Non-Price Competition, Risk Selection, and Heterogeneous Costs in Provider Networks

Natalia Serna

Stanford University[*]

April 22, 2024

Abstract

Health insurers typically compete on the breadth of their provider networks. This paper shows that insurers' decision to offer network breadth depends on two forces: risk selection and cost incentives. To decompose the relative importance of these forces, I estimate a structural model of insurer competition in networks applied to data from Colombia. I find that insurers risk-select by providing narrow networks in services that unprofitable patients require. Despite selection incentives, some insurers choose to offer broad networks because of heterogeneity in their cost structure. I discuss implications for the design of risk adjustment and network adequacy rules.

Keywords: Health Insurance; Provider networks; Risk selection; Cost Structure.

JEL codes: I11, I13, I18, L13.

[*]e-mail: nserna@stanford.edu. I want to thank the Colombian Ministry of Health for providing the data for this research. I am deeply grateful to Alan Sorensen, Ken Hendricks, Corina Mommaerts, JF Houde, and Dan Quint, as well as Naoki Aizawa, Lorenzo Magnolfi, Christopher Sullivan, and Ashley Swanson for their mentorship and thoughtful advice. This research has benefited substantially from discussions with Ignacio Cuesta, Kate Ho, Maria Polyakova, Mark Shepard, and Pietro Tebaldi; as well as from comments from participants at BFI's 2024 Women in Empirical Microeconomics, NBER IO 2023 Summer Institute, 2022 Stanford's SITE, IO Workshop at UW-Madison, ASHEcon 2022, IIOC 2023, MEA 2022, Central Bank of Colombia Workshop 2023, and the seminars at Boston University, California Institute of Technology, Kelley School of Business, London School of Economics and Political Science, Ohio State University, Olin School of Business, Quantil Colombia, Ross School of Business, Stanford Health Policy, University of Toronto, University College London, and University of Pennsylvania. Findings do not represent the views of any institution involved. All errors are my own.

1 Introduction

Health insurers respond to different incentives when crafting the various elements of their insurance contracts, the two most salient of which are risk selection and fixed costs. Policies attempting to increase insurance coverage for patients should therefore consider the interplay of these incentives. In this paper I use a structural model of insurer competition on one of the elements of the contract, namely provider networks, to offer a complete characterization of how risk selection and fixed costs impact equilibrium network breadth. This is an important question given the proliferation of narrow-network insurers in different health systems and the popularity of policies like network adequacy standards that impose a minimum level of coverage.

Most prior studies have focused on how risk selection affects premiums, holding provider networks fixed (e.g, Cabral et al., 2018; Ho and Lee, 2017). Conversely, Shepard (2022) laid foundational work on how risk selection impacts provider networks conditional on premiums. Finally, Dafny et al. (2017); Polsky et al. (2016); Dafny et al. (2015) described the relation between network breadth and other elements of the insurance contract. In line with this prior work, my paper provides new evidence of how risk selection affects coverage choices. But importantly I show that this incentive interacts with average and fixed costs resulting in ambiguous effects on equilibrium network breadth: risk selection induces insurers to offer narrow networks, but scale and scope economies induce them to offer broad networks.

My empirical setting is Colombia, where private insurers provide a national health insurance plan in a system similar to Medicare Advantage (MA) in the US. A key difference relative to MA is that almost all aspects of the insurance contract are closely regulated: premiums and cost-sharing rules are all set by the government. The only element of the public health insurance plan that is unregulated is provider networks, making it an ideal setting for my purpose.

Health insurers in Colombia decide over which services to cover at which providers. Insurers can then use their service-level provider networks as a mechanism to select risks and minimize costs. This kind of non-price, service-level risk selection has been studied from a theoretical perspective by Cao and McGuire (2003) and Frank et al. (2000), and documented by Park et al. (2017) in the context of MA.[1] However, whether cost incentives play a role in determining service-level provider networks on top of risk selection remains an open question.

I start by documenting basic evidence that insurers use their service-level provider networks to risk-select. First, I show that the coarseness of the government's risk adjustment formula generates risk selection incentives, because it leaves significant variation in expected patient profitability depending on the types of services the patient is likely to need. I then give evidence that provider networks tend to be narrower for less profitable services. Finally, I show that patients tend to select insurers that have broad networks in services they are likely to need. For example, patients with cardiovascular disease are more likely to choose insurers with broad networks for cardiac care services.

Motivated by these descriptive facts, I develop and estimate a model of insurer competition in service-level provider network breadth. The model allows me to quantify the relative importance of risk selection and cost incentives, as well as how network breadth, health care costs, and consumer welfare respond to policies that change the magnitude of these incentives. This model builds on prior empirical work that uses Horn and Wolinsky (1988)'s Nash-in-Nash bargaining solution to endogenize negotiated prices between buyers and suppliers holding outside option prices and networks fixed (e.g., Grennan, 2013; Gowrisankaran et al., 2015; Ho and Lee, 2017). I redefine the problem of which providers to include in the network and at what prices as a problem of choosing network breadth, where network breadth is the fraction of providers in a market that deliver a service and are covered by the insurer. This redefinition is useful to endogenize network breadth and insurer costs in a tractable way, relying on

[1]Related patterns have been shown for drug coverage. Geruso et al. (2019) find that in the context of the ACA Exchanges, drugs commonly used by predictably unprofitable individuals appear on higher tiers of an insurer's drug formulary. Lavetti and Simon (2018) report similar results in the context of Medicare Part D.

the assumption that providers are homogeneous conditional on the service. In fact, Ericson and Starc (2015) and Starc and Swanson (2021) use similar approximations.

On the demand side, I model new enrollees' discrete choices of insurer as a function of service network breadth and out-of-pocket costs. This function captures the cost-coverage trade-off that consumers face when making enrollment decisions: consumers may have strong preferences for broader networks, but enrolling with a broad-network insurer is associated with higher out-of-pocket costs. On the supply side, I model insurers' heterogeneous cost structures in their average and fixed costs. Average costs per enrollee are a nonlinear function of service network breadth and enrollee characteristics, that allows for potential economies of scope across services. Fixed costs capture administrative costs associated with a choice of network breadth. Insurers maximize profits by choosing their vector of network breadths conditional on rivals' choices. I assume insurers make a one-time choice of service network breadth, recognizing that this choice will affect both current and future profits as patients age and transition between diagnoses.

To estimate the model, I use a novel administrative dataset that encompasses all enrollees to the contributory health care system in Colombia during 2010 and 2011, which represents nearly half of the population in the country (25 million individuals) and their medical claims (650 million). My data also contain the set of providers that insurers cover for every service. Demand estimates show that, conditional on sex and age, willingness-to-pay for network breadth varies substantially across diagnoses and services, consistent with adverse selection. Insurers' average cost function exhibits economies of scope, and both average and fixed costs are heterogeneous across insurers. The estimates imply that if an insurer unilaterally increases network breadth for general medicine, roughly half of the resulting cost increase is explained by cost heterogeneity and the other half is due to adverse selection (attracting sicker patients).

To quantify the relative importance of risk selection and cost incentives for equilibrium network breadth, I conduct two counterfactual exercises that each eliminates one of these incentives. To get at risk selection, I examine whether network breadth responds to changes in the government's risk adjustment formula. Without any risk adjustment, I find that mean network breadth would fall 24% (dropping 9 providers in the average network) and long-run consumer surplus would fall 3% for those with chronic diseases (nearly a 2/3 reduction in the monthly minimum wage). In contrast, if risk adjustment were made more granular by compensating for diagnoses, mean network breadth would increase 22% with effects being larger among services that sick patients tend to claim. These results suggest that risk selection drives the choice of narrow networks.

To get at cost incentives, the second counterfactuals examine whether the heterogeneity in insurers' cost structure can explain why some of them choose to offer broad networks despite risk selection. My main finding is that with homogeneous fixed costs, service network breadth collapses. Mean network breadth decreases 7.6% relative to the observed scenario, and the decline is larger and economically meaningful in services that sick individuals tend to claim. This finding indicates that absent network adequacy rules, a market with universal health insurance coverage can produce broad-network insurers in equilibrium provided insurers are sufficiently heterogeneous in their costs.

Contributions and literature. This paper makes three contributions to the literature: first, it shows that cost incentives have opposite effects relative to risk selection on the decision to offer network breadth. I build on insights from Shepard (2022) who first explicitly modelled selection through star hospital coverage in the context of Massachusetts Health Exchange. Papers that study alternative selection mechanisms include Shapiro (2020); Aizawa and Kim (2018) for insurer advertising, Geruso et al. (2019) for drug formulary design, and Decarolis and Guglielmo (2017) for insurance generosity and premiums. Second, it develops a new model of insurer competition that endogenizes network breadth across several insurers and services in a tractable way, while maintaining a relation with the underlying bargaining game.

Gowrisankaran et al. (2015) estimate a bargaining game assuming that hospitals' disagreement payoffs are zero. Ho and Lee (2019) endogenize prices under alternative networks in the context of a monopolist insurer, although more recently Fleitas et al. (2024) extend their theoretical framework to an empirical application with several insurers and providers. Ghili (2022) and Liebman (2022) also provide groundwork on endogenous networks. The third contribution is in quantifying the relative importance of risk selection and costs for equilibrium network breadth. Existing literature has studied the relation between network breadth and premiums in the health insurance Marketplaces (Dafny et al., 2017; Polsky et al., 2016; Dafny et al., 2015), while others have focused on the impact of risk adjustment on selection efforts (Brown et al., 2014; McWilliams et al., 2012; Nicholson et al., 2004) and premiums (Cabral et al., 2018; McGuire et al., 2013; Pauly and Herring, 2007).

The remainder of this paper is structured as follows: section 2 describes the institutional background and data, section 3 provides descriptive evidence of adverse selection and cost variation, section 4 presents the structural model, section 5 shows estimation results, section 6 studies the impact of risk adjustment, section 7 investigates the importance of insurers' cost structure, and section 8 concludes.

2 Institutional Background and Data

The Colombian health care system was established in 1993 and is divided into a "contributory" and a "subsidized" regime. The first covers formal employees and independent workers who pay monthly payroll taxes (nearly 51% of the population). The second covers individuals who are poor enough to qualify and are unable to contribute (the remaining 49%). The national health care system has almost universal coverage, which implies that insurer competition for enrollees is zero-sum.

Private insurers provide the national insurance plan. This plan covers a comprehensive list of more than 7,000 services or procedures and 673 medications as of 2010.

The government sets premiums for the national plan to zero and sets cost-sharing rules as a function of the enrollee's monthly income level, but they are standardized across insurers and providers.[2] Provider networks are the only dimension in which insurers can differentiate. Insurers form provider networks separately for each health service offered in the national health insurance plan. For example, insurers can choose to offer a broad network for orthopedic care but a narrow network for cardiology.[3]

At the end of every year, insurers report to the government all the health claims made through the national insurance plan that they reimbursed providers in their network for. The data for this paper are: enrollment files of all enrollees to the contributory system during 2010 and 2011 (25 million), insurers' claims reports to the government (650 million), and insurers' provider network data per health care specialty between 2010 and 2011 from the National Health Superintendency.

I focus on the sample of individuals aged 19 or older, of whom 2/3 have continuous enrollment spells or no gaps in enrollment. Of the continuously enrolled, 2/3 are *current enrollees*, that is, individuals who are enrolled throughout 2010 and 2011. The remaining 1/3 are *new enrollees* or individuals who enroll for the first time in 2011. Because there is near universal coverage, new enrollees to the contributory system can be individuals who move from the subsidized system after they find a job, turn 18 and choose a different insurer from their parents', or for some reason

[2] Cost-sharing in the national insurance plan follows a three-tiered system. As of 2010, for individuals earning less than 2 times the minimum monthly wage (MMW) the coinsurance rate equals 11.5 percent, the copay equals 2,100 pesos, and the maximum out-of-pocket amount in a year equals 57.5% times the MMW. This corresponds to an actuarial value of 92%. For those with incomes between 2 and 5 times the MMW, the coinsurance rate is 17.3 percent, the copay is 8,000 pesos, and the maximum out-of-pocket amount is 230% times the MMW. The associated actuarial value is 84%. Finally, for people with incomes above 5 times the MMW, the coinsurance rate equals 23%, the copay 20,900 pesos, and the maximum out-of-pocket amount is 460% times the MMW, corresponding to an actuarial value of 78%. The average exchange rate during 2011 was $1,847 COP/USD.

[3] Although the government does stipulate a set of network adequacy rules to guarantee appropriate access to health services, these rules are very coarse and apply only to the provision of primary care, urgent care, and oncology. These rules are described in https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PSA/Redes-Integrales-prestadores-servicios-salud.pdf

were uninsured for 12 continuous months.[4] Consumer inertia in this market is also substantial: conditional on staying within the contributory system, less than 6 percent of all enrollees switch their insurer from 2010 to 2011.

The enrollment files have basic demographic characteristics like sex, age, municipality of residence, and enrollment spell length in the year. Although I do not observe individual income per month, using aggregate income data from enrollees to the contributory system I assign to each individual the average income of his or her municipality, sex, and age cell. The health claims data report date of provision, procedure code, procedure price, provider, insurer, and ICD-10 diagnosis code. These claims come from the 23 private insurers that participated in the contributory health care system during my sample period. I focus on the 10 largest insurers that account for 95% of enrollees. Insurers compete in each of the 33 Colombian states or markets, which are similar in size to a Metropolitan Statistical Area in the US.

Every claim is associated to a 6-digit procedure code from the national insurance plan. These codes can be mapped to the health care specialties that insurers and providers bargain over, contained in the provider network data. This data report 150 unique specialties, which I aggregate up to 20 "services" and corroborate with network inclusions inferred from claims. Some examples of specialties in the data are cardiology, pediatric cardiology, and cardiovascular surgery, which I aggregate to cardiac care services. Other specialties are intensive care unit, intermediate care unit, neonatal intensive care unit, and hospitalization, which I aggregate to hospital admission services. Appendix 2 provides the final list of services and an excerpt from the provider network data.

Service-level negotiations. The fact that insurers and providers negotiate net-

[4] Even if new enrollees in 2011 had enrollment before the start of my sample period in 2010, decree 806 of 1998 and decree 1703 of 2002 established that after three continuous months of non-payment of tax contributions, a person would be disenrolled and lose any information so far reported to the system. Enrollment after non-payment is therefore a "fresh-start" in the contributory system. Moreover, in 2011 only around 500 thousand enrollees switched from an insurer in the subsidized system that also had presence in the contributory system.

work inclusions for every health service could generate clinical chaos if, for example, a patient with diabetes is admitted to a hospital for a cardiac episode but the insurer does not cover diabetes care at that hospital. This kind of care fragmentation has received substantial media attention for an issue that came to be called "the rounds of death," where patients would have to go from clinic to clinic to be treated and died in the process.[5] To understand how insurers and providers bargain over services to avoid these issues, I interviewed three experts in Colombia who mentioned that insurers tend to cover all the services at large clinics and hospitals where inpatient admissions occur. But service coverage varies at smaller providers where outpatient care is delivered.[6] Appendix figure 2 corroborates the experts' information by showing that a large fraction of insurer-provider pairs cover all the services, but there is still substantial variation in service coverage within provider that motivates the following descriptive evidence.

3 Descriptive Evidence

Private insurers in the contributory system are reimbursed by the government at the beginning of every year (ex-ante) with capitated risk-adjusted transfers, and at the end of every year (ex-post) with the High-Cost Account. The ex-ante risk adjustment formula controls for sex, age group, and municipality of residence, but it does not include diagnoses. Appendix 1 describes how this risk-adjusted transfer is calculated. Because of the coarsely defined risk pools, the ex-ante formula poorly fits realized health care costs. Riascos et al. (2014, 2017) find that the R^2 of the government's formula is only 0.017. Using the demographic information contained in the enrollment files, I can recover the ex-ante transfers that each insurer received for each of its

[5] See `https://caracol.com.co/radio/2018/10/03/nacional/1538571677_077170.html` and `https://www.elespectador.com/tags/paseo-de-la-muerte/`.

