Non-Price Competition and Risk Selection Through Hospital Networks

Natalia Serna*

June 12, 2023

Click here for latest version

Abstract

Health insurers can engage in risk selection through the design of their hospital networks. I measure the impact of risk selection incentives on hospital network breadth using a model of insurer competition in networks applied to data from Colombia's health care system. I find that insurers risk-select by providing narrow networks in services that unprofitable patients require. Improving the risk adjustment formula increases median network breadth by 4.8 percent and consumer welfare by 2.8 percent. Simulations of the model with deregulated premiums show that the price and the non-price elements of insurance contracts are substitutes for risk selection.

Keywords: Hospital networks; Risk selection; Health Insurance; Risk adjustment.

JEL codes: I11, I13, I18, L13.

*University of Wisconsin-Madison. E-mail: nserna@wisc.edu. I want to thank the Colombian Ministry of Health and Alvaro Riascos 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. I also thank Victor Aguirregabiria, Zarek Brot-Goldberg, Ignacio Cuesta, Randy Ellis, Gastón Illanes, Haizhen Lin, Cici McNamara, Elena Prager, Stephen Ryan, Mark Shepard, and Benjamin Vatter, for their useful feedback. This research has benefited from comments from participants at the Stanford Institute for Theoretical Economics, IO Workshop at UW-Madison, ASHEcon, IIOC, MEA, Colombian Central Bank Workshop, and Quantil Seminar in Colombia.

1 Introduction

Risk selection is a first-order concern in the design of health insurance markets. Insurers may attempt to disproportionately enroll healthy (profitable) patients rather than sick (unprofitable) patients by carefully crafting various aspects of the insurance contract, such as cost-sharing rules, premiums, and hospital networks. This type of selection has been shown to reduce access to insurance (Shepard, 2022) and health care (Ellis and McGuire, 2007), and in principle it can lead a market to unravel altogether (Kong et al., 2022).

Most prior studies have focused on how risk selection affects premiums, holding other aspects of the insurance plan fixed (e.g, Cabral et al., 2018; Ho and Lee, 2017; Dafny et al., 2015). Less attention has been paid to whether insurers strategically choose the non-price characteristics of their plans to cream-skim the market. In this paper I study how insurers engage in risk selection through the design of their hospital networks. I measure the impact of risk selection incentives on the breadth of hospital networks, and simulate the effects of counterfactual policies like risk adjustment that aim to reduce insurers' incentives to risk-select.

My empirical setting is Colombia, where private insurers provide a national health insurance plan in a system similar to the Medicare Advantage program in the United States. However, a key difference is that almost all aspects of the insurance contract are closely regulated: premiums, coinsurance rates, copays, and maximum out-of-pocket amounts are all set by the government. The only element of the public health insurance plan that is unregulated is hospital networks. The strict regulation and near universal coverage of the national insurance plan makes the Colombian system an ideal setting to study risk selection through network breadth.

Importantly, health insurers in Colombia have discretion over which services to cover at which hospitals. Insurers can then use their service-specific hospital 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), who find that insurers in Medicare Advantage engage in risk selection by placing services that sick individuals need in higher cost-sharing tiers.[1]

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

I then develop and estimate a model of insurer competition in service-level hospital network breadth and consumer demand for insurance. The model allows me to quantify how hospital network breadth, health care costs, and consumer welfare respond to changes in the regulatory environment or to policies aimed at reducing insurers' incentives to risk select. This question has become increasingly relevant with the prevalence of network adequacy rules (Mattocks et al., 2021; Haeder et al., 2015). An important simplification I make is to

[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.

measure network breadth with a single index: the fraction of all hospitals in a market that can provide the service that are covered by the insurer. The cost of this approach is that it implicitly assumes hospitals' identities are not important–what matters is how many hospitals are in the network, not which hospitals–but the advantage is that this simplification allows me to tractably model insurers' equilibrium choices of service-specific network breadths.

For consumer demand, I model new enrollees' myopic and static discrete choices of insurance carrier. The enrollees' indirect utility is a function of insurers' network breadth per service and out-of-pocket costs. Out-of-pocket costs depend on network breadth to reflect 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 carrier is associated with higher out-of-pocket costs. On the supply side, I model insurers' average cost per enrollee as a nonlinear function of service-level network breadth and enrollee characteristics. The average cost function depends on the types of consumers that insurers enroll, allowing for potential economies of scope across services. The function represents an approximation to an equilibrium where insurers and hospitals bargain over service prices and then consumers make claims for those services.

Insurers maximize profits by choosing their vector of network breadths conditional on rivals' choices. To approximate a steady-state equilibrium, I assume insurers make a one-time choice of service-specific network breadth, recognizing that this choice will affect both current and future profits. Consumers are assumed to have infinite inertia: once they choose an insurer, they stay with that insurer. However, future profits from a given patient evolve as that patient ages and transitions between health states, with transition probabilities computed from the claims data.

I estimate the model on 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). As is usual in the literature on hospital networks (Gowrisankaran et al., 2015; Capps et al., 2003), I use the claims-level data to recover each insurer's network of hospitals in each of the service categories provided by the national insurance plan.

Demand estimates show that, conditional on sex and age, individuals with chronic conditions have higher willingness-to-pay for network breadth than healthy individuals, which is consistent with strong adverse selection. I find that insurers' average cost function exhibits economies of scope and that the network formation cost is heterogeneous across insurers, services, and markets. The estimates imply that if an insurer unilaterally increases network breadth for a service, roughly half of the resulting cost increase is due to adverse selection (attracting sicker patients).

I use my model to quantify how hospital networks respond to changes in the regulatory environment and how these network breadth changes affect health care costs and consumer welfare. In a first set of counterfactual simulations, I examine how equilibrium choices of network breadth depend on the government's risk adjustment mechanism. If there were no risk adjustment at all, insurers would reduce average network breadth by 6.7 percent and consumer welfare would fall by only 13,753 pesos ($7.2) per capita per year for those with chronic conditions. These small welfare effects are due to out-of-pocket cost savings partially compensating reductions in coverage. By contrast, if risk adjustment were made more granular, insurers would increase average network breadth by 4.6-28.0 percent, depending on how many risk factors are included in the formula. In the case of "perfect" risk adjustment, consumer welfare increases by 49,161 pesos ($25.9) per capita per year for those with chronic diseases.

Finally, I conduct a set of simulations to understand how deregulation of premiums would affect insurers' equilibrium choices of network breadth, given that in the current regulation premiums are zero. I assume insurers compete over premiums and network breadth per service and can discriminate premiums across sex and age group. Insurers receive premiums in addition to the risk adjusted transfers from the government. I find that premiums are U-shaped with respect to the enrollee's age and higher for males than for females, reflecting the correlation between these demographic characteristics and health status. Deregulating premiums incentivizes insurers to increase their network breadth per service by around 28 percent. Thus, the model illustrates that price and non-price competition are substitutes from the point of view of risk selection. The zero-premium policy adopted by the Colombian government generates narrow networks in equilibrium.

My paper contributes to the literature on risk selection in health insurance by identifying hospital networks as a selection mechanism and by quantifying the effect of risk adjustment and premium setting on service-specific hospital network breadth. Existing literature has focused on the impact of premiums on enrollment (Einav et al., 2019; Finkelstein et al., 2019; Tebaldi, 2017; Decarolis, 2015), of risk adjustment on selection effort (Brown et al., 2014; McWilliams et al., 2012; Nicholson et al., 2004), and of risk adjustment on premiums (Cabral et al., 2018; McGuire et al., 2013; Pauly and Herring, 2007). Other papers deal with alternative selection mechanisms such as insurer advertising (Aizawa and Kim, 2018) and drug formulary design (Geruso et al., 2019).

The most closely related paper is Shepard (2022), who shows that sick individuals' strong preferences for a star hospital incentivizes insurers to drop this hospital from their networks in the context of the Massachusetts Exchange. My paper builds on his intuition by showing that selection incentives exist on the multidimensional choice of service-specific network breadth. But differs by proposing an equilibrium model of insurer competition in network breadth that allows me to conduct counterfactual simulations where insurers compete mainly on the non-price characteristics of their insurance plans.

2 Institutional Background and Data

The Colombian health care system, established in 1993, is divided into a "contributory" and a "subsidized" regime. The first covers formal employees and independent workers who are able to pay their monthly taxes (nearly 51 percent of the population). The second covers individuals who are poor enough to qualify and are unable to contribute (the remaining 49 percent). The national health care system has almost universal coverage. Universal coverage implies that risk selection does not happen on the individual's decision of whether to enroll or not but on the decision of which insurer to enroll with.

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 functions of the enrollee's income level. These cost-sharing rules are standardized across insurers and providers.[2] Hospital

[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 expenditure amount in a year equals 57.5 percent times the MMW. This corresponds to an actuarial value of 92 percent. 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 expenditure is 230 percent times

networks are the only dimension in which insurers differ. Insurers can form networks separately for each of the services 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. Although the government does stipulate a set of network adequacy rules to guarantee appropriate access to health services, these rules are very coarse and relate only to the provision of primary care, urgent care, and oncology.[3]

At the end of every year, insurers report to the government all health claims made through the national insurance plan that they reimbursed hospitals in their network for. The data for this paper are the enrollment files of all enrollees to the contributory system during 2010 and 2011 (25 million), and their claims reports to the government (650 million). I focus on the sample of individuals aged 19 or older with continuous enrollment spells or no gaps in enrollment (8.5 million) and their associated claims (260 million). This distinguishes consumers whose choices are not conflated by variation in enrollment spells. Enrollment spells can vary due to variation in income across time, job loss, or informality.[4]

Of the continuously enrolled, 2/3 are *current enrollees* or 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

the MMW. The associated actuarial value is 84 percent. Finally, for people with incomes above 5 times the MMW, the coinsurance rate equals 23 percent, the copay 20,900 pesos, and the maximum expenditure amount is 460 percent times the MMW, all corresponding to an actuarial value of 78 percent. The average exchange rate during 2011 was $1,847 COP/USD.

[3]For more information visit https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PSA/Redes-Integrales-prestadores-servicios-salud.pdf

[4]Because the continuously enrolled represent only 36 percent of all enrollees to the contributory system, I conduct robustness checks on my descriptive analysis using all enrollees.

who move from the subsidized system after they find a job, or those who for some reason were uninsured for 12 continuous months and then enroll in the health care system.[5] Consumer inertia in this market is substantial. In the sample of current enrollees, only 0.06 percent switch their insurance carrier 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 each individual the average income for his or her municipality, sex, and age. The health claims data report date of provision, service description, service price, provider, insurer, and ICD-10 diagnosis code for each claim.

Every claim is associated to a 6-digit service code from the national insurance plan. I collapse the 6-digit codes to 2 digits resulting in 58 service categories ("service" for short). These services describe surgical and non-surgical procedures in parts of the body.[6] Services in my data are, for example, procedures in cardiac vessels, procedures in stomach, imaging, consultations, and hospital admissions. Each category, in turn, covers more detailed medical procedures. For instance, procedures in cardiac vessels includes angioplasty, pericardiotomy, heart transplant, and aneurysm excision. The complete list of

[5]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.

[6]The first two digits of the service codes (known as CUPS for its Spanish acronym) indicate the anatomical area where the procedure is performed, the third digit is the type of procedure, and the fourth and fifth digits define more specifically the methods used for the procedure. See Resolution 4678 of 2015 from the Ministry of Health.

services is provided in appendix 2.

Health claims reports to the government come from the 23 private insurers that participated in Colombia's contributory health care system during my sample period. I focus on the 10 largest insurers that account for 87 percent of enrollees. Insurers compete in every market which is a Colombian state; there are 32 markets in my data (similar in size to an MSA in the US).

Health claims can be provided either by in-network stand-alone doctors, clinics, or hospitals. In 2011, Colombia had around 11,200 hospitals and small clinics, which comprised 1/6 of all providers in the country. I focus on the sample of hospitals and clinics ("hospitals" for short) that provide inpatient, surgical, urgent care, and diagnostic services. These hospitals are in the upper tail of the distribution of health care costs, where variance is high and risk selection incentives are more salient.

My hospital sample selection criteria matters because I recover the insurers' service-level hospital network from observed claims. This can be problematic, particularly for small providers, because it may be the case that there are zero claims from a provider who is actually in-network. To avoid this type of measurement error, my sample focuses on relatively large hospitals for which there are sufficient claims in each service category to infer them as being part of an insurer's network. Appendix figure 1 shows the distribution of number of claims per hospital, insurer, and service.

I obtain the list of 1,663 hospitals in 2011 and 1,453 in 2010 that satisfy my sample definition from the Ministry of Health's Registry of Health Care Providers.[7] I match hospitals in my claims data to the registry and end up with a 97 percent match rate in 2010 and an 87 percent match rate in 2011.

[7] The registry can be accessed through the following website: https://prestadores.minsalud.gov.co/habilitacion/

The matched sample of hospitals, which represents 3 percent of all providers in the country, accounts for 32 percent of total health care costs and 27 percent of total claims in the contributory system.

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. The formula does not include information about a patient's previous diagnoses. For year t, the base un-adjusted capitated transfer is calculated using the claims data from all insurers from year $t-2$. This transfer is roughly equal to the present value of the average annual health care cost per enrollee. Then, for each risk pool defined by a combination of sex, age group, and municipality, the government calculates a risk adjustment factor that multiplies the base transfer. Appendix table 1 shows the national base transfer and its value for some specific municipalities. Appendix table 2 shows the risk group multipliers for 2011.

Because of the coarsely defined risk pools, the ex-ante risk adjustment 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 risk-adjusted transfer that each insurer received for each of its 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 300 million pesos.

The High-Cost Account compensates insurers that enroll an above-average share of people with any of the following chronic diseases: 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. The per-patient reimbursement equals the average cost of treatment for each disease. These reimbursements come from insurers that enroll a below-average share of individuals with those diseases.[8] 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 during the sample period, suggesting these ex-post transfers do not provide much risk adjustment.

Selection incentives in this system exist because annual health care costs exhibit enormous variation across patients conditional on government risk-adjusted transfers. Figure 1 shows that the mean and the variance (as reflected in the difference between 90th and 10th percentiles) of health care costs increase with the government's reimbursement or the individual's risk score. The rising trend in total costs by risk score suggests that insurers have incentives to engage in selection against old individuals. The rising trend in variance suggests that there is scope to select consumers in the upper tail of the distribution who are more likely to be overcompensated by the risk adjustment formula (Brown et al., 2014).

The coarse nature of the risk adjustment formula and the high variance in health care costs generate large variation in profits per enrollee that incentivizes risk selection. Table 1 presents the mean, 1st and 99th percentiles of profits per capita in the sample of current enrollees and new enrollees during 2011. If the risk adjustment formulas were able to completely eliminate

[8] See Resolution 000248 of 2014 from the Ministry of Health.

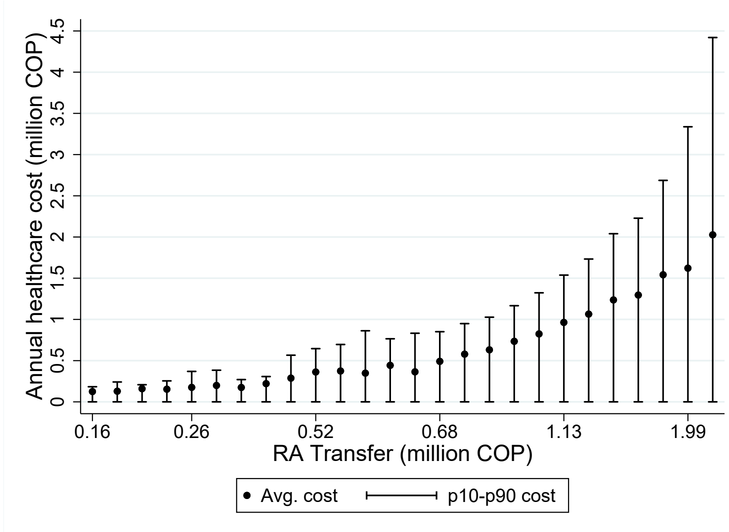


Figure 1: Health care cost by risk-adjusted transfer

Note: Mean, 10th, and 90th percentiles of annual health care cost by level of ex-ante risk-adjusted transfer from the government. The ex-ante transfer is the individual's risk score.

risk selection incentives, the variance in the distribution of profits per enrollee should be similar across insurers, but this is not the case. The table also shows that new enrollees' average profit is significantly higher than that of current enrollees, and their distribution of profits per capita less skewed to the left. Thus selection efforts should be stronger among new enrollees.

