C.1.

[14]Distances are computed using zip-code-level latitude and longitude estimates, taken from the SAS zipcode database (SAS Institute, 2009)

3.2 Consumer Demand for Insurance Plans

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

$$u_{Ijt} = \underbrace{-r_{Ijt}\alpha_I + EU^H_{Ijt}\beta_1 + \sum_s EU^s_{Ijt}\beta^s_{2I} + \mathbb{1}(Ij_t = Ij_{t-1})\beta_3 + x_{jt}\beta_4 + \eta_j}_{\delta_{Ijt}} + \omega_{Ijt} \tag{3}$$

Here, r_{Ijt} refers to the plan rate, or household 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 household's 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. These terms 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 Appendix subsection C.5. x_{jt} refer to time-varying plan characteristics, including an indicator for whether plan j tiered its PCPs in time t as well as the specific copay of the highest-tier specialists. η_j is the (time invariant) unobserved plan characteristics component, captured by a full set of plan fixed effects, 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.). I also 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 may therefore 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 co-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 $v^s_I \sim N(0, \sigma^s)$. 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.

[15] A network may, for instance, have fewer providers, and yet still yield a higher value of EU^s_{Ijt} for specialty group s if 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).

The market share of households of type I for plan j in market t is derived as the familiar logit 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 Appendix 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 households. These differences in expected utility stem from differences in household ages, locations (e.g. households that live closer to more prestigious doctors and hospitals than others), and illness histories (e.g. households with a higher disease burden).

The premium coefficient, α_I, 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. As discussed in section 2, this scaling is set by the GIC and remains fixed throughout my sample period, even as plans entered and plan characteristics changed. The expected utility term, however, increases linearly in the number of household members and, critically, increases differently depending on household characteristics. Households with children, for instance, will have a higher expected utility term when they are closer to highly desirable PCPs, but pay the same co-premium regardless of location.[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." I observe a substantial number of enrollees making choices for the first time, driven by new municipal entrants to the GIC. As shown in subsection 2.3, these "active choosers" make extremely different choices in plans, conditional on a rich set of observables. The second condition is that the choice set changes over the sample period. In my setting, I observe a panel of households making consecutive choices over time as plans and characteristics change, which I detail in Appendix subsection B.2. Moreover, 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).

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

[17] See Prager (2016) and Ho and Lee (2017) for additional discussion on this identification strategy. The primary identifying assumption in my context is that, controlling for age, income, and location, premium sensitivity does not vary across family type.

cost parameter. The logic is that if inertia to previous plans was driven by preference heterogeneity, we would not expect such considerable differences in choices between these two groups of enrollees when we include random coefficients in the model. The parameters describing unobserved persistent preferences can therefore be estimated from the choices made from new enrollees alone. Similar in spirit to Handel (2013) and Polyakova (2016), 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).

Estimates: Table 2 reports the results for the insurance plan demand model. Due to the high dimensionality of the data, 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 time-varying non-premium characteristics, plan fixed effects, as well as premium and network utility interactions with observables (e.g. age and location).

The first column presents results from a model that only includes hospital utility (EU^H_{Ij}), while the second column includes utility from both hospital and physician networks. Columns 3 and 4 repeat these specifications but add in the switching cost parameter. Finally, Column 5 presents the preferred specification: it includes hospital and physician utility, switching costs, and random coefficients on the utility of each of the three physician specialties.

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

There is, however, significant heterogeneity across the models in estimated premium elasticities and network preferences. These differences highlight three key insights. First, preferences for physicians—as opposed to just hospitals—are an important driver of demand for health plans. This is consistent with the patterns seen in subsection 2.3. Indeed, including physician networks nearly triples the estimated premium disutility, while significantly reducing the estimaeted coefficient on hospital utility (Models 1-2). This 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 \$213 per month to have their network reduced.[18] Moving to a model that includes physician networks, however, drastically reduces the implied WTP for hospital networks to merely \$18 per month, while yielding an implied WTP for physician networks of \$49 per month. This suggests that omitting physician networks causes WTP for hospitals to be greatly overstated, as it absorbs a significant amount of heterogeneity and dampens households' true premium sensitivity.

Second, the fact that households are inertial to their previously chosen health plans, conditional

[18] 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)	-0.0020^{**}	-0.0056^{***}	-0.0213^{***}	-0.0243^{***}	-0.0251^{***}
	(0.0009)	(0.0009)	(0.0017)	(0.0017)	(0.0017)
EU^{H}_{Ijt}	7.7752^{***}	1.8040^{***}	5.8329^{***}	1.8818^{**}	2.0542^{**}
	(0.6307)	(0.5058)	(0.8730)	(0.8589)	(0.8972)
EU^{PCP}_{Ijt}		0.1561^{***}		0.0954^{***}	0.2193^{***}
		(0.0215)		(0.0250)	(0.0582)
EU^{CAR}_{Ijt}		0.4214^{***}		0.4684^{***}	0.7283^{***}
		(0.0840)		(0.1222)	(0.2182)
EU^{ORS}_{Ijt}		1.2579^{***}		0.4456^{**}	1.2719^{***}
		(0.1311)		(0.1702)	(0.3376)
σ_{PCP}					0.1691^{***}
					(0.0398)
σ_{CAR}					0.5261^{***}
					(0.1796)
σ_{ORS}					1.0233^{***}
					(0.2329)
Prior Plan			4.9396^{***}	4.8962^{***}	4.9316^{***}
			(0.0886)	(0.0903)	(0.0944)
Plan FE	Yes	Yes	Yes	Yes	Yes
Obs.	41,673	41,673	41,673	41,673	41,673
Pseudo R^2	0.29	0.33	0.79	0.80	–
			Panel B: Willingness-to-Pay for Harvard Broad v. Harvard Narrow		
	No Switching Costs		Switching Costs		Random Coef.
WTP Hosp	\$213	\$18	\$15	\$4	\$4
WTP PCP		\$17		\$2	\$5
WTP CAR		\$14		\$4	\$6
WTP ORS		\$18		\$2	\$4
Switching Cost			\$232	\$202	\$197

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.

on network utility, explains a large share of enrollment into broad-network plans. The estimated plan switching cost is \$232 per month when households only have preferences over hospital utility (Model 3). When households are allowed to have preferences over physician networks as well, the switching cost estimate declines to \$202, again highlighting the important role of heterogeneity in physician network preference.[19] In both of these models, the magnitude of the premium disutilities further increases, and the overall implied WTP for Harvard Narrow v. Broad declines to about \$12 per month, a much more reasonable magnitude given observed premium differences. This large decline in WTP for provider networks relative to Models 1-2 demonstrates that switching costs are an important factor in determining plan choice.

Finally, to the extent that some of this persistence in plan choice is driven by unobserved preference heterogeneity, this ought to be captured by the inclusion of random coefficients. Indeed, the estimated standard deviations, σ, on all three physician specialties are large and significant, suggesting there is considerable variation in unobserved preferences (Model 5). Relative to a model without random coefficients, this *increases* the implied WTP for moving from Harvard Narrow to Harvard Broad by about 60% (about a \$7 per month increase), with the largest increase coming from WTP for PCPs. The estimated switching cost parameter, however, only declines by about \$5 per month to \$197, suggesting that unobserved preference heterogeneity had a fairly marginal impact on the estimated inertia parameter.[20]

Note that these averages are 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. These figures show that there is clearly significant heterogeneity in network preferences, with the average WTP being higher for the Fallon plans and certain household willing to pay nearly \$100 per month for access to the broader network. These distributions are reported in Appendix Figure C.4.

Taken together, 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 yields premium elasticities that are likely underestimated. In fact, the results imply 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 household I in plan j therefore be given by:

[19] 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 \$202 per month is approximately 1.2 times the individual premium for a broad network and approximately 50% of the family premium for the same network.

[20] This also suggests that a portion of the additional WTP may have been previously onto plan fixed effects.

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

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 household I in plan j be given by:

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

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:

$$\pi_{mt} = \sum_{j\in J_m}\sum_{I}\left(s_{Ijt}(\delta_{Jt},\theta)\left[R_{jt}(\delta_{Jt},\theta)\theta^R_I - \underbrace{c^o_{IjtH}(N^H_{jt}) - c^o_{IjtS}(N^S_{jt})}_{c^o_{Ijt}(N_{jt})} - c^u_{jt}(N_{jt})\theta^c_I \right]\right) \tag{8}$$

In the equation above, J_m refers to the set of products offered by insurer 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^R_I (recall this is 2.4 times the base premium if the household is a family). Finally, c^u_{jt} refers to to a base "unobserved cost" of health care for plan j in time t.[21] I assume that these costs scale linearly across household type, i.e.:

$$c^u_{Ijt} = c^u_{jt}\theta^c_I$$

where θ^c_I 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, 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. 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.[22] Finally, plans in Massachusetts are bound by state medical-loss-ratio (MLR) regulations 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

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

[22] 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.

(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:

$$\sum_I s_{Ijt}(\delta_{Jt}, \theta) R_{jt} \theta_I^R = 1.10 \sum_I (c_{Ijt}^o(N_{jt}) + c_{jt}^u(N_{jt}) \theta_I^c) \tag{9}$$

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 Appendix subsection C.6.

3.4 Employer Objective Function and Network Design

I assume that the GIC itself behaves as an employer. Specifically, in selecting product quality and setting prices, I assume it 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 ought to be willing to expand their product menu.

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).[23] However, they also include non-monetary opportunity or switching costs. For instance, fixed costs may reflect employers' beliefs that offering more plans might contribute to consumer confusion or suboptimal plan choice among employees.[24]

Formally the employer objective is:

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

where:

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

[24] 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.

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

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 employee 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.[25] The second term in the equation represents the payment in premiums to insurers the 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 the 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 a dollar of consumer surplus from *the network* by more than a dollar in savings from lower spending. Such plan designs would, therefore, be indicative of a mismatch between the employer and the average employee preferences to the extent 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 δ_{Jt}.

Estimation: I closely follow work by Ho (2009), Pakes et al. (2015), and Pakes (2010) in

[25] 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%.

constructing moment inequalities to bound the estimates of ρ and FC_j, rather than imposing an equilibrium through distributional assumptions on the parameters. Such moment inequality approaches have been 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.

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

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

Let δ^a_{Jt} be an 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 for 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:

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

where $m(\delta_J, \delta^a_J, \theta, z)$ is the moment for estimation and $g(z)$ is any positive function of instruments z.

For simplicity, the equations above omit any unobserved heterogeneity in demand and, therefore, condition on a fixed set of parameters, 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:

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

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:

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

Expectation Over New Enrollees: As described in subsection 2.3, the GIC sees new municipalities enter the market each year. As part of my estimation procedure, I therefore 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 are in Appendix subsection C.7.

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, is nearly infinite, making estimation largely infeasible. I therefore make several restrictions on the potential choice set for the GIC. First, I assume that the GIC cannot cease contracting with any insurer, but can adjust the number of plans offered by any insurer.[26] Second, I assume that the GIC can only alter the *number* of plans offered by smaller insurers (Fallon, Health New England, and Neighborhood Health Plan), but not the networks of their plans.[27] 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. Finally, I assume that the GIC can offer five potential networks for Harvard and Tufts, which are detailed in section 5. These networks were all, at some point, offered by at least one insurer in Massachusetts (and, as such, they satisfy network adequacy restrictions) and span a considerable range of network breadth, both in terms of hospitals and physicians. In total, this leaves permutations of 14 potential products.

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

[26] The GIC engages in long-term, five-year contracts with insurers. During my sample period of 2009-2013, insurers were under such a contract. As such, the assumption that the GIC could not drop an insurer during this time is reasonable.

[27] 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.

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. 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.[28] The estimate of the mismatch parameter, ρ, is 3.70, 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 (e.g. wages), we ought to expect this parameter to be close to 1, as the employer could pass back any savings from a move to narrow networks. 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. I also test robustness of these estimates in Appendix E. Overall, the results consistently suggest that the GIC "overweights" consumer utility from provider networks by between 3-4 times how the average employee values those networks.

The point estimate of fixed costs is \$3.98 million for each plan. Though this estimate appears quite high, it is actually a very small fraction of the GIC's overall net spending (row 2 of Panel A). One caveat here is that some of the estimates of $CS(\delta_{Jt}, \theta)$ may be driven by the addition and removal of logit error shocks. These logit shocks may bias parameter estimates, particularly of the fixed costs, if the model overestimates the gains to plan introduction. To test the sensitivities of the parameter estimates, Panel B reports ρ and FC_j when I set the logit shock to zero. Indeed, doing so yields a mismatch estimate that is similar to the baseline estimate, but significantly reduces the estimate of FC_j to approximately \$1.4 million per plan. We can interpret this, then, as the GIC spending between \$1.4 and \$4 million per plan depending on the specification. Although

[28] This is not surprising, given the large number of inequalities.

Table 3: Results of Employer Objective Function Estimation

	ρ	FC_j ($Millions)
Panel A: Estimating ρ and FC_j		
GIC/Employer	3.70	3.98
Percentage of Net Spending		0.41%
Percentage of Net Surplus		0.89%
Panel B: No logit error		
GIC/Employer	3.98	1.37
Percentage of Net Spending		0.14%
Percentage of Net Surplus		0.31%

Notes: Results from ρ and FC_j estimation for 2009-2013. Panel A reports parameters from the full model. Panel B reports parameter estimates assuming no logit error. 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.

these estimates, in theory, contain both tangible and non-tangible fixed costs, the magnitudes are actually in line with reported administrative costs estimates by insurers in Massachusetts.[29]

My main specification throughout the rest of the paper allows for the presence of logit shocks, but in Appendix F, I report counterfactuals 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.

4 What Drives Employer-Employee Mismatch?

The high estimate of ρ in Table 3 indicates that the employer weighs total consumer utility from a plan menu by nearly four times what it spends on health care and premiums. The implication of this estimate is that, when selecting plan menus for employees, employers overweight consumer valuation for broad-network plans relative to 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 reduce wages, modulo the tax exclusion on health insurance (Summers, 1989; Gruber and Krueger, 1991; Bundorf, 2002).[30]

There are several potential drivers of this mismatch between employer and employee preferences. The first is that there is considerable heterogeneity in both employee medical costs and preferences for broad networks. The ability to precisely target a plan menu to match these diverse preferences

[29] 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).

[30] It should be noted that part of my estimate on ρ might be driven by the well known tax exclusion for health insurance (Gruber and Lettau, 2004). A back-of-the-envelope calculation suggests that, even using extremely conservative assumptions, the tax exclusion would only explain a small portion of my estimated mismatch. Specifically, even assuming that each state employee had a marginal tax rate of 33%, employees would need to value a dollar in wages at 67% of a dollar in health insurance coverage. If there were no true employer "overweighting" of preferences, the expected ρ under this tax rate would be equal to about 1.5, which is considerably smaller than my estimate.

is limited. 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 vocal segments of employees who have high expected medical costs or otherwise value comprehensive insurance highly. Alternatively, employers may place emphasis on attracting and retaining employees who have high labor market bargaining power or those who they believe will be most productive at the firm. This may include, 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. While it does bear the responsibility of determining plan benefits for all state employees of Massachusetts, it 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).[31] Further, public-sector benefits are often subject to voter influence, as they are funded by taxpayers. The full cost of fringe benefits is 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 simply 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 et al., 2010). It might also be the case that employers *misattribute* employee switching costs as genuine preferences for broad-network plans.