[6] These experts were the former National Quality Coordinator for Coomeva (insurer), the Chief Contracting Officer for Sanitas (insurer), and the Director of Marketing for Medicina Integral en Casa (provider).

enrollees. Ex-ante reimbursements range from 162.2 thousand pesos (males aged 15-18) to 2.2 million pesos (for females aged 75 or older), while realized costs range from 0 to over 300 million pesos.

The High-Cost Account compensates insurers that enroll an above-average share of people with certain diagnoses, and reimbursements come from insurers that enroll a below-average share.[7] My data contain total High-Cost Account transfers that each insurer received per year. Total ex-post transfers represent only 0.4 percent of total ex-ante transfers per insurer, suggesting these ex-post transfers do not provide much risk adjustment.

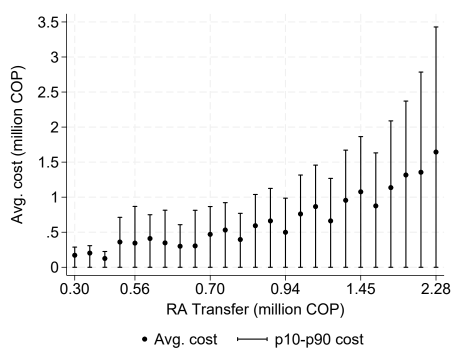


Figure 1: Health care cost by risk-adjusted transfer

Note: Figure presents mean and 10th and 90th percentiles of annual health care cost conditional on the government's ex-ante risk-adjusted transfer.

Selection incentives in this system exist because annual health care costs exhibit enormous variation across patients conditional on the risk-adjusted transfers. Figure 1 shows that the mean and the variance (as reflected in the difference between 10th and 90th percentiles) of health care costs increase with the government's reimbursement. The rising trend in average costs suggests that insurers can have incentives to enroll costly individuals because they can receive higher government reimbursements. The rising trend in variance suggests that there is scope to select consumers in the

[7]Diseases compensated by the High-Cost Account include: cervical cancer, breast cancer, stomach cancer, colon cancer, prostate cancer, lymphoid leukemia, myeloid leukemia, hodgkin lymphoma, non-hodgkin lymphoma, epilepsy, rheumatoid arthritis, and HIV-AIDS. See Resolution 000248 of 2014 from the Ministry of Health.

upper tail of the distribution who are more likely to be overcompensated by the risk adjustment formula (as in Brown et al., 2014).

3.1 Measuring Network Breadth

If insurers respond to selection incentives using their provider networks, then differences in health care costs should appear as differences in service network breadth. I define service network breadth as the fraction of providers in a market offering a particular service that are covered by the insurer.[8]

Table 1: Distribution of service network breadth per insurer in 2011

Insurer	mean	p25	p75
EPS001	0.14	0.03	0.21
EPS002	0.29	0.00	0.45
EPS003	0.15	0.00	0.25
EPS005	0.33	0.18	0.43
EPS008	0.04	0.00	0.04
EPS009	0.10	0.00	0.12
EPS010	0.07	0.00	0.13
EPS012	0.10	0.01	0.11
EPS013	0.51	0.33	0.70
EPS016	0.32	0.15	0.49
EPS017	0.14	0.00	0.20
EPS018	0.21	0.04	0.31
EPS023	0.04	0.00	0.04
EPS037	0.52	0.34	0.73

Note: Table presents mean and 25th and 75th percentiles of service network breadth per insurer during 2010 and 2011.

Table 1 shows that there is substantial variation in this measure of coverage across insurers. Even with poor risk adjustment some insurers like EPS013 and EPS037 choose to offer broad service networks, while others like EPS001 and EPS002 choose narrow networks. Although by US standards some of these insurers would have ultra-narrow networks, these standards are based on the coverage of large hospitals.[9] My

[8] Although it is mandatory that insurers cover at least one provider for every service in the national insurance plan, coverage choices can be determined by the type of consumers that insurers want to risk select upon.

[9] See https://www.mckinsey.com/industries/healthcare/our-insights/

measure of network breadth is instead defined over coverage of hospitals, small clinics, and even smaller physician practices. As long as a provider is certified by the Ministry of Health, this provider will be included in my measure. Dafny et al. (2017) report that provider networks in the US based on a similar definition also tend to be much narrower than hospital networks.

3.2 Network breadth as a Means of Risk Selection

Variation in service network breadth is consistent with differences in selection efforts and costs *across insurers*. In this subsection I characterize selection incentives *across services* by replicating figures in Geruso et al. (2019) with data from *all* enrollees in the contributory health system.

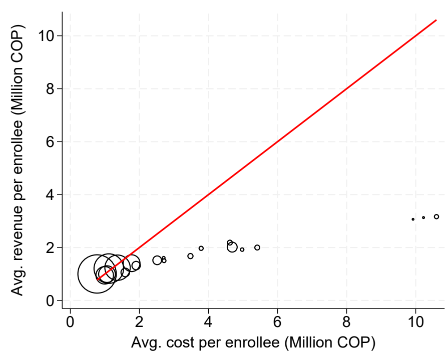


Figure 2: Service-level selection incentives after risk adjustment

Note: Figure presents a scatter plot of average revenue and average cost per enrollee. Each dot is a service weighted by the number of individuals who make claims for the service. Revenues are calculated as government ex-ante and ex-post risk-adjusted transfers, plus revenues from copays and coinsurance rates. The red line is a 45 degree line. One enrollee can be represented in several dots if she makes claims for different services. Enrollees who make zero claims are not represented in this figure.

Figure 2 shows whether the current risk adjustment systems are effective at neutralizing service-level risk selection. The figure plots the average cost per enrollee against the average revenue per enrollee conditional on patients who make claims for each service. Every circle represents a service weighted by the number of patients who

`hospital-networks-updated-national-view-of-configurations-on-the-exchanges/` for a definition of ultra narrow networks in the ACA marketplaces.

make claims for it. Patients who make claims for several services will be represented in several circles, while patients who make zero claims (and are the most profitable) are not represented in this figure. The red line is the 45 degree line, which splits the space into services that are overcompensated by the risk-adjusted transfers (above the line) and those that are undercompensated (below the line). The main takeaway is that patients who make any claim are likely to be unprofitable; but this is especially true for patients who have claims in certain services such as cardiac care, renal care, and hospital admissions, which are located toward the right of this figure.

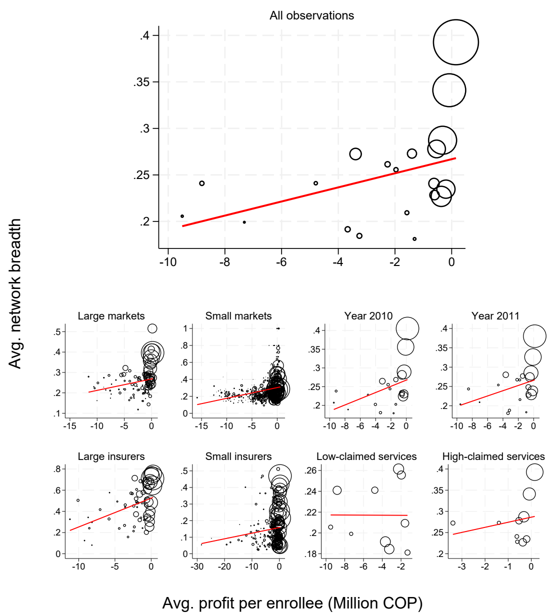


Figure 3: Correlation between network breadth and service profitability

Note: Figure presents a scatter plot of average service network breadth and average profit per enrollee. Each dot is a service weighted by the number of individuals who make claims for the service. Profits are calculated as government's ex-ante and ex-post risk-adjusted transfers, plus revenues from copays and coinsurance rates, minus total health care costs. The red line corresponds to a linear fit. One enrollee can be represented in several dots if she makes claims for different services. Enrollees who make zero claims are not represented in this figure.

The existence of services that are outliers in terms of profits per enrollee suggests a scope for insurers to engage in service-level risk selection or cost minimizing strategies through their choice of provider networks. One way to test whether the data are consistent with selection at the service level is to show whether network breadth covaries with the profitability of a service, a version of the positive correlation test in

Chiappori and Salanie (2000).

Figure 3 plots the average profit per enrollee against average service network breadth across insurers and markets. Average profits are calculated conditional on patients who make claims for each service. The red line corresponds to a linear fit and shows that relatively profitable services, such as general medicine and laboratory, tend to have broader networks than relatively unprofitable services, such as cardiac care and renal care. This correlation holds along several dimensions considered in the bottom panels of the figure and is not necessarily driven by services with few claims.

Figure 4 unpacks some of the variation in network breadth across services to give an example of those that are likely to be under-covered. The distribution of network breadth for services related to primary care, such as general medicine and laboratory testing, is shifted to the right and has a wider right tail compared to the distribution of complex services such as hospital admissions, neurological care, cardiac care, and renal care. These patterns suggest that relatively complex services tend to be under-covered, which we would expect if insurers want to avoid the patients who need those services. Appendix figure 4 also shows that even though the distribution of network breadth per service is heterogeneous across markets, which services are under-covered is common across markets.

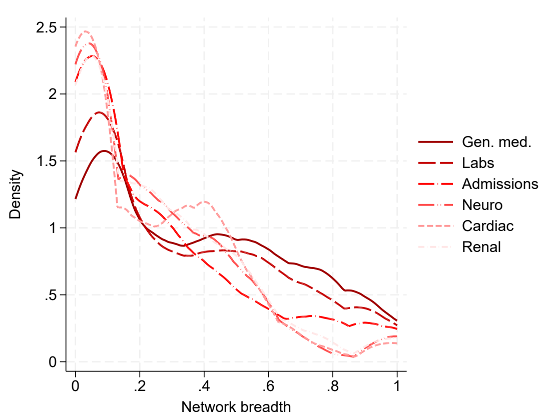


Figure 4: Distribution of network breadth per service

Note: Figure presents kernel density estimates for the distribution of network breadth conditional on six services: general medicine, laboratory testing, hospital admissions, neurological care, cardiac care, and renal care.

Another explanation for why complex services tend to be under-covered is that total demand for those services is relatively low. This explanation would rule out risk selection as a driver of insurers' network breadth choices. To test the importance of risk selection I estimate the correlation between insurer market shares in the number of new enrollees with chronic diseases and network breadth for the services those patients are most likely to need. Table 2 shows that market shares in initial choices are positively correlated with network breadth, suggestive of adverse selection. For example, insurers with relatively broad networks for renal care have a higher share of new enrollees with renal disease.

Table 2: Market share in initial choices and network breadth

	(1) Healthy	(2) Cancer	(3) Diabetes	(4) Cardio	(5) Renal
Network breadth	0.411 (0.022)	0.332 (0.042)	0.409 (0.022)	0.297 (0.046)	0.327 (0.064)
Observations	312	312	312	312	312

Note: Table presents OLS regression of insurer market share on service network breadth. Column (1) uses the sub-sample of individuals without diagnoses and network breadth for general medicine. Column (2) uses the sub-sample of individuals with cancer and network breadth for chemotherapy. Column (3) uses the sub-sample of individuals with diabetes and network breadth for laboratory. Column (4) uses the sub-sample of individuals with cardiovascular disease and network breadth for cardiac care services. Column (4) uses the sub-sample of individuals with renal disease and network breadth for renal care services. All specifications include market fixed effects. Standard errors in parenthesis are clustered at the market level.

Switching decisions. Prior papers on selection in health insurance markets leverage enrollees' switching decisions to test for adverse selection (e.g., Shepard, 2022; Gruber and McKnight, 2016; Newhouse et al., 2015; Brown et al., 2014; Einav et al., 2013). This type of analysis would be under-powered in my setting since, conditional on staying in the contributory system, only 6 percent of enrollees switch their insurer between 2010 and 2011. Despite these limitations, appendix table 5 shows evidence of adverse selection on switching decisions in my setting. For instance, findings show that healthy enrollees are more likely to switch out of insurers with broad networks for primary care, while patients with cardiovascular disease are more likely to switch out of insurers with narrow networks for cardiac care.

4 Model

Motivated by the descriptive evidence, I develop a structural model of the Colombian insurance market to decompose risk selection and cost incentives as potential mechanisms for network breadth. I limit my analysis sample moving forward to individuals who have continuous enrollment spells, which distinguishes consumers whose choices are not conflated by variation in income, job loss, or informality. Appendix 4 shows some summary statistics and replicates all the descriptive evidence presented earlier for this sample.

4.1 Foundations and Relation to Prior Work

My model is specified over insurers' decision to offer service network breadth. This measure of provider coverage allows me to endogenize networks across insurers and services in a tractable way, but it loses the identities of in-network providers. Yet, substantial research in the US suggests that patients care strongly about whether their preferred provider is included in the network (e.g., Ho, 2006; Shepard, 2022).

Models of provider choice that allow for preference heterogeneity across patients can be used to derive an alternative measure of coverage given by consumers' expected utility for the network following (McFadden, 1996). This measure is usually fed into models of insurer and provider Nash bargaining with the goal of endogenizing negotiated prices (Gowrisankaran et al., 2015; Ho and Lee, 2017). The structural unobservable of the Nash-in-Nash surplus function is typically the provider's marginal cost, but off-equilibrium prices in the event that the insurer and the provider disagree are assumed to be fixed at their equilibrium values.

When trying to endogenize the insurer's decision of which providers to include in its network in addition to negotiated prices, the assumption that disagreement payoff prices are fixed is far too strong. Nash-in-Nash proves to be infeasible in this case because both the provider's marginal cost and the off-equilibrium price are unobserved, resulting in a system with more unknowns than equations.[10] Ho and Lee (2019) provide one solution to this problem assuming that the off-equilibrium price is the price that makes the insurer indifferent between keeping the hospital in the network or replacing it for another hospital at its reservation price. Here I derive an alternative solution by redefining the problem of *which* providers to include in the network and at what *price* as a problem of *how* many providers to include and at what *cost*.

Appendix 5 shows that my model can be micro-founded with these traditional models of provider choice and bargaining. The limitations of my approach are that it relies on providers being homogeneous conditional on the service and on the underlying provider demand being orthogonal to prices (as in Ho and Lee, 2017). The model allows me to quantify by how much would network breadth and insurer average costs change under counterfactual policies, which are both objects of interest for policymakers; but, it does not allow me to measure how would negotiated prices change under different networks.[11] I move now to describing my econometric implementation.

4.2 Insurer Demand

I model insurer demand in the sample of new enrollees in 2011, who do not experience inertia when making their first enrollment choice. Assume that a new enrollee i living in market m is of type θ. With probability $q_{\theta k}$, such that $\sum_k q_{\theta k} = 1$, the consumer will need each of the $k = \{1, ..., K\}$ services. An individual's type is given by a combination of sex, age category (19-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, ≥75), and diagnosis $d \in D = \{$cancer, diabetes, cardiovascular disease, pulmonary disease, renal disease, other chronic disease, no

[10] Crawford and Yurukoglu (2012) discuss some of the limitations of allowing for endogenous networks in Nash-in-Nash in the context of television markets.