3.1 Measuring network breadth

Insurers in Colombia have discretion over how many hospitals to cover for each service. But it is mandatory that they cover at least one hospital for every service in the national insurance plan. This means that network breadth is defined over the number of hospitals *conditional on the service*, but not over services. Although insurers' coverage choices are in part determined by differences in hospital specialty and available capacity, these choices also depend on the type of consumers that insurance companies want to risk select upon.

Table 1: Distribution of profit per enrollee

Insurer	Current			New		
	mean	p1	p99	mean	p1	p99
EPS001	0.15	-7.76	2.06	0.49	-0.73	2.11
EPS002	0.09	-5.61	1.82	0.35	-1.12	1.99
EPS003	0.07	-5.94	1.88	0.42	-0.82	2.11
EPS005	0.14	-8.37	1.99	0.39	-1.62	2.11
EPS008	0.08	-6.54	1.88	0.30	-2.40	2.02
EPS009	-0.36	-15.25	1.87	0.26	-3.32	1.99
EPS010	0.10	-5.68	1.87	0.39	-1.03	2.11
EPS012	0.08	-6.35	1.80	0.36	-1.00	1.90
EPS013	0.08	-5.30	1.86	0.34	-1.00	1.99
EPS016	0.08	-6.25	1.90	0.38	-1.27	1.99
EPS017	0.04	-6.23	1.83	0.31	-1.60	1.99
EPS018	0.04	-6.22	1.83	0.28	-1.63	1.68
EPS023	0.10	-4.51	1.68	0.33	-0.82	1.68
EPS037	0.13	-17.16	2.15	0.66	-1.34	3.12
Total	0.08	-7.77	1.99	0.41	-1.28	2.11

Note: Table presents mean, 1st and 99th percentiles of profit per enrollee for each insurer in the sample of current enrollees and new enrollees. Profit per enrollee is calculated as the risk-adjusted transfers (ex-ante and ex-post), plus revenues from copays and coinsurance rates, minus total health care cost. Profits are measured in millions of 2011 Colombian pesos.

If insurers use their service-level hospital networks to select risks, then differences in risk selection efforts should appear as differences in service network breadth. I define service network breadth as the fraction of all hospitals in a market offering a particular service that are covered by the insurer. Table 2 shows that there is substantial heterogeneity in network breadth per service across insurers and markets. Of the total variation in service-level network breadth, 30 percent is explained by the insurer, 10 percent by the service, and 4 percent by the market.

Network breadth defined as a continuous measure in the unit interval is my primary object of interest in the rest of this paper. Enrollee satisfaction surveys conducted by the Colombian Ministry of Health show that narrow networks are one of the main reasons for dissatisfaction with an insurance company.

Table 2: Distribution of network breadth per service

	2010		2011	
Insurer	mean	sd	mean	sd
EPS001	13.1	12.5	12.1	11.6
EPS002	29.6	23.7	30.3	23.6
EPS003	22.4	20.5	20.4	19.4
EPS005	27.2	24.9	27.7	23.7
EPS008	10.8	14.8	12.6	16.3
EPS009	11.8	20.2	9.5	14.6
EPS010	10.0	14.4	9.4	13.2
EPS012	17.1	19.1	13.8	14.9
EPS013	52.7	32.6	49.7	31.5
EPS016	46.5	27.1	55.0	26.2
EPS017	16.5	21.9	16.3	20.0
EPS018	14.2	21.6	12.2	18.6
EPS023	11.5	17.5	10.0	15.7
EPS037	37.0	29.7	34.2	27.9

Note: Mean and standard deviation of service network breadth per insurer across markets in 2010 and 2011.

Patients enrolled with insurers that have low network breadth typically have to travel longer distances to seek care. Network breadth can thus be interpreted as a measure of proximity to hospitals (see appendix figure 2).

Implications on hospital quality. By collapsing networks to an index per service, I am effectively assuming that, conditional on the service, hospital quality is constant. This simplification is useful to explain the existence of narrow networks in equilibrium, but it could be losing important information if it matters *which* hospitals are included in the network, and not just how many. One reason why this information might be important is if some hospitals are star hospitals. Or, more generally, if some hospitals have higher quality than others. However, I show using various tests that this is not an issue. First, star hospitals are not as common in Colombia as they are in other countries like the United States. Second, I find that hospital quality and network breadth are positively correlated (see appendix table 3), thus high-quality hospitals are more likely to be included in a broad network as they are in a narrow network. The information about hospital quality can then be subsumed in the information about network breadth. I provide additional evidence of the robustness of my model to hospital quality in sections 5 and 6.

3.2 Network breadth as a means of risk selection

The descriptive statistics show that there is substantial variation in service-level network breadth and profits per enrollee that is consistent with differences in selection efforts. In this subsection I link profits per enrollee with service utilization to characterize selection incentives at the service level by replicating figures in Geruso et al. (2019) with my data.

In figure 2 I show 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 dot in the figure represents a service weighted by the number of patients who make claims for it. Patients who make claims for several services will be represented in several dots, 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 procedures in heart valves, cardiac vessels, and pancreas, which are located toward the right of this figure. In the case of procedures in heart valves, average costs are almost 5

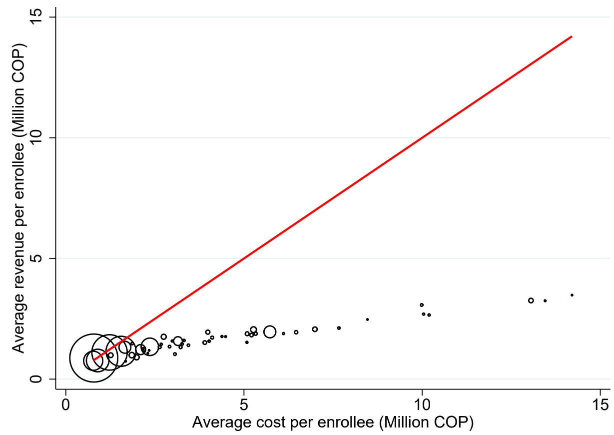


Figure 2: Service-level selection incentives after risk adjustment

Note: 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 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.

times larger than average revenues per enrollee.[9]

The striking differences between revenues and costs per service arise from the simple fact that government payments do not compensate for enrollee characteristics that predict service usage. But insurers can set up their hospital networks separately per service. 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 through their choice of hospital 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 network breadth per service across insurers and markets. Average profits are calculated conditional on patients who make claims for the service. The red line

[9]These findings hold when using information from all individuals enrolled to the contributory system without constraining enrollment to be continuous.

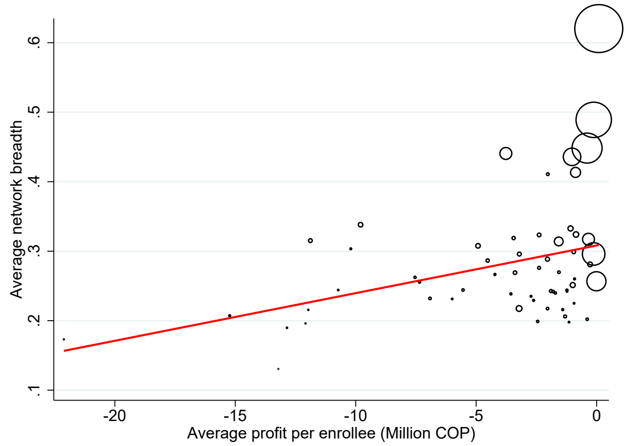


Figure 3: Correlation between network breadth and service profitability

Note: 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. 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.

corresponds to a linear fit and shows that relatively profitable services, such as consultations and procedures in teeth and tongue, tend to have broader networks than relatively unprofitable services, such as procedures in heart valves and cardiac vessels. This positive correlation is consistent with service-level network breadth being a mechanism for risk selection.

Health care cost variation across services can be due either to selection or moral hazard, which the previous figures conflate. In appendix 4 I assess how much of the variation in costs can be explained by adverse selection and cream-skimming alone. The appendix shows that broad networks attract less profitable patients.[10] So, if insurers cannot charge premiums to cover the higher costs resulting from a broad network, why don't all insurers choose to have narrow networks? Appendix 5 explores different mechanisms that explain insurers' asymmetric choices of network breadth.

[10]This is similar to Liebman and Panhans (2021) who find that narrow network plans are cheaper because they are better able to steer patient to low-cost hospitals.

4 Model

In this section I put together the descriptive findings to model insurer competition in hospital networks. In the model, insurers first simultaneously set their vector of service-specific network breadths in every market. Then, given service-level network breadths, consumers make enrollment decisions.

I model insurer demand in the sample of new enrollees in 2011, who do not experience inertia when making their *first* enrollment decision. The demand model captures consumers' cost-coverage trade-off by allowing out-of-pocket costs to depend on network breadth. To the extent that healthy consumers care more about out-of-pocket costs than network breadth relative to sick consumers, insurers will be able to screen healthy individuals by choosing narrow service-level networks. I also assume that, after making their first insurer choice, enrollees do not switch as seen in the data.

On the supply side, I assume insurers are forward looking and compete for the set of new enrollees every period. Insurers internalize the dynamic incentives introduced by the zero switching rate in demand. With zero switching, the dynamic programming problem of network formation every period can be approached as a static problem. Insurers maximize the sum of current and future discounted profits by simultaneously choosing their vector of network breadths once. The solution concept is a Nash equilibrium.

I allow insurers to have heterogeneous costs. Together with preference heterogeneity, cost differences across insurers can explain their asymmetric choices of network breadth. The supply model is thus needed to assess the relative importance of these two factors in generating the observed equilibrium.

4.1 Insurer Demand

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 the 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, $\geq$75), and diagnosis $d \in D = \{$cancer, cardiovascular disease, diabetes, renal disease, long-term pulmonary disease, arthritis, asthma/smoking, other, no diseases$\}$. Diagnoses in the list are groupings of ICD-10 codes following Riascos et al. (2014). For individuals with several comorbidities, I assign the most expensive disease. 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.

I assume that the individual knows her diagnoses before making her first enrollment choice. This could be either because of medical family history or because, prior to enrolling in the contributory system, she went to the doctor and received a diagnosis. Enrollees know their health condition because selection occurs on observable, un-reimbursed (or poorly reimbursed) consumer characteristics such as those associated with health status. The consumer observes each insurer's network breadth in service k and market m, H_{jkm}; weights each service by the probability of claiming it $q_{\theta k}$, conditional on the diagnosis; and then makes a one-time myopic choice of carrier. Denote by u_{ijm} the indirect utility of a new enrollee i in market m for insurer j, which takes the following form:

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

where

$$\beta_i = x_i'\beta$$

$$\alpha_i = x_i'\alpha$$

The vector x_i includes consumer demographics such as sex, dummies for each age category, indicators for the diagnoses in D, type of municipality indicators, an indicator for individuals making less than two times the monthly minimum wage, and an intercept.[11] $c_{\theta jm}$ is the average out-of-pocket cost of consumer type θ at insurer j, which depends on the insurer's vector of network breadth $H_{jm} = \{H_{jkm}\}_{k=1}^{K_m}$, either because negotiated prices are higher or because patients tend to consume more health services the broader is the network. 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 T1EV.

Average out-of-pocket costs are the sum of coinsurance payments, copays, and tax contributions to the system:

$$c_{\theta jm} = \text{Coins}_{\theta jm} + \text{Copay}_{\theta jm} + \text{Tax}_{\theta}$$

Tax contributions are a function of the enrollee's income level. While coinsurance payments and copays depend on the individual's income level, on

[11] In the government's risk adjustment formula, municipalities are grouped into urban, normal, and peripheral. Urban municipalities belong to metropolitan areas, normal municipalities are those adjacent to metropolitan areas, and peripheral municipalities are those characterized by difficult geographical access.

the insurer's negotiated service prices with hospitals, and on the individual's health care utilization. Prices and utilization are correlated with the insurer's service-level network breadth because of insurer-hospital price bargaining and because of patient moral hazard.[12] To capture this correlation, I assume that out-of-pocket costs are a linear function of the insurers' average cost per enrollee $AC_{\theta jm}$, which in turn depends on network breadth as follows:

$$c_{\theta jm} = \mu_y AC_{\theta jm}(H_{jm}) + \epsilon_{\theta jm} \tag{2}$$

where $\epsilon_{\theta jm}$ is a standard normal error term. I estimate equation (2) separately by income group to recover μ_y; results are presented in appendix table 8. A more detailed description of insurers' average cost per enrollee is provided in the next subsection.

The first term on the right side of equation (1) can be interpreted as a reduced-form approximation to the consumer's expected utility for the network obtained from a 2-step model, in which first the individual chooses an insurer and then chooses an in-network hospital, as in Ho and Lee (2017). In the case of Ho and Lee (2017), the insurer offers the same network of hospitals to consumers of different medical conditions. In my case, variation in network breadth across services and variation in the likelihood of making claims for those services, together imply that the network can also be disease-specific.

The probability of making a claim, $q_{\theta k}$, is calculated from the claims data in a preliminary step as the average prediction per consumer type and service, of a fractional logistic regression estimated at the consumer-type level given

[12] Appendix figure 4 shows that out-of-pocket costs vary substantially across insurers and across consumer types.

by:

$$\text{Fraction claims}_{\theta k} = \psi_k + \psi_\theta + \psi_{ik} \tag{3}$$

The dependent variable is the fraction of type-θ patients that make 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 observable characteristics. Even though new consumers are myopic when choosing their insurance carriers, I assume that their expectations over the type of services they will need conditional on their initial health condition are correct on average. These expectations do not depend on the insurer they enroll with. I estimate equation (3) on data from both current and new enrollees in 2010 and 2011. Appendix figure 3 presents the distribution of $q_{\theta k}$ separately for healthy and sick individuals, and for a few service categories.

I allow preferences for network breadth to vary across demographic characteristics and diagnoses to capture the extent of service-specific adverse selection documented in the descriptive section. However, I do not explicitly model unobserved heterogeneity with inclusion of random coefficients. Instead I include preference shocks ε_{ijm} that are independent across choice alternatives. This means that the only way in which risk selection can arise in my model is through the observable characteristics.

The second term on the right of equation (1) captures differences in prices and utilization across insurers, giving rise to consumer sorting based on out-of-pocket costs. This sorting is needed to rationalize the existence of narrow network carriers in the *observed* equilibrium since myopic, healthy new enrollees disproportionately choose narrow network insurers with lower implied out-of-pocket costs. I allow out-of-pocket costs to depend on network breadth to reflect the cost-coverage trade-off that consumers face when making enrollment decisions.[13]

The probability that consumer i in market m enrolls with insurer j is:

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

Identification. To identify the parameters associated with network breadth in the utility function, 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 procedures in lungs in two different markets, but one of these markets has a higher prevalence of respiratory diseases, then we should observe higher insurer demand in the market where people are relatively sicker.

There are however several threats to identification. The first is that network breadth may be correlated with unobserved insurer quality or unobserved consumer characteristics. For instance, network breadth could reflect how good the insurer is in processing health claims. This type of unobserved insurer quality potentially does not vary across markets conditional on consumer types. So, the inclusion of insurer-by-market fixed effects allows me to correct for this potential source of endogeneity.

To identify the parameters associated with the out-of-pocket cost, I use the variation in income across markets, which generates variation in coinsurance rates. But because out-of-pocket costs are a function of insurer-hospital nego-

[13]Out-of-pocket costs are aggregated across services with weights given by the claim probabilities because a service-level specification would require imputing costs for consumer-insurer combinations for which I do not observe claims being provided for a service, so measurement errors and mechanical bias would be much more likely.

tiated service prices, these may be correlated with unobserved hospital quality. For example, if an insurer covers a star hospital, demand and negotiated prices for that insurer will be relatively high, and my model would interpret consumers as having low sensitivity to out-of-pocket costs. This endogenous variation in negotiated prices occurs within insurer and across markets, therefore the inclusion of insurer-by-market fixed effects accounts for such type of variation. I also conduct additional robustness checks in section 5 to verify that variation in hospital quality is not a meaningful source of bias.

Incorporating hospital quality. While network breath is positively correlated with measures of hospital quality, it begs the question: can hospital quality be directly incorporated in the model? One way to do this would be to specify network breadth as a weighted average across services of the expected value for the insurer's network. This approach however would be very demanding of the data; it would require estimating a hospital demand function for each service category. Moreover, for some services having data about distance from the patient to the hospital is crucial to get accurate measures of the value of the network, which I unfortunately do not observe.

