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

Natalia Serna*

November 29, 2023

Click here for latest version

Abstract

Health insurers typically compete on the breadth of their hospital networks. In this paper I show that insurers' decision to offer network breadth depends on two forces: risk selection and cost incentives. To decompose the relative importance of these forces, I estimate a structural model of insurer competition in networks applied to data from Colombia. I find that insurers risk-select by providing narrow networks in services that unprofitable patients require. Despite selection incentives, some insurers choose to offer broad networks because of heterogeneity in their cost structure. Broad networks can further be promoted by allowing insurers to compete on premiums. Findings suggest that markets with universal coverage can produce broad-network insurers without network adequacy rules and that price and non-price elements of insurance contracts are substitutes for risk selection.

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

JEL codes: I11, I13, I18, L13.

*Stanford University. E-mail: nserna@stanford.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. This research has benefited from comments from participants at BFI's Women in Empirical Microeconomics, Stanford's SITE, IO Workshop at UW-Madison, ASHEcon, IIOC, MEA, Colombian Central Bank Workshop, and the seminars at Boston University, Kelley School of Business, London School of Economics and Political Science, Ohio State University, Olin School of Business, Quantil Colombia, Ross School of Business, Stanford Health Policy, University of Toronto, University College London, and University of Pennsylvania. Findings do not represent the views of any institution involved. All errors are my own.

1 Introduction

Health insurers respond to different incentives when crafting the various elements of their insurance contracts. These incentives include risk selection and fixed costs. While there is extensive literature documenting the effects of risk selection, the impact of fixed costs is much less explored. In this paper I quantify the relative importance of these incentives on insurers' decision to offer hospital network breadth.

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). Conversely, a few studies analyze the impact of risk selection on hospital networks conditional on premiums (Shepard, 2022). In addition to providing evidence of how risk selection affects network breadth, the novelty of my paper is in showing how this incentive interacts with fixed costs when insurers decide on how many hospitals to include in their networks and how to set premiums.

My empirical setting is Colombia, where private insurers provide a national health insurance plan in a system similar Medicare Advantage (MA) in the US. A key difference however 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.

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 MA engage in risk selection by placing services that sick individuals need in higher cost-sharing tiers.[1] However, whether cost incentives play a role in determining

[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

service-level hospital networks on top of risk selection remains an open question.

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 risk selection incentives, because it leaves significant variation in expected patient profitability depending on the types of services the patient is likely to need. I then 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 they are likely to need. For example, patients with cardiovascular disease are more likely to choose insurers with broad networks for cardiac care services.

Motivated by these descriptive facts, I develop and estimate a model of insurer competition in service-level 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 policies that change the magnitude of risk selection and cost incentives. This question has become increasingly relevant with the prevalence of network adequacy rules that either force insurers to cover specific hospitals or require minimum hospital-to-enrollee ratios (Mattocks et al., 2021; Haeder et al., 2015).

For consumer demand, I model new enrollees' static discrete choices of insurer. Enrollees' utility is a function of insurers' service network breadth and out-of-pocket costs. Network breadth is defined as the fraction of hospitals in a market that provide a service and are covered by the insurer. 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 insurer is associated with higher out-of-pocket costs. On the supply side, I model insurers' heterogeneous cost structures in their average cost and network formation cost. Average costs per enrollee are a nonlin-

the context of Medicare Part D.

ear function of service network breadth and enrollee characteristics, that allows for potential economies of scope across services. Network formation costs capture fixed, administrative costs associated with a choice of service network breadth.

Insurers maximize profits by choosing their vector of network breadths conditional on rivals' choices. I assume insurers make a one-time choice of service network breadth, recognizing that this choice will affect both current and future profits. 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 diagnoses.

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

I then use my model to quantify the relative importance of risk selection and cost incentives for determining service network breadth. In a first set of counterfactual simulations, I examine whether the heterogeneity in insurers' cost structure can explain why some of them choose to offer broad networks despite risk selection incentives. My main finding is that with homogeneous fixed costs, service network breadth collapses. Mean network breadth decreases 7.6 percent relative to the observed scenario, and the decline is larger in services that sick individuals tend to claim. Instead, homogeneity in average costs has a negligible impact on insurers' network breadth decisions. These findings suggest that absent network adequacy rules, a market with universal health insurance coverage can produce broad-network insurers in equilibrium.