[11] Estimating a richer model of provider choice and Nash-in-Nash bargaining is also infeasible in my setting as the size of consumer choice sets for relatively common services such as general medicine equals 179 in the largest market. This choice set would imply solving $179 \times |J_m|$ bilateral negotiations in this market and for general medicine alone.

diseases}. Diagnoses in the list are groupings of ICD-10 codes following Riascos et al. (2014).[12]

I assume that the individual knows his or her diagnoses before making the first enrollment choice. This could be either because of medical family history or because, prior to enrolling in the contributory system, they went to the doctor and received a diagnosis. Private information on their diagnosis, which I observe in the data, implies that selection in my model will occur on observable, un-reimbursed (or poorly reimbursed) characteristics such as those associated with health status.[13]

Denote by u_{ijm} the indirect utility of a new enrollee i in market m for insurer j, which takes the following form:

$$u_{ijm} = \beta_{ij} \sum_k q_{\theta k} H_{jkm} - \alpha_i c_{\theta jm}(H_{jm}) + \phi_{jm} + \varepsilon_{ijm} \tag{1}$$

where $\beta_{ij} = (x_i \quad x_j)'\beta$ and $\alpha_i = x_i'\alpha$. The vector x_i includes consumer characteristics such as dummies for sex, age category, diagnosis, living in a rural market, and having low income. x_j is an indicator for relatively large, high-quality insurers based on quality rankings constructed by the Ministry of Health for 2013. The average out-of-pocket cost of consumer type θ at insurer j is given by $c_{\theta jm}$ and depends on the insurer's vector of network breadth $H_{jm} = \{H_{jkm}\}_{k=1}^{K_m}$. The coefficient ϕ_{jm} is an insurer-by-market fixed effect that captures unobserved insurer quality that varies across markets. Finally, ε_{ijm} is an *iid* unobserved shock to preferences assumed to be distributed type-I extreme value.

Average out-of-pocket costs are the sum of coinsurance payments, copays, and tax

[12] These diagnoses were chosen for being the most expensive in Colombia and thus the most likely to be undercompensated by the current risk adjustment formula. For example, the most expensive patients with renal disease had annual health care cost of over 55 million pesos in 2011, more than 100 times the monthly minimum wage. For individuals with several comorbidities, I assign the most expensive disease.

[13] Results are also robust to a version of the model where individuals are uncertain about their diagnoses (see appendix table 11).

contributions to the system:

$$c_{\theta jm} \equiv \text{Coins}_{\theta jm} + \text{Copay}_{\theta jm} + \text{Tax}_{\theta} = r_{\theta} AC_{\theta jm}(H_{jm})$$

Coinsurance payments and copays are indexed by j because they are a function of the insurer's negotiated service prices and of the individual's health care utilization. Negotiated prices and utilization may be correlated with service network breadth for two reasons: first, an insurer's bargaining position depends on how many providers it has included in the network; second, individuals may consume more services the broader is the network. I capture this correlation by noting the pass-through of insurers' costs to consumers' out-of-pocket costs via cost-sharing. Out-of-pocket costs equal the individual's coinsurance rate times the insurers' average cost per enrollee, which in turn depends on network breadth. This dependence is needed to rationalize the existence of narrow network insurers in equilibrium since myopic, healthy new enrollees will disproportionately choose narrow-network insurers with lower implied out-of-pocket costs.[14]

The probability of making a claim, $q_{\theta k}$, is calculated outside of the model as the average prediction of a logistic regression. Appendix 6.1 explains this procedure. Given the distribution of the preference shock, the probability that consumer i in market m enrolls with insurer j is:

$$s_{ijm}(H_m) = \frac{\exp\left(\beta_{ij} \sum_k q_{\theta k} H_{jkm} - \alpha_i c_{\theta jm}(H_{jm}) + \phi_{jm}\right)}{\sum_{j' \in \mathcal{J}_m} \exp\left(\beta_{ij} \sum_k q_{\theta k} H_{j'km} - \alpha_i c_{\theta j'm}(H_{j'm}) + \phi_{j'm}\right)}$$

Identification. To identify the parameters associated with network breadth, I rely on variation in market demographics across markets, which generates exogenous variation in the claim probabilities. For example, if an insurer offers the same network breadth for cardiac care in two different markets, but one of these markets has a

[14] My specification for out-of-pocket costs per consumer type and insurer is equivalent to aggregating service-level out-of-pocket costs with weights given by the claim probabilities.

higher prevalence of cardiovascular conditions, then we should observe higher insurer demand in the market where people are relatively sicker.

Identification is threatened if network breadth is correlated with unobserved insurer quality or unobserved consumer characteristics, such as their valuation for specific providers. Network breadth could also be correlated with how good the insurer is in processing health claims. These types of unobserved insurer characteristics potentially do not vary across markets conditional on the consumer type. Therefore, inclusion of insurer-by-market fixed effects allows me to identify preferences for network breadth off of exogenous variation in market demographics.

To identify the parameters associated with the out-of-pocket cost, I use variation in income across markets, which generates exogenous variation in coinsurance rates. This variation may not be sufficient for identification if negotiated service prices are correlated with unobserved provider quality. For example, if an insurer covers a star hospital, demand and negotiated prices for that insurer will be relatively high across all income groups, and my model would interpret consumers as having low sensitivity to out-of-pocket costs. This endogenous variation is specific to an insurer-market combination, hence the inclusion of insurer-by-market fixed effects help isolate the variation in out-of-pocket costs that is exogenous. I also conduct robustness checks in section 5 to verify that differences in provider quality conditional on the service are not significant.

4.3 Insurer Average Costs per Enrollee

I approximate the expected cost of type-θ individuals as the average cost across all consumers i that are of type θ. Then, I model the logarithm of average cost per consumer type as a quadratic function of network breadth:

$$\log(AC_{\theta jm}(H_{jm})) = \tau_0\Big(\sum_{k}^{K_m} q_{\theta k}A_k\Big) + \tau_1\Big(\sum_{k}^{K_m} q_{\theta k}H_{jkm}\Big) + \frac{1}{2K_m}\tau_2\sum_{k}^{K_m}\sum_{l\neq k}^{K_m} q_{\theta k}q_{\theta l}H_{jkm}H_{jlm} + \lambda_\theta + \eta_m + \delta_j \quad (2)$$

where K_m is the number of services available in market m, A_k is the government's reference price for service k which guides service-level negotiations (explained in more detail in appendix 7), and λ_θ, η_m, and δ_j are consumer type, market, and insurer fixed effects, respectively. Equation (2) captures the fact that consumers of different types will imply different average costs to the insurer conditional on network breadth. This reduced-form approach is very much in the spirit of Tebaldi (2024) who also models insurers' expected costs as an exponential function of consumer characteristics and generosity of coverage. The quadratic approximation is supported by the underlying bargaining game as seen in appendix 5 and by the empirical relationship between log average costs and network breadth as seen in appendix figure 9.

The coefficient τ_0 captures whether insurers bargain higher or lower prices than the reference price with the average provider in their network. τ_1 represents the direct effect of network breadth for service k on average costs. τ_2 captures the average degree of complementarity between pairs of services. If $\tau_2 < 0$, insurers have economies of scope across services, thus greater coverage for service $l \neq k$ makes it more attractive to provide higher coverage for service k. This measure of scope economies helps rationalize the fact that insurers with broad networks in one service, tend to offer broad networks in other services as well (see appendix figure 10).

Identification. The parameters of equation (2) are identified from variation in average costs within consumer types and across insurers that are identical except for their service network breadth. My source of identification does not rely on different consumers implying different costs for similar insurers as in Tebaldi (2024) but, *conditional* on the composition of enrollee pools, for different service coverage levels to imply different costs to the insurer. In this case, variation in network breadth across insurers is exogenous conditional on the rich set of fixed effects. However, one worry is that consumers may select into insurers based on their unobservables. One way to check this is to test whether estimates are robust to more granular definitions of consumer types. I conduct robustness checks of this style in appendix table 12.

4.4 Competition in Network Breadth

Insurers compete separately in every market choosing their service network breadths after taking expectations of demand and costs. Let $\pi_{ijm}(H_m, \theta)$ be insurer j's annual per-enrollee profit in market m, which depends on j's network breadth and its rivals' $-j$, all collected in the vector $H_m = \{H_{jm}, H_{-jm}\}$. The annual per-enrollee profit is given by:

$$\pi_{ijm}(H_m, \theta) = (R_{\theta m} - (1 - r_\theta)AC_{\theta jm}(H_{jm}))s_{ijm}(H_m)$$

where $R_{\theta m}$ is the per-capita revenue (including ex-ante and ex-post risk-adjusted transfers plus average copayments), $AC_{\theta jm}$ is the average cost of a type-θ consumer net of patients' coinsurance payments with r_θ denoting the coinsurance rate, and s_{ijm} is consumer i's choice probability for insurer j in market m.

I focus on a Nash equilibrium in which insurers choose networks simultaneously to maximize the sum of current profits and future discounted profits minus fixed costs:

$$\Pi_{jm}(H_m) = \sum_{\theta} \Bigg(\underbrace{\pi_{ijm}(H_m, \theta)N_{\theta m}}_{\text{current profit}} + \underbrace{\sum_{s=t+1}^{T} \zeta^s \sum_{\theta'} (1 - \rho_{\theta' m})\mathcal{P}(\theta'|\theta)\pi_{ijm}(H_m, \theta')N_{\theta' m} \Bigg)}_{\text{future profit}}$$
$$- \underbrace{\sum_{k} (\omega H_{jkm} + \xi_{jkm})H_{jkm}}_{\text{fixed cost}}$$

Insurers take into account the future profits associated with each enrollee since, after making their first enrollment choice, individuals do not switch as seen in the data.

Insurers therefore maximize the net present value of their profits. $N_{\theta m}$ is the fixed market size of consumers type θ. In the expression for future profits, $\rho_{\theta m}$ represents the probability that a type-θ consumer drops out of the contributory system. This probability is (assumed) exogenous to the choice of network breadth as it is mostly governed by the event of being unemployed. $\mathcal{P}(\theta'|\theta)$ is the transition probability from type θ in period t to type θ' in period $t+1$. Future profits at year t are discounted by a factor of ζ^t, which I set to 0.95 and forward simulate this profit function for 100 periods.[15]

In addition to its indirect effect on insurer profits through expected costs and demand, I assume network breadth involves a direct fixed cost to the insurer. This is an administrative cost associated with the inclusion of an additional provider to the network. The fixed cost is non-linear in network breadth and heterogeneous across insurers with $\xi_{jkm} = \xi_j + \vartheta_{jkm}$. In this specification, ξ_j represents the observed insurer-specific cost component and ϑ_{jkm} represents the idiosyncratic cost shock that is observed by insurance companies but unobserved to the econometrician. The multiplicative structure of the unobserved cost is needed to obtain a first-order condition that is linear in ϑ_{jkm}.

Profit maximization involves a set of $|J| \times |K|$ FOCs in each market, which assuming an interior solution in network breadth, is given by:

$$\sum_i \left(\frac{\partial \pi_{ijm}}{\partial H_{jkm}} N_{\theta m} + \sum_{s=t+1}^{T} \zeta^s \sum_{\theta'} (1 - \rho_{\theta' m}) \mathcal{P}(\theta'|\theta) \frac{\partial \pi'_{ijm}}{\partial H_{jkm}} N_{\theta' m} \right) = \omega H_{jkm} + \xi_{jkm} \quad (3)$$

The left-hand side of equation (3) represents the marginal variable profit MVP_{jkm}, and the right-hand side is the marginal cost of network formation. The derivative of

[15] In the formulation of insurer profits, I use θ to denote sex-age-diagnosis combinations as opposed to sex-age group-diagnosis, for simplicity in notation, but to be consistent between transition probabilities and periods over which future profits are calculated (years).

the short-run per enrollee profit, which enters MVP_{jkm}, is:

$$\frac{\partial \pi_{ijm}}{\partial H_{jkm}} = \underbrace{\overbrace{R_{\theta m}\frac{\partial s_{ijm}}{\partial H_{jkm}}}^{\substack{\text{Selection}\\ \text{incentives}}} + \overbrace{R_{\theta m}\frac{\partial s_{ijm}}{\partial AC_{\theta jm}}\frac{\partial AC_{\theta jm}}{\partial H_{jkm}}}^{\text{Cost incentives}}}_{\text{Marginal revenue}} \tag{4}$$

$$- \underbrace{(1-r_\theta)\Big(\overbrace{AC_{\theta jm}\frac{\partial s_{ijm}}{\partial H_{jkm}} + s_{ijm}\frac{\partial AC_{\theta jm}}{\partial H_{jkm}}}^{\text{Selection incentives}} + \overbrace{AC_{\theta jm}\frac{\partial s_{ijm}}{\partial AC_{\theta jm}}\frac{\partial AC_{\theta jm}}{\partial H_{jkm}}}^{\text{Cost incentives}}\Big)}_{\text{Marginal cost}}$$

Equation (4) shows how selection and cost incentives affect insurers' network breadth choices. If an insurer unilaterally increases its network breadth for a particular service, marginal revenues will increase because demand from individuals with high willingness-to-pay for that service is higher (selection effect). Marginal costs also increase because patients with high willingness-to-pay for the service are the most expensive in that service, and because changes in network breadth increase the cost of the marginal consumer (selection effect). Cost incentives have opposite effects on marginal revenues and marginal costs. Expanding networks for a particular service increases consumers' out-of-pocket costs and thus puts a downward pressure on marginal revenues. An increase in network breadth also reduces marginal costs because if relatively sicker consumers disenroll due to higher out-of-pocket payments, then the marginal consumer is cheaper.

Identification. Rewriting the FOC as

$$\text{MVP}_{jkm}(H_{jkm}) = \omega H_{jmk} + \xi_j + \vartheta_{jkm}, \quad \forall\ H_{jkm} \in (0,1) \tag{5}$$

makes explicit the endogeneity problem between H_{jkm} and the network formation cost shocks, ϑ_{jkm}. Insurers observe ϑ_{jkm} before or at the same time as they are deciding on their service network breadths. For instance, if an insurer hires a highly trained manager to bargain with providers or if an insurance company is vertically integrated with its network, then $E[\vartheta_{jkm}|H_{jkm}] < 0$.[16] Identification of the network formation cost shock thus relies on insurer fixed effects given by ξ_j, which capture the endogenous variation in marginal variable profits across insurers. I estimate the FOC via OLS since only 1% of observations correspond to corner solutions in H_{jkm} in my estimation sample.[17]

5 Estimation

5.1 Insurer Demand

The insurer demand model is a conditional logit estimated by maximum likelihood. To reduce the computational burden, I estimate equation (1) on a random sample of 500,000 new enrollees. Results in table 3 show that insurer demand is decreasing in out-of-pocket costs and increasing in network breadth. A 10 thousand pesos increase in out-of-pocket costs reduces the choice probability by 24%, corresponding to an average elasticity of -0.26.[18] A ten percentage point increase in network breadth across all services increases the choice probability by 23%.[19] These results suggest not only that there is selection on network breadth but also that consumers prefer broad service networks overall.

Interactions between consumer and insurer characteristics matter for enrollment decisions. Sensitivity to out-of-pocket costs is decreasing with income. Patients aged 65 or older are both more likely to enroll in broad-network insurers and more sensitive to out-of-pocket costs compared to younger patients. One explanation for this result is that old individuals need more expensive care. Individuals with cancer

[16]Vertical integration is restricted by the Colombian government to up to 30% of an insurance company's assets. So, endogeneity stemming from integration is unlikely.

[17]Alternatively, the parameters of the network formation cost can be estimated using moment inequalities (Pakes et al., 2015).

[18]The elasticity with respect to out-of-pocket costs is $\frac{\partial s_{ijm}}{\partial c_{\theta jm}} \frac{c_{\theta jm}}{s_{ijm}}$, which is averaged across consumers and insurers.

[19]Calculated as $\beta_{ij} \sum_k q_{\theta k}$ and averaged across consumers and insurers.

and renal disease have stronger preferences for broader networks than their healthy peers. Consumers with chronic conditions are also significantly less responsive to out-of-pocket costs. Appendix 6.2 presents some measures of in-sample model fit.