I investigate the potential for each of these mechanisms in the subsections below.

4.1 Heterogeneity in Employer Preferences

Older and Sicker Employees: 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 ages.

To test this, I reweight the distribution of employees such that 90% of the employee pool is greater than 55 years old.[32] The results of this simulation are depicted in Panel 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 declines: ρ decreases to 2.97, while FC_j decreases to \$3.54 million. This conforms to expectation: the GIC

[31] This dovetails nicely with the first channel noted above, as union membership tends to traditionally skew towards older employees (https://www.bls.gov/news.release/union2.nr0.htm)

[32] At baseline, the share of employees 55 and older represent only about 30-35% of the population, depending on the year.

indeed places an outsized 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 such that the average employee is sicker (Column 3).[33] Interestingly, the estimated mismatch does not decline as much as for the older population, suggesting that either households who value broad-network plans the most are not necessarily the costliest employees or that the GIC weights the preferences of older—though not necessarily sicker—employees more.

Table 4: Employer Objective Function Parameters For Different Populations

	(1) Baseline	(2) Older	(3) Chronic	(4) Regions 1,4,6	(5) Region 4	(6) Prior Doc
Panel A: Estimation with All Moments, GIC						
ρ	3.70	2.97	3.12	2.46	1.62	2.19
FC_j ($Millions)	3.98	3.54	2.75	2.48	2.50	2.20
Panel B: Estimation with Fixed Number of Plans, GIC						
ρ	3.71	2.97	3.10	2.46	1.59	2.20
FC_j ($Millions)	–	–	–	–	–	–
Panel C: Estimation on Private Employer Sample						
ρ	2.83	1.69	3.05	1.53	1.92	2.09
FC_j ($Millions)	–	–	–	–	–	–

Notes: Results from ρ and FC_j estimation for 2009-2013. Column 1 presents estimates for the baseline population of GIC enrollees. Columns 2-6 present estimates that reweight the population such that 90% of households are older (age 55 and older), have a member with a chronic condition, reside in certain rating regions, or have previously used a provider only covered by a broad-network plan. Panel A reports the estimates 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. FC_j reported in millions of dollars.

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

Employees in Less Competitive Regions: 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: Western Massachusetts (Rating Region 1), the North Shore (Rating Region 4), and Bristol/Plymouth (Rating Region 6).[34] The weight on consumer surplus falls again, this time to 2.46, a 34% decline from the baseline estimate. The effect

[33] I define this as the member having an "unstable" chronic condition, i.e. the member is diagnosed with a chronic condition, according to the AHRQ definition, and the member previously experienced at least one inpatient hospital stay.

[34] In aggregate, enrollees in these areas make up 17% of the GIC population. For details on network coverage by rating region, see Appendix B.

is the most prominent in Region 4, as reported in Column 5. Here, the implied employer-employee mismatch falls dramatically, to just 1.62, a 56% decline. This is noticeably closer to 1, lending credence to the theory that the employer is, at least in part, choosing to offer broad-network 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.

These regions 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.[35] 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 reach providers. Finally, these rating areas are close to the state border and may therefore cater to households who desire access to providers in neighboring states.[36] Taken together, the characteristics of these regions imply a stronger loss for employees from removing a provider relative to areas in the state with a higher volume of nearby options. Indeed, the high variation in market dynamics across the regions highlights the difficult task the GIC faces in designing a single menu of options to satisfy all its members.[37]

Employees with Prior Provider Relationships: I also consider simulations in which I reweight the population of the employee pool such that most households have had a prior relationship with a provider. Column 6 reports estimates such that 90% of the population had previously used any provider only covered by a broad-network plan. Appendix E reports additional estimates by specific health system. The results do show that the mismatch parameter declines substantially for households with prior provider relationships, implying that employers may overweight these populations in their network design. However, this is not the case for all health systems, even ones that have a large presence in the state (e.g. Atrius). Moreover, the estimates of ρ from the regional analyses in Column 4 are comparable, and in the case of the North Shore (Region 4) they are substantially lower. This suggests that while prior provider relationships are likely to play a role, regional factors potentially play a larger role in employer benefit design.

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. To do this, I use the APCD to construct a sample of private employers. I then simulate narrow-network offer distributions by supplementing the APCD with microdata from the Kaiser/HRET annual employer survey (Kaiser Family Foundation, 2016). Overall, this produced 123 large firms, approximately 8% of which offered a narrow-network plan in 2013. Details of the construction of the private employer sample are given in Appendix D.

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

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

[37] 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 2.83, suggesting that private employers still substantially "overweight" consumer preferences for health insurance relative to the average, though to a lesser extent 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.[38]

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. Most notably, the implied employer-employee mismatch among the older population drops to just 1.69, a 40% decline, compared with just a 20% decline for the GIC. Interestingly, the mismatch for the sicker population barely budges from the baseline. The mismatch term does decline further when considering the same regions that drove the result in the GIC sample, but by a lesser extent.[39] Taken together, this highlights some potentially important differences in incentives between private and public firms. Whereas regional equity appears to play a large role for the GIC, the weight firms place on the preferences of older employees may account for a larger share of the mismatch for private firms. This could potentially indicate the influence that managers or executives wield over benefit design.

4.3 Employer Misperceptions and Other Frictions

Another possible driver of employee-mismatch is employer misperceptions of employee preferences. For example, employers may mistake switching costs for genuine network utility or they may otherwise assume that the preferences of a vocal segment of employees to reflect the preferences of the average employee. To test whether there is evidence that the observed plan menus reflect such mistakes, I estimate a model where I load the entirety of plan switching costs onto network utility. These tests do suggest a potential role for employer misperceptions, though I still estimate a high mismatch parameter. Moreover, the reweighting exercises in subsection 4.1 continue to produce similar patterns. I therefore conclude that, even in the presence of these frictions, employer weighting of certain population segments still may be a significant driver of the mismatch. Details of these checks are presented in Appendix E.

Finally, estimates of ρ might be sensitive to model specification choices, including assumptions on physician inertia, active plan choice frictions, and moments used for estimation. In the appendix, I also estimate several additional specifications on the employer objective function addressing these potential frictions.

[38] I note two caveats about these estimates. The first is that I am using the same estimated demand parameters here as I did for the GIC sample. As such, the assumption is that employees working for private firms have similar valuations of networks and similar premium sensitivity—conditional on the observables in my demand model—as those working for the state. Second, I assume the same fixed markup pricing rule as the GIC. To the extent that private firms deviate from this, my premium predictions from alternate networks might be somewhat biased.

[39] The weighting for Boston among private employers was 2.70, 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.

4.4 Summary

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 three plausible phenomena. First, employers may be attempting to attract highly productive workers at the margin, even at the expense of the inframarginal employee. Second, employers may be favorably weighing employees with higher labor market bargaining power, such as those in managerial roles or those with strong union protections. These two mechanisms may play a particularly large role for private employers, where I show a substantial reduction in the estimated employer-employee mismatch for older, but not sicker workers. Finally, employer plan designs might be motivated out of equity concerns, particularly for workers living in less competitive provider markets. That ρ declines by a large magnitude for GIC employees living in certain regions is highly suggestive that this may be the operative mechanism for public employers. Regardless of the type of employer, this overweighting of the average employee's preferences leads to significant welfare distortions, as I address in section 5.

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 five 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 (equivalent to Harvard Primary Choice)
- N2: A "narrow" network where Partners *and* Atrius are removed (equivalent to Tufts Select)
- N1: A "narrow" network where Partners and Atrius are removed, along with other hospital and physician groups (equivalent to Tufts Spirit)
- VN: A "very narrow" network where many hospital and physician groups are removed (equivalent to Fallon Direct)

Here, I allow the employer to completely sever a relationship with an insurer. However, I make a "no uninsurance" assumption such that all households must have access to at least one plan that has network coverage in their county. For example, the GIC cannot offer a menu that just contains Health New England, since this plan only operates in counties in Western Massachusetts, and hence this menu would leave households in Eastern Massachusetts without a network. 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 (defined as consumer surplus minus health spending minus fixed costs) changes. The change in consumer surplus, ΔCS, is defined as:

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

where δ^a_{Ijt} is the counterfactual plan menu.

Table 5 reports the equilibrium predicted products/networks offered from the simulations (Panel A). 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.[40] Model 3 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. Overall, the number of plans falls from 8 to 7. Under this scenario, consumer surplus significantly declines, by about \$67 per household per month (Panel C). However, this loss is more than compensated for by a significant decrease in health spending of approximately \$115 per household per month and an additional \$4 decrease in fixed costs. This implies that the employer-employee mismatch decreases total surplus by about \$620 per household per year.

For comparison, I also report predicted networks and surplus changes when I fix ρ to the estimated mismatch parameter for the private employer sample from Table 4. Under this scenario, the employer still preserves the Fallon and Tufts broad-network plans, suggesting that the employer-employee mismatch is playing a significant role in network design. However, here the employer drops the Harvard Broad plan, which alone yields substantial social surplus gains from the reduction in health spending (see Model 4).

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 subsidizes 75% 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

[40] 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.

Table 5: Counterfactuals: Equilibrium Networks Chosen Under Uniform Pricing

Insurer	Network	(1) Observed	(2) Pred.	(3) $\rho = 1$	(4) $\rho = \rho_{prv}$	(5) Enthoven
Panel A: Equilibrium Plan Menus/Networks						
Fallon	VN	x	x	x	x	x
Harvard	VN			x		
Tufts	VN			x		x
Harvard	N1					x
Tufts	N1	x		x		x
Harvard	N2		x		x	
Tufts	N2					
HNE	N	x	x	x	x	x
NHP	N	x	x	x	x	x
Harvard	M	x	x	x	x	x
Tufts	M					
Fallon	B	x	x		x	x
Harvard	B	x	x			
Tufts	B	x	x		x	
Total Plans		8	8	7	7	8
Panel B: Welfare/Spending with Fixed Menu						
ΔCS				–	–	$-\$139.19$
$\Delta Costs$				–	–	$-\$180.06$
ΔFC				–	–	–
$\Delta Surplus$				–	–	\$40.87
Panel C: Welfare/Spending with Menu Changes						
ΔCS				$-\$66.70$	$-\$13.18$	$-\$121.17$
$\Delta Costs$				$-\$114.68$	$-\$43.22$	$-\$180.06$
ΔFC				$-\$3.72$	$-\$3.72$	\$0.00
$\Delta Surplus$				\$51.70	\$33.76	\$58.89

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. "$\rho = \rho_{prv}$" refers to predicted networks when ρ is set to the estimated ρ for private employers. "Enthoven" refers to predicted networks when employers are moved to a uniform pricing mechanism that fixes employer subsidies to the current lowest-priced plan. Panel B reports the welfare and cost changes assuming plan menus remain fixed. Panel C reports these quantities allowing the employer changes to menus. "ΔCS" refers to change in consumer surplus per-household-per-month. "$\Delta Costs$" refer to the change in GIC health costs per-household-per-month."ΔFC" refer to changes in fixed costs from the new menus. $\Delta Surplus$ refers to the change in total surplus.

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.[41] I conduct a simulation of one such scenario: moving the GIC to an "Enthoven"-style managed competition approach, wherein the employer retains its current subsidy for the lowest-cost plan, but passes on the full premium differential for all other plans.[42]

Results of this simulation are presented in Model 5 of Table 5. Under this new pricing structure, if plan menus are held fixed, then consumer surplus declines significantly by \$140 per household per month (Panel B). This is intuitive as the premium differential between broad and narrow networks widens considerably. However, this is more than offset by a substantial decrease in employer costs, driven by consumers switching away from broad-network plans towards lower-cost narrow networks. The counterfactual market shares and premiums under the Enthoven approach are reported in Appendix Table F.1. Indeed, they show enrollment in broad-network plans falling by more than 50%, with consumers shifting into more narrow or medium-sized networks (e.g. NHP or HNE). Overall, this move leads to total surplus gains of about \$40 per household per month, about \$10 less than if the employer-employee mismatch were fully eliminated.

Enthoven Approach with Endogenous Menus: If the GIC is permitted to adjust its plan menu in response to this policy change, while maintaining the same flat subsidy it offers for the lowest-cost plan on the existing menu, then it is predicted to switch to a combination of plans that is very close to the scenario when setting $\rho = 1$. In particular, the GIC still maintains 8 offered plans, but reduces the network breadth of many, including the Tufts and Harvard broad-network products. Unlike the scenario where the employer-employee mismatch is eliminated, however, there is not a full unraveling of broad networks. Rather, the GIC preserves access to one broad-network plan: Fallon Broad. Total surplus from this change increases by nearly \$60 per household per month, a full \$20 per month more than when holding plan menus fixed. By definition, the decrease in employer medical spending is identical to the decrease predicted in the scenario where the menus are held fixed, as the employer subsidy is also held fixed by assumption. This \$20 surplus gain, then, comes from consumer utility (row 1 of Panel C). Specifically, it comes from the employer shifting consumers with low WTP but high switching costs into narrower network plans. In the scenario where menus are fixed, about 30% of GIC enrollees remain on Harvard and Tufts broad plans, even though many would benefit from moving to a plan with a medium-sized network. By removing these options, the employer is able to overcome these switching costs. The combination of these two forces: altering the pricing schedule *and* adjusting the plan menu ultimately results in surplus gains that even exceed fully eliminating the employer-employee mismatch.

[41] 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.

[42] This exercise is similar to Bundorf et al. (2012).

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, while still maintaining the fixed markup assumption from subsection 3.3.[43] The results are reported in Table 6. When the employer holds plan menus fixed at their observed level, but varies prices by region, 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 alter its plan menus, there are significant effects to plan design, welfare, and spending. In Panel A, 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). In Region 4, the employer preserves access to most of the existing plans, but drops Health New England (which does not operate in the region) and Fallon Narrow. In Region 1, however, the employer drops all but four plans and *only* retains broad-network products, as well as Health New England. Finally, for the Boston region, the employer significantly *narrows* the networks in its menu. Specifically, the GIC drops three plans from the menu, including both Harvard Broad and Fallon Broad. In fact, it drops *all* plans by Fallon and HNE, and instead adds a narrow-network plan 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. The employer continues to offer eight distinct plans, so fixed costs do not increase. Overall total surplus increases by \$25 per household per month.

5.4 Distributional Consequences

While the Enthoven approach yields significant surplus gains *on average*, employers may find it undesirable to impose a policy that results in the elimination of several flagship broad-network products, particularly if the distributional consequences are severe. In Panel A of Figure 4, I plot the predicted surplus changes from this move across households by age and region. I focus here on households enrolled in family plans. The results for individuals are similar and shown in Appendix F. For the purposes of this exercise, I assume that the employer passes the cost savings from the approach evenly across households.