4.2 Insurer Average Costs per Enrollee

Following Shepard (2022), I assume that the average cost of type-θ individuals can be calculated as the average annual health care cost across all consumers i that are of type θ. I model the logarithm of average cost per consumer type as a function of network breadth, as follows:

$$\log(AC_{\theta jm}(H_{jm})) = \tau_0\Big(\sum_{k}^{K_m} q_{\theta k}A_k\Big) + \tau_1\Big(\sum_{k}^{K_m} q_{\theta k}H_{jkm}\Big) + \frac{1}{2K_m}\tau_2\sum_{k}^{K_m}\sum_{l\neq k}^{K_m} q_{\theta k}q_{\theta l}H_{jkm}H_{jlm} + \lambda_\theta + \eta_m + \delta_j \quad (4)$$

where K_m is the number of service categories available in market m, that is, services that existing hospitals in the market can provide. A_k is the government's reference price for service k (explained in more detail in appendix 8), λ_θ is a consumer type fixed effect, η_m is a market fixed effect, and δ_j is an insurer fixed effect. Equation (4) can be interpreted as an approximation to an equilibrium where insurers and hospitals bargain over service prices and then consumers make claims for those services.

The coefficient τ_0 captures whether insurers bargain higher or lower prices than the reference price with the average hospital in their network. τ_1 represents the elasticity of average costs with respect to insurer j's network breadth. τ_2 captures the average degree of complementarity between pairs of services. If $\tau_2 < 0$, then insurers have economies of scope across services, so greater coverage for service $l \neq k$ makes it more attractive to the insurer to provide higher coverage for service k. I include this measure of scope economies to rationalize the fact that insurers with broad networks in one service, tend to offer broad networks in other services as well (see appendix figure 6).

Identification. The parameters of equation (4) are identified from variation in average costs within consumer types across insurers that are identical except for their network breadth per service. My source of identification does not rely on different consumers implying different costs for similar insurers as in Tebaldi (2017) but, *conditional* on the composition of enrollee pools, for different coverage levels per service to imply different costs to the insurer. Variation in network breadth across insurers is thus exogenous conditional on the rich set of fixed effects.

If consumer selection into insurers happens mostly on observables, then the consumer-type fixed effects in equation (4) help correct for the endogenous variation in network breadth across enrollees. If selection happens mostly on unobservables, then it should be the case that there is unobserved cost variation within consumer types. One way to check this is to test whether estimates are robust to a more granular definition of consumer-type. I conduct robustness checks of this style in appendix table 14 using patient-level data.[14]

4.3 Competition in Network Coverage

Insurers compete separately in every market, choosing their service-specific network breadth 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', all collected in the vector $H_m = \{H_{jm}, H_{-jm}\}$, as well as on the enrollee's type θ. The annual per-enrollee profit is given by:

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

where $R_{\theta m}$ is the per-capita revenue including ex-ante and ex-post risk-adjusted transfers from the government and average copayments, $AC_{\theta jm}$ is the average cost of a type-θ consumer net of patients' coinsurance payments with r_i 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

[14] My average cost model aggregates total healthcare cost to the consumer type level to avoid econometric issues that arise from the overwhelming amount of zeros in health care costs at the patient-service level.

minus the cost of network formation:

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

Insurers take into account the future profits associated with each enrollee since, after making their first enrollment choice, individuals experience infinite inertia. $N_{\theta m}$ is the fixed market size of consumers type θ. In the expression for future profits, $\rho_{\theta m}$ represents the probability that type θ 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 falling into unemployment. $\mathcal{P}(\theta'|\theta)$ 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. I set ζ equal 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 cost to the insurer, which can be interpreted as an administrative cost associated with inclusion of an additional hospital to the network, where:

$$\xi_{jkm} = \xi_k + \vartheta_{jkm}$$

This network formation cost is nonlinear in network breadth, ω capturing whether the cost function is convex. The network formation cost is also

[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).

heterogeneous across insurers, services, and markets. ξ_k represents the observed service-specific cost component and ϑ_{jkm} represents the idiosyncratic cost shock that is observed by insurance companies but unobserved by the econometrician. I assume the cost shock is *iid* across insurers, services, markets, and time, and that it is mean independent of insurers' network formation cost shifters. 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 $J \times K$ first-order conditions (FOC) in each market, which assuming an interior solution in network breadth, is given by:

$$\sum_i \left(\frac{\partial \pi_{ijm}}{\partial H_{jkm}} N_{\theta m} + \sum_{s=t+1}^{T} \zeta^s \sum_{\theta'} (1 - \rho_{\theta' m}) \mathcal{P}(\theta'|\theta) \frac{\partial \pi'_{ijm}}{\partial H_{jkm}} N_{\theta' m} \right) = 2\omega H_{jkm} + \xi_{jkm} \tag{5}$$

The left-hand side of equation (5) represents the marginal variable profit MVP_{jkm} and the right-hand side is the marginal cost of network formation. The derivative of the short-run per enrollee profit, which enters MVP_{jkm}, is:

$$\begin{aligned} \frac{\partial \pi_{ijm}}{\partial H_{jkm}} = & \underbrace{\overbrace{R_{\theta m} \frac{\partial s_{ijm}}{\partial H_{jkm}}}^{\text{Selection}} + \overbrace{R_{\theta m} \frac{\partial s_{ijm}}{\partial AC_{\theta jm}} \frac{\partial AC_{\theta jm}}{\partial H_{jkm}}}^{\text{Cost heterogeneity}}}_{\text{Marginal revenue}} \qquad (6) \\ & - \underbrace{(1 - r_i)\Big(\overbrace{AC_{\theta jm} \frac{\partial s_{ijm}}{\partial H_{jkm}} + s_{ijm} \frac{\partial AC_{\theta jm}}{\partial H_{jkm}}}^{\text{Selection}} + \overbrace{AC_{\theta jm} \frac{\partial s_{ijm}}{\partial AC_{\theta jm}} \frac{\partial AC_{\theta jm}}{\partial H_{jkm}}}^{\text{Cost heterogeneity}} \Big)}_{\text{Marginal cost}} \end{aligned}$$

Equation (6) shows the effect of adverse selection and cost heterogeneity on insurers' network breadth choices. If an insurer unilaterally increases its network breadth for a particular service, revenues will increase because demand from individuals with high willingness-to-pay for that service is higher (selection effect). Insurers' costs also increase because patients with high willingness-to-pay for the service are the most expensive in that service category, and because changes in network breadth increase the cost of the average consumer (selection effect). Cost heterogeneity has opposite effects on revenues and costs. Broadening networks for a particular service, increases consumers' out-of-pocket costs and thus puts a downward pressure on insurer demand and revenues. An increase in network breadth also reduces insurers' costs because the average consumer is now cheaper.

Model discussion. My model of insurer competition extends and complements the work in Shepard (2022), who models the binary decision of an insurer to include or exclude a star hospital from its network in the context of the Massachusetts Health Exchange. In my case, I allow for insurer heterogeneity in network breadth across different services and model the dynamic incentives that insurers face when setting up their networks, which are introduced by infinite consumer inertia.

This model is robust to different assumptions about consumer behavior. Notice that if consumers were forward looking rather than myopic and could anticipate their future diagnoses, the equilibrium would be one where all insurers choose broad networks. However, the equilibrium implications of myopia are similar to a model where consumers are forward looking but (wrongly) believe that switching costs are zero, so they can re-optimize every period. Equilibrium implications of myopia are also similar to a model where consumers heavily discount the future and therefore choose their insurer based on current preferences and characteristics. This means that even without myopia, the model would generate adverse selection on network breadth and co-existence of broad and narrow network carriers in equilibrium.

Identification. Rewriting the FOC as

$$\text{MVP}_{jkm}(H_{jkm}) = 2\omega H_{jkm} + \xi_k + \vartheta_{jkm}, \quad \forall\ H_{jkm} \in (0,1) \tag{7}$$

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 network breadth per service. For instance, if an insurer hires a highly trained manager to bargain with hospitals or if an insurance company is vertically integrated with its network, then $E[\vartheta_{jkm}|H_{jkm}] < 0,\ \forall\ H_{jkm} \in (0,1)$ and OLS estimation of (7) would result in ω that is biased towards zero.[16] Identification of network formation cost parameters thus relies on instrumental variables Z_{jkm} that are correlated with network breadth but not with the cost shock, and that are correlated with marginal variable profits only through network breadth, such that $E[\vartheta_{jkm}Z_{jkm}] = 0, \quad \forall\ H_{jkm} \in (0,1)$.

The instrument set is populated as follows. First, I include the set of service fixed effects in equation (7). Second, because I use data from 2011 in estimating the model and ϑ_{jkm} is assumed *iid* over time, I use the service-specific network breadth in 2010. Third, I include the average probability that a person aged 19-24, 25-29, and 30-34 make a claim for service k in market m. These probabilities are calculated as the average prediction of equation (3) across consumers that share the demographic traits. Finally, I include the interactions between 2010 network breadth and the average service claim probabilities.

The moment conditions at an interior solution, given by $E[\vartheta_{jkm}Z_{jkm}] = 0$,

[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.

can only rationalize the observed equilibrium in markets where no insurer chooses a corner solution in any of the services. To estimate the parameters of the network formation cost, I restrict my sample to the four largest markets in the country (Antioquia 05, Atlántico 08, Bogotá 11, and Valle de Cauca 76) that cover 60% of the population in the contributory regime, In this final sample, all insurers choose an interior solution.

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, suggestive of positive selection into health insurance. A 10 thousand pesos increase in out-of-pocket costs reduces insurer demand by 11.5 percent, corresponding to an average elasticity of -0.88.[17] A ten percentage point increase in network breadth across all services increases the choice probability by 22.6 percent on average.[18]

Interactions between consumer and insurer characteristics matter for enrollment decisions. Males are less sensitive to out-of-pocket costs than females but have a stronger taste for network breadth. Sensitivity to out-of-pocket costs is also decreasing in the consumer's income level. Patients aged 75 or

[17]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.

[18]This marginal effect of network breadth is calculated as $\beta^D \sum_k q_{\theta lkm}$ and averaged across consumers and insurers.

older are both less likely to enroll in broad network carriers and more sensitive to out-of-pocket costs compared to younger patients. One explanation for this is that old individuals have had more contact with the health care system and are more likely to concentrate their care in a few providers. Given that old consumers tend to have higher out-of-pocket costs, the findings also imply that the average demand elasticity for patients aged 75 or older (-1.42) is almost twice that of patients aged 19-24 (-0.72).

Findings show that individuals with cancer and renal disease have stronger preferences for network breadth than their healthy peers. But the preference for network breadth is similar between consumers with diabetes, arthritis, or asthma and consumers without diagnoses. Individuals with chronic conditions are all significantly less responsive to out-of-pocket costs than healthy ones. The interactions between diagnosis indicators with out-of-pocket costs overcompensate the interactions with network breadth. The implied average elasticity for individuals without diseases (-0.87) is larger, for instance, than for patients with renal disease (-0.85). Appendix 7.2 presents some measures of the in-sample model fit.

With my estimates of the preference for network breadth and out-of-pocket costs, I calculate patient willingness-to-pay for an additional percentage point of network breadth for each service as $\frac{1}{|\alpha_i|}\frac{\partial s_{ijm}}{\partial H_{jkm}}$. Differences in willingness-to-pay across consumer types will be suggestive of patient sorting based on networks. Consumers with relatively high willingness-to-pay for a particular service will tend to enroll with insurers that have a high network breadth for that service.

Table 4 presents the average willingness-to-pay for a few services relative to healthy individuals. Patients with chronic conditions have a higher willingness-to-pay for network breadth than individuals without diagnoses, consistent with

Table 3: Insurer demand

Variable		Network breadth	OOP spending (million)
Mean		2.26 (0.19)	-11.5 (0.26)
Interactions			
Demographics	Male	0.37 (0.02)	0.83 (0.13)
	Age 19-24	1.81 (0.06)	-0.24 (0.47)
	Age 25-29	2.58 (0.07)	2.46 (0.26)
	Age 30-34	2.17 (0.06)	1.59 (0.31)
	Age 35-39	1.78 (0.06)	0.43 (0.41)
	Age 40-44	1.58 (0.06)	1.49 (0.37)
	Age 45-49	1.30 (0.06)	1.14 (0.30)
	Age 50-54	0.99 (0.06)	1.29 (0.32)
	Age 55-59	0.94 (0.07)	1.50 (0.30)
	Age 60-64	0.66 (0.07)	1.01 (0.29)
	Age 65-69	0.56 (0.07)	0.55 (0.29)
	Age 70-74	0.47 (0.07)	0.93 (0.29)
	Age 75 or more	(ref)	(ref)
Diagnoses	Cancer	0.08 (0.07)	5.85 (0.25)
	Cardiovascular	-0.25 (0.05)	4.79 (0.23)
	Diabetes	-0.11 (0.12)	5.60 (0.43)
	Renal	0.24 (0.27)	8.28 (0.17)
	Pulmonary	-0.27 (0.18)	7.63 (0.33)
	Arthritis	-0.13 (0.12)	7.79 (0.28)
	Asthma	-0.16 (0.24)	8.61 (0.50)
	Other	-0.81 (0.15)	7.26 (0.25)
	Healthy	(ref)	(ref)
Location	Normal	3.70 (0.04)	1.99 (0.16)
	Special	5.47 (0.08)	0.94 (0.32)
	Urban	(ref)	(ref)
Income	Low	0.30 (0.03)	-1.13 (0.22)
	High	(ref)	(ref)
N		5,852,405	
N enrollees		500,000	
Pseudo-R^2		0.23	

Note: Conditional logit for the insurer choice model estimated 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

	Cardiac vessels	Dialysis	Imaging	Consult	Laboratory	Nuclear Medicine	Hospital Admission
Cancer	7.3	7.2	4.4	1.4	2.7	7.1	6.4
Cardiovascular	4.6	4.5	3.1	1.1	2.1	4.5	4.2
Diabetes	6.8	6.8	4.5	1.6	2.9	6.7	6.2
Renal	130.4	129.8	79.0	25.4	48.4	126.4	115.7
Pulmonary	47.8	47.6	29.1	9.4	17.9	46.3	42.5
Arthritis	24.5	24.4	15.1	5.0	9.4	23.8	21.8
Asthma/Smoking	35.1	35.0	24.1	9.1	16.1	34.3	32.2
Other	22.3	22.2	15.0	5.5	9.9	21.8	20.4
Healthy	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Note: Willingness-to-pay for an additional percentage point of network breadth for the service in the column relative to healthy individuals. Willingness-to-pay is calculated as $\frac{1}{-\alpha_i} \frac{\partial s_{ijm}}{\partial H_{jkm}}$.

strong adverse selection. For instance, patients with renal disease are willing to pay 129.8 times more than a healthy individual for an additional hospital in the network for dialysis. Patients with cardiovascular disease are also willing to pay 4.6 times more than a healthy individual for an additional hospital in the network for procedures in cardiac vessels.[19]

Robustness checks. Although insurer-by-market fixed effects in the demand function absorb insurer-level unobserved quality that may be correlated with network breadth or out-of-pocket costs, this unobserved quality could vary within insurers, across consumer types in ways that are not captured by the fixed effects. In the appendix I conduct several robustness checks to provide encouraging evidence that this type of unobserved quality does not pose a threat to identification. In appendix table 10, I estimate a demand function that includes an indicator of star hospital coverage. In appendix table 11, I estimate demand in the subsample of markets excluding the capital city that

[19]The measure of willingness-to-pay can also be interpreted in terms of travel times to the nearest hospital. For example, the estimates imply that patients with cancer are willing to pay 7.1 times more than a healthy individual for a reduction of approximately 10 minutes per visit in travel time to the nearest hospital that offers nuclear medicine services.

has several large hospitals. In appendix table 12, I use different provider samples to construct my measure of network breadth. Finally, because requiring that new enrollees know their diagnoses before enrolling can create mechanical bias, in appendix table 13, I identify a new enrollee's diagnoses using only the information from claims made in January. The results of these exercises show that my estimates do not vary in ways that significantly affect subsequent counterfactual simulations.

5.2 Insurer Average Costs Per Enrollee

I estimate equation (4) in the sample of new and current enrollees, conditional on observed choices in 2010 and 2011. Table 5 shows the results, and appendix figure 8 presents the estimated consumer type fixed effects with their corresponding 95 percent confidence intervals. Average costs are increasing in network breadth and decreasing in the interaction between network breadth for different pairs of services. This suggests that insurer coverage decisions are characterized by economies of scope. A 1 percent increase in network breadth for service k reduces the average cost of providing service $l \neq k$ by 0.88 percent per enrollee.[20] However, scope economies are smaller in magnitude than the direct effect of network breadth on average costs. My estimates show that a 1 percent increase in network breadth raises average costs by 3.13 percent per enrollee.[21] My average cost model is robust to more granular definitions of consumer type as seen in appendix table 14. This provides suggestive evidence that there is no relevant source of unobserved cost heterogeneity within consumer types.