With my estimates of the preference for network breadth and out-of-pocket costs, I calculate patient willingness-to-pay (wtp) for an additional percentage point of network breadth in each service as $\frac{1}{-\alpha_i}\frac{\partial s_{ijm}}{\partial H_{jkm}}$. Differences in wtp across consumer types will be suggestive of patient sorting based on network breadth. Table 4 presents the average wtp for some services among patients with chronic diseases, normalizing healthy individuals to 1. Patients with chronic conditions have a significantly higher wtp for network breadth across all services compared to individuals without diagnoses. For example, patients with renal disease are willing to pay 27 times more than a healthy individual for an additional provider in the network for renal care services.[20] This variation in wtp implies that, in principle, insurers can avoid unprofitable patients by offering narrow networks in the services they require, and that for some services insurers can find it profitable to offer broad networks.

Robustness checks. I conduct several robustness checks to provide encouraging evidence of my identification arguments. Appendix table 9 presents a demand function that includes an indicator of star hospital coverage, showing that it is insignificant conditional on network breadth. Because requiring that new enrollees know their diagnoses before enrolling can create mechanical bias, in appendix table 10 I identify new enrollees' diagnoses using only the information from claims made in January 2011. Finally, appendix table 11 shows a version of demand where consumers are uncertain about their diagnoses.

[20]The measure of willingness-to-pay can also be interpreted in terms of travel times to the nearest provider as seen in appendix figure 3. For example, the estimates imply that patients with renal disease are willing to pay 27 times more than a healthy individual for a reduction of approximately 10 minutes in travel time per visit to the nearest provider that offers renal care services.

Table 3: Insurer demand

Variable		Network breadth	OOP spending (million)
Mean		2.34 (0.42)	-2.41 (0.11)
Interactions			
Demographics	Male	0.15 (0.02)	0.06 (0.07)
	Age 19-24	-0.60 (0.05)	1.51 (0.12)
	Age 25-29	-1.19 (0.05)	0.70 (0.12)
	Age 30-34	-1.46 (0.05)	0.56 (0.15)
	Age 35-39	-1.50 (0.05)	0.30 (0.18)
	Age 40-44	-1.31 (0.05)	0.49 (0.17)
	Age 45-49	-1.17 (0.05)	0.51 (0.14)
	Age 50-54	-0.95 (0.05)	0.69 (0.12)
	Age 55-59	-0.88 (0.06)	0.39 (0.14)
	Age 60-64	-0.43 (0.06)	0.16 (0.14)
	Age 65 or more	(ref)	(ref)
Diagnoses	Cancer	0.55 (0.05)	0.46 (0.09)
	Diabetes	-0.11 (0.08)	0.41 (0.12)
	Cardio	-0.50 (0.04)	0.19 (0.08)
	Pulmonary	-0.60 (0.11)	1.11 (0.14)
	Renal	1.87 (0.14)	1.52 (0.08)
	Other	-0.43 (0.06)	0.88 (0.09)
	Healthy	(ref)	(ref)
Insurer	High-quality	1.07 (0.31)	—
Location	Rural	4.08 (0.04)	-0.21 (0.09)
	Urban	(ref)	(ref)
Income	Low	0.28 (0.03)	-1.72 (0.14)
	High	(ref)	(ref)
N		5,544,805	
N enrollees		500,000	
Pseudo-R^2		0.15	

Note: Table presents conditional logit model of insurer choice estimated by maximum likelihood on a random sample of 500,000 new enrollees. Includes insurer-by-market fixed effects. Robust standard errors in parenthesis.

Table 4: Average willingness-to-pay per service and diagnosis

Diagnosis	Cardiac care	Renal care	Imaging	General medicine	Laboratory	Hospital admissions
Cancer	3.78	3.78	2.20	1.16	1.80	3.50
Diabetes	3.93	3.93	2.41	1.33	2.00	3.67
Cardio	2.85	2.85	1.77	0.98	1.47	2.67
Pulmonary	6.20	6.20	3.21	1.60	2.57	5.62
Renal	27.24	27.25	12.46	5.83	9.72	24.04
Other disease	6.15	6.15	3.55	1.87	2.90	5.69
Healthy	1.00	1.00	1.00	1.00	1.00	1.00

Note: Table presents average willingness-to-pay for a percentage point increase in network breadth for the service in the column relative to healthy individuals. Willingness-to-pay is calculated as $\frac{1}{-\alpha_i}\frac{\partial s_{ijm}}{\partial H_{jkm}}$.

5.2 Insurer Average Costs Per Enrollee

I estimate equation (2) in the sample of new and current enrollees, conditional on observed choices in 2010 and 2011. Table 5 shows the results and appendix figure 11 presents the estimated consumer type fixed effects with their corresponding 95% confidence intervals. Average costs are increasing in network breadth and decreasing in the interaction between network breadth for different pairs of services. Insurer coverage decisions are thus characterized by economies of scope: a 1% increase in network breadth for service k reduces the average cost of providing service $l \neq k$ by 4.8% per enrollee.[21] Moreover, the estimate for τ_1 indicates that a 1% increase in network breadth raises average costs by 2.2% per enrollee.[22]

A potential mechanism for why insurers enjoy economies of scope from a bargaining perspective is that insurers enjoy price discounts when they cover several services. For example, if provider h is dropped from the network of laboratory testing, then demand for other diagnostic services like imaging is more likely to increase at lower-priced providers the broader is the network for imaging. This implies that the equilibrium price that provider h can charge to the insurer for laboratory testing is lower than it would be without the interaction with imaging providers.

[21] Calculated as the average of $100 \times \frac{1}{2K_m}\hat{\tau}_2 \sum_{l \neq k} q_{\theta k} q_{\theta l} H_{jlm}$

[22] Calculated as the average of $100 \times \hat{\tau}_1 q_{\theta k}$

Table 5: Insurer average costs per enrollee

Variable	Coefficient	Std. Error
Network breadth	0.44	(0.08)
Scope economies	-93.0	(45.0)
Reference price	40.9	(6.63)
Insurer		
EPS001	-0.02	(0.05)
EPS002	-0.16	(0.04)
EPS003	-0.14	(0.04)
EPS005	-0.24	(0.04)
EPS008	0.17	(0.05)
EPS009	0.20	(0.04)
EPS010	-0.06	(0.06)
EPS012	-0.02	(0.04)
EPS013	-0.13	(0.03)
EPS016	-0.01	(0.03)
EPS017	-0.11	(0.04)
EPS018	0.06	(0.06)
EPS023	-0.18	(0.04)
EPS037	(ref)	(ref)
N	8,662	
R^2	0.66	

Note: Table presents OLS regression of logarithm of average costs per consumer type on network breadth, economies of scope, and service reference prices. Includes insurer, market, and consumer type fixed effects. Robust standard errors in parenthesis.

Table 5 also shows substantial heterogeneity across insurers. EPS008, EPS009, and EPS018, have average costs per enrollee that are between 6% and 20% higher than the average cost of EPS037. My estimates fit the data for observed log average costs as seen in appendix figure 12, which suggests that an average cost function that is quadratic in network breadth is a good approximation to insurers' equilibrium log costs. Appendix tables 12 and 13 show that my model is robust to more granular definitions of consumer type and to explicitly modelling hospital quality with inclusion of a star hospital indicator. These exercises provide evidence of no relevant unobserved cost heterogeneity within consumer types.

Table 6: Model of insurer network formation costs

$\log(MVP_{jmk})$	coef	se
Network breadth	6.69	(0.40)
Insurer		
EPS001	-1.04	(0.31)
EPS002	-0.15	(0.28)
EPS003	-0.55	(0.28)
EPS005	-0.64	(0.27)
EPS008	0.09	(0.41)
EPS009	0.63	(0.31)
EPS010	0.71	(0.30)
EPS012	0.29	(0.36)
EPS013	0.52	(0.23)
EPS016	0.77	(0.23)
EPS017	-0.51	(0.29)
EPS018	-1.26	(0.38)
EPS023	0.07	(0.27)
EPS037	(ref)	(ref)
Constant	6.68	(0.25)
N	1,060	
R^2	0.35	

Note: Table presents OLS regression of log marginal variable profit on network breadth and insurer fixed effects with data from markets 25, 11, 76, 05, and 13. Robust standard errors in parenthesis.

5.3 Competition in Network Breadth

The third piece of the insurers' profit function left to estimate is the fixed cost, for which I use insurers' FOCs. Demand and average cost estimates allow me to compute MVPs in the left-hand side of equation (3). Dropout and transition probabilities are calculated off-line non-parametrically from the data. Appendices 9 and 10 present summary statistics of these probabilities and of MVPs, respectively. The fact that MVPs are positive for all insurer-services suggests a role for fixed costs in explaining the profit maximizing choices of network breadth.

Table 6 presents the results of equation (5) for the log of marginal variable profits. I find that fixed costs are increasing in network breadth and are substantially heterogeneous across insurers. The unobserved cost component explains nearly 65% of the variation in MVPs. This fixed cost function fits untargeted moments coming from insurers' public income statements, such as the ratio of total costs to total revenues, as seen in appendix figure 13.

Table 7: Decomposition of short-run variable profits

Insurer	Demand	Total avg. cost
EPS001	0.17	0.18
EPS002	0.28	0.31
EPS003	0.29	0.31
EPS005	0.22	0.24
EPS008	0.17	0.18
EPS010	0.37	0.38
EPS013	0.48	0.50
EPS016	0.40	0.42
EPS017	0.26	0.29
EPS018	0.15	0.16
EPS023	0.28	0.31
EPS037	0.20	0.22

Note: Table presents percentage change in demand and total average costs after the insurer in the row unilaterally increases network breadth for general medicine by 1 percent, while holding its rivals' choices fixed.

Magnitude of adverse selection and cost incentives. The heterogeneity in costs across insurers suggests that the decision to offer narrow (broad) service networks need not respond to selection incentives coming from demand but to services being associated with higher (lower) fixed and average costs. To see how important each of these components are for determining network breadth, I conduct a partial equilibrium exercise where I allow an insurer to unilaterally deviate and increase network breadth for general medicine by 1%. Table 7 presents the percentage change in short-run demand and total average costs from this exercise. Changes in demand or risk selection incentives explain 48% of the variation in insurers' total variable profits, while average cost incentives explain the remaining 52%.

6 Addressing Risk Selection

Risk selection and cost incentives weigh equally in insurers' decision to offer network breadth. In this section I use my model estimates to assess whether risk adjustment can affect provider networks through its impact on selection incentives. Risk adjustment has been evaluated extensively in settings like the US and on outcomes such as premiums and enrollment (e.g., Geruso and Layton, 2017; Brown et al., 2014), however there is no previous evidence of how network breadth responds to this policy.

I conduct two counterfactual simulations: first, I eliminate risk adjustment by imposing the same transfer across consumer types. Then, I improve the government's formula by compensating for a list of 14 health conditions listed in appendix table 17. In these analyses I hold long-run government spending fixed across all markets, so that changes in networks are determined only by changes in how resources are redistributed across insurers, but not by the level of the transfer itself. The effect of risk adjustment on network breadth is ambiguous under fixed government spending, and will depend on the relative magnitude of selection vis-à-vis cost incentives.

One concern in the counterfactual analyses is that the model may admit multiple equilibria. For instance, my measure of scope economies can make it such that every firm choosing complete networks or no coverage at all are both feasible equilibria. While a direct proof of uniqueness is challenging, in appendix 11 I provide suggestive evidence of uniqueness by computing the second partial derivative of the insurers' profit function with respect to network breadth, all else equal. Findings show that the rich preference and cost heterogeneity prevent multiple equilibria from arising. In computing the counterfactual analyses, I also use several different starting values for the vector of service network breadth to confirm that they all converge to the same equilibrium. For tractability, I conduct all my counterfactuals with data from the largest market, Bogotá, where 29% of all continuously enrolled individuals reside and where all private insurers compete.

Panel A of table 8 presents the percentage change in mean network breadth and long-run consumer surplus for sick and healthy individuals under no risk adjustment in column (1) and with improved risk adjustment in column (2). I find that without risk adjustment, mean network breadth falls 24% (roughly a reduction of 9 providers in the network of the average market). The reduction in network breadth generates large

Table 8: Networks, costs, and welfare under no risk adjustment

	Variable	No RA (1)	RA (2)
A. Overall	Mean network breadth	-23.79	21.70
	Consumer surplus (sick)	-2.97	-0.52
	Consumer surplus (healthy)	-3.14	-0.58
B. Service network breadth	Otorhinolaryngologic care	-24.42	22.78
	Cardiac care	-25.22	22.36
	Gastroenterologic care	-24.99	22.45
	Renal care	-26.70	23.67
	Gynecologic care	-25.74	23.14
	Orthopedic care	-25.42	22.58
	Imaging	-15.34	15.20
	General medicine	-9.97	9.77
	Laboratory	-13.49	13.65
	Hospital admission	-21.63	19.56

Note: Panel A presents the percentage change in mean network breadth and long-run consumer surplus for sick and healthy individuals, in the scenario without risk adjustment in column (1), and the scenario with improved risk adjustment in column (2). Panel B presents the percentage change in mean network breadth by service category.

welfare effects despite decreases in out-of-pocket costs. Eliminating risk adjustment results in a 3% decrease in long-run consumer surplus for individuals with and without chronic conditions (nearly a 2/3 reduction in the monthly minimum wage). Panel B shows that insurers reduce coverage of relatively expensive services by a greater magnitude than coverage of relatively cheap services. For instance, mean network breadth for hospital admissions decreases approximately 22% relative to the observed scenario, while mean network breadth for general medicine decreases 10%.

In column (2) I find qualitatively opposite results. With improved risk adjustment, mean network breadth increases 22%. Effects are larger for services that mostly sick patients claim, which is consistent with weakened selection incentives and with adverse selection being a determinant of narrow networks. Panel B shows that mean network breadth for cardiac care increases 22% (roughly equal to adding 5 providers to this service network in the average market), while mean network breadth for general medicine increases nearly 10%. Despite the substantial changes in network coverage, I find essentially no variation in consumer surplus. This is because consumers in this counterfactual make higher out-of-pocket payments relative to the observed scenario, which compensate the welfare gains from having greater coverage.

7 The Importance of Insurers' Cost Structure

The previous counterfactuals show that risk selection drives the use of narrow networks. However, my model provides two explanations for why insurers can choose to offer broad networks despite selection incentives. The first is that consumers on average prefer to have broad networks. Although willingness-to-pay for service network breadth is lower for healthy individuals relative to those with chronic diseases, it is not zero. The second is that insurers are sufficiently heterogeneous in their average and fixed costs. If some insurers enjoy economies of scope or have fixed costs that decline in network size, these insurers may have incentives to offer broad networks.

In this section I explore the importance of cost heterogeneity in producing broad networks by computing new market equilibria making costs homogeneous across insurers. This exercise is important to understand whether there is scope for commonly used network adequacy rules requiring insurers to have minimum hospital-to-enrollee ratios or forcing coverage of essential community providers.[23] I start by eliminating average cost heterogeneity imposing the median insurer fixed effect ($\bar{\delta}$) to all insurers. Then I remove fixed cost heterogeneity by assigning the median insurer-specific fixed cost component ($\bar{\xi}$) to every insurer.