Finally, I conduct a set of simulations to understand whether competition on premiums can further incentivize insurers to offer broad networks, given that in the current regulation premiums are zero. I assume that insurers compete over premiums and service network breadth and price discriminate across sex, age group, and income group. I find that premiums are hump-shaped with respect to the enrollee's age, higher for males than for females, and higher for higher income individuals. These patterns reflect cross-subsidization from relatively profitable consumers to relatively unprofitable ones. Deregulating premiums incentivizes insurers to increase mean service network breadth between 16 and 35 percent, depending on how sensitive consumers are to premiums. The model thus illustrates that price and non-price elements of insurance contracts are substitutes from the point of view of risk selection.

My paper contributes to the literature on health insurance market design in two ways: first, by quantifying the importance of insurers' cost structure and premium setting on service network breadth; and second, by identifying service-level hospital networks as a risk selection mechanism. 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 efforts (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 study alternative selection mechanisms such as star hospital coverage (Shepard, 2022), insurer advertising (Aizawa and Kim, 2018), and drug formulary design (Geruso et al., 2019).

The remainder of this paper is structured as follows: section 2 describes the institutional background and data, section 3 provides descriptive evidence of adverse selection, section 4 presents the structural model, section 5 show estimation results, section 6 investigates the relative importance of insurers' cost structure, section 7 studies the effect of premiums on service network breadth, and section 8 concludes.

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, which implies that insurer competition for enrollees is zero-sum.

Private insurers provide the national insurance plan. This plan covers a comprehensive list of more than 7,000 services or procedures and 673 medications as of 2010. The government sets premiums for the national plan to zero and sets cost-sharing rules as functions of the enrollee's monthly income level. These cost-sharing rules are standardized across insurers and providers.[2] Hospital networks are the only dimension in which insurers can differentiate. Insurers can form hospital networks separately for each service offered in the national health insurance plan. For example, insurers can choose to offer a broad network for orthopedic care, but a narrow network for cardiology. Although the government does stipulate a set of network adequacy rules to guarantee appropriate access to health services, these rules are very coarse and apply only to the provision of primary care, urgent care, and oncology.[3]

[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 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/

At the end of every year, insurers report to the government all the health claims made through the national insurance plan that they reimbursed hospitals in their network for. The data for this paper are:

- Enrollment files of all enrollees to the contributory system during 2010 and 2011 (25 million),
- Insurers' claims reports to the government (650 million), and
- Insurers' hospital network data per health care specialty between 2010 and 2011 from the National Health Superintendency.

I focus on the sample of individuals aged 19 or older, of whom 2/3 have continuous enrollment spells or no gaps in enrollment. Of the continuously enrolled, 2/3 are *current enrollees*, that is, individuals who are enrolled throughout 2010 and 2011. The remaining 1/3 are *new enrollees* or individuals who enroll for the first time in 2011. Because there is near universal coverage, new enrollees to the contributory system can be individuals who move from the subsidized system after they find a job, turn 18 and choose a different insurer from their parents', or for some reason were uninsured for 12 continuous months.[4] Consumer inertia in this market is also substantial: the data shows that less than 1 percent of current enrollees switch their insurer from 2010 to 2011. This fraction increases to only 6 percent when considering all enrollees regardless of their enrollment spell lengths.

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

BibliotecaDigital/RIDE/VS/PSA/Redes-Integrales-prestadores-servicios-salud.pdf

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

municipality, sex, and age. The health claims data report date of provision, procedure code, procedure price, provider, insurer, and ICD-10 diagnosis code. These claims come from the 23 private insurers that participated in the contributory health care system during my sample period. I focus on the 10 largest insurers that account for 95 percent of enrollees. Insurers compete in each of the 33 Colombian states or markets (similar in size to a Metropolitan Statistical Area in the US).