Model-based evidence of adverse selection. With my demand and

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

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

Table 5: Insurer average costs per enrollee

Variable	Coefficient	Std. Error
Network breadth	1.81	(0.21)
Scope economies	-134.3	(24.9)
Reference price	20.5	(6.43)
Insurer		
EPS001	0.11	(0.04)
EPS002	-0.29	(0.02)
EPS003	-0.22	(0.02)
EPS005	-0.08	(0.02)
EPS008	0.02	(0.06)
EPS009	0.06	(0.05)
EPS010	-0.08	(0.03)
EPS012	-0.70	(0.13)
EPS013	-0.07	(0.02)
EPS016	-0.12	(0.02)
EPS017	-0.25	(0.03)
EPS018	-0.18	(0.04)
EPS023	-0.45	(0.03)
EPS037	(ref)	(ref)
N	40,989	
R^2	0.39	

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

average costs estimates I can test for the presence of adverse selection implied by the model by looking at the correlation between the insurers' marginal cost with respect to service-level network breadth and the patients' willingness-to-pay for service-level network breadth along the lines of Einav et al. (2010). Marginal costs for service k are given by $\frac{\partial(AC(H_{jkm})s_{ijm}(H_{jkm}))}{\partial H_{jkm}}$, and willingness-to-pay for network breadth over service k is $\frac{1}{|\alpha_i|}\frac{\partial s_{ijm}(H_{jkm})}{\partial H_{jkm}}$. Averaging across services, I find that these two variables are strongly positively correlated as seen in appendix figure 9, suggesting that the endogenously selected patients with the highest willingness-to-pay for network breadth are also the most expensive to the insurer, a classic adverse selection result.

5.3 Competition in Network Breadth

The third piece of the insurers' profit function left to estimate is the network formation cost. To recover the network formation cost I use the first order condition from the insurers' profit maximization problem. Demand and average cost estimates allow me to compute the left-hand side of equation (5) denoting marginal variable profits (MVP). Appendix 10 presents some summary statistics of this variable as well as of dropout probabilities and transition probabilities, which are calculated off-line non-parametrically from the data. The fact that MVPs are positive for all insurer-service-market triplets suggests a role for network formation costs in explaining the profit maximizing choices of network breadth.

Table 6: Model of insurer network formation costs

asinh(MVP_{jmk})	coef	se
Network breadth	6.86	(0.16)
<u>Service</u>		
Cardiac Vessels	1.47	(0.20)
Stomach	1.25	(0.20)
Intestines	4.77	(0.20)
Imaging	6.64	(0.20)
Consultation	6.37	(0.21)
Laboratory	7.35	(0.20)
Nuclear Medicine	4.67	(0.20)
Hospital Admission	4.90	(0.20)
First stage F-stat	1,718.5	
N	2,262	
R^2	0.76	

Note: 2-step GMM estimation of equation (7) on the subsample of markets 05, 08, 11, 76, and the subsample of the 10 largest insurers in these markets. Excluded instruments are described in section 4.3. Robust standard errors in parenthesis and first-stage F-statistic for the endogenous variable, network breadth reported.

Table 6 presents the results of a 2-step GMM estimation for the inverse hyperbolic sine of MVP in equation (7).[22] The specification includes fixed effects for each of the 58 services, but only 8 are reported for exposition. I find that network formation costs are strictly convex in network breadth. The elasticity of marginal variable profits with respect to network breadth for service k equals 6.86. Lagged network breadth and average market demographics are strong instruments for observed network breadth as seen in appendix table 19. As a measure of out-of-sample fit, appendix figure 10 compares the model's predicted ratio of total costs to total revenues per insurer to the ratio obtained from insurers' public income statements.

Magnitude of adverse selection. Changes in network breadth generate profit variation that can be decomposed into its portions explained by variations in demand, average costs, and network formation costs. The variation in profits that is explained by changes in demand evidence the magnitude of adverse selection. To quantify this magnitude, I decompose profit changes that result from a partial equilibrium exercise where an insurer unilaterally increases network breadth for service k by 10%, while holding its rivals fixed.

Table 7 presents the average percentage change in short-run demand (s_{ijm}), total revenues ($R_{\theta m} s_{ijm}$), total average costs ($AC_{\theta jm} s_{ijm}$), average cost per enrollee ($AC_{\theta jm}$), and network formation costs (F_{jm}) across insurers and markets, following a 10% increase in network breadth for the service in the row. I find that the change in demand explains on average 46% of the change in insurer total costs, while cost heterogeneity explains the remaining 54%. Most of the variation in demand following an increase in coverage of consultations comes from healthy individuals. In the case of hospital admissions demand variations are explained mostly by patients with renal disease. These heterogeneous ef-

[22] I use the inverse hyperbolic sine transformation to allow for negative values of the marginal variable profit in counterfactuals.

Table 7: Decomposition of short-run profit changes after network breadth increase

Service	s_{ijm}	$R_{\theta m} s_{ijm}$	$AC_{\theta jm} s_{ijm}$	$AC_{\theta jm}$	F_{jm}
Cardiac vessels	0.00	0.00	0.00	0.00	0.01
Stomach	0.00	0.00	0.00	0.00	0.01
Intestines	0.12	0.12	0.13	0.01	0.32
Imaging	1.57	1.57	1.79	0.22	3.42
Consultations	14.14	14.14	19.12	4.29	14.88
Laboratory	4.40	4.40	5.21	0.77	7.79
Nuclear medicine	0.05	0.05	0.06	0.01	0.16
Hospital admissions	0.32	0.32	0.36	0.04	0.68

Note: Average percentage change in demand (s_{ijm}), total revenues ($R_{\theta m} s_{ijm}$), average costs per enrollee ($AC_{\theta jm}$), total average costs ($AC_{\theta jm} s_{ijm}$), and network formation costs (F_{jm}), following an insurer unilaterally increasing network breadth for the service in the row by 10%, while holding its rivals' choices fixed.

fects on demand across services and consumer types provide further evidence of the extent of adverse selection.

6 The Effect of Risk Adjustment on Network Breadth

In this section I use my model estimates to conduct two counterfactual exercises that reveal how risk adjustment affects network breadth and consumer welfare. While incentivizing insurers to broaden their networks might seem desirable to improve access to care, broader networks are also associated with higher health care costs. The goal of my counterfactual analysis is to quantify the extent to which hospital networks respond to risk adjustment and the resulting pass-through to health care costs. In my counterfactual analyses, I hold long-run government spending, dropout probabilities, and transition probabilities fixed. Keeping government spending fixed allows changes in networks to be determined only by changes in how resources are redistributed across insurers but not by the level of the transfer itself.

For computational tractability, I conduct my counterfactual analyses in a single market: Bogotá, which is the capital city of Colombia and the largest market in the country. This market represents 29 percent of all continuously enrolled individuals in the contributory regime and has presence of all 10 insurers.

One concern in the counterfactual analyses is that the model may admit multiple equilibria in insurers' choices of network breadths. 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. Whether there are multiple equilibria in this market depends on the shape of the insurers' profit functions. While a direct proof of uniqueness is challenging, in appendix 12 I provide intuition for the sign of the second partial derivative of the insurers' profit function with respect to network breadth, all else equal. 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-level network breadth and confirm that they all converge to the same equilibrium.

6.1 No Risk Adjustment

I start by describing the effect of eliminating the risk adjustment systems. In this counterfactual scenario, the per capita transfer to each insurer equals the national base transfer from appendix table 1 times an adjustment factor λ calibrated to match observed long-run government spending:

$$R^{cf}_{\theta m} = \lambda \times \text{UPC}_{\text{National}}, \quad \forall(\theta, m)$$

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

	Variable	% change
A. Overall	Median network breadth	-1.9
	Avg. cost per enrollee	2.4
	Total avg. cost	2.7
	Consumer surplus (sick)	-1.1
	Consumer surplus (healthy)	-0.8
B. Service network breadth	Skull, spine, nerves, glands	-2.5
	Eyes, ears, nose, mouth	-2.3
	Pharynx, lungs	-1.5
	Heart and cardiac vessels	-0.5
	Lymph nodes, bone marrow	-1.5
	Esophagus, stomach and intestines	-2.6
	Liver, biliary tract	-2.1
	Abdominal wall	-2.7
	Urinary system	-1.0
	Reproductive system	-2.9
	Bones and facial joints	-3.4
	Joints, bones, muscles, tendons	-2.9
	Skin	-1.9
	Imaging, lab, consultation	0.3
	Hospital admission	2.5

Note: Panel A presents the percentage change in median network breadth across insurers, insurer total average costs, short-run average cost per enrollee, and short-run consumer welfare for the healthy and sick, between the no-risk adjustment scenario and the observed scenario. Panel B presents the percentage change of median network breadth by service category. I collapse the 58 original categories into 15 broader groups. The counterfactual is calculated with data from Bogotá only.

Failure to compensate for individuals' health risk should exacerbate risk selection, incentivizing insurers to drop coverage in services that unprofitable patients require. Panel A of table 8 presents the percentage change in median network breadth, insurer total average costs, short-run average costs per enrollee, and long-run consumer surplus for sick and healthy individuals between the counterfactual scenario and the observed scenario.[23] I find that under no

[23]Insurer total average costs are calculated as:
$\sum_{ij}\left(AC_{\theta jm}s_{ijm}+\sum_{s=t+1}^{T}\zeta^{s}\sum_{\theta'}(1-\rho_{\theta'})\mathcal{P}(\theta'|\theta)AC_{\theta' jm}s'_{ijm}\right)$,
Short-run average cost per enrollee is $(1/N_{\theta m})\sum_{ij}AC_{\theta jm}$,
and long-run consumer surplus is $\sum_i(-\alpha_i)^{-1}\log(\sum_j\exp(\beta_i\sum_k q_{\theta k}H_{jkm}-\alpha_i c_{\theta jm}+\phi_{jm}))+\sum_{s=t+1}^{T}\zeta^{s}\sum_{\theta'}(1-\rho_{\theta'})\mathcal{P}(\theta'|\theta)(-\alpha_i)^{-1}\log(\sum_j\exp(\beta_i\sum_k q_{\theta' k}H_{jkm}-\alpha_i c_{\theta' jm}+$

risk adjustment median network breadth falls by 1.9 percent from a baseline of 0.36. The reduction in coverage results in a 2.4 percent increase in average costs per enrollee and in a 2.7 percent increase in total average costs. The fact that average costs per enrollee rise indicates that the direct effect of network breadth is overcompensated by the effect of scope economies.

The reduction in network breadth generates large welfare effects despite decreases in out-of-pocket costs. The average elasticity with respect to out-of-pocket costs goes from -0.88 in the observed scenario to -0.95 in the counterfactual. Eliminating risk adjustment results in a 1.1 percent decrease in long-run consumer surplus for individuals with chronic conditions and in a 0.8 percent decrease in long-run surplus for healthy individuals. While the percentage changes seem rather small, a back-of-the-envelope calculation indicates that the reduction in welfare for sick individuals equals approximately one minimum wage per capita. The welfare effect on patients with diseases is greater in magnitude because access to and quality of care worsen, in particular for services that these patients a more likely to claim.

Panel B of table 8 shows that the reduction in coverage happens across relatively unprofitable services. For exposition purposes, I collapse the 58 service categories into 15 broader groups. When they are not compensated for their enrollees' health risk, insurers reduce coverage of relatively expensive services like procedures in skull by 2.5 percent and procedures in stomach by 2.6 percent. For cheaper services, such as imaging, lab, and consultations, median network breadth remains virtually unchanged.

These counterfactual results represent a lower bound (in absolute value) of the effect of eliminating risk adjustment when there is no significant vari-

---

$\phi_{jm}))\Big)$

ation in hospital quality in a market. In appendix table 22 I compare results from using my estimates to predict counterfactual outcomes in a market without variation in hospital quality, against predictions of a model that is both estimated and evaluated in a market without variation in hospital quality.

6.2 Improved Risk Adjustment

I now move to the opposite exercise where I improve the current risk adjustment formula by compensating for a list of diagnoses ex-ante. If allowing for variation in per capita transfers across diagnoses helps better predict health care costs, then risk selection incentives should decrease, resulting in broader networks. Results in this counterfactual will be suggestive of how strong demand-side selection incentives are relative to cost incentives in generating narrow networks in equilibrium, since improved risk adjustment effectively eliminates demand-side incentives.

I assume that the counterfactual risk-adjusted transfer is given by the annualized average cost per consumer type θ. More formally, this is:

$$R^{cf}_{\theta m} = \lambda \times a_m \times 360 \times \frac{\sum_{\theta(i)=\theta} T_i}{\sum_{\theta(i)=\theta} b_i}$$

where T_i is the total health care cost of individual i of type θ, b_i is the number of days enrolled to the contributory system during the year, a_m is the market multiplier from appendix table 1, and λ is an adjustment factor calibrated to match observed long-run government spending. I use two sets of diseases to compensate for ex-ante. The first is the list of 9 diseases used in the model of section 4. The second is a more granular list of 20 diseases presented in appendix table 21.[24] These conditions are obtained by grouping the ICD-10

[24]The more granular list of 20 diseases still imperfectly compensates insurers for an indi-

codes accompanying an individual's claims following Riascos et al. (2014).[25]

Table 9 presents the percentage change in median network breadth, insurer costs, and long-run consumer surplus under the improved risk adjustment formula with 9 diseases in column (1) and with 20 diseases in column (2). Effects on each of these variables are greater the more granular the risk adjustment formula is. Median network breadth increases 4.2 percent relative to the observed scenario in column (1) and 4.8 percent in column (2). Network breadth increases across all services as seen in panel B of the table, and effects are larger for services that mostly sick patients claim. This is consistent with weakened selection incentives and with demand-side adverse selection being a main factor in determining narrow networks.

Not only does a more granular risk-adjusted compensation generate broader networks, but the distribution of this effect is more even across services compared to a coarser formula. For instance, network breadth for procedures in abdominal wall increases 8.9 percent in column (1) and 4.6 percent in column (2), while network breadth for procedures in joints, bones, muscles, and tendons increases 1.8 percent in column (1) and 5.4 percent in column (2). This is because a more granular compensation requires more detailed estimates of dropout probabilities and transition probabilities across health states. However, implementing this type of compensation introduces a trade-off between more detailed estimates but less precise ones.

Insurers' total average cost increase 0.3 percent in column (2), suggesting that the direct effect of network breadth on costs overcompensates cost savings from scope economies. Consumer out-of-pocket costs also increase due to

viduals' health care cost compared, for example, to CMS' Hierarchical Conditions Categories risk adjustment formula in the Medicare program in the US, which controls for 79 conditions.

[25] See https://www.alvaroriascos.com/researchDocuments/healthEconomics/CLD_xCIE10.tab

Table 9: Networks, costs, and welfare under improved risk adjustment

	Variable	(1) 9 diseases	(2) 20 diseases
A. Overall	Median network breadth	4.2	4.8
	Avg. cost per enrollee	-0.9	-0.1
	Total avg. cost	-0.4	0.3
	Consumer surplus (sick)	2.2	2.8
	Consumer surplus (healthy)	2.0	2.7
B. Service network breadth	Skull, spine, nerves, glands	8.1	5.4
	Eyes, ears, nose, mouth	2.5	5.9
	Pharynx, lungs	5.0	4.7
	Heart and cardiac vessels	6.7	4.9
	Lymph nodes, bone marrow	5.1	5.4
	Esophagus, stomach and intestines	4.4	4.5
	Liver, biliary tract	6.8	5.5
	Abdominal wall	8.9	4.6
	Urinary system	4.3	5.7
	Reproductive system	5.2	5.5
	Bones and facial joints	4.1	4.9
	Joints, bones, muscles, tendons	1.8	5.4
	Skin	2.4	3.8
	Imaging, lab, consultation	4.5	3.2
	Hospital admission	4.4	2.9

Note: Panel A presents the percentage change in median network breadth across insurers, insurer total average costs, short-run average cost per enrollee, and short-run consumer welfare for the healthy and sick, between the improved risk adjustment scenarios and the observed scenario. Panel B presents the percentage change of median network breadth by service category. I collapse the 58 original categories into 15 broader groups. The counterfactual is calculated with data from Bogotá only. Column (1) shows results from an improved formula that compensates insurers ex-ante for sex, age, location, and a list of 9 diseases. Column (2) uses a list of 27 diseases in addition to sex, age group, and location.

changes in insurers' average cost. These changes imply an average elasticity of -0.65, which represents a 33 percent increase (in absolute value) over baseline.