Table 9 presents the results. The main takeaway is that, absent fixed cost heterogeneity, network breadth collapses, but average cost heterogeneity has very little impact on service network breadth. Column (2) shows that if insurers had homogeneous fixed costs, mean network breadth would decrease 7.6% relative to the observed scenario. Insurers' total average cost would increase 2.6% because they can no longer take advantage of scope economies. Consumer surplus for individuals with and without diagnoses would increase by a moderate amount, suggesting that welfare losses

[23] https://www.kff.org/health-reform/issue-brief/network-adequacy-standards-and-enforcement/

Table 9: Networks, costs, and welfare under homogeneous costs

	Variable	(1) Avg cost	(2) Fixed cost
A. Overall	Mean network breadth	0.17	-7.59
	Avg. cost per enrollee	1.22	0.40
	Total avg. cost	1.85	2.56
	Consumer surplus (sick)	0.92	1.64
	Consumer surplus (healthy)	0.60	1.32
B. Service network breadth	Otorhinolaryngologic care	0.20	-9.47
	Cardiac care	0.18	-8.32
	Gastroenterologic care	0.19	-8.70
	Renal care	0.19	-9.95
	Gynecologic care	0.19	-8.62
	Orthopedic care	0.18	-9.12
	Imaging	0.09	-5.04
	General medicine	0.09	-7.21
	Laboratory	0.04	-5.19
	Hospital admission	0.16	-6.53

Note: Panel A presents the percentage change in mean network breadth, insurer total average costs, short-run average cost per enrollee, and long-run consumer welfare for sick and healthy individuals, in the scenario with homogeneous average costs in column (1), and the scenario with homogeneous average and network formation costs in column (2). Insurer fixed effects in average costs and network formation costs are set to the median fixed effect. Panel B presents the percentage change in mean network breadth by service category.

due to lower network coverage are slightly overcompensated by welfare gains from lower out-of-pocket costs.

Panel B of column (2) shows that the reduction in network breadth is larger for services that mostly sick individuals tend to claim. Network breadth for general medicine decreases approximately 7%, while network breadth for renal care falls around 10% relative to the observed scenario (a reduction of nearly 2 providers from this service network in the average market). These results are robust to different ways of imposing cost homogeneity, such as using the average rather than the median insurer fixed effect as seen in appendix table 18.

8 Conclusions

Private health insurers respond to different incentives when crafting the various elements of their insurance contracts. This paper shows that risk selection and cost incentives are the main drivers of insurers' decision to offer provider network breadth. Risk selection induces insurers to offer narrow networks, while fixed cost heterogeneity induces insurers to offer broad networks despite selection incentives. I use a structural model of insurer competition in provider network breadth to decompose the relative importance of these incentives in counterfactuals. The empirical setting is Colombia, where the government regulates premiums and cost-sharing, and allows insurers to choose only which and how many providers to cover for each health service.

To quantify the equilibrium impact of risk selection on network breadth I modify the risk adjustment formula. Without risk adjustment, mean network breadth would decrease 24%, consistent with increased selection incentives. Instead, improving the risk adjustment formula by compensating for a granular list of diagnoses would increase mean network breadth by 22%. To quantify the equilibrium impact of cost incentives, I impose a homogeneous fixed cost structure across insurers. Results show that mean network breadth falls 7.6%, with reductions being larger in services that sick individuals require the most.

The findings of this paper provide new evidence of selection on provider networks and speak to the increasing use of network adequacy rules in markets where narrow-network plans have proliferated. When health systems have universal coverage, insurer competition and heterogeneity in insurers' costs structure can generate broad hospital networks. Hence maintaining healthy levels of competition is crucial to improve access to health care for those most in need.

References

Aizawa, N. and Kim, Y. (2018). Advertising and Risk Selection in Health Insurance Markets. *American Economic Review*, 108(3):828–867.

Brown, J., Duggan, M., Kuziemko, I., and Woolston, W. (2014). How Does Risk Selection Respond to Risk Adjustment? New Evidence from the Medicare Advantage Program. *American Economic Review*, 104(10):3335–3364.

Cabral, M., Geruso, M., and Mahoney, N. (2018). Do Larger Health Insurance Subsidies Benefit Patients or Producers? Evidence from Medicare Advantage. *American Economic Review*, 108(8):2048–2087.

Cao, Z. and McGuire, T. (2003). Service-Level Selection by HMOs in Medicare. *Journal of Health Economics*, 22(6):915–931.

Chiappori, P. and Salanie, B. (2000). Testing for Asymmetric Information in Insurance Markets. *Journal of Political Economy*, 108(1):56–78.

Crawford, G. S. and Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review*, 102(2):643–685.

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*, 105(5):110–14.

Dafny, L. S., Hendel, I., Marone, V., and Ody, C. (2017). Narrow networks on the health insurance marketplaces: Prevalence, pricing, and the cost of network breadth. *Health Affairs*, 36(9):1606–1614.

Decarolis, F. and Guglielmo, A. (2017). Insurers' response to selection risk: Evidence from Medicare enrollment reforms. *Journal of Health Economics*, 56(1):383–396.

Einav, L., Finkelstein, A., Ryan, P., Schrimpf, D., and Cullen, M. (2013). Selection on Moral Hazard in Health Insurance. *American Economic Review*, 103(1):178–219.

Ericson, K. M. and Starc, A. (2015). Measuring consumer valuation of limited provider networks. *American Economic Review*, 105(5):115–19.

Fleitas, S., Gowrisankaran, G., Starc, A., and Swanson, A. (2024). The power of exclusion: Pharmacy networks and bargaining in medicare part d.

Frank, R., Glazer, J., and McGuire, T. (2000). Measuring adverse selection in managed health care. *Journal of Health Economics*, 19(6):829–854.

Geruso, M. and Layton, T. (2017). Selection in Health Insurance Markets and Its Policy Remedies. *Journal of Economic Perspectives*, 31(4):23–50.

Geruso, M., Layton, T., and Prinz, D. (2019). Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges. *American Economic Journal: Applied Economics*, 11(2):64–107.

Ghili, S. (2022). Network formation and bargaining in vertical markets: The case of narrow networks in health insurance. *Marketing Science*, 41(3):501–527.

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.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review*, 103(1):145–77.

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):219–250.

Ho, K. (2006). The Welfare Effects of Restricted Hospital Choice in the US Medical Care Market. *Journal of Applied Econometrics*, 21(7):1039–1079.

Ho, K. 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):473–522.

Horn, H. and Wolinsky, A. (1988). Bilateral monopolies and incentives for merger. *The RAND Journal of Economics*, pages 408–419.

Lavetti, K. and Simon, K. (2018). Strategic Formulary Design in Medicare Part D Plans. *American Economic Journal: Economic Policy*, 10(3):154–92.

Liebman, E. (2022). Bargaining in Markets with Exclusion: An Analysis of Health Insurance Networks.

McFadden, D. (1996). Computing willingness-to-pay in random utility models. *University of California at Berkeley, Econometrics Laboratory Software Archive, Working Papers*.

McGuire, T., Glazer, J., Newhouse, J., Normand, S., Shi, J., Sinaiko, A., and Zuvekas, S. (2013). Integrating risk adjustment and enrollee premiums in health plan payment. *Journal of Health Economics*, 32(6):1263–1277.

McWilliams, J., Hsu, J., and Newhouse, J. (2012). New Risk-Adjustment System Was Associated With Reduced Favorable Selection In Medicare Advantage. *Health Affairs*, 31(12):2630–2640.

Newhouse, J. P., Price, M., McWilliams, J. M., Hsu, J., and McGuire, T. G. (2015). How much favorable selection is left in medicare advantage? *American journal of health economics*, 1(1):1–26.

Nicholson, S., Bundorf, K., Stein, R., and Polsky, D. (2004). The Magnitude and Nature of Risk Selection in Employer-Sponsored Health Plans. *Health Services Research*, 39(6):1817–1838.

Pakes, A., Porter, J., Ho, K., and Ishii, J. (2015). Moment inequalities and their application. *Econometrica*, 83(1):315–334.

Park, S., Basu, A., Coe, N., and Khalil, F. (2017). Service-level Selection: Strategic Risk Selection in Medicare Advantage in Response to Risk Adjustment. *NBER Working Paper*, (24038).

Pauly, M. and Herring, B. (2007). Risk Pooling and Regulation: Policy and Reality in Today's Individual Health Insurance Market. *Health Affairs*, 26(3):770–779.

Polsky, D., Cidav, Z., and Swanson, A. (2016). Marketplace plans with narrow physician networks feature lower monthly premiums than plans with larger networks. *Health Affairs*, 35(10):1842–1848.

Riascos, A., Alfonso, E., and Romero, M. (2014). The performance of risk adjustment models in colombian competitive health insurance market. *Documentos CEDE*, (33).

Riascos, A., Romero, M., and Serna, N. (2017). Risk Adjustment Revisited using Machine Learning Techniques. *Documentos CEDE*, (27).

Ruiz, F., Amaya, L., Garavito, L., and Ramírez, J. (2008). *Precios y Contratos en Salud: Estudio Indicativo de Precios y Análisis Cualitativo de Contratos*. Ministerio de la Protección Social.

Shapiro, B. T. (2020). Advertising in health insurance markets. *Marketing Science*, 39(3):587–611.

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

Starc, A. and Swanson, A. (2021). Preferred pharmacy networks and drug costs. *American Economic Journal: Economic Policy*, 13(3):406–46.

Tebaldi, P. (2024). Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA. *Forthcoming, Review of Economic Studies.*

Online Appendix

Appendix 1 Current risk adjustment system

For year t, the base un-adjusted capitated transfer is calculated using the claims data from year $t-2$. The per-enrollee transfer is roughly equal to the average annual health care cost in the population multiplied by a risk adjustment factor that is specific to a combination of sex, age group, and municipality. Appendix table 1 shows the national base transfer and appendix table 2 shows the risk adjustment multipliers.

Appendix Table 1: Base capitated transfer for the Contributory System during 2011

Department/city		Transfer
National (pesos)		525,492
Market multiplier a_m	Amazonas	$\times$ 1.10
	Arauca, Arauca	$\times$ 1.10
	Yopal, Casanare	$\times$ 1.10
	Florencia, Caquetá	$\times$ 1.10
	Chocó	$\times$ 1.10
	Riohacha, Guajira	$\times$ 1.10
	Guainía	$\times$ 1.10
	Guaviare	$\times$ 1.10
	Villavicencio, Meta	$\times$ 1.10
	Putumayo	$\times$ 1.10
	San Andrés y Providencia	$\times$ 1.10
	Sucre, Sincelejo	$\times$ 1.10
	Vaupés	$\times$ 1.10
	Vichada	$\times$ 1.10
	Soacha, Cundinamarca	$\times$ 1.06
	Bello, Antioquia	$\times$ 1.06
	Itaguí, Antioquia	$\times$ 1.06
	Envigado, Antioquia	$\times$ 1.06
	Sabaneta, Antioquia	$\times$ 1.06
	Soledad, Antioquia	$\times$ 1.06
	Bogotá	$\times$ 1.06
	Medellín, Antioquia	$\times$ 1.06
	Barranquilla, Atlántico	$\times$ 1.06

Note: Table reports national base risk-adjusted transfer which includes payments for promotion and prevention programs. Table also reports risk-adjustment multipliers for each market.

Appendix Table 2: Risk Adjustment Factors in the Contributory System during 2011

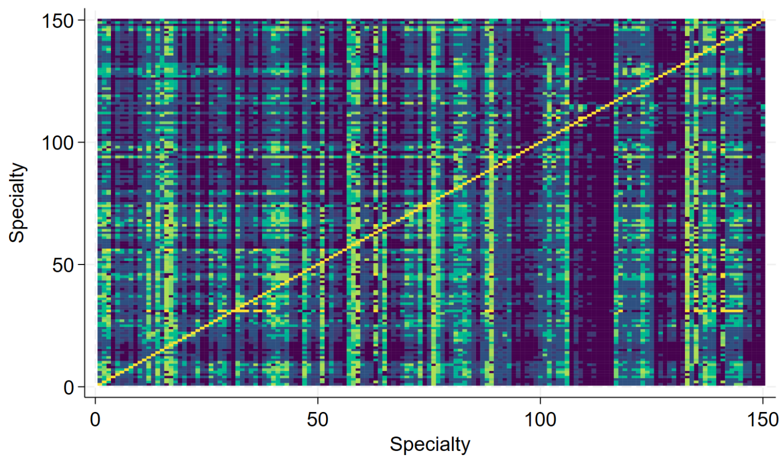
Age group	Sex	Multiplier
Less than 1	—	3.0000
1-4	—	0.9633
5-14	—	0.3365
15-18	M	0.3207
15-18	F	0.5068
19-44	M	0.5707
19-44	F	1.0588
45-49	—	1.0473
50-54	—	1.3358
55-59	—	1.6329
60-64	—	2.1015
65-69	—	2.6141
70-74	—	3.1369
More than 74	—	3.9419

Note: Table reports government risk-adjustment multipliers by sex and age group.

Appendix 2 Service categories

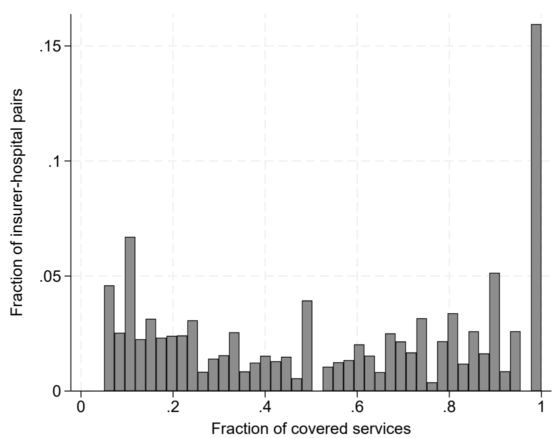
The service-level provider network data reports 150 unique specialities over which insurers and providers bargain. Some of these specialties are highly correlated in the sense that insurers tend to include them together at a particular provider. Appendix figure 1 presents a heatmap of the fraction of insurer-provider pairs that include the specialty in the horizontal axis, and also include the specialty in the vertical axis. Light colors represent higher fractions of insurer-provider pairs. The heatmap shows that (i) there are very common specialties such as general medicine and internal medicine seen in the vertical light-colored lines, and (ii) some specialties are correlated along the diagonal.

Appendix figure 2 shows that most insurers cover all the services at a particular provider, but there is still substantial variation in service coverage within insurer-provider pair. I group the different specialties of the network data into a final list 20 service categories, which can be mapped to the claims data based on the 6-digit service code reported for each claim. Appendix table 3 provides the final list of services and appendix table 4 provides a data excerpt for three hospitals and three services.



Appendix Figure 1: Heatmap of specialty pairs network inclusions

Note: Figure presents a heatmap of the fraction of insurer-hospital pairs in the network data that include the specialty in the horizontal axis and the specialty in the vertical axis. Lighter colors represent higher fractions.



Appendix Figure 2: Service inclusions within hospital

Note: Figure presents the distribution of the fraction of services that the provider can deliver which are covered by the insurer.

Appendix Table 3: List of services

Service code	Description
01	**Neurosurgery:** Procedures in skull, brain, and spine
02	**Other neurologic care:** Procedures in nerves and glands
03	**Otorhinolaryngologic care:** Procedures in face and trachea
04	**Pneumologic care:** Procedures in lungs and thorax
05	**Cardiac care:** Procedures in cardiac system
06	**Angiologic care:** Procedures in lymphatic system and bone marrow
07	**Gastroenterologic care:** Procedures in digestive system
08	**Hepatologic care:** Procedures in liver, pancreas, and abdominal wall
09	**Renal care:** Procedures in urinary system
10	**Gynecologic care:** Procedures in reproductive system
11	**Orthopedic care:** Procedures in bones and joints
12	**Other orthopedic care:** Procedures in tendons, muscles, and breast
13	**Diagnostic aid:** Diagnostic procedures in skin and subcutaneous cellular tissue
14	**Imaging:** Radiology and non-radiology imaging
15	**Internal and general medicine:** Consultations
16	**Laboratory:** Laboratory and blood bank
17	**Nuclear medicine:** Nuclear medicine and radiotherapy
18	**Rehab and mental health:** Rehabilitation, mental health care, therapy
19	**Therapy (chemo and dialysis):** Prophylactic and therapeutic procedures
20	**Hospital admissions:** Inpatient services

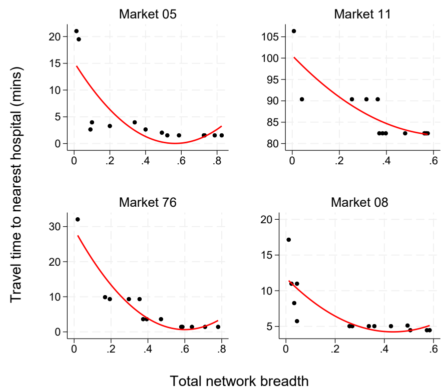
Note: Table presents the final list of 20 services and their description.