I predict that changing to an Enthoven approach would result in net surplus increases for most households, consistent with the results seen in Table 5. However, these surplus gains diminish rapidly with age. This conforms to expectation: younger households rarely interact with the health care system and would therefore see substantial surplus gains from being compensated for a menu of narrower networks. Households at the upper end of the age distribution see much smaller gains from this approach, with households at around age 67 starting to see surplus losses.

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

Table 6: Counterfactuals: Equilibrium Networks Chosen Under Region-Based Pricing

Insurer	Network	(1) Observed	(2) R1	(3) R4	(4) R5
Panel A: Equilibrium Plan Menus/Networks					
Fallon	VN	x			
Harvard	VN				
Tufts	VN				
Harvard	N1				x
Tufts	N1	x			
Harvard	N2			x	x
Tufts	N2				
HNE	N	x	x		
NHP	N	x		x	x
Harvard	M	x		x	x
Tufts	M				
Fallon	B	x	x	x	
Harvard	B	x	x	x	
Tufts	B	x	x	x	x
Total Plans		8	4	6	5
Panel B: Welfare/Spending with Fixed Menu					
ΔCS				$-\$0.46$	
$\Delta Costs$				$-\$0.01$	
ΔFC				–	
$\Delta Surplus$				$-\$0.45$	
Panel C: Welfare/Spending with Menu Changes					
ΔCS				$-\$7.28$	
$\Delta Costs$				$-\$32.46$	
ΔFC				\$0.00	
$\Delta Surplus$				\$25.18	

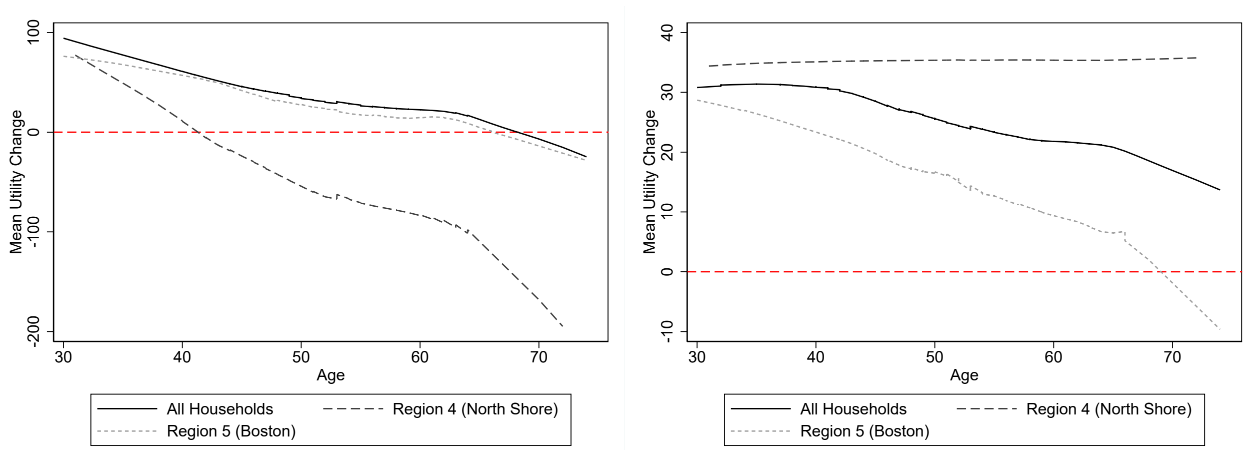
Notes: GIC observed and predicted products offered under region-based rating. "R1" refers to plan networks for Region 1, etc. Panel B reports the welfare and cost changes assuming plan menus remain fixed. Panel C reports these quantities allowing the employer changes to menus. "ΔCS" refers to change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

However, these aggregate numbers mask significant heterogeneity by region. Households in the North Shore (Region 4) begin seeing surplus *losses* at age 40, with households standing to lose about \$100 per month, on average, by age 60. Even in the Boston area (Region 5), where there is more competition among health care providers, households start seeing modest surplus losses at around age 65 from the move to narrow networks. Notably, these consumers represent a tiny share of the overall pool. Most younger households and even most consumers in the Boston region would see large overall gains from an Enthoven policy. Indeed, this coincides precisely with the results in section 4. Namely, the older households—particularly those in less competitive regions—are the ones who value broad networks the most. As such, the removal of those networks represents a significant loss for these consumers. These distributional concerns may therefore explain why firms do not adopt Enthoven-style policies more uniformly: these are precisely the households that the employers appear to favor when designing plan menus.

Figure 4: Total Surplus Changes by Age

(a) Enthoven Approach

(b) Region-Rating Approach



Notes: This figure plots the average utility change across households with family plans by age from implementing an Enthoven pricing approach (Panel A) and a region-rating approach (Panel B). All estimates allow 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 change in spending incurred by the GIC.

Turning to the region-based rating approach, while the aggregate surplus gains are only a third of that the size of moving towards an Enthoven-style pricing approach, the distributional consequences are considerably less severe (Panel B of Figure 4). Indeed, averaged across all households, utility from the menu change declines as consumers age. This is 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, nearly every household's net utility change remains positive, regardless of location or age. The exception is households in Boston where the eldest member is at least 70, who see very modest utility declines that are less than \$5 per month. All other households in Boston do stand to lose the most from the change, both due to reduction in the number of plans and to the loss of Harvard Broad, but they can be more than compensated with spending savings. Meanwhile, households residing in the North Shore—the ones the GIC appears to value the most—see their utility remain flat, a stark contrast with the Enthoven approach. As seen in Table 6, this is due to the GIC preserving all the broad-network plans in the region.

5.5 Summary

The results of these simulations shed light on an important tradeoff between equity and efficiency in plan design. To achieve efficiencies through uniform benefit design, the employer would likely need to offer products that impose severe utility costs on precisely the consumers who they value the most. Indeed, by removing broad-network products from the menu and shifting towards narrower networks, most younger employees living in dense, urban markets see large gains, while older workers living in less dense markets see losses. This is driven by the fact that the urban markets in the state also have the highest cost providers, in spite of the fact that the households residing in those areas have lower WTP for broad networks. As such, a non-uniform benefit scheme might be more tenable for employers. Indeed, by offering different networks in different regions, employers can shift these consumers with low valuations of broad networks into lower-cost products, while using the savings to compensate consumers in regions with high valuations. I show that this yields far less surplus in aggregate—about a third of the Enthoven approach—but ultimately minimizes adverse distributional consequences for workers who the employer weighs heavily in its benefit design.

6 Conclusion

The rollout of the ACA 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 accounts, 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 individual marketplaces, 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 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 persistence in offering broad networks—even conditional on possible misperceptions and frictions—may be driven by firms placing high weight on the preferences of older consumers and those in select geographic markets. The weight on older consumers is consistent with employers perhaps seeking to design health plans that attract more productive employees, or otherwise employers responding to employees in positions with higher bargaining power. 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 partly 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).

Abaluck, J. and Gruber, J. (2020). When less is more: Improving choices in health insurance markets. *Working Paper*.

American Hospital Directory (2013). Free hospital profiles. https://www.ahd.com/.

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.

Bhargava, S., Loewenstein, G., and Sydnor, J. (2017). Choose to lose: Health plan choices from a menu with dominated options. *Quarterly Journal of Economics*, 132(3).

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.

Center for Medicare and Medicaid Services (2009). Pfs relative value files. https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeeSched/PFS-Relative-Value-Files.

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.

Group Insurance Commission (2009-2013). State employee basic life and health plan rates and limited network hospital list. https://www.mass.gov.

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 Lettau, M. (2004). How elastic is the firm's demand for health insurance? *Journal of Public Economics*, 88.

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).

Handel, B. R. and Kolstad, J. T. (2015). Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8).

Healthcare Cost and Utilization Project (2015). Clinical classifications software for icd-9-cm. Technical report, Agency for Healthcare Research and Quality.

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*.

IMS Health and Quintiles (2009-2013). Sk&a office-based physician database.

Kaiser Family Foundation (2014-2016). Employer health benefits survey.

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*.

Massachusetts Center for Health Information and Analysis (2009-2013). Massachusetts all payer claims database.

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*.

SAS Institute (2009). Sashelp zip code file. https://support.sas.com/resources/papers/proceedings/proceedings/sugi31/143-31.pdf.

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.*

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.

United States Census Bureau (2015). Cartographic boundary files - shapefile. https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html.

ONLINE APPENDIX

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

Nicholas Tilipman[1]

A Data Descriptions

A.1 APCD Sample Creation

Hospital Admissions: I first create a sample of hospital admissions, which I use to estimate patient demand for hospitals. To do so, I limit the APCD to any facility claim flagged as an inpatient admission within my 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 use the organization's National Provider Identification (NPI) number to match the hospital to a set of hospital characteristics from the American Hospital Directory (AHD) database (American Hospital Directory, 2013). 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 are aggregated to the hospital admission level and the "allowed amounts" are summed over all service lines for that particular admission 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 (Healthcare Cost and Utilization Project, 2015).

Table A.1 contains the hospital sample summary statistics for hospital admissions from 2009-2013. On average, patients admitted to Massachusetts hospitals are 45 years old, and about half of the patients suffer from a chronic condition. Approximately 7% of patients are admitted with a primary cardiac condition, while about 31% are admitted with an obstetrics-related diagnosis. Patients are, on average, willing to travel approximately 13 miles to visit a hospital, and visit teaching hospitals approximately 80% of the time, while visiting academic medical centers approximately 37% 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

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

Table A.1: Hospital Sample Summary Statistics

	Mean	Std Dev
Patient Characteristics		
Age	45.08	22.94
Female	0.67	0.47
Chronic	0.53	0.50
Neurological	0.02	0.12
Cardiac	0.07	0.26
Obstetrics	0.31	0.46
Imaging	0.26	0.44
Hospital Characteristics		
Distance	12.62	14.07
NICU	0.89	0.31
Neuro	0.95	0.22
MRI	0.94	0.24
Critical Access	0.01	0.10
Teaching	0.80	0.40
Specialty	0.05	0.22
Academic Medical Center	0.37	0.48

Notes: Hospital sample summary statistics 2009-2013. Diagnosis characteristics (e.g. Neurological, Cardiac, etc.) are derived from AHRQ's Clinical Classifications Software categories and Chronic Conditions Indicators

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. PCP practices are defined as any medical group that contains at least one physician who 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 components 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, and surgical care.

For each service line, I merge in Medicare Part B physician fee schedules from the Center for Medicare and Medicaid Services (CMS) (Center for Medicare and Medicaid Services, 2009). These data contain annual federal updates to each procedure (CPT) code's "Relative-Value-Unit" (RVU) weight, which are constructed to assign each service a measure capturing its relative resource intensity to other procedures. These weights are used by CMS to determine Medicare payment rates for physicians. As such, I use them both a proxy for procedure intensity and in the construction of insurer-physician negotiated rates, described further in subsection C.6. 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.

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 that are more resource-intensive and thus assigned higher RVUs. About 85% of primary care patients saw a doctor between 2009 and 2013 that they also have seen previously, while this number was about 64% for cardiologists and about 61% for orthopedists. Distance traveled to PCPs was about 6 miles, on average, and about 10 miles for cardiologists or orthopedists. When seeing a PCP, patients on average visit practices with 41 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 owned by hospitals or owned by health systems.

Table A.2: Physician Sample Summary Statistics

	PCPs	Cardiologists	Orthopedists
Age	47.92	54.12	44.36
	(15.59)	(13.87)	(18.52)
Female	0.57	0.43	0.52
	(0.50)	(0.49)	(0.50)
RVU	2.61	2.96	5.55
	(1.64)	(4.90)	(12.56)
Used Doc Previously	0.85	0.64	0.61
	(0.36)	(0.48)	(0.49)
Used Med Grp Previously	0.86	0.70	0.65
	(0.35)	(0.46)	(0.48)
Used System Previously	0.86	0.74	0.67
	(0.34)	(0.44)	(0.47)
Distance	5.57	9.54	9.69
	(5.55)	(10.99)	(10.42)
Doctors on Site	41.48	116.86	65.25
	(105.00)	(180.55)	(143.49)
Number of Locations	8.89	9.96	5.51
	(8.63)	(9.29)	(8.31)
Part of Medical Group	0.72	0.72	0.63
	(0.45)	(0.45)	(0.48)
Owned by Hospital	0.26	0.43	0.20
	(0.44)	(0.49)	(0.40)
Owned by System	0.52	0.59	0.32
	(0.50)	(0.49)	(0.47)

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.

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 five-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. Finally, I merge this list of GIC members to external, publicly available data on GIC annual plan premiums and hospital networks. 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 ensured 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 merge the data by provider name. 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 a 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 provider 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 (e.g. they may 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 I code this "ownership" variable as that group. If the highest level of ownership is a particular hospital (i.e. a hospital-owned physician practice), then this "ownership" variable is coded as that hospital. Finally, if the highest level of ownership is reported as a health system (e.g. Partners Health Care, Steward Health System), then this "ownership" variable is coded as that system. This variable is used primarily in constructing networks (see below).

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 (e.g. Fallon), others only report the list of hospitals. I therefore use the "ownership" variable defined above. I 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), then any physicians or groups that are owned by that system are also excluded.[2] For 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 in a particular year is considered in that plan's network. For robustness, I also construct networks that default each of these small practices to being in the plan's network *unless* a majority of claims from a particular plan in a particular year are explicitly flagged as being "out of network."

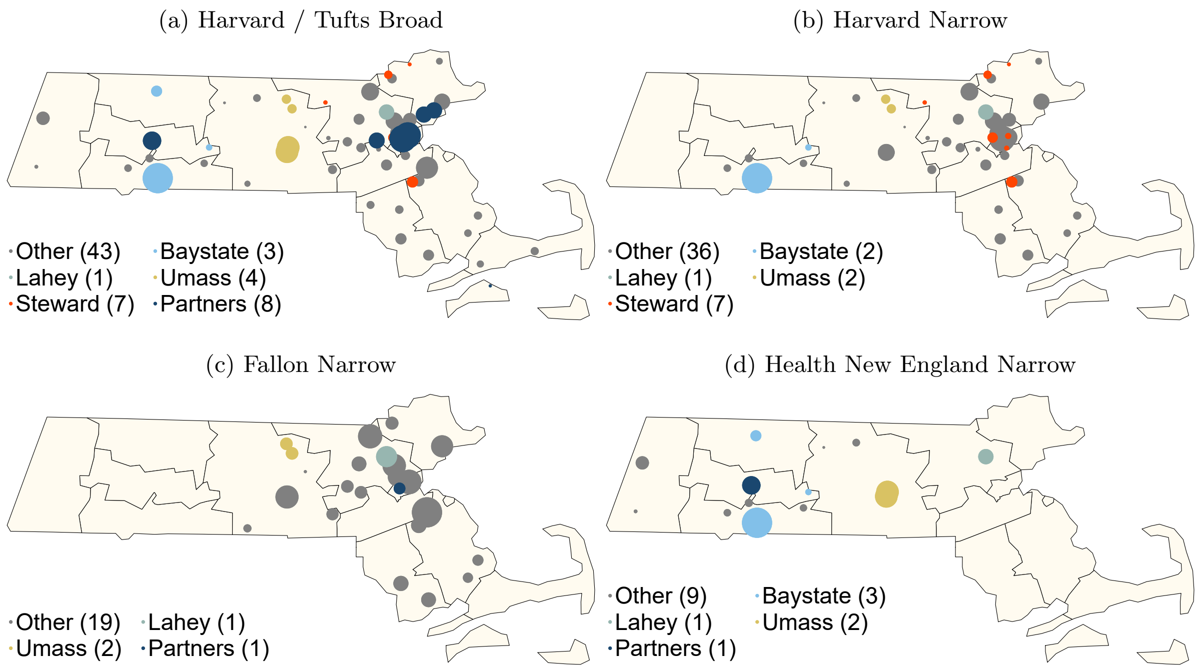
[2]In practice, this is a close approximation of contracts observed on the GIC. Harvard Primary Choice and Tufts Spirit, for instance, cease contracting with all Partners-owned medical groups as well as Partners hospitals.