With the improved risk adjustment formula that uses the list of 20 diseases, I find that long-run consumer surplus increases 2.8 and 2.7 percent for patients with any chronic condition and for healthy consumers, respectively. Both welfare changes are significantly larger than the ones estimated in column (1) using a risk adjustment formula that compensates for 9 diseases. Welfare effects in these counterfactuals are greater for patients with chronic diseases because their willingness-to-pay for network breadth is higher than that of healthy individuals. Note that even though greater network breadth increases out-of-pocket costs, the resulting marginal disutility is averaged-out with insurer competition. The direct effect of network breadth on welfare therefore dominates the effect of out-of-pocket costs.

Results in column (1) are qualitatively similar to the ones obtained from a model that is estimated and evaluated in a market without variation in hospital quality as seen in appendix table 22. Ignoring hospital quality therefore does not introduce significant bias in my predictions.

If allowing diagnoses to enter the ex-ante risk adjustment formula results in greater network coverage and welfare for patients most at need of care, at no extra cost for the government, why hasn't this formula been implemented in Colombia? First, there are information frictions that prevent a diagnosis-specific risk adjustment to have positive hospital network effects. Recall that risk-adjusted transfers for year t are calculated using claims data from year $t-2$, which might not be informative about the prevalence of diseases in t. Second, allowing for variation across diagnoses could incentivize insurers to engage in upcoding practices, which are difficult to observe and penalize.

7 The Effect of Premiums on Network Breadth

One outstanding question is whether regulating premiums –and hence forcing insurers to compete on other dimensions of the health insurance plan– effectively generates competition in risk selection. This is an interesting economic question since little is known about the interaction between price and non-price competition in health insurance markets, and since selection incentives can exist because of price regulation. In this section I study how hospital network breadth responds to premiums by simulating market outcomes under premium deregulation. I assume insurers compete Nash-Bertrand on premiums and are allowed to discriminate premiums based on the enrollee's sex, age group, and income group in each market. Insurers receive premiums in addition to the government's risk-adjusted transfers.

More formally, let $\theta = (\theta_1, \theta_2)$, where θ_1 corresponds to combinations of sex, age group, and income group, and where θ_2 represents diagnoses. Denote by $P_{\theta_1 jm}$ insurer j's premium in market m for consumer type θ_1. As with tax contributions, I assume the individual pays 1/3 of the premium and her employer pays 2/3. Let $\hat{P}_{\theta_1 jm} = (1/3) \times P_{\theta_1 jm}$. Individual i's choice probability for insurer j in market m is:

$$s_{ijm}(H_m) = \frac{\exp\left(\beta_i \sum_k q_{\theta k} H_{jkm} - \alpha_i \hat{P}_{\theta_1 jm} - \alpha_i c_{\theta jm} + \phi_{jm}\right)}{\sum_{j' \in \mathcal{J}_m} \exp\left(\beta_i \sum_k q_{\theta k} H_{j'km} - \alpha_i \hat{P}_{\theta_1 j'm} - \alpha_i c_{\theta j'm} + \phi_{j'm}\right)} \tag{8}$$

Although equation (8) implicitly assumes that the sensitivity of demand to premiums is the same as to coinsurance payments and copays, I conduct analyses allowing the average component of α_i across patients to be greater for premiums than for out-of-pocket costs. This is following Abaluck and Gruber (2011) who find that consumers are more responsive to premiums than to other measures of cost-sharing in the context of prescription drug coverage in Medicare Part D.[26]

Under counterfactual premiums, the short-run per enrollee profit is given

[26] I calibrate the fixed component of α_i associated to premiums in a way that generates similar average premium elasticities as in Abaluck and Gruber (2011), equal to -1.17, and as in Shepard (2022), equal to -1.48.

by:

$$\pi_{ijm}(H_m, P_m) = (R_{\theta m} + P_{\theta_1 jm} - (1 - r_i)AC_{\theta jm}(H_{jm}))s_{ijm}(H_m, P_m)$$

where $P_m = \{\{P_{\theta_1 jm}\}_{\theta_1(i)=\theta}\}_{j=1}^{\#\mathcal{J}_m}$. Insurers simultaneously choose premiums and network breadth per service taking into account the future profits associated to each new consumer that enrolls with it. The solution concept is a Nash equilibrium. Insurers choose premiums and networks to maximize:

$$\Pi_{jm}(H_m, P_m) = \sum_{\theta} \Big(\pi_{ijm}(H_m, P_m, \theta)N_{\theta m} + \sum_{s=t+1}^{T} \zeta^s \sum_{\theta'} (1 - \rho_{\theta' m})\mathcal{P}(\theta'|\theta)\pi_{ijm}(H_m, P_m, \theta')N_{\theta' m}\Big)$$
$$- \sum_{k} \Big(\omega_0 H_{jkm} + \xi_{jkm}\Big)H_{jkm}$$

The FOC with respect to premiums is:

$$\frac{\partial \Pi_{jm}}{\partial P_{\theta_1 jm}} = \sum_{i \in \theta_2} \left(\frac{\partial \pi_{ijm}}{\partial P_{\theta_1 jm}} N_{\theta m} + \sum_{s=t+1}^{T} \zeta^s \sum_{\theta'} (1 - \rho_{\theta' m})\mathcal{P}(\theta'|\theta) \frac{\partial \pi'_{ijm}}{\partial P_{\theta_1 jm}} N_{\theta' m} \right) = 0$$

where

$$\frac{\partial \pi_{ijm}}{\partial P_{\theta_1 jm}} = \Omega\Big(R_{\theta m} + \tilde{P}_{\theta_1 jm} - (1 - r_i)AC_{\theta jm}\Big) + s_{ijm}(H_m, P_m)$$

and

$$\Omega(j, j') = \begin{cases} -s_{ijm}(1 - s_{ijm})\alpha_i & \text{if } \; j = j' \\ 0 & \text{if } \; j \neq j' \end{cases}$$

The annual premium can be solved for from the FOC above, which determines a fixed point iteration in the vector of premiums.

The FOC with respect to the service-level network breadth is given by the equation below, from which H_{jkm} can also be solved for in a fixed point:

$$\frac{\partial \Pi_{jm}}{\partial H_{jkm}} = \sum_i \left(\frac{\partial \pi_{ijm}}{\partial H_{jkm}} N_{\theta m} + \sum_{s=t+1}^{T} \beta^t \sum_{\theta'} (1-\rho_{\theta' m}) \mathcal{P}(\theta'|\theta) \frac{\partial \pi'_{ijm}}{\partial H_{jkm}} N_{\theta' m} \right) - \left(2\omega_0 H_{jkm} + \xi_{jkm} \right) = 0$$

As before, I simulate this counterfactual with data from Bogotá for simplicity. Computation proceeds as a nested fixed point. For every guess of the equilibrium vector of network breadth, I solve for the fixed point in premiums in an inner loop. Then, in the outer loop, I solve for the fixed point in network breadth.

Table 10 presents the average annual premium in thousands of pesos conditional on demographics and insurer. The average component of α_i for premiums is set to 1.0 and 1.5 times the average component for out-of-pocket costs in columns (1) and (2), respectively. Focusing on column (2), I find that premiums are higher for males than for females, and hump-shaped in age. This suggests cross-subsidization from younger, relatively healthier individuals to older, relatively sicker individuals. Premium deregulation is a progressive policy since the annual premium is increasing with income. I also find a negative correlation between premiums and market share, and a large pass-through of hospital coverage to premiums, a finding similar to Cabral et al. (2018).[27]

Panel A of table 11 presents the percentage change in median network breadth, insurer total average costs, short-run average cost per enrollee, insurer total revenue, and long-run consumer welfare for sick and healthy individuals between the premium deregulation counterfactual and the observed

[27] For example, EPS005 charges a premium that is 171 thousand pesos higher than the premium for EPS037, but has a market share that is 10 percentage points lower.

scenario. The panel also shows the implied average elasticity with respect to premiums.[28] Consumer welfare in the counterfactual is monetized using the estimate of consumer sensitivity to out-of-pocket costs rather than the estimate for premiums to allow for a fair comparison with the observed scenario. The fixed component of demand sensitivity to premiums is set to 1.0 and 1.5 times the fixed component of out-of-pocket costs in columns (1) and (2), respectively.

Deregulating premiums incentivizes insurers to offer nearly complete networks. Median network breadth increases 24.2 percent in column (1) and 21.2 percent in column (2). The model thus implies that the price and the non-price elements of insurance contracts in this market are substitutes from the point of view of risk selection. If allowed to charge premiums, insurers would cream-skim the market using premiums rather than service-level networks. The effect of premium competition on network breadth decreases with higher values of α_i (in absolute value). The average elasticity with respect to premiums is higher (in absolute value) in column (2) compared to column (1). The increase in coverage happens across all services as seen in panel B of the table. While relatively expensive services such as hospital admissions see smaller increases in median network breadth compared to relatively cheap services such imaging, lab, and consultations, the effect is still substantial.

I find that deregulating premiums generates a significant transfer of surplus from consumers to insurers. Insurer total revenues increase nearly 33 percent in this counterfactual at the expense of consumers with and without diagnoses, for whom welfare falls 6.7 and 5.5 percent, respectively, in column (1). In the case of healthy individuals, who have a low willingness-to-pay for

[28] The individual elasticity with respect to annual premiums is calculated as $\frac{\partial s_{ijm}}{\partial P_{\theta jm}} \frac{P_{\theta jm}}{s_{ijm}}$ and then averaged across individuals.

network breadth and a high sensitivity to out-of-pocket costs, welfare losses due to higher out-of-pocket payments overcompensate welfare gains from having broader networks in every service. For healthy consumers and consumers with chronic diseases, welfare effects depend on how the sensitivity of demand to premiums is calibrated.

The fall in consumer welfare allowing for premiums is problematic. If healthy individuals could drop out of the health care system, a premium deregulation policy could unravel the market. By making the healthy disproportionately choose uninsurance, broad network carriers would face significant uninsurable costs from the remaining enrollees with chronic diseases. Market equilibria under premium deregulation could be restored not only by making enrollment mandatory, but also by having the government pay a fraction of premiums. These types of premium subsidy policies have been studied in the context of the Health Insurance Exchanges in the United States (e.g Tebaldi, 2017), but are out of the scope of this paper.

Table 10: Average annual premium

		(1) Low	(2) High
Sex	Female	127	120
	Male	221	269
Age group	19-24	173	154
	25-29	171	161
	30-34	183	220
	35-39	288	298
	40-44	280	305
	45-49	243	284
	50-54	212	232
	55-59	187	232
	60-64	132	192
	65-69	121	172
	70-74	53	49
	75 or more	45	34
Income group	Low	81	61
	Medium	178	200
	High	—	—
Insurer	EPS001	111	192
	EPS002	105	136
	EPS003	94	119
	EPS005	484	437
	EPS010	57	92
	EPS013	223	237
	EPS016	120	173
	EPS017	185	176
	EPS018	70	116
	EPS037	291	266

Note: Table presents average annual premium in thousands of pesos conditional on demographic characteristics and insurer. Columns (1) and (2) set the average component of α_i for premiums equal to 1 and 1.5 times the average component for out-of-pocket costs, respectively.

Table 11: Networks, costs, and welfare under premium deregulation

		(1) Low	(2) High
A. Overall	Median network breadth	24.2	21.2
	Avg. cost per enrollee	9.2	8.1
	Total avg. cost	8.2	6.4
	Consumer surplus (sick)	-6.7	-9.6
	Consumer surplus (healthy)	-5.5	-8.4
	Elasticity	-1.8	-2.0
B. Service network breadth	Skull, spine, nerves, glands	29.3	23.7
	Eyes, ears, nose, mouth	30.8	31.6
	Pharynx, lungs	20.7	15.5
	Heart and cardiac vessels	21.7	11.6
	Lymph nodes, bone marrow	26.8	18.7
	Esophagus, stomach and intestines	20.7	17.1
	Liver, biliary tract	26.5	30.7
	Abdominal wall	24.4	22.7
	Urinary system	23.1	15.0
	Reproductive system	28.1	23.5
	Bones and facial joints	26.5	21.4
	Joints, bones, muscles, tendons	26.3	25.8
	Skin	20.1	16.7
	Imaging, lab, consultation	20.1	23.1
	Hospital admission	11.0	7.1

Note: Panel A presents the percentage change in median network breadth across insurers, insurer total average costs, short-run average cost per enrollee, and short-run consumer welfare for the healthy and sick, between the premium deregulation scenario and the observed scenario. Panel B presents the percentage change of median network breadth by service category. I collapse the 58 original categories into 15 broader groups. The counterfactual is calculated with data from Bogotá only. The average component of α_i for premiums is set to 1 and 1.5 times the average component for out-of-pocket costs in columns (1) and (2), respectively.

8 Conclusions

Risk selection is a main concern in health insurance systems where governments make risk-adjusted payments to insurers. In this paper I show that health insurers can engage in risk selection using their hospital networks. I model insurer competition in service-level hospital networks in the context of the Colombian health care system.

In Colombia, the government sets premiums to zero and compensates private insurers with per-capita risk-adjusted payments that control only for sex, age, and location. Insurers in this market have discretion over which services to cover at which hospitals, making hospital networks service-specific. Therefore, conditional on risk adjustment, insurers engage in risk selection by offering narrow networks in unprofitable services. I use my model of insurer competition to measure the effect of typical policies used to combat risk selection, such as risk adjustment and premium setting, on service-level network breadth.

I find that eliminating risk adjustment makes insurer competition a race to the bottom in terms of network breadth. Median network breadth falls 1.9 percent, and the reduction is largest in services that sick individuals require the most. As a result consumer surplus falls by 1.1 percent relative to the observed scenario. Improving the risk adjustment formula by compensating insurers for sex, age, location, and a list diagnoses results in broader service-level networks. With an improved risk adjustment formula, median network breadth increases between 4.2 and 4.8 percent, and consumer surplus increases between 2.2 and 2.8 percent depending on how many risk factors are included in the formula. Allowing insurers to compete Nash-Bertrand in premiums results in substantially broader hospital networks for every insurer and service. Thus the price and the non-price elements of insurance contracts are substitutes from the point of view of risk selection.

In quantifying the extent to which networks respond to risk adjustment and premiums, the findings of this paper speak to the trade-off between having better access to care and containing health care costs. Findings can help policymakers in the design of public health systems with private provision of health insurance. But policy implications extend beyond these types markets to ones where insurers compete on the non-price characteristics of their health insurance plans.

References

Abaluck, J. and Gruber, J. (2011). Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. *American Economic Review*, 101(4):1180–1210.

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.

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

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

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.

Decarolis, F. (2015). Medicare Part D: Are Insurers Gaming the Low Income Subsidy Design? *American Economic Review*, 105(4):1547–1580.

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

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.

Einav, L., Finkelstein, A., and Tebaldi, P. (2019). Market Design in Regulated Health Insurance Markets: Risk Adjustment vs. Subsidies.

Ellis, R. and McGuire, T. (2007). Predictability and Predictiveness in Health Care Spending. *Journal of Health Economics*, 26(1):25–48.

Finkelstein, A., Hendren, N., and Shepard, M. (2019). Subsidizing Health Insurance for Low-Income Adults: Evidence from Massachusetts. *American Economic Review*, 109(4):1530–1567.

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., 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.

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.

Haeder, S., Weimer, D., and Mukamel, D. (2015). California hospital networks are narrower in marketplace than in commercial plans, but access and quality are similar. *Health Affairs*, 34(5):741–748.

Ho, K. and Lee, R. (2017). Insurer Competition in Health Care Markets. *Econometrica*, 85(2):379–417.

Kong, E., Layton, T., and Shepard, M. (2022). Adverse Selection Pricing and Unraveling of Competition in Insurance Markets.

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. and Panhans, M. T. (2021). Why do narrow network plans cost less? *Health Economics*, 30(10):2437–2451.

Mattocks, K., Elwy, A., Yano, E., Giovannelli, J., Adelberg, M., Mengeling, M., Cunningham, K., and Matthews, K. (2021). Developing network adequacy standards for va community care. *Health Services Research*, 56(3):400–408.

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.

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.

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.

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.

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

Tebaldi, P. (2017). Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA.