Every claim is associated to a 6-digit procedure code from the national insurance plan. These codes can be mapped to the health care specialties that insurers and hospitals bargain over, contained in the hospital network data. The network data reports 150 unique specialties, which I aggregate up to 20 "services" and corroborate with network inclusions inferred from the claims data. The aggregation from specialties to services is made on the basis of the fraction of insurer-hospital pairs including specialty x in the network, that also include specialty y. Some examples of specialties in the data are cardiology, pediatric cardiology, and cardiovascular surgery, which I aggregate to cardiac care services. Other specialties are intensive care unit, intermediate care unit, neonatal intensive care unit, and hospitalization, which I aggregate to hospital admission services. Appendix 2 provides a description of the aggregation from specialties to services, as well as the final list of services and an excerpt from this data.

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

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 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 treating each disease. These reimbursements come from insurers that enroll a below-average share of individuals with those diagnoses.[5] 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. This rising trend in average costs suggests that insurers have incentives to engage in selection against old individuals. The rising trend in variance suggests

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

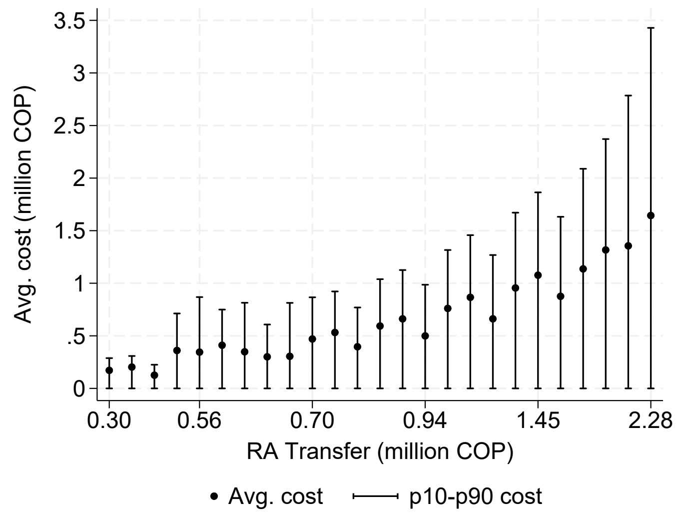


Figure 1: Health care cost by risk-adjusted transfer

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

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 risk adjustment in Colombia and the high variance in health care costs also generate substantial variation in profits per enrollee as seen in appendix table 6.

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. Although insurers' coverage choices are in part determined by differences in hospital specialties and available capacity, these choices also depend on the type of consumers that insurers want to risk select upon and on their cost structure.

If insurers use their service-level hospital networks to select risks or minimize costs, then differences in these demand and cost incentives should appear as differences in service network breadth. I define service network breadth as the fraction of hospitals in a market offering a particular service that are covered by the insurer. Table 1 shows that there is substantial variation in this measure of coverage across insurers. In particular, the table shows that even absent premiums, some insurers like EPS013 and EPS037 choose to offer broad service networks while others like EPS001 and EPS002 choose narrow service networks.

Table 1: Distribution of network breadth per service

	2010			2011		
Insurer	mean	p25	p75	mean	p25	p75
EPS001	0.13	0.00	0.20	0.14	0.03	0.21
EPS002	0.31	0.11	0.44	0.29	0.00	0.45
EPS003	0.15	0.00	0.25	0.15	0.00	0.25
EPS005	0.34	0.19	0.44	0.33	0.18	0.43
EPS008	0.06	0.00	0.06	0.04	0.00	0.04
EPS009	0.09	0.00	0.07	0.10	0.00	0.12
EPS010	0.08	0.00	0.14	0.07	0.00	0.13
EPS012	0.10	0.00	0.10	0.10	0.01	0.11
EPS013	0.54	0.35	0.73	0.51	0.33	0.70
EPS016	0.28	0.09	0.42	0.32	0.15	0.49
EPS017	0.13	0.00	0.17	0.14	0.00	0.20
EPS018	0.10	0.00	0.09	0.21	0.04	0.31
EPS023	0.05	0.00	0.03	0.04	0.00	0.04
EPS037	0.58	0.42	0.80	0.52	0.34	0.73

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

Service network breadth 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. Patients enrolled with insurers that have low network breadth typically travel longer distances to seek care (see appendix figure 2). Network breadth is also strongly correlated with the average quality of in-network hospitals as measured by their mortality rates (see appendix figure 3).