B Additional Descriptives

B.1 Additional Network Figures

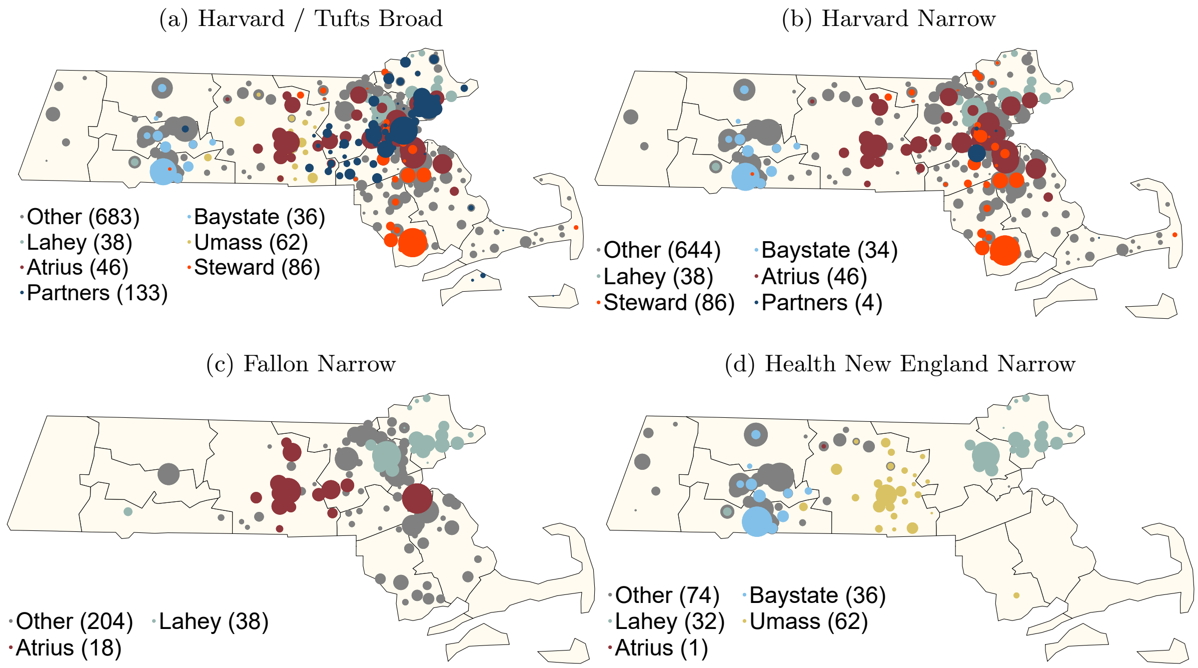
In Figure B.1, Figure B.2, Figure B.3, and Figure B.4, I present additional maps depicting the hospital, PCP, cardiology, and orthpedic practice network coverage across Massachusetts of Harvard and Tufts Broad, Harvard Narrow, Fallon Narrow, and HNE Narrow.

Figure B.1: Hospital Networks by Plan, 2013



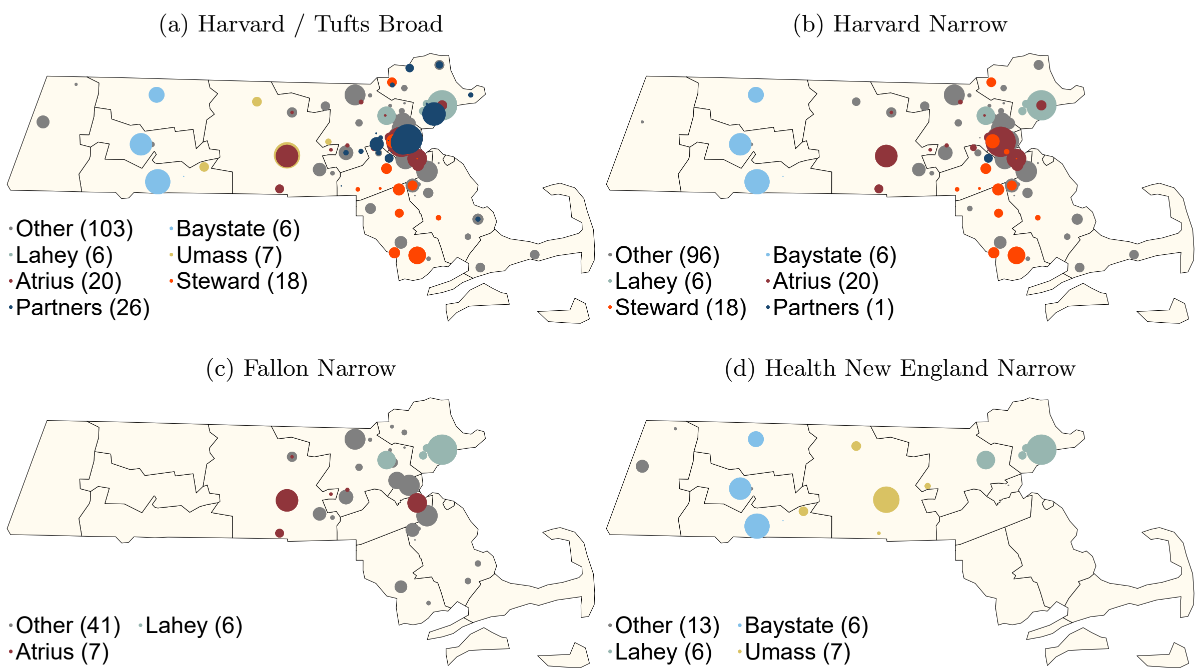
Notes: This figure plots the hospital networks of specified plans on the GIC in 2013. Sizes of the data points reflect relative market shares of the hospitals. Colors reflect ownership status (which health systems owns which hospital).

Figure B.2: Primary Care Practice Networks by Plan, 2013



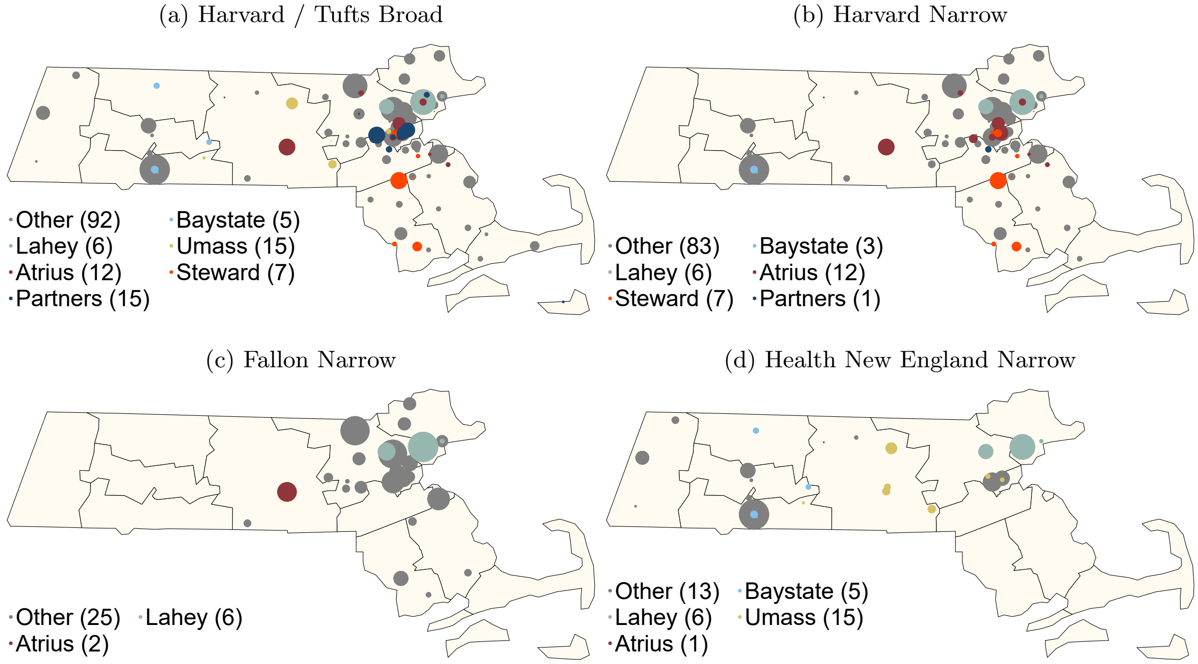
Notes: This figure plots the PCP practice networks of specified plans on the GIC in 2013. Sizes of the data points reflect relative market shares of the practices. Colors reflect ownership status (which health systems owns which practice).

Figure B.3: Cardiology Networks by Plan, 2013



Notes: This figure plots the cardiology practice networks of specified plans on the GIC in 2013. 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, 2013



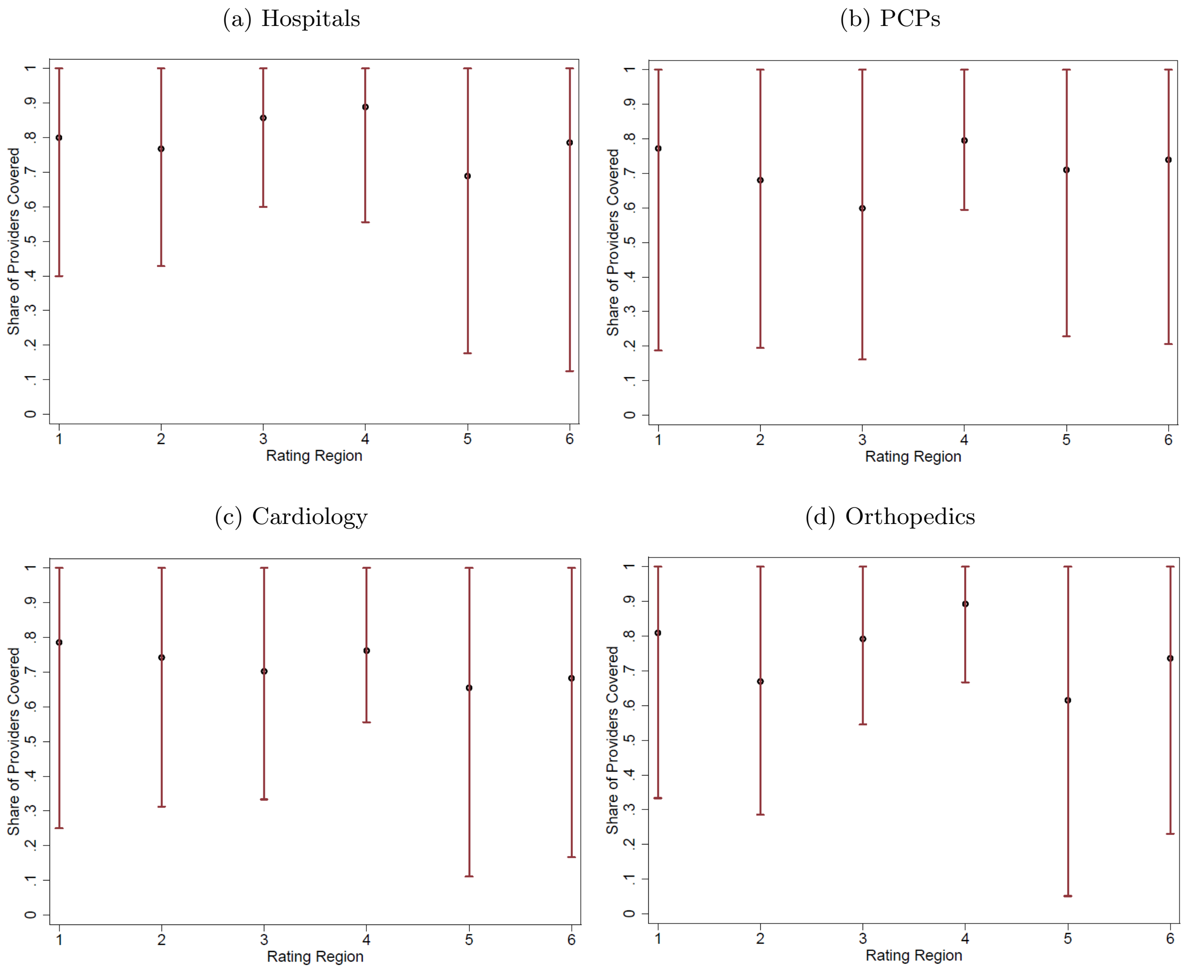
Notes: This figure plots the orthopedic practice networks of specified plans on the GIC in 2013. Sizes of the data points reflect relative market shares of the practices. Colors reflect ownership status (which health systems owns which practice).

Figure B.5 plots the variation in hospital and physician networks across plan and rating region in Massachusetts in 2011.[3] The y-axis represents the share of providers operating in the rating region that each plan covers as an in-network provider (hereafter referred to as "network breadth"). The dots represent the unweighted 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* unweighted network breadth for PCPs, for instance, ranges from about 60% to about 80%. Within rating region, the broadest plans cover virtually all the top 50 practices and hospitals, while the narrowest plans cover only about 20% of the providers. In Rating Region 5, which includes 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.

Figure B.6 displays the unweighted average network breadth over time. The y-axis here represents the share of the state's hospitals and physician 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

[3] 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.