For Online Publication

Appendix 1 Current risk adjustment system

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

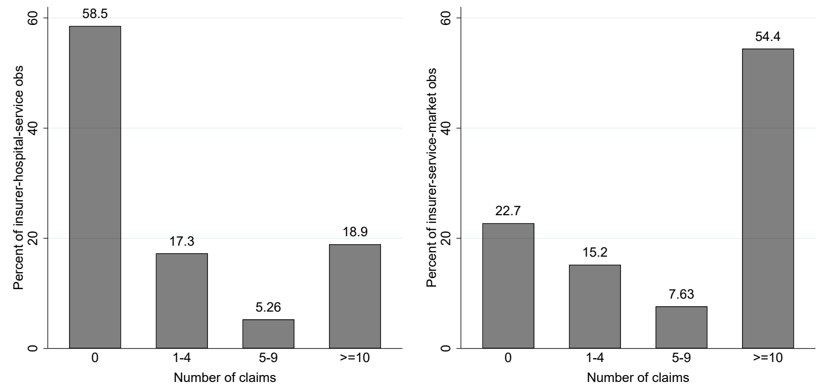
Department/city		Transfer
National (pesos)		505,627.2
Market multiplier a_m	Amazonas	$\times$ 1.10
	Arauca, Arauca	$\times$ 1.10
	Yopal, Casanare	$\times$ 1.10
	Florencia, Caquetá	$\times$ 1.10
	Chocó	$\times$ 1.10
	Riohacha, Guajira	$\times$ 1.10
	Guainía	$\times$ 1.10
	Guaviare	$\times$ 1.10
	Villavicencio, 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

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

Appendix 2 Service categories

Service code	Description
01	Procedures in skull, brain, and cerebral meninges
03	Procedures in spinal cord and structures of spine
04	Procedures in peripheral and skull nerves
05	Procedures in nerves or sympathetic ganglia
06	Procedures in thyroid and parathyroid gland
08	Procedures in eyelids and lacrimal apparatus
10	Procedures in conjunctive, cornea, iris, retina, orbit
18	Procedures in ear
21	Procedures in nose and paranasal sinuses
23	Procedures in teeth, tongue, salivary glands
27	Procedures and interventions in mouth and face
28	Procedures in tonsils and adenoids
29	Procedures in pharynx, larynx, trachea
32	Procedures in lung and bronchus
34	Procedures in thoracic wall, pleura, mediastinum, diaphragm
35	Procedures in heart valves
36	Procedures in cardiac vessels
37	Procedures in heart and pericardium
38	Procedures in blood vessels
40	Procedures in lymphatic system
41	Procedures bone marrow and spleen
42	Procedures in esophagus
43	Procedures in stomach
45	Procedures in intestines
47	Procedures in appendix
48	Procedures in rectum, rectosigmoid, perirectal tissue
50	Procedures in liver
51	Procedures in gallbladder and biliary tract
52	Procedures in pancreas
53	Procedures in abdominal wall
55	Procedures in kidney
56	Procedures in ureter
57	Procedures in bladder
58	Procedures in urethra and urinary tract
60	Procedures in prostate, seminal vesicles, scrotum, testicles, penis
65	Procedures in ovaries, fallopian tubes, cervix, uterus
70	Procedures in vagina and cul-de-sac
72	Procedures and interventions in vaginal delivery
76	Procedures in bones and facial joints
79	Reduction of fracture and dislocation
80	Procedures in joint structures
81	Repair procedures and plasties in joint structures
82	Procedures in tendons, muscles, and hand fascia
83	Procedures in muscle, tendon, fascia, bursa except hand
85	Procedures in breast
86	Diagnostic procedures in skin and subcutaneous cellular tissue
87	Radiology and non-radiology imaging
89	Consultation, anatomic measures, physiology, manual tests, and pathology
90	Laboratory
91	Blood bank and transfusion medicine
92	Nuclear medicine and radiotherapy
93	Procedures and interventions in functional development and rehabilitation
94	Procedures related to mental health
95	Non-surgical procedures and interventions related to eye and ear
97	Substitution and extraction of therapeutic devices
98	Non-surgical extraction of kidney stones
99	Prophylactic and therapeutic procedures
S1	Inpatient services



Appendix Figure 1: Distribution of number of claims

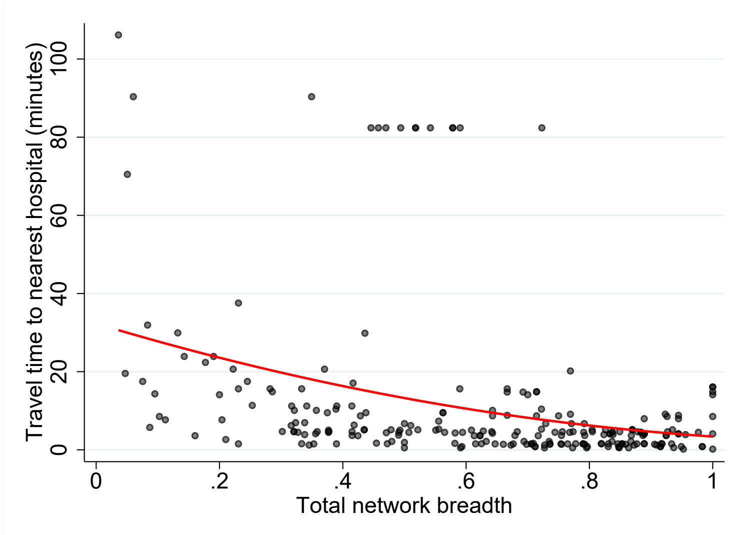
Note: Distribution of number of claims by hospital-insurer-service in the left panel and by insurer-service in the right panel.

Appendix 3 Variation in hospital quality

Appendix Table 3: Network breadth and hospital quality

	Star hospital	Patient satisfaction	Inverse mortality
Network breadth	0.15	0.002	0.16
	(0.01)	(0.003)	(0.01)
R^2	0.67	0.43	0.43
N	13,572	10,195	9,949

Note: Regression of star hospital coverage indicator, patient satisfaction, and inverse inpatient mortality rate after 48 hours on service-level network breadth. Patient satisfaction levels and mortality rates per hospital are obtained from the National Health Superintendency's 2011 hospital quality measures (https://docs.supersalud.gov.co/PortalWeb/SupervisionInstitucional/IndicadoresCalidadEPS/Indicadores-Calidad-IPS-consolidado-2011.xlsx). All regressions include insurer, market, and service fixed effects. Robust standard errors in parenthesis.



Appendix Figure 2: Correlation between total network breadth and travel time

Note: Scatter plot of network breadth over all services and travel time from the municipality centroid to the nearest in-network hospital in minutes. The red line represents a quadratic fit.

Appendix 4 Selection vs. Moral Hazard

To separate selection from moral hazard, I start by estimating a regression on the sample of current enrollees who received a diagnosis in 2010. The dependent variable is the number of claims or an indicator for making a claim during 2011 in a service that is associated with treatment of their health condition diagnosed in 2010. The regression specification is given by:

$$y_{ijm}^{k,2011} = \beta_0 + \beta_1 H_{jm}^{k,2011} + \mathbf{d}_i^{2010}\beta_2 + \gamma_m + \varepsilon_{ijm},$$

where $H_{jm}^{k,2011}$ is insurer j's network breadth for service k during 2011, $\mathbf{d}_i^{2010}$ is a vector of demographics and diagnoses received during 2010, and γ_m is a market fixed effect. This specification captures the extent of selection into moral hazard as in Einav et al. (2013). A positive correlation between service-specific network breadth and number of claims or probability of a claim for that service in $t+1$ would be suggestive of patients enrolling in carriers with more generous coverage for services that they anticipate needing given their health conditions in t.

Results presented in table 4 are suggestive of this source of selection. Column (1) shows that the probability of childbirth in 2011 among women who were in childbearing age during 2010 is increasing in network breadth for delivery services. The probability of dialysis claims, antirheumatic drug claims, and chemotherapy claims are also positively correlated with network breadth for dialysis, procedures in bones and joints, and chemotherapy, respectively. Column (2) uses the full sample of individuals enrolled in 2011 as a robustness check.

Appendix Table 4: Service-specific network breadth and types of claims

Dep var	Indep var	(1) Current	(2) Full
Any childbirth claim	H_{jm} Delivery	1.66	1.45
		(0.12)	(0.07)
Any dialysis claim	H_{jm} Dialysis	2.83	2.51
		(0.37)	(0.30)
log(Dialysis claim)	H_{jm} Dialysis	8.52	7.47
		(0.87)	(0.69)
Any anti-rheumatic drug claim	H_{jm} Therapy	0.16	0.18
		(0.10)	(0.08)
log(Anti-rheumatic drugs)	H_{jm} Therapy	0.27	0.28
		(0.16)	(0.12)
Any chemotherapy claim	H_{jm} Therapy	0.29	-0.19
		(0.17)	(0.12)
log(Chemotherapy claim)	H_{jm} Therapy	-0.16	-0.31
		(0.31)	(0.21)

Note: OLS regressions of the probability of childbirth, any dialysis claim, any antirheumatic drug claim, and any chemotherapy claim during 2011, on service-level network breadth, conditional on the sample of individuals who received a diagnosis during 2010. Column (1) uses the sample of current enrollees and column (2) uses the full sample without constraining enrollment to be continuous. All regressions include market fixed effects and control for sex and age group. Robust standard errors in parenthesis. Coefficients and standard errors are multiplied by 100.

While the positive correlation between network breadth and number of claims conditional on 2010 health status could be explained by moral hazard rather than by selection, the correlation with the likelihood of making claims is unlikely the result of moral hazard. For example, given that all patients with renal disease make at least one dialysis claim, the positive correlation between network breadth and likelihood of a dialysis claim is not due to moral hazard. The positive correlation may also be unlikely the result of adverse selection as it requires individuals knowing if and when they will develop a disease. However, focusing on the admittedly small sample of switchers, appendix table 5 shows that consumers whose health status changes over time tend to switch towards the insurer that has the broadest network for services they need given their newly diagnosed conditions relative to their incumbent insurer.

Appendix Table 5: Insurer choice among switchers with changes in health status

Sample	Variable	Insurer choice
Women in childbearing ages	$H_{jm}^{2010} - H_{j'm}^{2011}$ Delivery	-2.87
		(0.27)
	N	2,676
New renal patient	$H_{jm}^{2010} - H_{j'm}^{2011}$ Dialysis	-1.19
		0.74
	N	108
New cancer patient	$H_{jm}^{2010} - H_{j'm}^{2011}$ Therapy	-2.98
		0.32
	N	2,363
New arthritis patient	$H_{jm}^{2010} - H_{j'm}^{2011}$ Procedures in bones	-2.29
		0.63
	N	494

Note: Table presents results of a conditional logit estimated by maximum likelihood on the sample of switchers that are newly diagnosed in 2011. The main explanatory variable is the difference in network breadth between the incumbent insurer j and all other insurers j'. Robust standard errors in parenthesis.

To isolate the effect of risk selection or cream-skimming, I explore whether insurers' network breadth choices are correlated with their enrollees' baseline costs and risk scores. I estimate a regression in the spirit of Brown et al. (2014). I compare baseline costs of switchers into insurers that reduce their network breadth over time to baseline costs of stayers in insurers that expand their network breadth. By focusing on baseline costs rather than current costs as outcome, this analysis also separates risk selection from moral hazard.

Given that the fraction of switchers in my data is very small, these exercises will only be suggestive of the effectiveness of risk selection. The regression specification is as follows:

$$y_{ikm}^{2010} = \beta_0 + \beta_1 (H_{jkm}^{2010} - H_{jkm}^{2011}) + \beta_2 S_{im} + \beta_3 S_{im} \times (H_{jkm}^{2010} - H_{jkm}^{2011}) + \mathbf{d}_i \beta_4 + \lambda_k + \delta_j + \eta_m + \varepsilon_{ikm}$$

Here y_{ikm}^{2010} is either the logarithm of total health care cost of individual i in service k and market m during 2010 or an indicator for having non-zero claims. S_{im} is an indicator for whether the consumer switched carriers from one year to the other. The subscript j denotes the insurer chosen in 2011, so H_{jkm}^{2010} is the 2010 network breadth of insurer j' and H_{jkm}^{2011} is the 2011 network breadth of insurer j. $\mathbf{d}_i$ is a vector of demographics and diagnoses, λ_k is a service fixed effect, δ_j is an insurer fixed effect, and η_m is a market fixed effect. The coefficient of interest is β_3.

The choice of service-specific network breadth is an effective risk selection mechanism on enrollee's baseline costs. Column (1) of table 6 shows that individuals who switch into carriers that reduce their network coverage over time tend to be less costly in that service than individuals who do not switch. Results in column (2) for the probability of making a claim in each service are consistent with this finding.[29]

[29] Results in column (1) of table 6 are robust to alternative modelling specifications such

Appendix Table 6: Selection on baseline costs and risk

	(1) $\log(\text{total cost}^{2010}_{ijkm}+1)$	(2) any $\text{claim}^{2010}_{ijkm}$	(3) $\log(\text{risk transfer}^{2011}_{new})$
$H^{2010}_{j'km} - H^{20101}_{j'km}$	0.44	-0.00	-16.8
	(0.24)	(0.02)	(0.78)
Switch	-8.61	-0.74	
	(1.52)	(0.14)	
Switch$\times(H^{2010}_{j'km} - H^{20101}_{j'km})$	-23.5	-2.05	
	(8.02)	(0.73)	
Demog+Diag	Y	Y	—
Market	Y	Y	Y
Service	Y	Y	—
Insurer	Y	Y	Y
N	14,457,009	14,457,009	2,653,415
R^2	0.50	0.51	0.06

Note: Column (1) presents a regression of log 2010 service costs on a switching indicator and the difference in network breadth across years for the insurer chosen in 2011. Column (2) presents a regression of making any service claims. Both columns include demographics and diagnoses indicators, as well as insurer, service, and market fixed effects. Column (3) presents a regression of log of new enrollees' risk-adjusted transfer. Robust standard errors in parenthesis. Coefficients and standard errors are multiplied by 100.

Appendix 5 Trade-offs to Network Breadth

While the asymmetric equilibrium in network breadth in the observed scenario can be explained by possible heterogeneity in the costs of network formation, it is important to note that broad networks attract more of every kind of patient. In Appendix Table 7 I estimate the correlation between consumer choice and network breadth during 2011 using the following linear regression at the insurer-market level:

$$s^k_{jm} = \beta_0 + \beta_1 H^k_{jm} + \gamma_m + \varepsilon^k_{jm}$$

as using a two-part model for baseline costs, with a first stage logit for the probability of having non-zero cost, and a second stage log-linear regression conditional on having non-zero cost. Results also hold when defining changes in network breadth between the insurer chosen in 2010 (j') and the one chosen in 2011 (j).

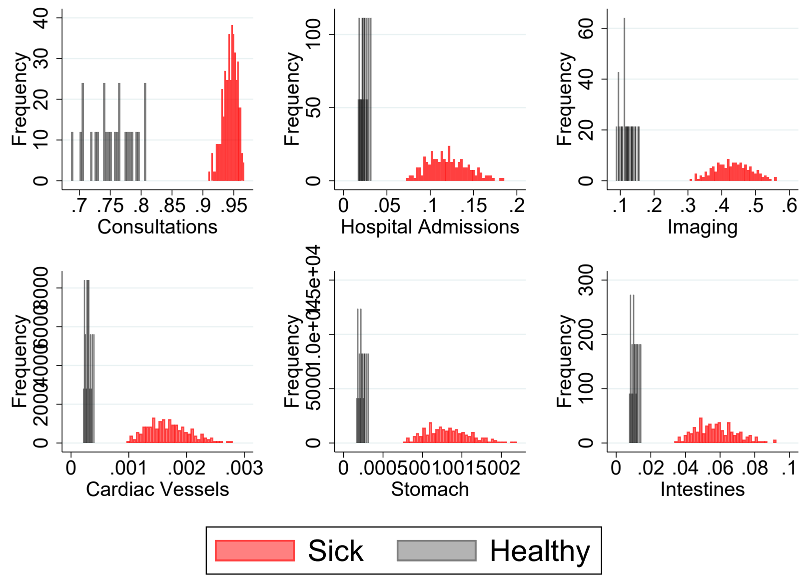
Here s_{jm}^k is insurer j's market share in the number of patients with any disease, no diseases, renal disease, cancer, arthritis, pregnancy, or cardiovascular disease. H_{jm}^k is either average network breadth across all services, or network breadth for dialysis, chemotherapy and radiotherapy, procedures in bones and facial joints, delivery, and procedures in heart, respectively. Results show that, relative to narrow network carriers, insurers that offer broad networks have higher demand from patients with chronic diseases who are usually unprofitable, but also higher demand from healthy individuals who are profitable. Thus, consumers have strong preferences for broader networks and this preference is heterogeneous across individuals.