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. I show later on that my measure of network breadth relates directly to consumers' expected utility for in-network hospitals. I also note that hospital quality and network breadth are highly positively correlated (see appendix figure 3), thus high-quality hospitals are more likely to be included in a broad network as they are in a narrow network. Lastly, I provide evidence of the robustness of my model to hospital quality in section 5.

3.2 Network breadth as a means of risk selection

The descriptive statistics show that there is substantial variation in service network breadth and profits per enrollee that is consistent with differences in selection efforts and costs. 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 data from *all* enrollees in Colombia's contributory health system.

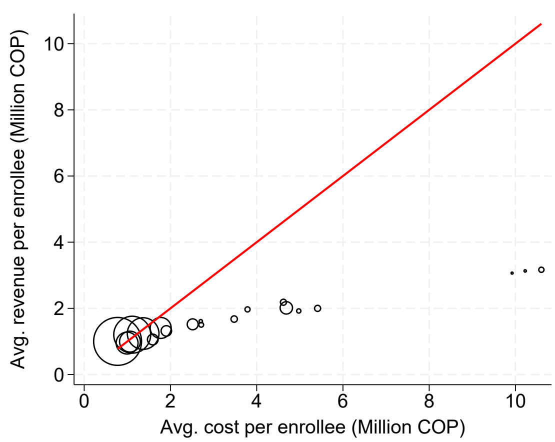


Figure 2: Service-level selection incentives after risk adjustment

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

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

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, although insurers can set up their hospital networks 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 or to minimize costs 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 service network breadth across insurers and markets. Average profits are calculated conditional on patients who make claims for the service. The red line corresponds to a linear fit and shows that relatively profitable services, such as general medicine and laboratory, tend to have broader networks than relatively unprofitable services, such as cardiac care and renal care. This correlation holds along several dimensions considered in the bottom panels of the figure and, in particular, is not necessarily driven by services with few claims.

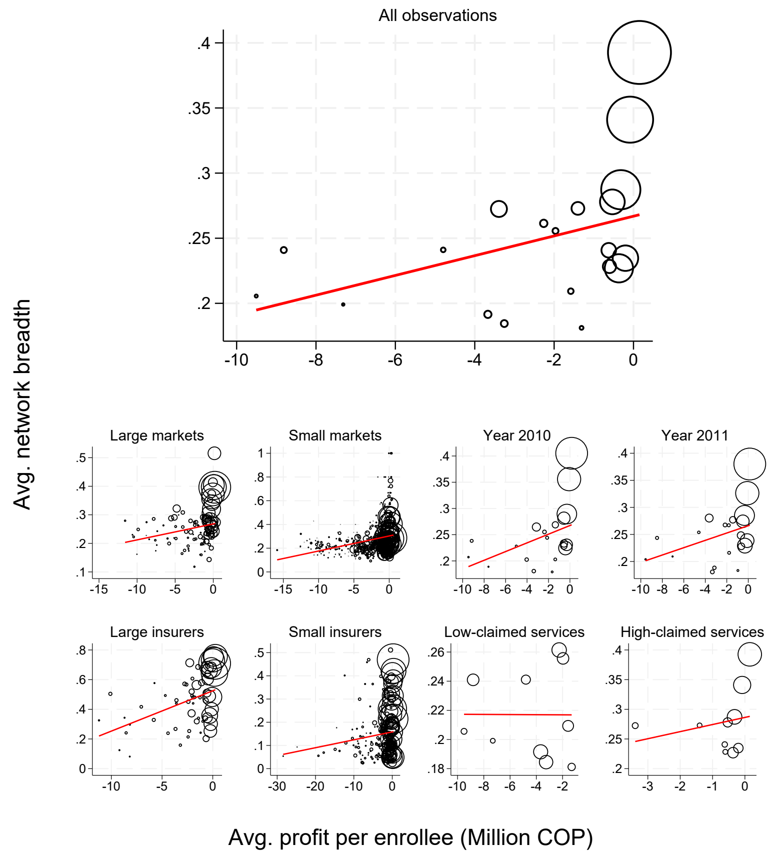


Figure 3: Correlation between network breadth and service profitability

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

Switching patterns. While the previous figures conflate selection and moral hazard as sources of variation in service-level profits, I turn now to assessing how much of this variation can be explained by adverse selection and cream-skimming alone. I leverage switching decisions over time using information from *all* enrollees in the contributory system. This analysis is identified from the 6 percent of individuals in the system that switch their insurer between 2010 and 2011.