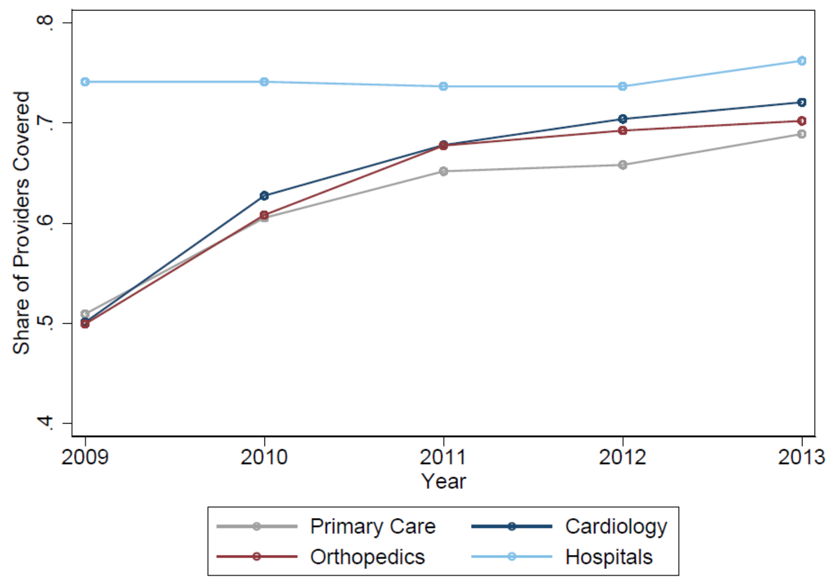
Figure B.5: Share of Providers Covered by Rating Region and Specialty, 2011



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 unweighted 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.

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.[4] Second, during this period there were significant hospital acquisitions of physician practices. Third, certain narrow plans grew more generous in coverage over time.[5]

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 the probability of enrollment in a narrow-network plan against 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, 8% less likely to be enrolled in a narrow-network plan than new members, suggesting that plan choice inertia may play a large role in explaining broad-network enrollment.

To see more evidence that new cohorts behave differently than existing cohorts, I also report 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. However,

[4] As an example, the Atrius Health system gradually purchased several prominent medical groups, including Harvard Vanguard and the Fallon Clinic, which were previously separate entities.

[5] For example, Fallon Health Plan did not cover the Partners system until 2013.

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

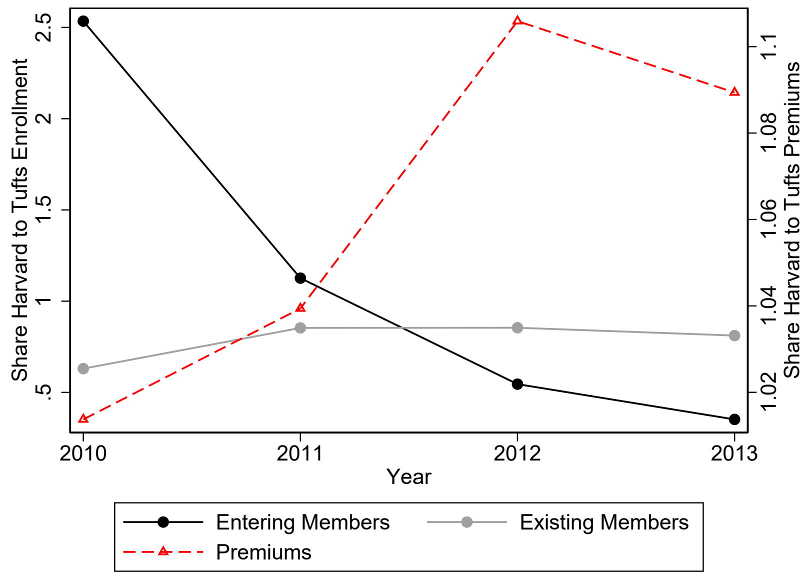
Variable	Coefficient	Standard Error
Existing GIC Member	-0.0807^{***}	0.0015
Age	-0.0004^{***}	0.0000
Female	-0.0267^{***}	0.0015
Chronic Condition	-0.0318^{***}	0.0018
Members in HH	-0.0103^{***}	0.0004
Constant	0.2773^{***}	0.0024
Year FE	Yes	
County FE	Yes	
Obs.	426,288	
Adj R^2	0.27	

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

beginning in 2011, the premium difference between the two plans began to rise, with Harvard Broad growing significantly more expensive than Tufts Broad.

Figure B.7 shows the ratio of enrollment in Harvard Broad versus Tufts Broad over time, along with 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 Harvard declines dramatically relative to Tufts among new members. By 2012, Tufts' premiums were about 10% less than Harvard's (amounting to about $30 per month for families). Enrollment in Harvard among new members, meanwhile, declined from almost three times that of Tufts in 2009 to about 50% 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 Harvard Broad vs. Tufts Broad by Whether New to GIC



Notes: This figure plots the ratio of members selecting Harvard's broad-network plan over Tufts' broad-network plan as well as the ratio of Harvard to Tufts premiums. 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 with diagnosis l will choose hospital h in time t is given by:

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

where N_{jt}^H refers to the hospitals in plan j'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:

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

where N_{jt}^S is the network of practices of type s in plan j'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 rating regions and by specialty group (PCP, cardiology, and orthopedics).

All models include patient characteristics interacted with provider characteristics, travel time interacted with both patient and provider characteristics, and a full set of provider fixed effects 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 Classifications 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.[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 interact hospital fixed effects with the CCS categories.

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

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 patient RVU as well as 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.

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

[6] The construction of these weights follow closely to work by Shepard (2016); a discussion of their construction can be found in subsection C.6.

across specialties and between physician groups and hospitals.

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). I address these potential selection concerns in subsection C.4.

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 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 individual practices are location-specific, this allows me to include a larger span of the full physician practice space in my estimation. In addition, it allows for estimation of flexible parameters that vary by region.

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. $\{T_{id} \in b\} = T_{ib} = b^{mid}$. As an example, $b^{mid} = 2.5$ for distance band b =0 to 5 miles. Given these assumptions, and dropping the region and time subscripts for convenience, the model in Equation 1 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 γ^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, two empirical facts support this claim. First, the top 50 practices by market share in a given region account for most of patient claims.[7] Second, most practices outside the top 50 are included in all plans' networks. As a result, most of the variation in networks across plans comes from variation in coverage of 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 tractable, but also is likely a reasonable assumption given this context.

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 the outside option to 0 in all 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_5^s, 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 in the model. I abstract away from decomposing the extent to which λ_5^s 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: that λ_5^s entirely reflects physician-patient capital. In counterfactual exercises, when patients lose access to their previously used physicians or practices, I treat this as a "welfare-relevant" utility loss.[8] However, to test the robustness of this assumption, in Appendix E I also present estimates of the key parameters of my employer objective function that treat this inertia term as being driven by the two other aforementioned 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 between 2009 and 2013. These models incorporate 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 reported are

[7] 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.

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

for Barnstable County (the base county), Worcester (Central Massachusetts), Hampden (Western Massachusetts), and Suffolk (Eastern Massachusetts). Consistent with prior literature, the distance coefficients are negative and significant in all reported counties, implying that patients prefer to go to hospitals that are close to where they live. 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). While these coefficients are difficult to interpret (the measure is in utils instead of a dollarized amount), comparing them with other parameter estimates shed some light on their practical magnitude. For instance, the estimates imply that hospital patients in Suffolk are, on average willing, to travel approximately 25 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.

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 R^2	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.

A second important finding concerns the large positive and significant coefficient on individuals who have used the hospital previously. 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 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 MRIs.

It is worth mentioning that this model so far 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 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. 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. I omit year 2009, the earliest year of data in the claims, as I cannot observe patients' prior use of physicians in that year. 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 analogous parameter estimates for the Worcester region, for comparison.

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

Variable	PCP Practices	Cardiology Practices	Orthopedic Practices
Distance	−0.4168***	−0.2994***	−0.2575***
	(0.0186)	(0.0147)	(0.0159)
Owned by Hosp. or System	0.2813***	1.4316***	0.6287***
	(0.0977)	(0.0895)	(0.0867)
Used Prac Previously	3.6494***	1.0508***	2.4033***
	(0.0401)	(0.0354)	(0.0406)
x Length of Relationship	0.3751***	0.1947***	−0.1962***
	(0.0097)	(0.0104)	(0.0118)
x RVU	−0.0643***	0.0646***	0.0090***
	(0.0094)	(0.0053)	(0.0019)
Used Med Grp Previously	1.4447***	1.6987***	1.5252***
	(0.0401)	(0.0345)	(0.0431)
Used System Previously	0.5677***	0.8538***	1.0928***
	(0.0362)	(0.0303)	(0.0356)
	Interactions with Patient Characteristics		
DistxFemale	0.0004	−0.0046	−0.0064*
	(0.0015)	(0.0029)	(0.0034)
DistxAge	−0.0008***	0.0000	0.0001
	(0.0000)	(0.0001)	(0.0001)
DistxChronic	−0.0015	0.0297***	0.0261***
	(0.0019)	(0.0079)	(0.0055)
	Interactions with Provider Characteristics		
DistxNumDocs	0.0004***	−0.0004***	−0.0002***
	(0.0000)	(0.0000)	(0.0000)
DistxNumLocs	0.0036***	−0.0036***	−0.0054***
	(0.0009)	(0.0006)	(0.0007)
DistxMedGrp	0.0260***	−0.0381***	−0.0507***
	(0.0075)	(0.0080)	(0.0074)
AgexNumDocs (00s)	−0.0020***	0.0045***	0.0000
	(0.0000)	(0.0000)	(0.0000)
AgexNumLocs (00s)	0.0192	0.1423***	0.0791***
	(0.0165)	(0.0133)	(0.0149)
AgexMedGrp	0.0027*	0.0102***	0.0243***
	(0.0014)	(0.0015)	(0.0014)
Practice FE	Yes	Yes	Yes
Obs.	3,289,932	1,853,631	1,634,164
Pseudo R^2	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, and gender interacted with additional practice characteristics: number of unique services at 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.

Consistent with the results of the hospital demand model, distance plays an extremely important role in patient choice of physician. Across the three specialist groups, distance has a negative and significant effect on utility.[9] Patients, on average, prefer visiting practices owned by hospitals or health systems, though the effect is considerably stronger for cardiology practices.[10]

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 latter is consistent with the result from hospital demand, namely that

[9] 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 scaling of the logit error.

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

conditional on risk, older individuals prefer to travel smaller distances to seek care, particularly for routine primary care treatment. 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 farther to access practices with more physicians on site. In addition, they are willing to travel farther 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.[11] This is 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 an individual, on average, would be willing to travel an additional 8.5 miles to access the same PCP practice, 5.8 miles to access the same cardiology practices, and 9.5 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 are on average willing to travel an additional 8.5 miles to access the same PCP practice, this figure is approximately 30 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

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

requires driving more so than walking.[12] 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.[13] Much like in Boston, older patients seeking specialist care significantly prefer doctors that are part of medical groups and that work for practices which have multiple locations.

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

Variable	PCP Practices	Cardiology Practices	Orthopedic Practices
Distance	−0.1546***	−0.1595***	−0.2210***
	(0.0118)	(0.0104)	(0.0107)
Owned by Hosp. or System	−0.1576**	0.1139	−0.2913***
	(0.0690)	(0.0970)	(0.0902)
Used Prac Previously	4.6802***	1.3632***	3.1657***
	(0.0446)	(0.0525)	(0.0657)
x Length of Relationship	0.2562***	0.0940***	−0.3311***
	(0.0118)	(0.0185)	(0.0248)
x RVU	−0.1352***	0.0192***	0.0106***
	(0.0114)	(0.0053)	(0.0031)
Used Med Grp Previously	0.6712***	1.2439***	1.0773***
	(0.0416)	(0.0532)	(0.0629)
Used System previously	0.7626***	0.8348***	1.0247***
	(0.0411)	(0.0512)	(0.0594)
		Interactions with Patient Characteristics	
DistxFemale	0.0018	−0.0004	−0.0029
	(0.0011)	(0.0025)	(0.0028)
DistxAge	−0.0003***	−0.0002*	0.0002**
	(0.0000)	(0.0001)	(0.0001)
DistxChronic	0.0071***	0.0182***	0.0519***
	(0.0017)	(0.0056)	(0.0046)
		Interactions with Provider Characteristics	
DistxNumDocs	−0.0001*	0.0000	−0.0002***
	(0.0001)	(0.0000)	(0.0000)
DistxNumLocs	−0.0099***	−0.0018***	0.0005
	(0.0013)	(0.0006)	(0.0007)
DistxMedGrp	0.0340***	0.0033	0.0021
	(0.0069)	(0.0059)	(0.0056)
AgexNumDocs (00s)	−0.0036**	−0.0026	0.0149***
	(0.0018)	(0.0024)	(0.0026)
AgexNumLocs (00s)	0.3956***	0.1837***	0.0188
	(0.0373)	(0.0405)	(0.0421)
AgexMedGrp	−0.0024	0.0109***	0.0098***
	(0.0023)	(0.0032)	(0.0022)
Practice FE	Yes	Yes	Yes
Obs.	2,662,897	686,687	560,253
Pseudo R^2	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, and gender interacted with additional practice characteristics: number of unique services at 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.

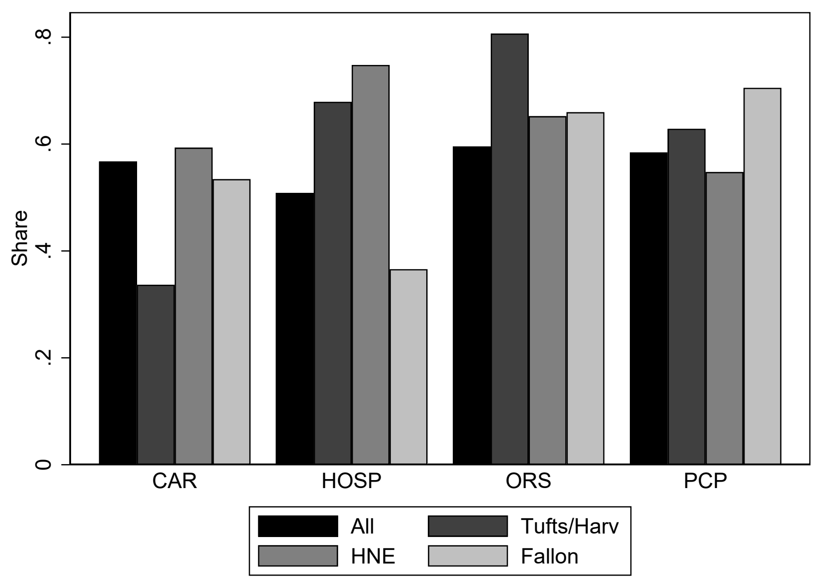
[12] The average distance traveled for PCPs in Boston is about half that of Worcester.

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

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_{jt}^H and N_{jt}^S, the expected utility of a particular hospital and physician network, EU_{Ijt}^H and EU_{ijt}^S (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. "Tufts/Harv" refers to Tufts Narrow and Harvard Narrow.

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, the model accurately predicts about 60% of individual choices. The model also predicts hospital choices quite well, with a particularly good fit for patients in

Health New England and the Tufts/Harvard narrow networks. The model does extremely well for orthopedic surgeons, predicting nearly 80% of choices accurately for Tufts/Harvard. However, the model performs slightly worse for cardiologists for Tufts/Harvard, and somewhat worse for hospitals for patients enrolled in Fallon.

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 for every year after 2010.[14] 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.