Appendix Table 7: Correlation between market share and service-level network breadth

Market share in	Variable	(1) Stayers	(2) New	(3) Full
Any disease	H_{jm} average	0.57	0.56	0.59
		(0.03)	(0.03)	(0.05)
Healthy	H_{jm} average	0.58	0.57	0.58
		(0.04)	(0.03)	(0.04)
Renal disease	H_{jm} dialysis	0.37	0.41	0.39
		(0.04)	(0.05)	(0.04)
Cancer	H_{jm} therapy	0.44	0.43	0.45
		(0.03)	(0.05)	(0.04)
Arthritis	H_{jm} procedures in bones	0.41	0.38	0.41
		(0.04)	(0.03)	(0.04)
Childbirth	H_{jm} delivery	0.50	0.48	0.50
		(0.04)	(0.04)	(0.04)
Cardiovascular	H_{jm} procedures in heart	0.46	0.46	0.45
		(0.04)	(0.04)	(0.05)
N		424	424	424

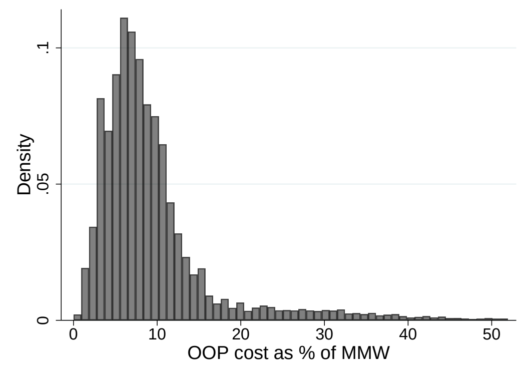
Note: Regressions of insurer market share in the number of patients with any disease, no diseases, renal disease, cancer, arthritis, childbirth, and cardiovascular disease, on service-level network breadth during 2011. Column (1) uses the sample of stayers from those who are continuously enrolled, column (2) uses new enrollees from those who are continuously enrolled, and column (3) uses the full sample without constraining enrollment to be continuous. All regressions include market fixed effects. Robust standard errors in parenthesis.

Appendix 6 Description of variables in demand model



Appendix Figure 3: Distribution of service claim probability

Note: Distribution of the probability of making a claim in 6 service categories separately for sick and healthy individuals. The 6 services are consultations, hospital admissions, imaging, and procedures in cardiac vessels, stomach, and intestines.



Appendix Figure 4: Out-of-pocket costs as percentage of monthly minimum wage

Note: Distribution of out-of-pocket costs conditional on observed insurer choices as percentage of the monthly minimum wage (MMW) in 2011. The average out-of-pocket cost equals 17% of the MMW. Of the total variation in out-of-pocket costs, 33% comes from consumer types and 3% from insurers.

Appendix 7 Additional demand results

7.1 Out-of-pocket cost and average cost per enrollee

Appendix Table 8: Pass-through of average costs to out-of-pocket costs

	Out-of-pocket cost		
	(1) Low income	(2) Middle income	(3) High income
$AC_{\theta jm}$	11.21	13.02	18.59
	(0.00)	(0.01)	(0.15)
Constant	-0.17	1.14	0.62
	(0.00)	(0.01)	(0.12)
N	178,533	318,588	2,879
R^2	0.97	0.88	0.84

Note: Regression of out-of-pocket costs on observed average cost per enrollee conditional on observed insurer choices. Column (1) uses the sample of individuals earning less than 2 times the monthly minimum wage (MMW), column (2) uses the sample of individuals earning between 2 and 5 times the MMW, and column (3) uses the sample of individuals earning more than 5 times the MMW. Robust standard errors in parenthesis.

Appendix Table 9: National market shares

Insurer	Observed	Predicted
EPS001	2.15	2.16
EPS002	7.23	7.28
EPS003	3.94	3.94
EPS005	4.39	4.41
EPS008	4.03	4.04
EPS009	2.10	2.09
EPS010	6.87	6.85
EPS012	1.17	1.19
EPS013	14.64	14.61
EPS016	19.51	19.47
EPS017	6.30	6.29
EPS018	3.91	3.86
EPS023	2.00	2.00
EPS037	21.75	21.79

7.2 In-sample demand model fit

7.3 Robustness checks

Appendix Table 10: Insurer demand with star hospital indicator

Variable		Network Breadth	OOP spending	Star hospital
Mean		2.69 (0.19)	-11.5 (0.26)	4.53 (0.37)
Interactions				
Demographics	Male	0.39 (0.02)	0.88 (0.13)	
	Age 19-24	1.86 (0.06)	-0.21 (0.47)	
	Age 25-29	2.62 (0.07)	2.51 (0.26)	
	Age 30-34	2.21 (0.06)	1.63 (0.31)	
	Age 35-39	1.82 (0.06)	0.45 (0.41)	
	Age 40-44	1.62 (0.06)	1.50 (0.37)	
	Age 45-49	1.34 (0.06)	1.14 (0.30)	
	Age 50-54	1.02 (0.06)	1.26 (0.32)	
	Age 55-59	0.97 (0.07)	1.49 (0.30)	
	Age 60-64	0.68 (0.07)	1.01 (0.29)	
	Age 65-69	0.58 (0.07)	0.56 (0.29)	
	Age 70-74	0.47 (0.07)	0.94 (0.29)	
	Age 75 or more	(ref)	(ref)	
Diagnoses	Cancer	-0.09 (0.07)	5.68 (0.26)	
	Cardiovascular	-0.42 (0.05)	4.63 (0.24)	
	Diabetes	-0.32 (0.12)	5.36 (0.45)	
	Renal	0.06 (0.27)	8.24 (0.17)	
	Pulmonary	-0.49 (0.19)	7.50 (0.32)	
	Arthritis	-0.32 (0.12)	7.68 (0.29)	
	Asthma	-0.32 (0.24)	8.56 (0.50)	
	Other	-0.99 (0.16)	7.16 (0.25)	
	Healthy	(ref)	(ref)	
Location	Normal	3.69 (0.04)	2.02 (0.16)	
	Special	5.46 (0.08)	0.97 (0.30)	
	Urban	(ref)	(ref)	
Income	Low	0.31 (0.03)	-1.15 (0.22)	
	High	(ref)	(ref)	
Pseudo-R2			0.23	
N			5,850,849	

Note: 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 11: Insurer demand excluding the capital city

Variable		Network Breadth	OOP spending
Mean		1.47 (0.20)	-11.9 (0.34)
Interactions			
Demographics	Male	0.29 (0.02)	0.59 (0.15)
	Age 19-24	1.32 (0.07)	0.58 (0.62)
	Age 25-29	2.05 (0.07)	3.69 (0.35)
	Age 30-34	1.72 (0.07)	2.67 (0.38)
	Age 35-39	1.37 (0.07)	0.97 (0.52)
	Age 40-44	1.14 (0.07)	2.82 (0.46)
	Age 45-49	0.97 (0.07)	2.31 (0.38)
	Age 50-54	0.74 (0.07)	2.59 (0.37)
	Age 55-59	0.73 (0.07)	2.65 (0.38)
	Age 60-64	0.50 (0.07)	2.02 (0.39)
	Age 65-69	0.44 (0.08)	1.20 (0.39)
	Age 70-74	0.46 (0.08)	1.41 (0.39)
	Age 75 or more	(ref)	(ref)
Diagnoses	Cancer	-0.11 (0.07)	5.83 (0.28)
	Cardiovascular	-0.28 (0.05)	4.56 (0.29)
	Diabetes	-0.10 (0.13)	4.94 (0.48)
	Renal	0.27 (0.28)	7.42 (0.22)
	Pulmonary	0.14 (0.20)	6.93 (0.41)
	Arthritis	-0.01 (0.12)	7.37 (0.31)
	Asthma	-0.14 (0.24)	8.18 (0.39)
	Other	-0.55 (0.16)	6.57 (0.27)
	Healthy	(ref)	(ref)
Location	Normal	3.69 (0.04)	2.30 (0.24)
	Special	5.43 (0.08)	1.10 (0.35)
	Urban	(ref)	(ref)
Income	Low	0.47 (0.03)	-1.27 (0.25)
	High	(ref)	(ref)
N		3,942,553	
Pseudo-R^2		0.27	

Note: Insurer choice model estimated in the sample of markets excluding the capital city where there is substantial variation in hospital quality. Specification includes insurer-by-market fixed effects. Robust standard errors in parenthesis.

Appendix Table 12: Insurer demand with alternative network measures

Variable		(1) Largest hospitals		(2) All providers	
		Network	OOP	Network	OOP
Mean		2.88 (0.22)	-10.6 (0.26)	1.35 (0.22)	-9.41 (0.26)
Interactions					
Demographics	Male	0.39 (0.03)	0.56 (0.13)	0.25 (0.03)	0.35 (0.13)
	Age 19-24	1.19 (0.08)	-0.44 (0.47)	0.64 (0.10)	-0.33 (0.44)
	Age 25-29	1.12 (0.09)	2.12 (0.26)	1.82 (0.11)	1.88 (0.25)
	Age 30-34	0.68 (0.08)	1.36 (0.31)	1.77 (0.11)	1.15 (0.29)
	Age 35-39	0.41 (0.08)	0.13 (0.41)	1.63 (0.11)	-0.83 (0.47)
	Age 40-44	0.40 (0.08)	1.30 (0.36)	1.34 (0.11)	0.77 (0.36)
	Age 45-49	0.60 (0.08)	1.02 (0.29)	1.26 (0.11)	0.79 (0.29)
	Age 50-54	0.46 (0.08)	1.17 (0.31)	0.97 (0.11)	1.08 (0.29)
	Age 55-59	0.58 (0.08)	1.30 (0.30)	0.96 (0.11)	1.22 (0.29)
	Age 60-64	0.62 (0.09)	0.87 (0.30)	0.8 (0.12)	0.80 (0.29)
	Age 65-69	0.66 (0.09)	0.43 (0.30)	0.79 (0.13)	0.38 (0.30)
	Age 70-74	0.83 (0.10)	0.69 (0.30)	0.83 (0.14)	0.60 (0.29)
	Age 75 or more	(ref)	(ref)	(ref)	(ref)
Diagnoses	Cancer	-0.04 (0.10)	5.33 (0.27)	0.26 (0.11)	4.95 (0.27)
	Cardiovascular	-0.77 (0.07)	4.41 (0.24)	-0.54 (0.08)	4.66 (0.25)
	Diabetes	-0.74 (0.15)	4.69 (0.48)	-0.56 (0.21)	4.85 (0.48)
	Renal	0.32 (0.38)	7.90 (0.17)	-0.24 (0.44)	7.70 (0.16)
	Pulmonary	-1.14 (0.20)	7.18 (0.29)	-0.89 (0.25)	7.07 (0.28)
	Arthritis	-0.51 (0.14)	7.58 (0.27)	-0.56 (0.17)	7.40 (0.26)
	Asthma	-1.05 (0.30)	8.19 (0.47)	-0.34 (0.38)	8.09 (0.45)
	Other	-1.11 (0.19)	7.02 (0.24)	-1.34 (0.26)	7.10 (0.21)
	Healthy	(ref)	(ref)	(ref)	(ref)
Location	Normal	4.98 (0.07)	1.58 (0.16)	5.49 (0.07)	0.97 (0.14)
	Special	12.4 (0.23)	0.67 (0.27)	9.07 (0.15)	-0.22 (0.56)
	Urban	(ref)	(ref)	(ref)	(ref)
Income	Low	0.98 (0.04)	-1.15 (0.21)	0.84 (0.05)	-1.10 (0.21)
	High	(ref)	(ref)	(ref)	(ref)
Pseudo-R^2		0.23		0.23	
N		5,850,849		5,850,849	

Note: Insurer choice model under alternative specifications of network breadth. Column (1) reports coefficients and standard errors in parenthesis of a model where network breadth is constructed from a sample of the largest hospitals in each market. Large hospitals are defined as having number of beds above the 70th percentile of the distribution of beds in each market. There are 314 hospitals under this definition. Column (2) presents coefficients and standard errors in parenthesis of a model where network breadth is constructed from the sample of all institutional providers of which there are 16,609. Specifications include insurer-by-market fixed effects.

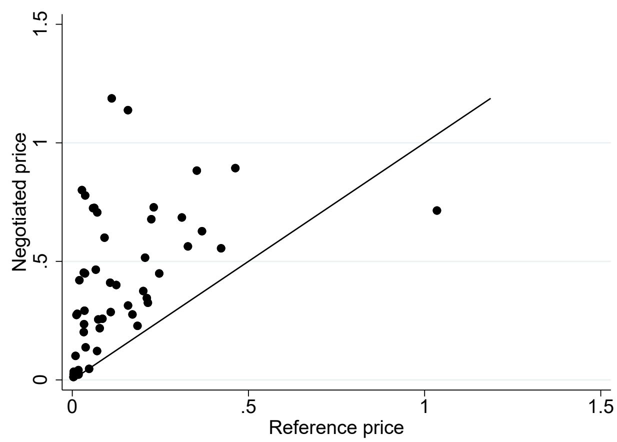
Appendix Table 13: Insurer demand with diagnosis in January

Variable		Network Breadth	OOP spending
Mean		4.25 (0.23)	-11.1 (0.31)
Interactions			
Demographics	Male	0.40 (0.02)	0.86 (0.14)
	Age 19-24	2.02 (0.06)	-1.11 (0.58)
	Age 25-29	2.76 (0.07)	2.92 (0.42)
	Age 30-34	2.47 (0.06)	2.25 (0.75)
	Age 35-39	2.15 (0.06)	-0.29 (0.48)
	Age 40-44	1.79 (0.06)	1.04 (0.48)
	Age 45-49	1.55 (0.06)	1.01 (0.36)
	Age 50-54	1.23 (0.06)	0.76 (0.42)
	Age 55-59	1.08 (0.07)	0.31 (0.46)
	Age 60-64	0.89 (0.07)	0.68 (0.41)
	Age 65-69	0.70 (0.07)	0.40 (0.41)
	Age 70-74	0.53 (0.07)	0.81 (0.39)
	Age 75 or more	(ref)	(ref)
Diagnoses	Cancer	0.02 (0.09)	5.12 (0.31)
	Cardiovascular	-0.47 (0.06)	4.03 (0.25)
	Diabetes	-0.12 (0.15)	5.66 (0.52)
	Renal	0.01 (0.34)	7.26 (0.27)
	Pulmonary	-0.62 (0.21)	7.02 (0.36)
	Arthritis	-0.30 (0.14)	6.84 (0.31)
	Asthma	-0.21 (0.3)	7.66 (0.54)
	Other	-1.10 (0.22)	6.18 (1.04)
	Healthy	(ref)	(ref)
Location	Normal	3.78 (0.04)	2.82 (0.30)
	Special	5.52 (0.08)	1.71 (0.65)
	Urban	(ref)	(ref)
Income	Low	0.32 (0.03)	-1.92 (0.57)
	High	(ref)	(ref)
N		5,849,169	
Pseudo-R^2		0.23	

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

Appendix 8 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 hospitals in the event of terrorist attacks, natural disasters, and car accidents.[30] Although they were not meant to guide price negotiations between insurers and hospitals, there is evidence that insurers use these reference prices as starting points in their negotiations (Ruiz et al., 2008). Appendix figure 5 shows that references prices are highly correlated with negotiated prices from the claims data.

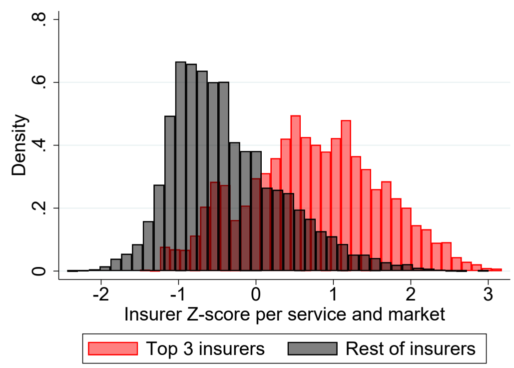


Appendix Figure 5: Negotiated prices and reference prices

Note: Scatter plot of negotiated prices and reference prices per service. The black line is a 45 degree line.