Appendix Table 4: Service coverage at hospitals

	Cardiac care			Renal care			Hospital admissions		
Insurer	Valle del Lili	Santa Fe	Pablo Tobón	Valle del Lili	Santa Fe	Pablo Tobón	Valle del Lili	Santa Fe	Pablo Tobón
EPS001	1	0	0	1	0	0	1	1	1
EPS002	1	0	1	1	0	1	1	1	1
EPS003	1	0	0	0	0	0	1	1	1
EPS005	1	1	1	1	1	1	1	1	1
EPS008	1	1	0	1	1	0	1	1	0
EPS009	0	0	1	0	0	1	0	0	1
EPS010	1	1	1	1	1	1	1	1	1
EPS012	1	1	0	1	1	0	1	1	0
EPS013	1	0	1	1	0	1	1	1	1
EPS016	1	1	1	1	1	1	1	1	1
EPS017	0	1	0	0	1	0	1	1	1
EPS018	1	1	1	1	1	1	1	1	1
EPS023	0	0	0	0	0	0	1	1	1
EPS037	1	1	1	1	1	1	1	1	1

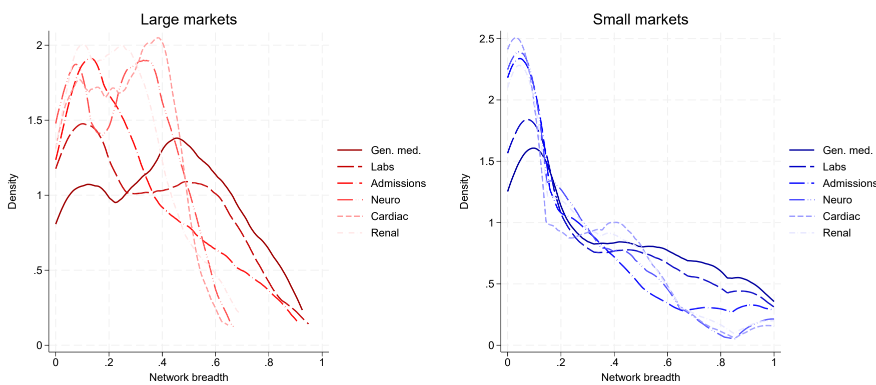
Note: Table presents service coverage per insurer at three hospitals in the country and for three services. Data comes from the National Health Superintendency.

Appendix 3 Correlates of Network Breadth



Appendix Figure 3: Average network breadth and travel times

Note: Figure presents a scatter plot of network breadth and travel time from the municipality centroid to the nearest in-network provider in minutes. The red line represents a quadratic fit.



Appendix Figure 4: Average network breadth and hospital deaths

Note: Figure presents kernel density estimates of the distribution of network breadth for general medicine, laboratory testing, hospital admissions, neurological care, cardiac care, and renal care in the 5 largest markets in the left panel and in the rest of markets in the right panel.

Appendix Table 5: Determinants of switching

	(1)	(2)	(3)	(4)
Sample:	Healthy	Cancer	Diabetes	Cardio
Service:	General medicine	Therapy	Laboratory	Cardiac care
Network breadth	7.21	-4.79	-1.34	-2.62
	(0.11)	(0.23)	(0.25)	(0.19)
Controls				
Demographics	x	x	x	x
Days enrolled	x	x	x	x
Market FE	x	x	x	x
N	10,703,261	771,447	346,022	1,723,168

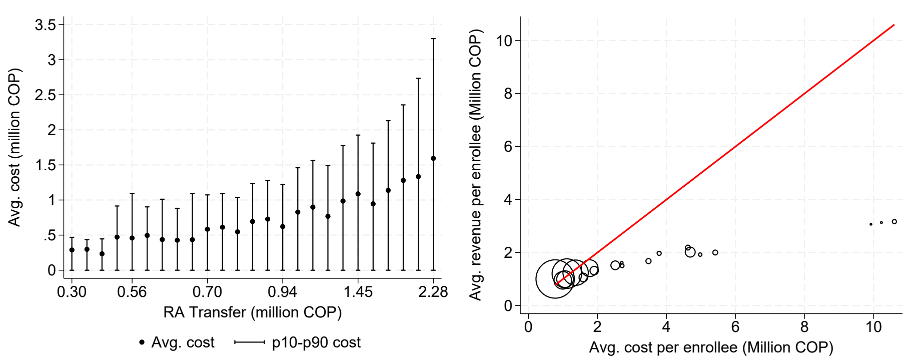
Note: Table presents OLS regression of a switching indicator on service network breadth for the 2010 insurer. Column (1) uses the sub-sample of individuals without diagnoses and network breadth for general medicine. Column (2) uses the sub-sample of individuals with cancer and network breadth for chemotherapy. Column (3) uses the sub-sample of individuals with diabetes and network breadth for laboratory. Column (4) uses the sub-sample of individuals with cardiovascular disease and network breadth for cardiac care services. All specifications control for enrollees' demographic characteristics, days enrolled, and market fixed effects. Robust standard errors in parenthesis.

Appendix 4 Descriptives in subsample

Appendix Table 6: Summary statistics in estimation sample

	mean	sd
Demographics		
Male	50.85	(49.99)
Age	41.70	(15.29)
Low income (%)	28.52	(45.15)
Rural (%)	24.73	(43.14)
Diagnoses (%)		
Cancer	5.55	(22.90)
Diabetes	1.77	(13.17)
Cardiovascular disease	9.55	(29.40)
Long-term pulmonary disease	0.99	(9.92)
Renal disease	0.64	(7.99)
Other disease	3.77	(19.04)
Health care utilization		
OOP spending	0.14	(0.12)
Weighted network breadth	0.51	(0.15)
Risk-adjusted transfer	0.67	(0.42)
Total health care cost	0.36	(2.26)

Note: Table presents mean and standard deviations in parenthesis of main analysis variables conditional on observed choices. OOP spending, risk-adjusted transfer, and total health care cost are measured in millions of COP.

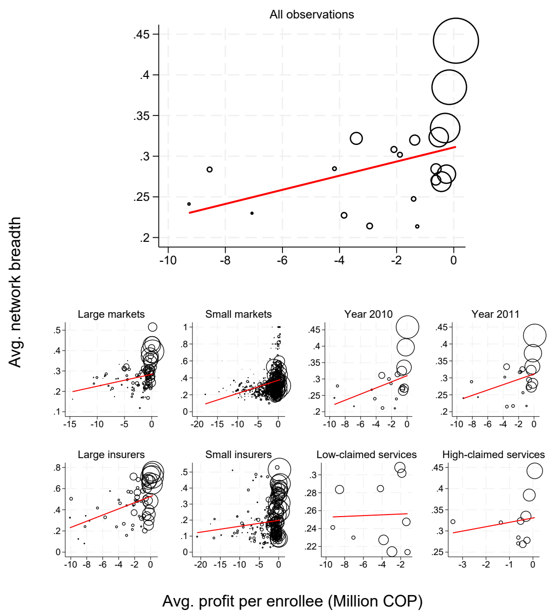


(a) Health care cost by risk-adjusted transfer

(b) Service-level selection incentives

Appendix Figure 5: Costs and selection incentives in the continuously enrolled

Note: Panel (a) of the figure presents mean, and 10th and 90th percentiles of annual health care cost by ex-ante government's risk-adjusted transfer in the sample of continuously enrolled. Panel (b) presents a scatter plot of average cost per enrollee against average revenue per enrollee in the sample of continuously enrolled. Each dot is a service weighted by the number of individuals who make claims for the service. The red line is a 45 degree line. One enrollee can be represented in several dots if she makes claims for different services. Enrollees who make zero claims are not represented in this figure.



Appendix Figure 6: Network breadth and service profitability in the continuously enrolled

Note: Figure presents a scatter plot of average service network breadth against average profit per enrollee in the sample of continuously enrolled. Each dot is a service weighted by the number of individuals who make claims for the service. Profits are calculated as government ex-ante and ex-post transfers, plus revenues from copays and coinsurance rates, minus total health care costs. The red line corresponds to a linear fit. One enrollee can be represented in several dots if she makes claims for different services. Enrollees who make zero claims are not represented in this figure.

Appendix 5 Micro-foundation

Insurer demand. For the demand side, take one market and consider a simple model of provider choice where individual i's indirect utility from choosing provider h for service k in the network of insurer j is:

$$u_{ijkh} = \overline{\xi_k} H_{jk} + \nu_{ijkh}$$

This model assumes that providers have identical quality conditional on the service which is equal to $\overline{\xi}_k$ weighted by the fraction of covered providers H_{jk}. Moreover, ν_{ijkh} is a preference shock distributed T1EV. Following McFadden (1996), individual i's value for insurer j's network of providers in service k, G_{jk}, is:

$$w_{ijk} = \log\Big(\sum_{h \in G_{jk}} \exp(\overline{\xi}_k H_{jk})\Big)$$

which simplifies to:

$$w_{ijk} = \log\Big(\sum_{h \in G_{jk}} \exp(\overline{\xi}_k H_{jk})\Big) = \log(|G_{jk}| \exp(\overline{\xi}_k H_{jk})) = \log(|G_{jk}|) + \overline{\xi}_k H_{jk} = \phi_{jk} + \overline{\xi}_k H_{jk}$$

where $|G_{jk}|$ is the number of providers in insurer j's network for service k and $\log(|G_{jk}|) = \phi_{jk}$. Summing across services yields:

$$\sum_k w_{ijk} = \phi_j + \sum_k \overline{\xi}_k H_{jk}$$

where $\phi_j = \sum_k \phi_{jk}$. This shows that insurer demand can be modelled as a function of $\sum_k \overline{\xi}_k H_{jk}$ and insurer fixed effects ϕ_j. It also shows that this approximation is correct under the assumption that providers are homogeneous conditional on the service.

The relation between network valuation and network breadth can be extended to a model where providers differ in quality and where consumers have heterogeneous preferences as follows. Suppose the utility function is:

$$u_{ijkh} = x_{\theta(i)}\xi_{hk} + \varepsilon_{ijkh}$$

where $x_{\theta(i)}$ is a vector of observed consumer characteristics describing a consumer type θ. Let γ_θ be the fraction of consumers type θ in the population and $|G_k|$ the total number of providers that deliver service k. Then:

$$\begin{aligned}
\sum_\theta \gamma_\theta w_{\theta(i)jk} &= \sum_\theta \gamma_\theta \log\Big(\sum_{h\in G_{jk}} \exp(x_{\theta(i)}\xi_{hk})\Big) \geq \sum_\theta \gamma_\theta \log\Big(\frac{1}{|G_k|}\sum_{h\in G_{jk}} \exp(x_{\theta(i)}\xi_{hk})\Big) \\
&\geq \sum_\theta \gamma_\theta \frac{1}{|G_k|}\sum_{h\in G_{jk}} \log(\exp(x_{\theta(i)}\xi_{hk})) = \sum_\theta \gamma_\theta \frac{1}{|G_k|}\sum_{h\in G_{jk}} x_{\theta(i)}\xi_{hk} \\
&= \sum_\theta \gamma_\theta \frac{|G_{jk}|}{|G_k|}\sum_{h\in G_{jk}} \frac{1}{|G_{jk}|} x_{\theta(i)}\xi_{hk} = \sum_\theta \gamma_\theta x_{\theta(i)}\overline{\xi}_{jk} H_{jk}
\end{aligned}$$

where the second inequality follows from Jensen's inequality and $\overline{\xi}_{jk} = \frac{1}{|G_{jk}|}\sum_{h\in G_{jk}} \xi_{hk}$ is the average quality of the hospitals in insurer j's network. This derivation suggests that when providers differ in quality conditional on the service and when consumers have heterogeneous preferences, a model of insurer demand defined over $\gamma_\theta x_{\theta(i)}\overline{\xi}_{jk}H_{jk}$ will be a lower bound of the demand function defined over $\gamma_\theta w_{\theta(i)jk}$.

Insurer costs. Moving to the supply side, suppose that insurer j and provider h engage in bilateral negotiations over service prices. Let $D_j(\cdot)$ be insurer j's demand, R the per-capita risk-adjusted transfer, $D_{jhk}(\cdot)$ provider h's demand for service k from j's enrollees, p_{jhk} the negotiated price, m_{hk} provider h's marginal cost of providing service k, H_{jk} the set of providers in insurer j's network for service k, and J_{hk} the set of insurers that cover provider h for service k. Insurer profits can be written as $\pi^j = D_j(\cdot)R - \sum_k \sum_{h\in H_{jk}} D_{jhk}(\cdot)p_{jhk}$ and provider profits as $\pi^h = \sum_k \sum_{j\in J_{hk}} D_{jhk}(\cdot)(p_{jhk} - m_{hk})$. Following the previous demand specification, suppose that $D_{jhk}(\cdot)$ does not depend on prices.

The log Nash surplus function is:

$$S_{jhk} = \beta \log(\pi_j - t_h^j) + (1-\beta)\log(\pi_h - t_j^h)$$

where β is the bargaining power of the insurer, and t_h^j and t_j^h are the insurer and provider disagreement payoffs, respectively. The insurer disagreement payoff is defined as the profit it would enjoy if it excludes provider h from the network, while reimbursing the rest of providers at their equilibrium prices. Provider disagreement payoffs are defined analogously. The FOC of the log Nash surplus function with respect to the negotiated price is:

$$\begin{aligned}\sum_k D_{jhk}p_{jhk} = \beta\Big(\sum_k D_{jhk}m_{hk} - \sum_k \sum_{n\in J_h\setminus j} \Delta D_{nhk}(\cdot)(p_{nhk} - m_{hk})\Big) \\ + (1-\beta)\Big(\Delta D_j(\cdot)R - \sum_k \sum_{l\in H_j\setminus h} \Delta D_{jlk}(\cdot)p_{jlk}\Big)\end{aligned}$$

Adding these FOCs across all providers in the market for service k, imposing symmetry across providers, and dividing on both sides by insurer j's demand, yields the following expression for the insurer's average cost per enrollee:

$$\begin{aligned}AC_j &= \frac{1}{D_j}\beta\Big(\sum_k \overline{D}_{jk}(\cdot)\overline{m}_k\ H_{jk} - \sum_k \Delta\overline{D}_{nk}(\cdot)(\overline{p}_{nk} - \overline{m}_k)\ (|J_k|-1)\ H_{jk}\Big) \\ &+ \frac{1}{D_j}(1-\beta)\Big(\frac{\Delta D_j(\cdot)}{|G_{jk}|}R\ H_{jk} - \sum_k \Delta\overline{D}_{jk}(\cdot)\overline{p}_{jk}\ (|G_{jk}|-1)\ H_{jk}\Big) \\ &= \frac{1}{D_j}\sum_k \Big(\beta\overline{D}_{jk}(\cdot)\overline{m}_k - \beta\Delta\overline{D}_{nk}(\cdot)(\overline{p}_{nk} - \overline{m}_k)(|J_k|-1)\Big)H_{jk} \\ &+ \Big(\frac{1}{D_j}(1-\beta)\frac{\Delta D_j(\cdot)}{|G_{jk}|}R\Big)H_{jk} + \frac{1}{D_j}(1-\beta)\sum_k \Big(\Delta\overline{D}_{jk}(\cdot)\overline{p}_{jk}\Big)H_{jk} \\ &- \frac{1}{D_j}(1-\beta)\sum_k \Big(\Delta\overline{D}_{jk}(\cdot)\overline{p}_{jk}\ |G_k|\Big)H_{jk}^2 \\ &= f(H_{jk}, H_{jk}^2)\end{aligned}$$

Here variables with over-lines denote the value for the average provider in service k. This derivation shows that an average cost function that is quadratic in network breadth is a correct simplification when providers are homogeneous conditional on the service. Hence, together with results on demand, my proposed model is internally consistent.