C.5 Plan Demand

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

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

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

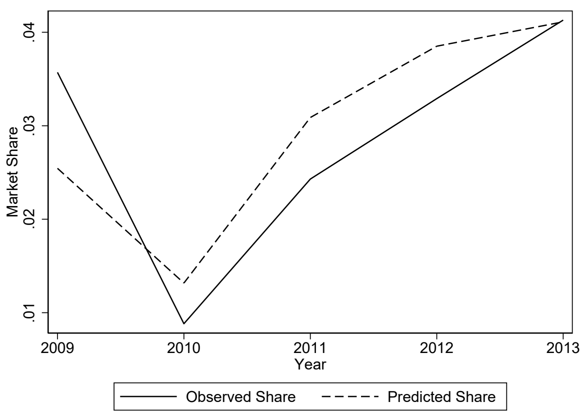
where, f_{il} and f_{ir} are the ex-ante probabilities that individual i contracts diagnosis l (requiring hospital care) or requires procedure r (requiring physician care). Note that, as demand for insurance plans is at the *household* level, the expected utility variables are also aggregated to the 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 into the probability of falling into discrete RVU bins within each specialty: 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

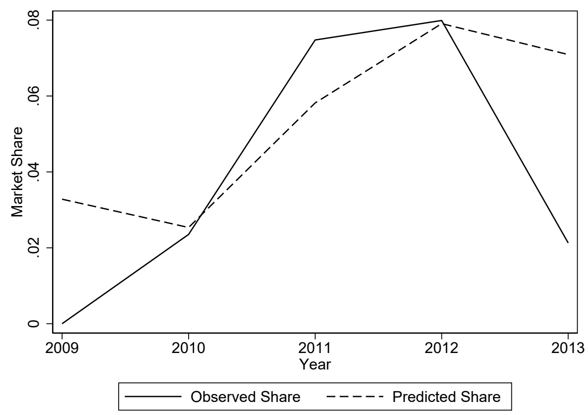
[14] This is likely due to small sample sizes of hospital admissions among narrow-network patients, which is particularly true in 2009 (prior to the introduction of the Tufts and Harvard narrow plans).

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

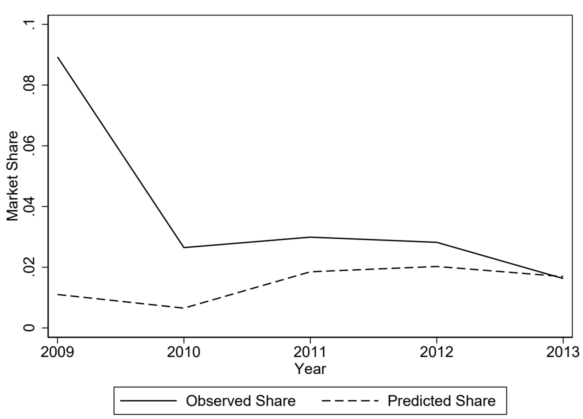
(a) Tufts Medical Center Share



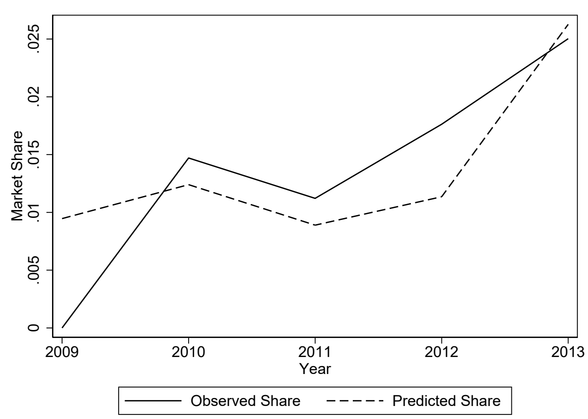
(b) Beth Israel Hospital Share



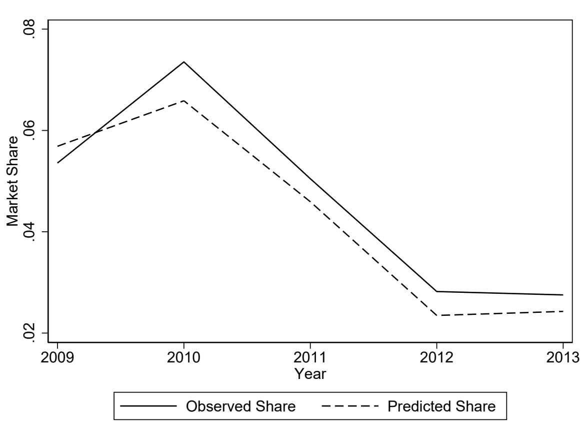
(c) Boston Medical Center Share



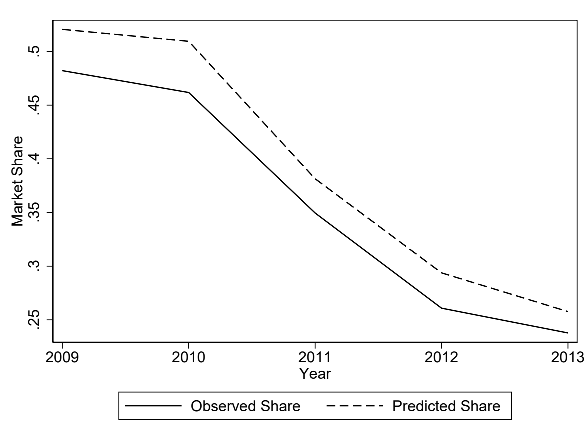
(d) Lahey Hospital Share



(e) Berkshire Medical Center Share



(f) Baystate Medical Center Share

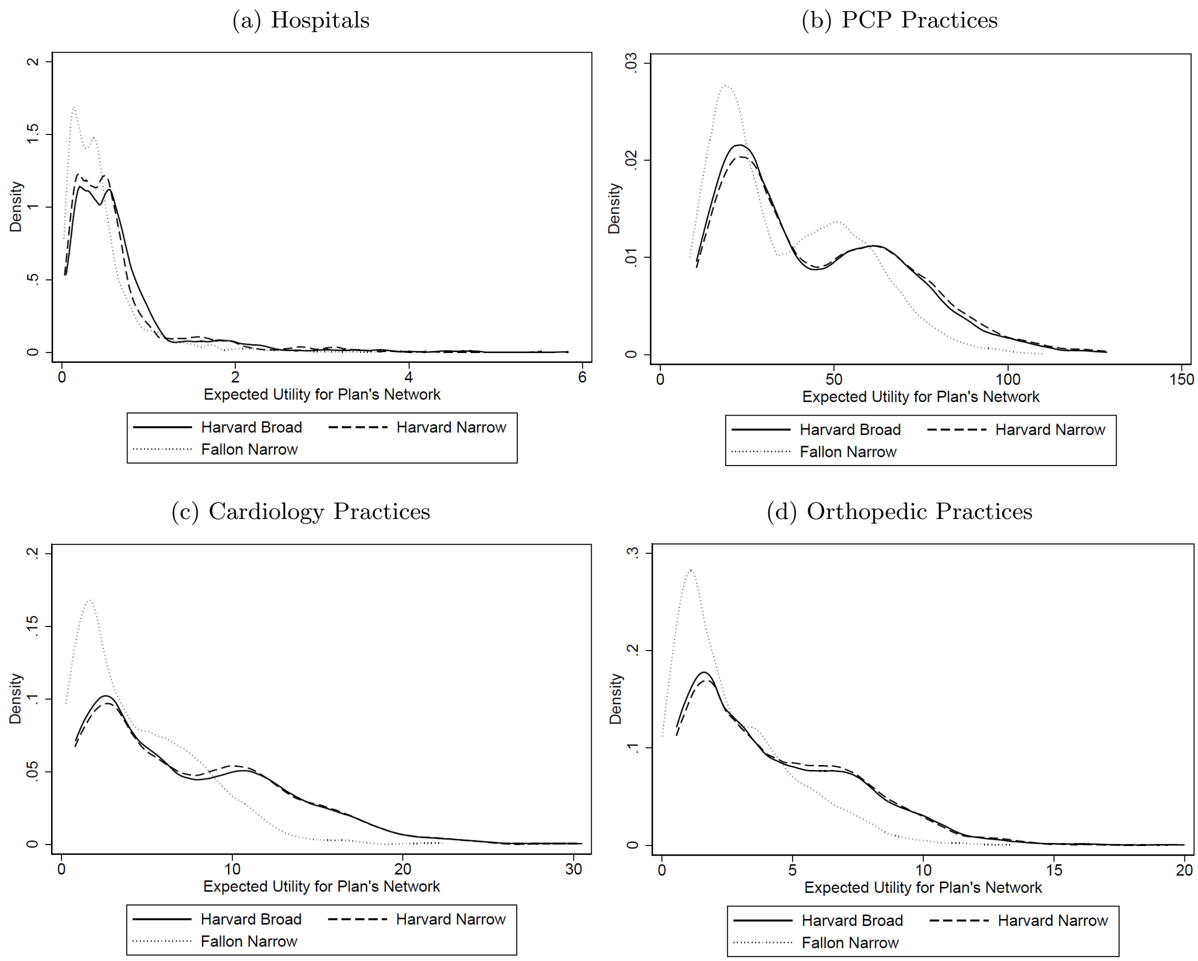


Notes: This figure plots actual market shares of select medical centers against the predicted market shares of those medical centers among consumers in narrow-network plans. Parameters used for prediction were estimated from a demand model among only individuals in broad-network plans.

certain specialists, but also of requiring treatment of varying levels of complexity.

Figure C.3 plots 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 plan yields lower utility than its broad plan, 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 (see Appendix B)—almost all excluding those owned by Partners—whereas Fallon covers significantly fewer providers in Boston.

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



Notes: This figure plots the distribution of EU^H_{Ijt} and EU^s_{Ijt} for hospitals and 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^H_{Ijt}. 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^{PCP}_{Ijt}, 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.

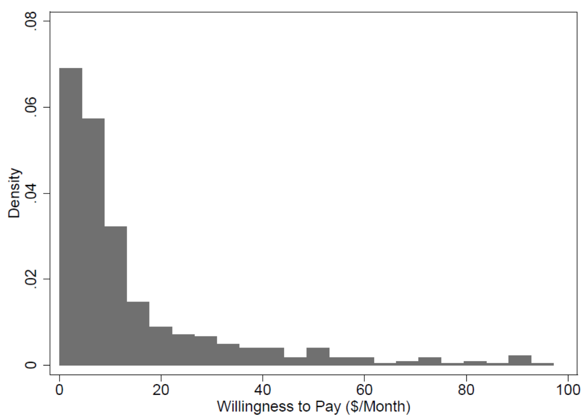
A full set of plan fixed effects are included. As with the provider demand model, I include an indicator variable for whether a particular plan matches an enrollee's plan choice from the previous year. This follows prior literature on plan inertia (Handel, 2013; Polyakova, 2016; Shepard, 2016) and is 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 implemented 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 from the corresponding municipalities who joined during the corresponding years.

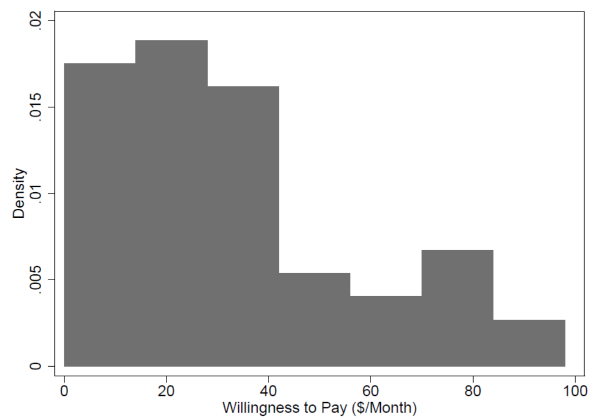
Distribution of WTP: Figure C.4 shows the distribution of estimated WTP from the plan demand model for Harvard Broad vs. Harvard Narrow, and Fallon Broad vs. Fallon Narrow. Two conclusions emerge from this figure. First, although the mean reported values for Harvard reported in Table 2 are around \$19 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.

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

(a) Harvard Broad vs. Harvard Narrow



(b) Fallon Broad vs. Fallon Narrow



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.

C.6 Premium Setting Stage

Construction 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 scale this price by a series of resource weights to arrive at a payment for each diagnosis and procedure. I use observed "allowed amounts" to specify a base rate for each insurer-provider combination.[15]

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:

$$A^s_{irjdt} = p^s_{jdt} * RVU_{rt} \tag{26}$$

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

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 resulting base price can therefore be interpreted as the negotiated rate between plan j and physician practice d for one RVU of care.

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. 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 insurer-hospital base prices, by running the following model:

$$ln(A_{iljht}) = \gamma_{jht} + \psi_{lt} + x_{ilt} + \epsilon_{iljh} \tag{27}$$

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, I proxy for diagnoses by including ψ_{lt}. These are a set of fixed effects for the 18 CCS diagnosis categories used in the hospital demand model. The model is therefore similar to the physician price construction, except that by including these fixed effects, I estimate weights for each diagnosis rather than using observed weights. The model also includes Elixhauser comorbidity indices 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. I scale the predicted price by this factor in order to arrive at the predicted base price for a standardized unit of care, 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.[16] The table suggests that negotiated prices do not vary considerably across medical specialties in Massachusetts. Within specialty, however, there is considerable variation. Facility-based cardiology practices, for instance, receive an average price-per-RVU of \$57, 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 about \$20,000 while the minimum was about \$3,000. 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 service, yielding an average of \$123 per procedure. Orthopedists, however, perform an average of 4 RVUs per service, implying an average payment of \$221 per procedure.

I next examine whether the preference for broad-network plans translates into higher negotiated

[16] I define practices that are "office-based" 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

	PCPs	Cardiologists	Orthopedists	Hospitals
		Office-Based		
Average Base Price	56.56	56.29	55.33	–
	(12.43)	(14.80)	(16.94)	–
Average Weight	2.19	2.74	3.99	–
	(0.60)	(1.28)	(2.46)	–
		Facility-Based		
Average Base Price	58.52	56.60	52.16	8,145.12
	(15.67)	(19.78)	(14.16)	(3,028.49)
Average Weight	2.52	2.05	6.44	1.02
	(1.44)	(1.69)	(5.17)	(0.12)

Notes: "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 $RVU_{rt} = 1$. In the case of hospitals, this refers to the case where $w_{lt} = 1$. Hospital weights are scaled so that the yearly average is one, meaning that hospital base prices refer to the price for an admission of average weight. "Office-based" settings are defined as practices where more than 70% of claims are flagged as in an office-based setting.