In table 2 I explore network breadth as a determinant of enrollees' switching behavior (risk selection on extensive margin), by regressing an indicator for whether

Table 2: Determinants of switching

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

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

the enrollee switched out of their incumbent insurer on that insurer's service network breadth in 2010. I estimate this regression separately for different diagnoses and the network breadth of the service that more closely relates to treatment of this health condition. A comparison of column (1) against columns (2)-(4) suggests that network breadth has different effects on switching decisions depending on the enrollee's health status. Healthy enrollees are more likely to switch out of insurers that have broad networks for general medicine, but patients with cancer and cardiovascular disease are less likely to switch out of insurers with broad networks for chemotherapy and cardiac care services, respectively.

To see how insurers' network coverage decisions affect the composition of consumers that they enroll (risk selection on the intensive margin), table 8 presents a specification in the spirit of Brown et al. (2014). I regress enrollees' (baseline) log total health care cost in 2010 on a switching indicator and mean service network breadth for their 2011 insurer. By focusing on baseline costs, this exercise separates selection from moral hazard on network breadth. Evidence shows that, conditional on their 2010 diagnosis, enrollees who switch have lower baseline health care costs compared

Table 3: Switchers' baseline costs

Sample:	(1) Healthy	(2) Cancer	(3) Diabetes	(4) Cardio
Switch	-0.12	-0.42	-0.75	-0.60
	(0.02)	(0.07)	(0.15)	(0.06)
Mean network breadth	-0.20	-0.17	0.20	0.04
	(0.01)	(0.02)	(0.03)	(0.01)
Switch x Mean network breadth	-0.54	-0.16	-0.22	-0.07
	(0.06)	(0.18)	(0.37)	(0.14)
Controls				
Demographic	x	x	x	x
Days enrolled	x	x	x	x
Insurer FE	x	x	x	x
Market FE	x	x	x	x
N	5,403,230	672,824	316,009	1,558,787

Note: Table presents OLS regression of total health care costs in 2010 on a switching indicator, mean service network breadth for the 2011 insurer, and their interaction. Column (1) uses the sub-sample of individuals without diagnoses in 2010. Column (2) uses the sub-sample of individuals with cancer in 2010. Column (3) uses the sub-sample of individuals with diabetes in 2010. Column (4) uses the sub-sample of individuals with cardiovascular disease in 2010. All specifications control for enrollees' demographic characteristics, days enrolled, insurer fixed effects, and market fixed effects. Robust standard errors in parenthesis.

to those who do not switch. Broad networks tend to attract low-cost healthy switchers, but not necessarily low-cost switchers with chronic conditions. These findings provide evidence of a trade-off associated with network breadth: broad networks may incentivize switching of *current* enrollees who are relatively healthy, but they also attract a larger volume of *new* consumers who are less costly at baseline.

Table 4 corroborates this intuition in the sample of new enrollees. The table presents a regression of new enrollees' risk score on insurers' service network breadth. Results show not only that broad-network insurers tend to enroll relatively riskier enrollees with chronic diseases, but also that they tend to enroll new healthy consumers with lower risk than narrow-network insurers.