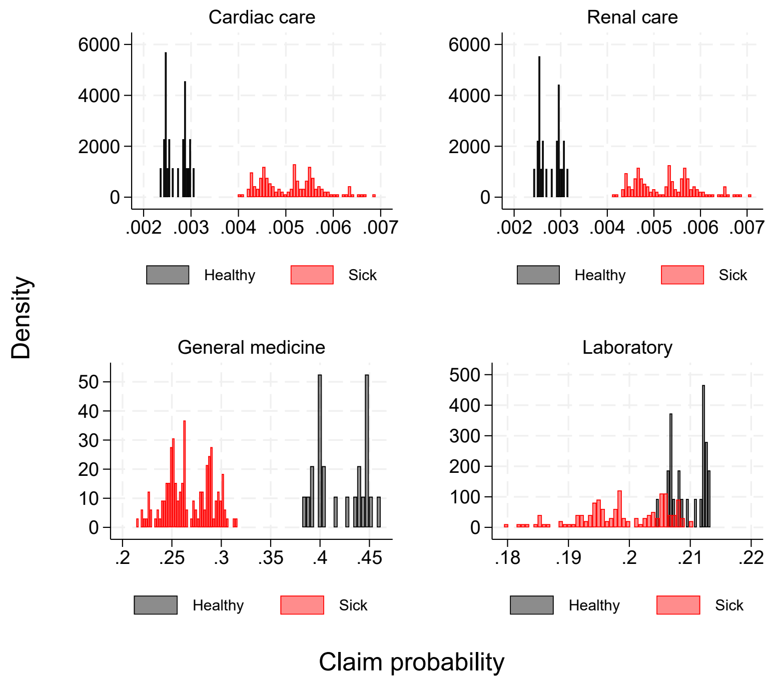
Appendix 6 Additional demand results

6.1 Estimating claim probabilities

I estimate the claim probabilities using the following logistic regression:

$$\text{logit}(\text{any claims})_{ik} = \psi_k + \psi_\theta + \psi_{ik} \tag{6}$$

The dependent variable is an indicator for whether patient i makes a claim for service k. On the right side, ψ_k and ψ_θ are service and consumer type fixed effects, respectively. ψ_{ikm} is a mean zero shock to the claim probability that is independent of network breadth conditional on consumer observables. I assume that new enrollees' expectations over the services they will need are correct on average, and that these expectations do not depend on the insurer they enroll with. I estimate equation (6) on data from both current and new enrollees in 2010 and 2011. Appendix figure 7 presents the resulting distribution of $q_{\theta k}$ for a few services such as cardiac care, renal care, general medicine, and laboratory.



Appendix Figure 7: Distribution of service claim probability

Note: Figure presents the distribution of the probability of making a claim in a sample of service categories separately for sick and healthy individuals. Services reported in the figure include cardiac care, renal care, general medicine, and laboratory.

6.2 In-sample demand model fit

This appendix shows the observed and predicted market shares in the country and in the three largest markets for every insurer, as a measure of the in-sample demand model fit.

Appendix Table 7: National market shares

Insurer	Observed	Predicted
EPS001	2.42	2.42
EPS002	7.50	7.45
EPS003	4.69	4.70
EPS005	6.19	6.23
EPS008	6.56	6.59
EPS009	2.42	2.45
EPS010	9.48	9.55
EPS012	2.12	2.09
EPS013	10.62	10.66
EPS016	15.23	15.19
EPS017	9.58	9.53
EPS018	4.66	4.65
EPS023	3.97	3.96
EPS037	14.57	14.54

Note: Table presents observed and model-predicted national market shares.

Appendix Table 8: Market shares in three largest markets

	Market 05		Market 11		Market 76	
Insurer	Obs	Pred	Obs	Pred	Obs	Pred
EPS001	0.81	0.82	4.29	4.28	1.14	1.12
EPS002	5.25	5.20	9.46	9.45	2.91	2.99
EPS003	3.21	3.22	8.18	8.16	0.82	0.82
EPS005	1.37	1.39	11.29	11.37	2.62	2.64
EPS008	—	—	14.72	14.78	—	—
EPS009	9.44	9.55	—	—	—	—
EPS010	26.79	27.06	3.26	3.23	4.39	4.43
EPS012	—	—	—	—	10.83	10.67
EPS013	11.51	11.48	9.15	9.21	7.58	7.68
EPS016	24.91	24.72	3.79	3.77	27.57	27.65
EPS017	—	—	16.72	16.64	—	—
EPS018	—	—	0.15	0.16	23.45	23.40
EPS023	2.29	2.31	6.63	6.59	1.85	1.81
EPS037	14.42	14.26	12.36	12.38	16.83	16.79

Note: Table presents observed and model-predicted market shares in the three largest markets.

6.3 Robustness checks

Appendix Table 9: Insurer demand with star hospital indicator

Variable		Network Breadth	OOP spending	Star hospital
Mean		2.32 (0.42)	-2.42 (0.11)	0.67 (0.45)
Interactions				
Demographics	Male	0.15 (0.02)	0.06 (0.07)	
	Age 19-24	-0.60 (0.05)	1.51 (0.12)	
	Age 25-29	-1.19 (0.05)	0.70 (0.12)	
	Age 30-34	-1.46 (0.05)	0.56 (0.15)	
	Age 35-39	-1.50 (0.05)	0.31 (0.18)	
	Age 40-44	-1.31 (0.05)	0.49 (0.17)	
	Age 45-49	-1.17 (0.05)	0.51 (0.14)	
	Age 50-54	-0.95 (0.05)	0.69 (0.12)	
	Age 55-59	-0.88 (0.06)	0.39 (0.14)	
	Age 60-64	-0.42 (0.06)	0.16 (0.14)	
	Age 65 or more	(ref)	(ref)	
Diagnoses	Cancer	0.54 (0.05)	0.46 (0.09)	
	Diabetes	-0.11 (0.08)	0.41 (0.12)	
	Cardio	-0.51 (0.04)	0.19 (0.08)	
	Pulmonary	-0.61 (0.11)	1.11 (0.14)	
	Renal	1.87 (0.14)	1.53 (0.08)	
	Other	-0.44 (0.06)	0.88 (0.09)	
	Healthy	(ref)	(ref)	
Insurer	High-quality	1.08 (0.31)	—	
Location	Rural	4.08 (0.04)	-0.21 (0.09)	
	Urban	(ref)	(ref)	
Income	Low	0.28 (0.03)	-1.72 (0.14)	
	High	(ref)	(ref)	
N			5,544,805	
N enrollees			500,000	
Pseudo-R^2			0.15	

Note: Table presents insurer choice model including a measure of star hospital coverage equal to $\sum_k q_{\theta k} Star_{jkm}$, where $Star_{jkm}$ is an indicator for insurer j covering a star hospital in market m for service k. Specification includes insurer-by-market fixed effects. Robust standard errors in parenthesis.

Appendix Table 10: Insurer demand with diagnosis in January

Variable		Network Breadth	OOP spending
Mean		1.67 (0.41)	-1.44 (0.1)
Interactions			
Demographics	Male	0.11 (0.02)	0.06 (0.07)
	Age 19-24	-0.52 (0.05)	1.63 (0.12)
	Age 25-29	-1.12 (0.05)	0.56 (0.14)
	Age 30-34	-1.38 (0.05)	0.46 (0.15)
	Age 35-39	-1.43 (0.05)	0.30 (0.19)
	Age 40-44	-1.24 (0.05)	0.64 (0.18)
	Age 45-49	-1.11 (0.05)	0.64 (0.16)
	Age 50-54	-0.92 (0.05)	0.81 (0.14)
	Age 55-59	-0.87 (0.06)	0.46 (0.15)
	Age 60-64	-0.43 (0.06)	0.01 (0.15)
	Age 65 or more	(ref)	(ref)
Diagnoses	Cancer	0.32 (0.13)	0.07 (0.23)
	Diabetes	-0.02 (0.16)	0.94 (0.24)
	Cardio	0.09 (0.07)	0.32 (0.16)
	Pulmonary	-1.31 (0.28)	1.52 (0.24)
	Renal	0.97 (0.39)	1.29 (0.11)
	Other	0.00 (0.14)	0.46 (0.19)
	Healthy	(ref)	(ref)
Insurer	High-quality	1.23 (0.31)	—
Location	Rural	4.08 (0.04)	-0.02 (0.1)
	Urban	(ref)	(ref)
Income	Low	0.26 (0.03)	-1.83 (0.16)
	High	(ref)	(ref)
N		5,544,805	
N enrollees		500,000	
Pseudo-R^2		0.15	

Note: Table presents insurer choice model defining diagnoses based on claims made in January. Specification includes insurer-by-market fixed effects. Robust standard errors in parenthesis.

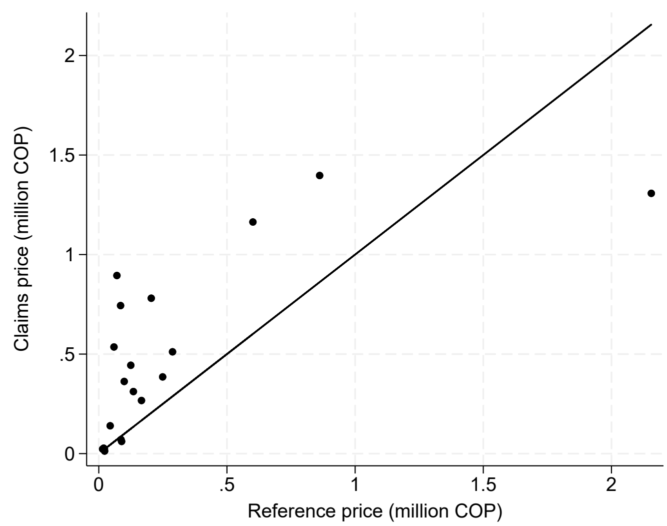
Appendix Table 11: Insurer demand with expectations over diagnoses

Variable		Network Breadth	OOP spending
Mean		27.11 (0.73)	-5.12 (0.18)
Interactions			
Demographics	Male	0.54 (0.02)	-0.11 (0.17)
	Age 19-24	0.53 (0.05)	10.43 (0.32)
	Age 25-29	-0.23 (0.05)	0.79 (0.24)
	Age 30-34	-0.53 (0.05)	-0.96 (0.32)
	Age 35-39	-0.59 (0.05)	-1.38 (0.41)
	Age 40-44	-0.49 (0.05)	-0.68 (0.46)
	Age 45-49	-0.49 (0.05)	1.99 (0.36)
	Age 50-54	-0.36 (0.05)	2.68 (0.32)
	Age 55-59	-0.5 (0.06)	0.75 (0.3)
	Age 60-64	-0.24 (0.06)	0.02 (0.24)
	Age 65 or more	(ref)	(ref)
Insurer	High-quality	-6.97 (0.52)	—
Location	Rural	4.06 (0.04)	-1.45 (0.24)
	Urban	(ref)	(ref)
Income	Low	0.17 (0.03)	-6.13 (0.24)
	High	(ref)	(ref)
N		5,544,805	
N enrollees		500,000	
Pseudo-R^2		0.15	

Note: Table presents insurer choice model where consumers have expectation over diagnoses and services. Specification includes insurer-by-market fixed effects. Robust standard errors in parenthesis.

Appendix 7 Service reference prices

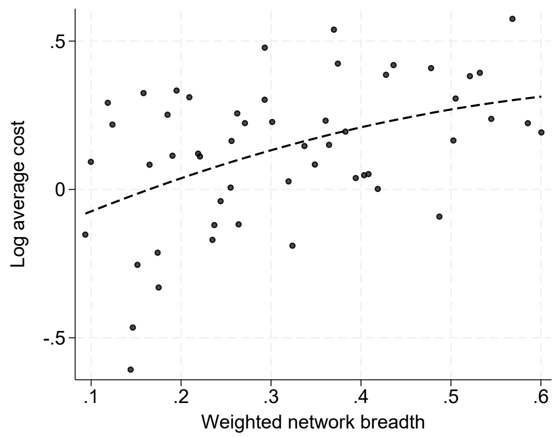
In 2005, the Colombian government published a list of reference prices for all the services included in the national health insurance plan. The list was created by a group of government officials and medical experts with the purpose of reimbursing healthcare providers in the event of terrorist attacks, natural disasters, and car accidents (Decree 2423 of 1996). Although they were not meant to guide price negotiations between insurers and providers, there is evidence that insurers use these reference prices as starting points in their negotiations with providers (Ruiz et al., 2008). Appendix figure 8 shows that references prices are highly correlated with negotiated prices from the claims data.



Appendix Figure 8: Negotiated prices and reference prices

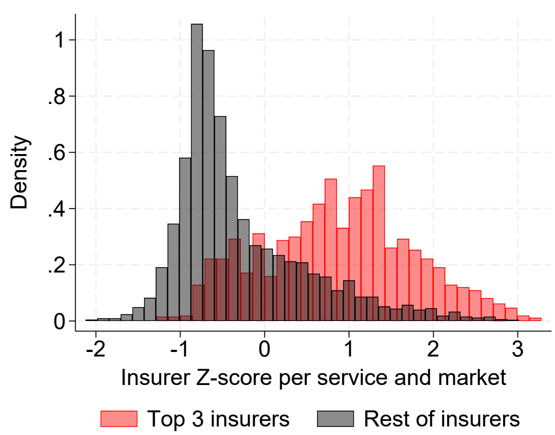
Note: Figure presents a scatter plot of average negotiated price obtained from the claims data and average reference price per service. The black line is a 45 degree line.

Appendix 8 Additional average cost results



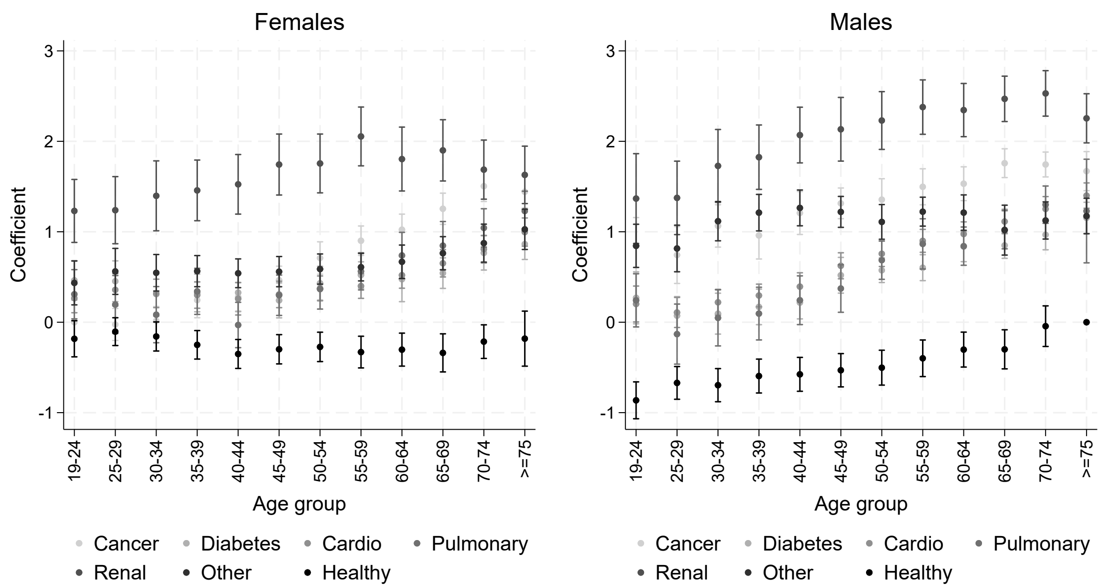
Appendix Figure 9: Empirical relation between average costs and network breadth

Note: Figure presents a scatter plot of the log of average costs per enrollee by average network breadth across services. The dashed line corresponds to a quadratic fit.