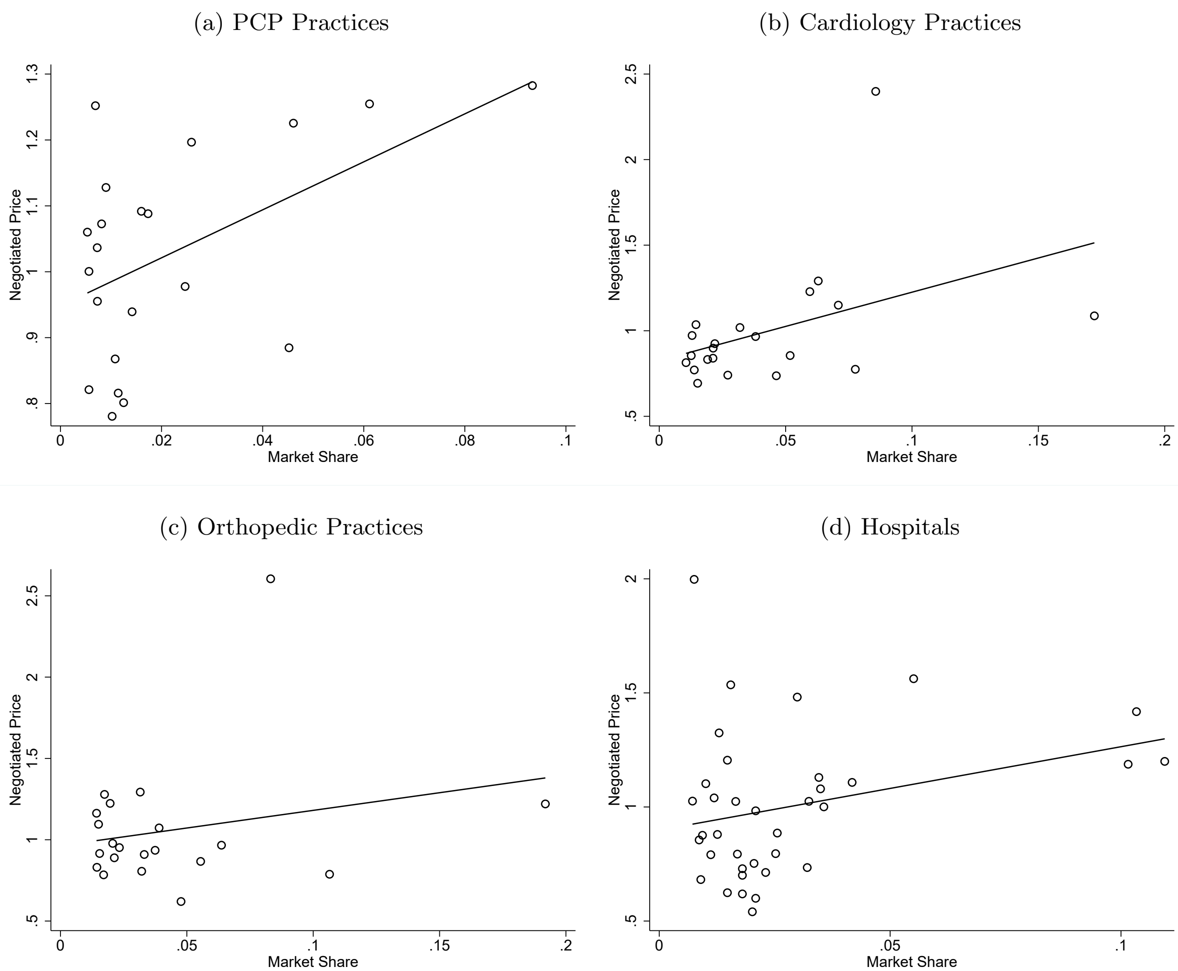
rates for those providers. Figure C.5 depicts the relationship between demand and negotiated provider price for one of the insurers on the GIC in the Boston rating region. Due to confidentiality concerns, I omit both the identity of the insurer and the actual negotiated rate. Instead, I report the negotiated rate relative to the insurer-specific average. The y-axis depicts this standardized rate, where the x-axis depicts the market share.

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 cardiologists. However, there is still a positive relationship for PCPs and orthopedists as well. 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 employers 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.

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 claims data, to recover *unobserved* marginal costs, I parameterize them as $c^u_{Ijt}(N_{jt}) = c^u_{jt}(N_{jt})\theta^c_I$, where θ^c_I scales base plan-specific unobserved costs, $c^u_{jt}(N_{jt})$, across household type I. I assume that unobserved marginal costs only vary by 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 orthopedists, and regressing the observed allowed amounts for these claims on household type.[17] This reduces the number of unknowns to JT,

[17] The critical assumption here is that all marginal costs that vary by more granular risk types are captured through

Figure C.5: Insurer Negotiated Price by Market Share, Boston Rating Region 2011



Notes: This figure plots the negotiated price for hospitals, p_{jht}, and for physician practices, p^s_{jdt}, against market share. Prices are reported for a single insurer and relative to the insurer-specific mean. Data is for year 2011. All plots are for Boston rating region only, except for Panel (d), which is reported for all of Massachusetts.

allowing for full recovery of $c^u_{jt}(N_{jt})$.

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:

$$c^u_{mjt}(N_{mjt}) = \kappa x_{mjt} + \gamma_m + \gamma_t + \varepsilon_{mjt} \tag{28}$$

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 of Equation 28, with the log of unobserved marginal costs as the dependent variable. Year 2012 is omitted due to potential bias from it being the year of the premium holiday.

Table C.5: Unobserved Marginal Cost Estimates

	Coefficient	Standard Error
Narrow Network	-0.164^{***}	0.019
Harvard Pilgrim	0.059^{**}	0.022
Health New England	-0.064^{**}	0.026
Neighborhood Health Plan	-0.039	0.026
Tufts Health Plan	0.057^{**}	0.022
2010	0.002	0.023
2011	0.035	0.022
2013	0.080^{***}	0.022
Constant	5.911^{***}	0.021
Observations	28	
R^2	0.93	

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 also omitted from the analysis due to concern of bias from the enactment of the premium holiday.

The results indicate that being a narrow-network plan reduces unobserved marginal costs of health care by approximately 16%. 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 a larger set of providers unaccounted for by the chosen specialties.[18] 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.

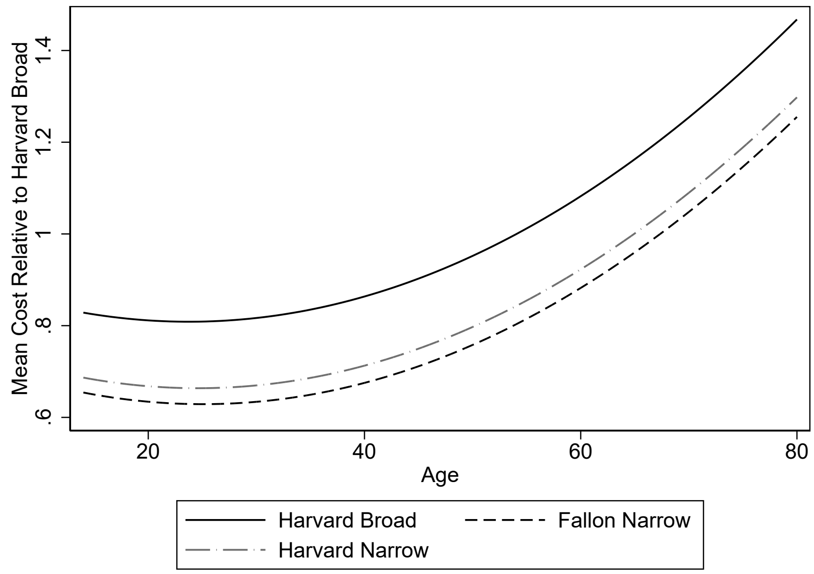
Figure C.6 plots the total estimated marginal costs of health care (hospital + PCP + cardiology + orthopedics + unobserved) against age for single-member households. I report estimated cost-curves for Harvard Broad, Harvard Narrow, and Fallon Narrow and, for confidentiality reasons, only report estimated costs relative to the Harvard Broad mean. 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

observed hospital and physician costs, whereas *unobserved* costs only vary by family type. While strong, this seems reasonable as a first-order approximation. I report robustness on this assumption in subsection E.3.

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

similar rates, although Harvard Broad does have a slight uptick after age 60 relative to the narrow products. This suggests the potential for selection on expensive providers, particularly among older individuals, conforming to the results of the hospital and physician demand models.

Figure C.6: 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 Details on Estimating the Employer Objective Function

I make several assumptions to proceed with the estimation of ρ and FC_j in Equation 11. 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.[19] More formally:

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

[19]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 entrants 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.

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:

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

In the estimation of Equation 17, I take the average of 10 disturbances of ω_k. That is, I estimate the moment inequalities assuming 10 different potential random samples of entrants in each year given the number of municipalities who entered the previous year. For each set of potential entrants, I also simulate 10 different distributions of demand coefficients, β^s_{2I}, to account for the presence of unobserved heterogeneity. This makes a total of 100 evaluations of each moment in each year.[20]

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. Prior work has treated such structural errors as disturbances in the fixed cost term. For instance, it could be that: $FC_j = FC + v_{2,j}$, 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 formulation if the error term varied by the type of product offered. In my setting, the GIC might choose to contract with certain insurers, offer certain products, or offer certain networks for which the fixed costs 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. $FC_j = FC$. Similar assumptions were made by Nosko (2014). 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. 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* an employer group are likely smaller. This assumption may be violated if, for instance, offering a product that was broader in network size than another product also meant an increase in 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 their maximum network, 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.

[20] In previous drafts of this paper, I have evaluated the employer objective function using 100 different samples of municipal entrants. However, this poses a major computational burden with the inclusion of random coefficients. The current results are encouragingly similar to higher-order simulations.

D Details on the 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 firms 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 the 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 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 using data on firm size, industry, and number of plans offered.

I make a few additional simplifying assumptions in creating the 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, it is nonetheless a sensible restriction to make.[21] 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, for which I have no estimated "brand effect."

The second simplifying assumption I make is in the network offerings of employers. For products outside the GIC, I do not observe the network breadth of each plan. One option would be to infer networks based on observed claims.[22] However, this approach is prone to significant measurement error, particularly for firms with fewer employees. Instead, I leverage institutional features of

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

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

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, known as "Harvard Focus" and "Tufts Select" (Office of the Attorney General Martha Coakley, 2013). I therefore assume that any firm in the private sample that I simulate offering a Harvard Pilgrim narrow-network product covers the same hospitals and physicians as Harvard Focus. Similarly, for firms that I simulate offering Tufts narrow products, I assign the network breadth of Tufts Select. 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).

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

Variable	Private Emp.	Private Emp.	GIC
	Firm Level	Employee Level	
Offered a Narrow Network	0.079	–	–
	(0.271)	–	–
Enrolled in a Narrow Network	–	0.018	0.118
	–	(0.134)	(0.323)
Age 55+	–	0.159	0.201
	–	(0.366)	(0.401)
Female	–	0.536	0.518
	–	(0.499)	(0.500)
Rating Area			
1	0.131	0.064	0.158
	(0.339)	(0.245)	(0.365)
2	0.071	0.087	0.127
	(0.259)	(0.282)	(0.333)
3	0.071	0.172	0.100
	(0.259)	(0.377)	(0.300)
4	0.143	0.171	0.218
	(0.352)	(0.376)	(0.413)
5	0.548	0.372	0.240
	(0.501)	(0.483)	(0.427)
6	0.036	0.125	0.135
	(0.187)	(0.331)	(0.342)
Industry			
Health Care	0.425	0.391	–
	(0.497)	(0.488)	–
Service	0.310	0.397	–
	(0.465)	(0.489)	–
Wholesale	0.023	0.023	–
	(0.151)	(0.148)	–
Transportation, Communications, Utilities	0.115	0.072	–
	(0.321)	(0.259)	–
Manufacturing	0.126	0.117	–
	(0.334)	(0.322)	–
Number of Employers	123		

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

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 8% 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 2 and 3).[23] The GIC sample is slightly older, with about 20% of employees being over age 55, compared to about 16% in the private employer sample. Together, the health care and service industry comprised 70% of the private firms. In terms of geographic distribution, most large private employers (55%) are headquartered in Boston (Rating Region 5). At the employee level, this translates to about 37% of all private employees working for firms in Boston, with the next largest share (17%) 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 22% 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 follows 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 employees 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 selection 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 two plans, then for that employer, I only consider alternate plan menus/networks in which that employer offers two 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:

$$m(\delta_J, \delta_J^a, \theta, z) = \sum_{s=1}^{10} \left(\frac{1}{F} \sum_f \left[(W(\delta_J, \theta_s) - W(\delta_J^a, \theta_s)) \otimes g(z) \right] \right) \geq 0 \tag{29}$$

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

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

E Robustness on Employer Objective Function

E.1 Households with Prior Provider Relationships

In Table E.1, I report estimates of ρ and FC_j where I reweight the population of the employee pool such that 90% of households have had a prior relationship with a provider. Here, I define a "prior relationship" as having previously (as of year t) visited a provider. In particular, I consider four different populations: employees with a prior relationship with a Partners provider (either a hospital or physician), a Umass provider, an Atrius physician, or any provider that is *only* covered by a broad-network plan. The results do show that the mismatch parameter declines for households with prior provider relationships, implying that employers may overweight these populations in their network design. In particular, ρ declines from 3.70 at baseline to 2.19 for the population with any prior provider relationship. The results are similar for households with a prior relationship with a Partners provider or Umass provider.

Notably this is not the case for households with a prior relationship with an Atrius provider. In addition, the estimates of ρ from the regional analyses in Table 4 are comparable, and in the case of the North Shore (Region 4) they are substantially lower. The combination of these facts suggests that, while employers do appear to be motivated by the preferences of households with prior provider relationships, these effects are likely strongest in regions where dominant providers (e.g. Partners) face less competition. This may explain why the mismatch parameter increases for households with prior Atrius relationships: the fact that Atrius primarily operates in dense regions (e.g. Boston) suggest that its removal from a network may not affect utility by as large a magnitude as the decrease in premium spending.

Table E.1: Employer Objective Function Parameters For Populations with Prior Provider Relationships

	(1) Baseline	(2) Any Provider	(3) Partners Provider	(4) Umass Provider	(5) Atrius Provider
ρ	3.70	2.19	2.23	2.08	4.04
FC_j ($Millions)	3.98	2.20	2.64	1.99	4.53

Notes: Results from ρ and FC_j estimation for 2009-2013. Column 1 presents estimates for the baseline population of GIC enrollees. Columns 2-5 present estimates that reweight the population such that 90% of the population have a prior relationship with a provider of a certain type. FC_j reported in millions of dollars.

E.2 Physician Inertia and Active Choice Frictions

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 driven by the formation of patient-physician capital. Under this assumption, if a physician were removed from the network and a patient had seen that physician previously, that patient would incur a substantial utility loss. However, the characteristics of that physician would not necessarily be informative as to which physician the patient would choose in his or her absence.

There are two alternate interpretations of physician inertia. 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 the 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 network offers.

Each of these interpretations, through their impact on consumer utility of a network change, can have significant impacts on the estimates of the employer objective function. In Table E.2, I report results on ρ and FC_j assuming that the entirety of the inertia term were driven by these various forces.[24] 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 3.70 to 6.39 (Column 2). This result makes sense: in this scenario, any potential narrowing of a network would result in a smaller utility loss, but a similar decline in 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.