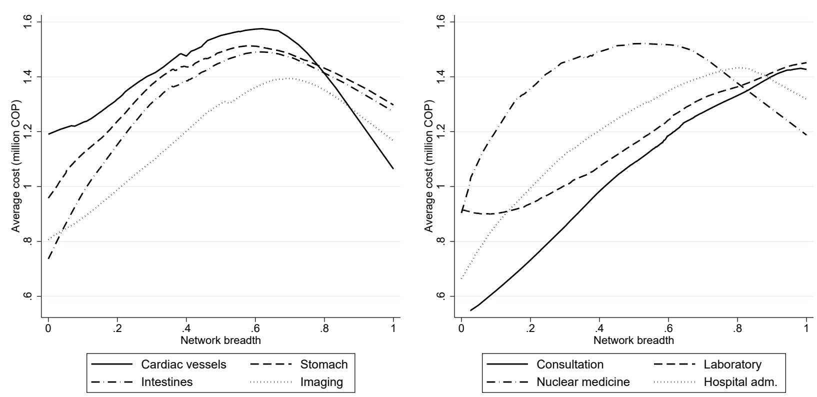
[30] Decree 2423 of 1996

Appendix 9 Additional average cost results



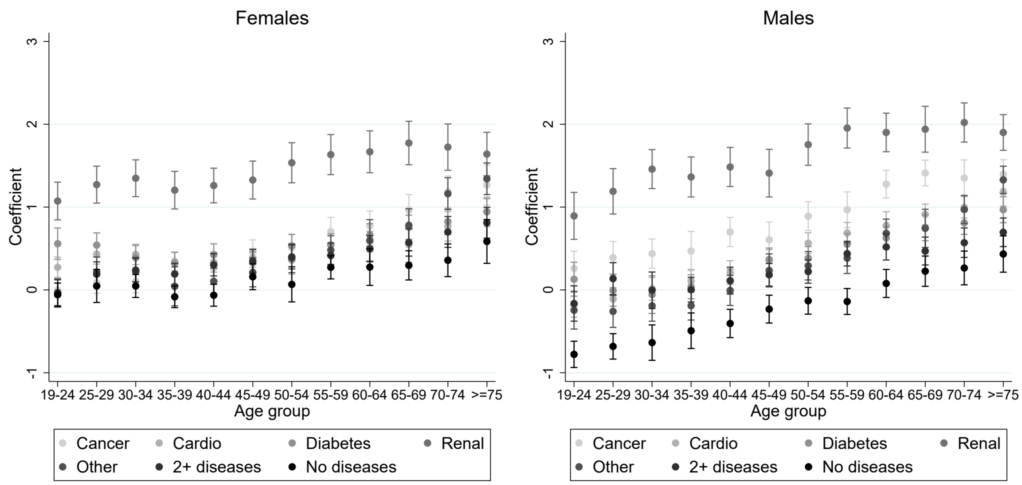
Appendix Figure 6: Standardized network breadth per service and market

Note: Distribution of network breadth standardized within service and market, separately for the top 3 insurers (EPS013, EPS016, and EPS037) and the rest of insurers. Standardized values of network breadth are obtained by subtracting the service-market level mean and dividing by the service-market level standard deviation. The top 3 insurers have consistently broad networks across services, while the rest tend to have narrow networks across services.



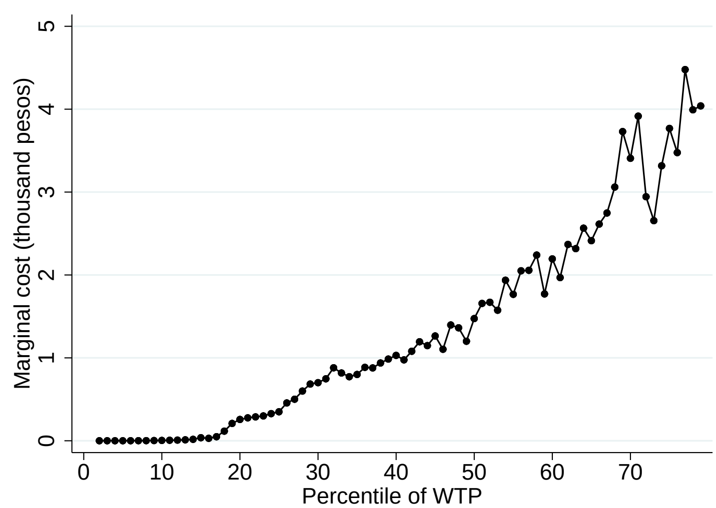
Appendix Figure 7: Average cost function per service

Note: Predicted average cost conditional on observed levels of network breadth separately for 8 service categories: procedures in cardiac vessels, stomach, intestines, and imaging, consultations, laboratory, nuclear medicine, and hospital admissions.



Appendix Figure 8: Consumer type fixed effects

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



Appendix Figure 9: Marginal cost and willingness-to-pay for network breadth

Note: Scatter plot of estimated insurer marginal cost and percentiles of the consumer willingness-to-pay for network breadth.

Appendix Table 14: Patient-level estimates of average cost

log(cost+1)	coef	se
Network breadth	3.33	(0.06)
Scope economies	-451.86	(7.12)
Reference price	—	—
<u>Insurer</u>		
EPS001	-0.58	(0.01)
EPS002	0.59	(0.01)
EPS003	0.63	(0.01)
EPS005	2.03	(0.01)
EPS008	1.83	(0.01)
EPS009	0.58	(0.01)
EPS010	0.76	(0.01)
EPS012	1.01	(0.01)
EPS013	1.33	(0.01)
EPS016	0.36	(0.01)
EPS017	1.10	(0.01)
EPS018	1.14	(0.01)
EPS023	1.36	(0.01)
EPS037	(ref)	(ref)
N	8,146,255	
R^2	0.23	

Note: Regression of log total health care cost (plus 1) 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 15: Average cost excluding the capital city

Variable	coef	se
Network breadth	1.04	(0.17)
Scope economies	-49.00	(20.14)
Reference price	18.25	(6.78)
<u>Insurer</u>		
EPS001	0.00	(0.04)
EPS002	-0.25	(0.02)
EPS003	-0.22	(0.02)
EPS005	-0.08	(0.02)
EPS008	-0.14	(0.08)
EPS009	0.03	(0.05)
EPS010	-0.14	(0.03)
EPS012	-0.17	(0.05)
EPS013	-0.06	(0.02)
EPS016	-0.11	(0.02)
EPS017	-0.33	(0.04)
EPS018	-0.21	(0.04)
EPS023	-0.53	(0.04)
EPS037	(ref)	(ref)
N	38,366	
R^2	0.38	

Note: Regression of log average cost per consumer type on markets excluding the capital city. Includes insurer, market, and consumer type fixed effects. Robust standard errors in parenthesis.

Appendix 10 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 16 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 17 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 16: Dropout probability

	mean	sd
Diagnosis		
Cancer	4.9	(3.2)
Cardio	3.1	(1.7)
Diabetes	3.1	(1.4)
Renal	4.7	(2.8)
Pulmonary	4.5	(2.9)
Arthritis	2.6	(1.4)
Asthma	3.3	(1.9)
Other	3.5	(2.1)
Healthy	46.1	(7.7)
Age		
19-24	10.8	(16.5)
25-29	7.6	(12.0)
30-34	7.0	(12.1)
35-39	7.2	(12.6)
40-44	7.2	(13.1)
45-49	7.2	(13.5)
50-54	7.6	(14.1)
55-59	7.6	(14.6)
60-64	7.7	(14.7)
65-69	8.0	(14.8)
70-74	8.6	(14.7)
75 or more	14.5	(14.4)
Sex		
Female	7.5	(12.1)
Male	9.3	(15.2)

Appendix Table 17: Transition probabilities

Diagnosis	Cancer	Cardio	Diabetes	Renal	Lung	Arthritis	Asthma	Other	Healthy
Cancer	30.0	13.3	1.7	0.7	1.4	2.6	0.4	1.4	48.6
	(7.4)	(8.5)	(1.5)	(0.6)	(1.3)	(1.9)	(0.2)	(0.5)	(17.9)
Cardio	4.1	53.8	2.7	1.1	1.4	1.7	0.4	1.1	33.8
	(3.4)	(20.9)	(1.7)	(0.9)	(1.4)	(0.9)	(0.3)	(0.5)	(23.3)
Diabetes	2.9	17.0	54.1	1.2	0.9	0.9	0.2	0.8	22.0
	(2.4)	(10.3)	(8.3)	(1.0)	(1.1)	(0.6)	(0.3)	(0.5)	(14.9)
Renal	4.7	21.9	3.7	27.2	1.3	2.0	0.3	2.9	36.1
	(3.6)	(13.3)	(3.0)	(4.4)	(1.3)	(1.7)	(0.4)	(2.0)	(17.8)
Lung	5.4	17.9	1.7	0.6	22.7	2.6	2.8	1.8	44.4
	(4.4)	(8.9)	(1.2)	(0.7)	(15.5)	(1.7)	(1.5)	(1.0)	(23.9)
Arthritis	5.8	15.8	1.5	0.6	1.6	23.6	0.5	2.1	48.6
	(4.4)	(10.5)	(1.2)	(0.4)	(1.6)	(5.7)	(0.3)	(1.1)	(16.4)
Asthma	4.5	13.4	1.2	0.4	8.9	2.4	28.5	1.2	39.4
	(3.9)	(9.5)	(1.3)	(0.6)	(8.3)	(2.0)	(9.2)	(1.0)	(16.2)
Other	5.4	15.2	1.6	1.0	2.5	3.6	0.4	33.3	37.1
	(3.6)	(11.7)	(1.5)	(0.7)	(3.2)	(2.8)	(0.3)	(11.8)	(8.9)
Healthy	5.5	12.8	1.5	0.6	1.6	2.9	0.4	1.0	73.6
	(4.1)	(9.4)	(1.3)	(0.7)	(1.8)	(2.0)	(0.2)	(0.2)	(14.5)

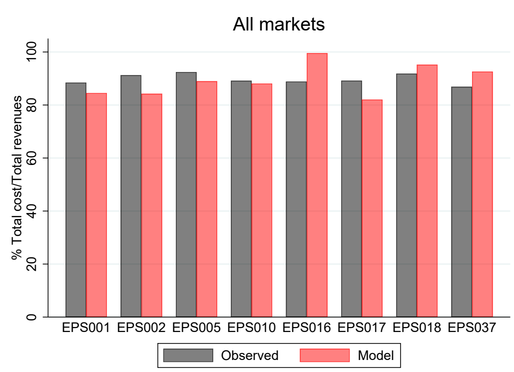
Note: Mean and standard deviation in parenthesis of transition probabilities across diagnoses. Summary statistics are calculated across sex-age combinations in each cell.

Appendix 11 Additional network formation cost results

Appendix Table 18: Summary statistics of marginal variable profits per insurer

Insurer	mean	sd
EPS001	1,046	6,746
EPS002	3,124	17,865
EPS003	2,414	16,812
EPS005	2,084	15,493
EPS010	1,995	9,205
EPS013	2,570	13,594
EPS016	3,459	12,803
EPS017	3,412	27,619
EPS018	1,406	9,073
EPS037	3,942	20,676

Note: Mean and standard deviation of marginal variable profits in the left-hand side of equation (5). Measured in millions of Colombian pesos per service per market.



Appendix Figure 10: Out-of-sample model fit

Note: Comparison of model-predicted ratio of total costs (total average costs plus network formation costs) to total revenues against 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

Appendix Table 19: First stage regression of network breadth

H_{jkm}	coef	se
H_{jkm}^{t-1}	0.85	(0.01)
$H_{jkm}^{t-1} \times \overline{q}_{\text{age 19-24, k}}$	-10.43	(10.01)
$H_{jkm}^{t-1} \times \overline{q}_{\text{age 25-29, k}}$	16.21	(37.19)
$H_{jkm}^{t-1} \times \overline{q}_{\text{age 30-34, k}}$	-5.19	(31.74)
<u>Service</u>		
Cardiac vessels	0.00	(0.02)
Stomach	0.02	(0.02)
Intestines	0.06	(0.02)
Imaging	-0.01	(0.02)
Consultation	-0.03	(0.05)
Laboratory	-0.01	(0.02)
Nuclear Medicine	0.03	(0.01)
Hospital Admission	0.06	(0.02)
F-statistic	1,718.5	
N	2,262	

Note: First stage of the GMM estimation of equation (7). H_{jkm}^{t-1} is network breadth in 2010. $\overline{q}_{i,k}$ is the average probability that a consumer with characteristic i makes a claim for service k. The specification includes service fixed effects. Robust standard errors in parenthesis and first-stage F-statistic reported.

Appendix Table 20: Network formation costs in markets without star hospitals

asinh(MVP_{jmk})	coef	se
Network breadth	3.96	(0.11)
<u>Service</u>		
Cardiac vessels	1.19	(0.15)
Stomach	0.67	(0.15)
Intestines	4.04	(0.15)
Imaging	5.78	(0.16)
Consultation	6.76	(0.17)
Laboratory	6.53	(0.16)
Nuclear Medicine	3.97	(0.15)
Hospital Admission	4.23	(0.16)
First stage F-stat	796.9	
N	3,190	
R2	0.77	

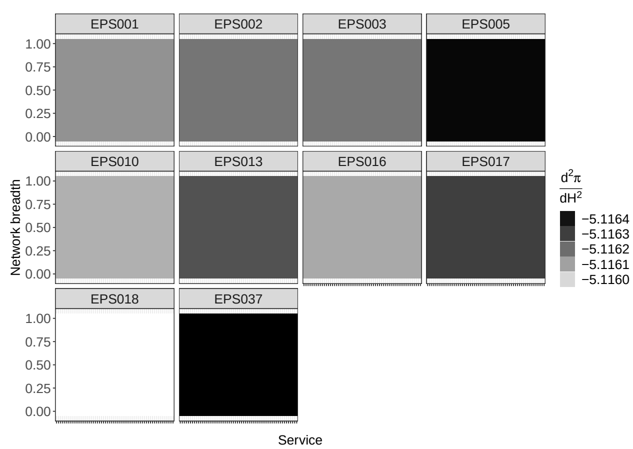
Note: 2-step GMM estimation of equation (7) on the subsample of markets without star hospitals, where there are at most 10 percent of corner solutions in network breadth, these are markets 15, 18, 25, 73, 50, 52, 47. Robust standard errors in parenthesis and first-stage F-statistic are reported.

Appendix 12 Concavity of the profit function

The second partial derivative of the short-run insurer profit function with respect to network breadth for service m, all else equal, 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, while holding its rivals' network breadth choices fixed at observed levels. I compute this exercise with data from Bogotá where my counterfactual simulations are conducted. Appendix figure 11 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 11: Second partial derivative of short-run profit function

Note: Figure presents the second partial derivative of the insurers' short-run profit function for every service.

Appendix 13 Additional counterfactual results

Appendix Table 21: Disease categories

Genetic anomalies	Lymph node cancer
Arthritis/Arthrosis	Diabetes
Asthma	Cardiovascular disease
Autoimmune disease	Long-term pulmonary disease
Cervical cancer	Renal disease
Breast cancer	Skin cancer
Cancer in digestive organs	HIV-AIDS
Lung cancer	Transplant
Other cancer	Tuberculosis
Epilepsy	Healthy

Appendix Table 22: Robustness of counterfactual results to variation in hospital quality

	No RA			Improved RA - 9 diseases		
Parameter estimation:	Main		No qual	Main		No qual
Counterfactual data:	Main Market 11 (1)	No qual Market 52 (2)	No qual Market 52 (3)	Main Market 11 (4)	No qual Market 52 (5)	No qual Market 52 (6)
Median network breadth	-1.9	-0.5	-21.7	4.2	17.2	21.6
Total avg. cost	2.7	-0.9	0.0	-0.4	-1.0	-1.5
Consumer surplus	-1.1	-0.2	-2.4	2.1	0.8	-3.3

Note: Main counterfactual results and robustness checks to variation in hospital quality. Columns (1) and (4) report main counterfactual results simulated in market 11 with parameters estimated in the full sample. Columns (2) and (3) report counterfactual results simulated in market 52 with parameters estimated in the full sample. Columns (3) and (6) report counterfactual results simulated in market 52 with parameters estimated in markets without variation in hospital quality. Market 52 corresponds to the state of Santander. This state does not have star hospitals nor significant variation in hospital quality. Comparisons of columns (2)-(3) and columns (5)-(6) show that my simulations provide a lower bound of the true effect of eliminating risk adjustment and improving risk adjustment on network breadth.