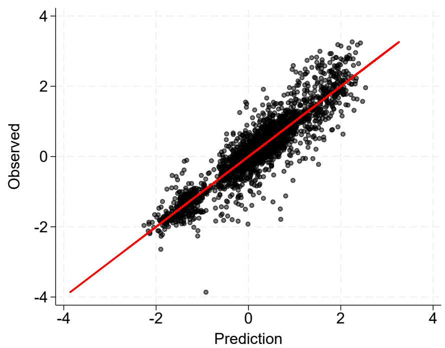
Appendix Figure 10: Standardized network breadth per service and market

Note: Figure presents distribution of network breadth standardized within service and market. The red histogram corresponds to the three largest insurers (EPS013, EPS016, and EPS037). The black histogram corresponds to the rest of insurers. Standardized values of network breadth are obtained by subtracting the service-market mean and dividing by the service-market standard deviation. The top 3 insurers have consistently broad networks across services, while the rest tend to have narrow networks across services.

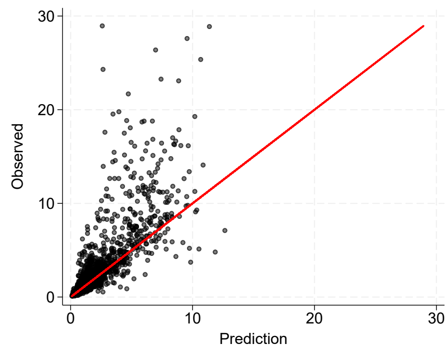


Appendix Figure 11: Consumer type fixed effects

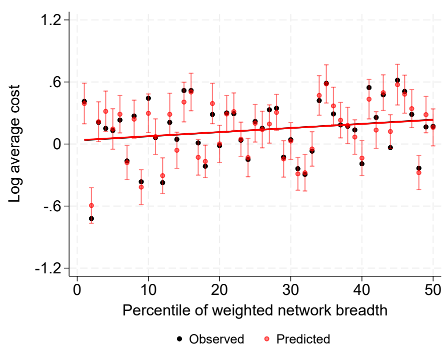
Note: Figure presents point estimate and 95 percent confidence interval of the consumer type fixed effects in the average cost function. The left panel presents fixed effects for females separately by disease category and age group. The right panel presents fixed effects for males separately by disease category and age group.



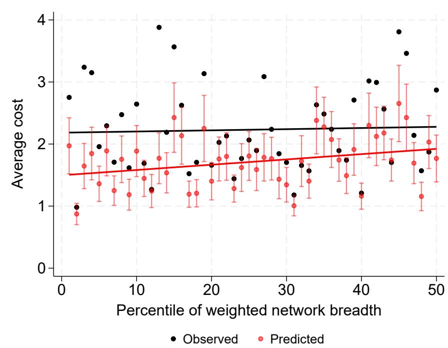
(a) Fit on observed log costs



(b) Fit on observed costs



(c) Log costs and network breadth



(d) Costs and network breadth

Appendix Figure 12: Average cost model fit

Note: Panel A presents a scatter plot of observed log average costs against predicted log average costs per insurer and consumer type. Panel B presents a scatter plot of observed and predicted average costs in millions of COP. The red line in panels A and B is the 45 degree line. Panel C presents a scatter plot of observed and predicted log average cost by percentile of $\sum_k q_{\theta k} H_{jkm}$. Panel D presents a scatter plot of exponentiated observed and predicted log average cost by percentile of $\sum_k q_{\theta k} H_{jkm}$.

Appendix Table 12: Patient-level estimates of average cost

log(cost+1)	coef	se
Network breadth	2.63	(0.06)
Scope economies	-146.53	(2.76)
Reference price	—	—
Insurer		
EPS001	-1.65	(0.01)
EPS002	0.50	(0.01)
EPS003	0.82	(0.01)
EPS005	2.06	(0.01)
EPS008	1.83	(0.01)
EPS009	0.98	(0.01)
EPS010	0.81	(0.01)
EPS012	1.31	(0.01)
EPS013	1.37	(0.01)
EPS016	0.21	(0.01)
EPS017	1.41	(0.01)
EPS018	1.32	(0.01)
EPS023	1.70	(0.01)
EPS037	(ref)	(ref)
N	9,976,897	
R^2	0.24	

Note: Table presents OLS regression of log health care cost (plus 1) per patient on network breadth, economies of scope, and service reference price. Uses a random sample of 500,000 patients. Includes insurer, market, and consumer type fixed effects. Reference price omitted due to multicollinearity. Robust standard errors in parenthesis.

Appendix Table 13: Average cost with star hospitals

Variable	coef	se
Network breadth	0.39	(0.09)
Star hospital	0.39	(0.19)
Scope economies	-105.52	(45.49)
Reference price	41.01	(6.64)
Insurer		
EPS001	-0.04	(0.05)
EPS002	-0.18	(0.04)
EPS003	-0.12	(0.04)
EPS005	-0.25	(0.04)
EPS008	0.15	(0.05)
EPS009	0.20	(0.04)
EPS010	-0.08	(0.06)
EPS012	-0.04	(0.04)
EPS013	-0.13	(0.03)
EPS016	-0.02	(0.03)
EPS017	-0.11	(0.04)
EPS018	0.05	(0.06)
EPS023	-0.16	(0.04)
EPS037	(ref)	(ref)
N	8,662	
R^2	0.66	

Note: Table presents OLS regression of log average cost per consumer type excluding the capital city, on network breadth, scope economies, and service reference prices. Includes insurer, market, and consumer type fixed effects. Robust standard errors in parenthesis.

Appendix 9 Dropout and transition probabilities

To estimate the marginal cost of network formation in the third step of my model, I first need to compute the probability that consumer type θ drops out of the contributory system and the probability that consumer type θ in period t transitions into θ' in period $t+1$. I use the data from *all* enrollees to the contributory system in 2010 and 2011, regardless of their enrollment spell length, to compute dropout probabilities. For each consumer type θ, I calculate the probability that she drops out of the system non-parametrically as the number of individuals of type θ observed only in 2010 but not 2011, divided by the total number of type θ individuals in 2010. Appendix table 14 presents the mean and standard deviation of the dropout probability conditional on diagnoses, sex, and age.

I use a non-parametric approach to compute transition probabilities as well, using data from continuously enrolled new *and* current enrollees in 2010 and 2011. The probability that type θ transitions into θ' equals the number of type θ in 2010 that end up with diagnosis l' in 2011, divided by the number of type θ individuals in 2010. Appendix table 15 presents the mean and standard deviation in parenthesis of transition probabilities from having cancer, cardiovascular disease, diabetes, renal disease, other diseases, 2 or more diseases, and no diseases in period t to having each of these 9 diagnoses in period $t+1$.

Appendix Table 14: Dropout probability

	mean	sd
Diagnosis		
Cancer	4.79	(2.40)
Diabetes	2.75	(0.83)
Cardio	2.79	(0.90)
Pulmonary	4.04	(1.51)
Renal	4.42	(1.79)
Other	2.62	(1.11)
Healthy	45.00	(7.29)
Age group		
19-24	12.00	(17.73)
25-29	8.72	(13.36)
30-34	8.13	(13.47)
35-39	8.47	(14.07)
40-44	8.47	(14.59)
45-49	8.51	(14.93)
50-54	8.88	(15.32)
55-59	9.09	(15.77)
60-64	9.20	(15.84)
65-69	9.63	(15.93)
70-74	10.37	(15.95)
75 or more	12.38	(16.43)
Sex		
Female	8.42	(13.07)
Male	10.55	(16.50)

Note: Mean and standard deviation in parenthesis of dropout probabilities conditional on diagnosis in the first panel. age group in the second panel, and sex in the third panel.

Appendix Table 15: Transition probabilities

Diagnosis	Cancer	Cardio	Diabetes	Renal	Pulmonary	Other	Healthy
Cancer	31.6	1.7	13.9	1.4	0.7	4.7	46.0
	(6.7)	(1.4)	(9.0)	(1.3)	(0.6)	(1.9)	(17.6)
Diabetes	3.0	55.7	17.0	0.9	1.3	2.1	20.0
	(2.6)	(7.8)	(10.0)	(1.0)	(1.1)	(1.0)	(14.0)
Cardio	4.3	2.8	55.4	1.4	1.1	3.4	31.6
	(3.6)	(1.8)	(20.5)	(1.2)	(1.0)	(0.9)	(22.4)
Pulmonary	5.5	1.9	19.1	23.4	0.7	7.8	41.6
	(4.6)	(1.4)	(8.9)	(15.2)	(0.6)	(3.4)	(23.1)
Renal	4.4	3.6	21.4	1.2	37.1	5.8	26.5
	(3.5)	(3.0)	(13.2)	(1.3)	(6.2)	(3.1)	(15.4)
Other	5.6	1.6	15.6	2.3	0.8	34.3	39.8
	(4.0)	(1.3)	(10.6)	(2.0)	(0.4)	(5.8)	(9.5)
Healthy	5.5	1.2	10.8	1.4	0.4	4.5	76.2
	(4.2)	(0.8)	(6.8)	(1.4)	(0.3)	(2.1)	(10.9)

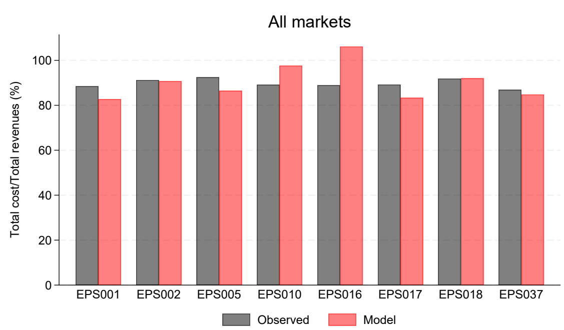
Note: Table presents mean and standard deviation in parenthesis of transition probabilities across diagnoses. Summary statistics are calculated across sex-age combinations in each cell.

Appendix 10 Additional fixed cost results

Appendix Table 16: Summary statistics of marginal variable profits

Insurer	mean	sd
EPS001	10,040	32,640
EPS002	49,683	119,549
EPS003	31,825	103,891
EPS005	38,409	128,447
EPS008	65,442	163,999
EPS010	55,998	164,682
EPS013	97,856	204,481
EPS016	121,612	271,489
EPS017	99,873	260,659
EPS018	86,641	218,653
EPS023	31,752	94,038
EPS037	102,637	191,252

Note: Table presents mean and standard deviation of marginal variable profits per insurer. Measured in millions of Colombian pesos per service.



Appendix Figure 13: Out-of-sample model fit

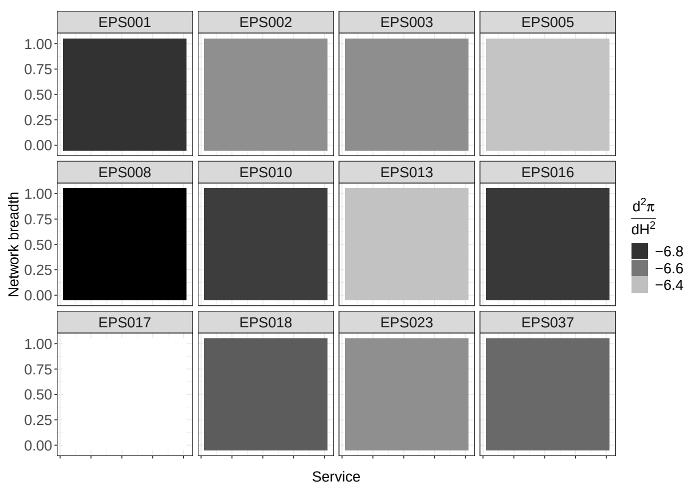
Note: Figure presents the model-predicted ratio of total costs (total average costs plus fixed costs) to total revenues and the observed ratio from insurers' public income statements. Public income statements are obtained from `https://docs.supersalud.gov.co/PortalWeb/SupervisionRiesgos/EstadisticasEPSRegimenContributivo/RC%20Estados%20financieros%20Dic%202011-CT2011.pdf`. Because my model is estimated on the sample of new enrollees with continuous enrollment and public income statements correspond to all enrollees, I scale up estimated insurer revenues and costs by multiplying by the total number of enrollees in the country and dividing by the number of new enrollees.

Appendix 11 Concavity of the profit function

The second partial derivative of the short-run profit function with respect to network breadth for service k is:

$$\frac{\partial^2 \Pi_{jm}}{\partial H_{jkm}^2} = \sum_i \left((R_{\theta m} - (1-r_i)AC_{\theta jm}) \frac{\partial^2 s_{ijm}}{\partial H_{jkm}^2} - 2(1-r_i)\frac{\partial s_{ijm}}{\partial H_{jkm}} \frac{\partial AC_{\theta jm}}{\partial H_{jkm}} - (1-r_i) s_{ijm} \frac{\partial^2 AC_{\theta jm}}{\partial H_{jkm}^2} \right) - 2\omega$$

To check whether this derivative is negative at all values of network breadth, I conduct a partial equilibrium exercise where each insurer is allowed to deviate and set $H_{jkm} = \{0, 0.1, 0.2, 0.3, ..., 1\}$ for each service k, holding its rivals' choices fixed at observed levels. I compute this exercise with data from Bogotá as in my counterfactuals. Appendix figure 14 presents the results. Each panel corresponds to the deviating insurer, and displays the value of the second partial derivative for each service in the horizontal axis and for each value of network breadth in the vertical axis. Results show that the second partial derivative of the short-run profit function is negative for all insurers and services.



Appendix Figure 14: Second partial derivative of short-run profit function

Note: Figure presents the second partial derivative of insurers' short-run profit function for every service. Each panel corresponds to an insurer, the horizontal axis is a service, and the vertical axis is the value of service network breadth.

Appendix 12 Additional counterfactual results

Appendix Table 17: Insurer demand

Diagnosis list
Healthy
Cardiovascular disease
Other Disease
Cervical Cancer
Breast Cancer
Other Renal Disease
Other Cancer
Chronic Kidney Disease
Diabetes
Skin Cancer
Lymphoma
Stomach Cancer
HIV-AIDS
Lung Cancer

Note: Table presents list of diagnoses used in the improved risk adjustment counterfactual.

Appendix Table 18: Networks, costs, and welfare under homogeneous costs

	Variable	(1) Avg cost	(2) Fixed cost
A. Overall	Mean network breadth	0.24	-7.56
	Avg. cost per enrollee	0.73	-0.08
	Total avg. cost	1.37	2.07
	Consumer surplus (sick)	0.95	1.67
	Consumer surplus (healthy)	0.62	1.35
B. Service network breadth	Otorhinolaryngologic care	0.27	-9.45
	Cardiac care	0.25	-8.30
	Gastroenterologic care	0.26	-8.68
	Renal care	0.26	-9.94
	Gynecologic care	0.26	-8.60
	Orthopedic care	0.25	-9.10
	Imaging	0.15	-5.02
	General medicine	0.15	-7.17
	Laboratory	0.09	-5.16
	Hospital admission	0.22	-6.52

Note: Panel A presents the percentage change in mean network breadth, insurer total average costs, short-run average cost per enrollee, and long-run consumer welfare for sick and healthy individuals, in the scenario with homogeneous average costs in column (1), and the scenario with homogeneous average and network formation costs in column (2). Insurer fixed effects in average costs and network formation costs are set to the average fixed effect. Panel B presents the percentage change in mean network breadth by service category.