Table E.2: Employer Objective Function Parameters Under Alternate Provider Inertia Assumptions

	(1) Inertia = Pat./Prov. Match	(2) Inertia = Switching Costs	(3) Inertia = Unobserved Quality
ρ	3.70	6.39	4.00
FC_j ($Millions)	3.98	7.90	4.55

Notes: Results from ρ and FC_j estimation for 2009-2013. Columns 1-3 presents estimates under different assumptions of the interpretation of physician inertia. FC_j reported in millions of dollars.

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. I then use these demand estimates in estimation of the employer objective function. 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 in the first-time choices made by brand new patients. Under this interpretation, the employer-employee mismatch again rises, though very slightly, from baseline. (Column 3). This suggests two things. First, removing any physician yields a smaller utility loss for patients than the baseline assumption, implying that patients are typically able to find close substitutes. Second, the baseline model does reasonably well

[24]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.

at estimating unobserved provider quality.[25] 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, likely lies somewhere between 3.70 and 6.39, but it always considerably greater than 1.

Active Choice Frictions: While the demand model estimated in subsection 3.2 incorporates plan inertia and switching costs, it does not explicitly model active choice frictions that might apply to both new and existing enrollees. Recent work has shown that choice complexity, information asymmetry, and choice overload drive enrollees to, for example, opt into dominated health plans (Bhargava et al., 2017; Handel and Kolstad, 2015; Abaluck and Gruber, 2020). For instance, it may be the case that employees select broad networks not out of a "true" valuation of the network, but rather out of a lack of full information about plan features and an aversion to potential out-of-network bills they might incur. My model treats these frictions as "welfare-relevant" in the sense that employers observe household plan selections ex-post and make their plan offer decisions assuming those choices reflect full information. However, to the extent that observed household choices reflect these frictions—and are therefore not welfare-relevant—estimates of $CS(\delta_{Jt}, \theta)$ (and therefore ρ) may be biased. Though I do not test robustness to this formally, the exercises with physician inertia shown affect the employer objective function through a similar mechanism as would active choice frictions. Part of the valuation of a broad-network plan may reflect frictions that may alter computations of consumer surplus if the employer treated them as welfare-irrelevant. In those simulations, I show that estimates of ρ *increase* when employers treat physician inertia as welfare-irrelevant. Intuitively, households lose less utility from a shift to narrow-network plans and, as such, the employer-employee mismatch rises. I therefore take the estimates of ρ in Table 3 as conservative estimates.

E.3 Employer Mistakes and Additional Robustness

The Role of Switching Costs: A possible explanation for employer persistence in offering 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.[26] 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 sub-

[25] 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.

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

segment 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 employees, by definition, have no inertia. However, unlike other new employees, these municipal entrants have previously lived and worked in the state, and have also been previously enrolled in private health insurance. Moreover, prior plans were uniformly generous, broad-network plans.[27]

Panel A of Table E.3 reports these estimates, focusing only on moments that fix the number of products to isolate the effects on ρ.[28] The employer-employee mismatch term drops to 1.89, about a 49% 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 same patterns across populations observed in Table 4 persist. In fact, assuming that the *entire* discrepancy between the baseline estimates and these estimates is driven by misperceptions, this still implies that up to a quarter of the mismatch can potentially be attributed to unequal weighting in household preferences.

Table E.3: Additional Specifications for the Employer Objective Function

	(1) Baseline	(2) Older	(3) Older, Regions 1,4,6	(4) Older, Region 4
Panel A: Estimation On New, Municipal Entrants				
ρ	1.89	1.49	1.52	1.25
FC_j ($Millions)	–	–	–	–
Panel B: Estimation With Alternate Marginal Costs				
ρ	2.94	2.51	2.05	1.65
FC_j ($Millions)	3.65	3.60	2.24	2.65
Panel C: Estimation Restricting Harvard Broad to Narrow				
ρ	3.67	3.07	2.64	1.93
FC_j ($Millions)	3.88	3.35	2.47	2.82
Panel D: Estimation Restricting Tufts Broad to Medium				
ρ	3.70	3.08	2.65	1.93
FC_j ($Millions)	3.98	3.39	2.50	2.83
Panel E: Estimation Restricting Tufts Narrow to Harvard Medium				
ρ	4.38	3.26	2.67	1.86
FC_j ($Millions)	4.68	3.44	2.61	2.73
Panel F: Estimation Restricting Harvard Broad to Medium				
ρ	1.97	1.67	1.30	0.98
FC_j ($Millions)	2.57	2.27	1.57	1.95

Notes: Results from ρ and FC_j estimation for 2009-2013. Column 1 presents estimates for the baseline population of GIC enrollees. Columns 2-4 present estimates that reweight the population such that 90% of households are older (age 55 and older) and/or reside in certain rating regions. Panel A reports estimates on only new, municipal entrants. Panel B reports estimates from alternate marginal cost assumptions. Panels C-E present estimates restricting the GIC from altering certain plans.

Alternate Assumptions on Unobserved Marginal Costs: I also test my model's sensitivity to the estimation of unobserved marginal costs, $c^u_{jt}(N_{jt})$, as detailed in subsection C.6. In particular, if the cost differential from switching to a narrow-network plan is lower than the cost differential I estimate, then the estimate of ρ may be inflated. I address this in two ways.

[27] 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 zero deductible.

[28] Because much of the switching cost is loaded onto network utility in this specification, the estimated upper bound on fixed costs would be exceedingly large.

First, rather than estimating the cost equation in Equation 28, I instead predict $c^u_{jt}(N_{jt})$ for each counterfactual network non-parametrically. Specifically, I take the ratio of estimated unobserved marginal costs to *observed* marginal costs, $c^o_{jt}(N_{jt})$, for each household and insurer. I then scale the observed marginal costs for each counterfactual network by this ratio. This has the effect of allowing unobserved marginal costs to vary not just by whether a counterfactual network is narrow or not, but by the size of its network relative to a broad-network plan. To the extent that small changes in networks result in smaller changes in unobserved marginal costs than currently estimated in Table C.5, this specification ought to address this. Second, I allow more flexibility in θ^c_I in Equation 9 by allowing unobserved marginal costs to vary not only by household type (individual or family), but also by age and rating region.

Note that these are fairly conservative assumptions. My model, for instance, assumes that provider prices are fixed in equilibrium. Prior literature on narrow networks have shown that more prevalent use of these plans may result in decreases in equilibrium provider prices (Ho and Lee, 2019; Ghili, 2020; Liebman, 2018). In addition, I am assuming a fixed *quantity* of care in equilibrium. To the extent that narrow networks induce reductions in this quantity, I may be underestimating the cost differential (Gruber and McKnight, 2016; LoSasso and Atwood, 2016). Finally, in Table C.5, the estimated differential for $c^u_{jt}(N_{jt})$ between a broad and narrow plan within the same insurer is about 16%, even though the premium differential between these plans is, on average, 20%. The assumptions here reduce these differentials further when the network changes are small.

Panel B of Table E.3 shows these results. Indeed, assuming that unobserved marginal costs decline by smaller magnitudes reduces the estimated ρ, though this is somewhat mechanical. Importantly, even under these conservative assumptions on costs, my estimates still imply employers overweight consumer surplus by about three times relative to the preferences of the average consumer.

Restricting Certain Moments: I run a series of robustness checks in which I restrict the estimation to certain moments to ascertain which plan changes drive the primary estimates of the employer-employee mismatch in Table 3. The results are displayed in Panels C-E of Table E.3. In Panel C, I restrict the GIC's ability to reduce Harvard's broad network to a "narrow" equivalent to network "N1" as described in section 5. This is a fairly network of both hospitals and physicians, with both Partners and Atrius removed, as well as additional hospitals and physician groups. In Panel D, I restrict the GIC from reducing Tufts' broad network to a "medium" sized network ("M."). In this network, the only major change is that Partners is removed. Atrius physicians, however, are preserved. The mismatch parameter is robust to each of these specifications.

In Panel E, I restrict the GIC from *broadening* Tufts' narrow network by limiting all moments in which it switches to a Harvard "M" network. Here, the baseline estimate of ρ increases slightly to 4.38. This is driven by the fact that these moments serve as relevant "upper bounds" on the parameter. Removing them therefore changes the bound on the estimate. Even still, the mismatch parameter is largely consistent with the baseline specification.

The one glaring exception is when I restrict the GIC's ability to reduce the network breadth of Harvard Broad to either its "M" network or a similarly-sized network on the small group market ("N2"). Again, these networks are somewhat larger than the others considered in the choice set, as they primarily remove Partners, but preserve a wide network of physicians in other systems. Panel F reports these estimates. Here, the mismatch parameter changes considerably: the estimate of ρ for the baseline population falls from 3.70 to 1.97.

The implication is that the mismatch appears to be largely driven by employers' unwillingness to make small network changes, yet those with significant implications on premium spending. 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. The cost implications, however, 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 Table 4 continue to persist. Reweighting the population to the older sample described above, the estimate of ρ falls further to 1.67, while it falls to 1.30 when rescaling towards the employees residing in regions 1,4, and 6.

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 to the same product space used in subsection 3.4 and detailed below. 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 matrix of $2^{14} = 16,384$ possible combinations of products offers.
2. For each product combination, 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 28.
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 menu offered.
7. Compute the employer's objective function using estimated $CS(\delta_{Jt}, \theta)$, total expenditures, the estimated mismatch parameter, ρ, and fixed costs, FC_j.
8. Repeat this procedure for each vector of possible plan menus, and take the max of all the computed employer objective functions.

F.2 Additional Counterfactual Tables and Figures

Premiums and Market Shares Under the Enthoven Counterfactual: Table F.1 shows the counterfactual premiums and market shares for each health plan under the Enthoven approach, assuming the product menu remains fixed. Under this counterfactual, individual co-premiums for broad-network plans increase substantially, while co-premiums for narrow-network plans remains relatively similar to their baseline. Harvard Broad, in particular, increases from its observed value of $152 per month to $311 per month. Not coincidentally, these plans also see substantial shifts in enrollment. The share of enrollees in Harvard Broad declines from 34% to just 11%, with many of those households shifting to Fallon, HNE, and NHP. As a result of these shifts, the overall spending burden for the GIC falls by *more* than the loss in consumer surplus.

Table F.1: Counterfactuals: Shares and Premiums for Enthoven Approach, 2011

Insurer	Network	Baseline	Counterfactual	Baseline	Counterfactual
		Market Shares		Co-Premiums	
Fallon	Very Narrow	0.02	0.08	$105	$111
Fallon	Broad	0.05	0.03	$126	$197
Harvard	Med	0.04	0.05	$122	$179
Harvard	Broad	0.34	0.11	$152	$311
HNE	Narrow	0.11	0.22	$105	$105
NHP	Narrow	0.03	0.29	$106	$109
Tufts	Narrow	0.01	0.02	$117	$159
Tufts	Broad	0.41	0.19	$147	$272

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

Policy Simulations Assuming No Logit Error: I re-estimate the region-rating counterfactuals presented in section 5 under the assumption of no logit error. Table F.2 reports these results. The results remain largely consistent with those reported in Table 6. 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 the GIC offering five plans in Rating Region 1 and Rating Region 4, while only offering four plans in Rating Region 5. Despite these changes, the welfare implications remain similar, but for a slight increase in total surplus (Panel C) relative to the estimates in Table 6. This is driven by the fact that the fixed cost estimates are smaller with the logit error removed (as in Table 3) and, as a result, social surplus is somewhat higher relative to the baseline scenario.

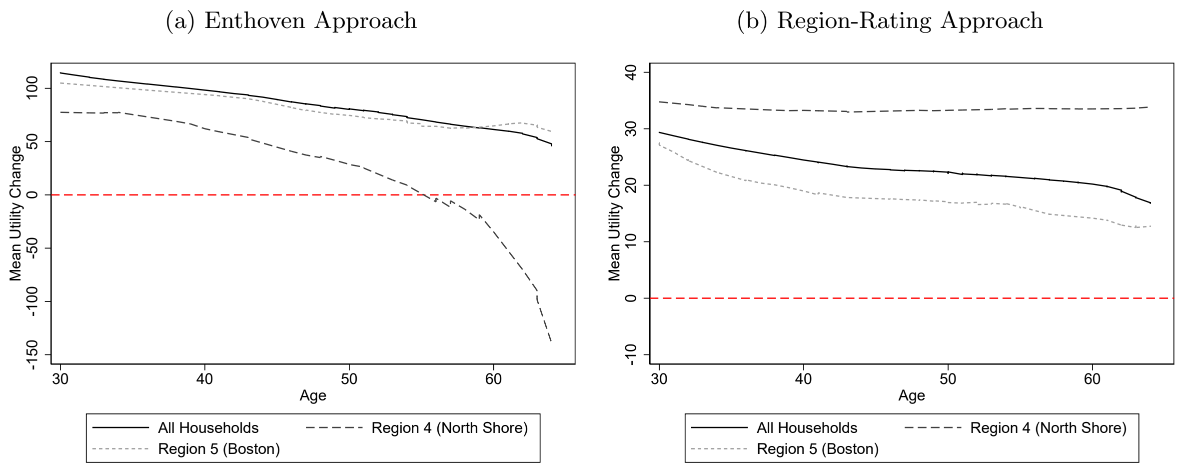
Table F.2: Counterfactuals: Equilibrium Networks Chosen Under Region-Based Pricing, No Logit Error

Insurer	Network	(1) Observed	(2) R1	(3) R4	(4) R5
Panel A: Equilibrium Plan Menus/Networks					
Fallon	VN	x		x	
Harvard	VN		x		
Tufts	VN				
Harvard	N1				
Tufts	N1	x			
Harvard	N2				x
Tufts	N2				
HNE	N	x	x		
NHP	N	x		x	x
Harvard	M	x	x		x
Tufts	M				
Fallon	B	x	x	x	
Harvard	B	x		x	
Tufts	B	x	x	x	x
Total Plans		8	5	5	4
Panel B: Welfare/Spending with Fixed Menu					
ΔCS				−\$0.55	
$\Delta Costs$				−\$1.26	
ΔFC				–	
$\Delta Surplus$				\$0.71	
Panel C: Welfare/Spending with Menu Changes					
ΔCS				−\$7.93	
$\Delta Costs$				−\$37.31	
ΔFC				\$3.85	
$\Delta Surplus$				\$25.54	

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. Panel C reports these quantities allowing the employer changes to menus. "ΔCS" refers to change in consumer surplus per-household-per-month. "$\Delta Costs$" refer to the change in GIC health costs per-household-per-month."ΔFC" refer to changes in fixed costs from the new menus. $\Delta Surplus$ refers to the change in total surplus.

Distributional Consequences for Individuals: Figure F.1 plots results from Figure 4, but for individuals rather than families. The results are quite similar, except that individuals see fewer surplus losses from either approach. For the Enthoven approach, individuals living in the North Shore do not see surplus losses until around age 55, while in the region-rating approach, *no* individuals see surplus losses.

Figure F.1: Total Surplus Changes by Age



Notes: This figure plots the average utility change across households with individual plans by age from implementing an Enthoven pricing approach (Panel A) and a region-rating approach (Panel B). All estimates allow 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 change in spending incurred by the GIC.