Table 4: Determinants of new enrollee risk

	(1)	(2)	(3)	(4)
Sample:	Healthy	Cancer	Diabetes	Cardio
Service:	General medicine	Therapy	Therapy	Cardiac care
Network breadth	-18.65	89.45	260.28	91.05
	(1.20)	(4.99)	(13.43)	(9.00)
Controls				
Demographics	x	x	x	x
Days enrolled	x	x	x	x
Market FE	x	x	x	x
N	4,772,159	172,075	48,055	281,966

Note: Table presents OLS regression of new enrollees' risk-adjusted transfer in 2011 on service network breadth for the 2011 insurer. Column (1) uses the sub-sample of individuals without diagnoses and network breadth for general medicine. Column (2) uses the sub-sample of individuals with cancer and network breadth for chemotherapy. Column (3) uses the sub-sample of individuals with diabetes and network breadth for laboratory. Column (4) uses the sub-sample of individuals with cardiovascular disease and network breadth for cardiac care services. All specifications control for enrollees' demographic characteristics, days enrolled, and market fixed effects. Robust standard errors in parenthesis.

4 Model

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

4.1 Insurer Demand

I model insurer demand in the sample of new enrollees in 2011, who do not experience inertia when making their first enrollment choice. Assume that a new enrollee i living in market m is of type θ. With probability $q_{\theta k}$, such that $\sum_k q_{\theta k} = 1$, the consumer will need each of the $k = \{1, ..., K\}$ services. An individual's type is given by 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, ≥75), and diagnosis $d \in D = \{$cancer, diabetes, cardiovascular disease, pulmonary disease, renal disease, other chronic disease, 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.[6]

I assume that the individual knows his or her diagnoses before making the first enrollment choice. This could be either because of medical family history or because, prior to enrolling in the contributory system, they went to the doctor and received a diagnosis. Enrollees know their health condition because selection occurs on observable, un-reimbursed (or poorly reimbursed) consumer characteristics such as those associated with health status.[7]

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

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

where $\beta_{ij} = (x_i \;\; x_j)'\beta$ and $\alpha_i = x_i'\alpha$. The vector x_i includes consumer demographics such as dummies for sex, age category, diagnosis, rural, and low income. x_j is an indicator for relatively large, high-quality insurers based on quality rankings constructed by the Ministry of Health for 2013 (EPS013, EPS010, and EPS016). The average out-of-pocket cost of consumer type θ at insurer j is given by $c_{\theta jm}$ and depends on the insurer's vector of network breadth $H_{jm} = \{H_{jkm}\}_{k=1}^{K_m}$. The coefficient ϕ_{jm} is an insurer-by-market fixed effect that captures unobserved insurer quality that varies across markets. Finally, ε_{ijm} is an *iid* unobserved shock to preferences assumed to be

[6] 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.

[7] Results are also robust to a version of the model where individuals are uncertain about their diagnoses (see appendix table 14).

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 are also a function of the insurer's negotiated service prices with hospitals, and of the individual's health care utilization. Negotiated prices and utilization may be correlated with service network breadth for two reasons: first, an insurer's bargaining position depends on how many hospitals it has included in the network; second, individuals may consume more services the broader is the network. I capture this correlation by noting the pass-through of insurers' costs to consumers' out-of-pocket costs via cost-sharing. Out-of-pocket costs equal the individual's coinsurance rate times the insurers' average cost per enrollee, which in turn depends on network breadth: $c_{\theta jm} = r_{\theta} AC_{\theta jm}(H_{jm})$. I provide a more detailed description of insurers' average cost per enrollee $AC_{\theta jm}(H_{jm})$ in the next subsection.

The first term on the right side of equation (1) can be interpreted as an approximation to the consumer's expected utility for the network obtained from a model of hospital choice as in Ho and Lee (2017). I expand on this approximation later on. The probability of making a claim, $q_{\theta k}$, is calculated outside of the model as the average prediction of a logistic regression given by:

$$\text{logit}(\text{any claims})_{ik} = \psi_k + \psi_{\theta} + \psi_{ik} \tag{2}$$

The dependent variable is an indicator for whether patient i makes a claim for service k. On the right side, ψ_k and ψ_{θ} are service and consumer type fixed effects, respectively. Moreover, ψ_{ikm} is a mean zero shock to the claim probability that is independent of network breadth conditional on consumer observables. I assume that new enrollees' expectations over the services they will need are correct on average, and that these expectations do not depend on the insurer they enroll with. I estimate equation (2) on data from both current and new enrollees in 2010 and 2011. Appendix figure 6 presents the resulting distribution of $q_{\theta k}$.

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 (to the econometrician, but not necessarily to the insurer) characteristics.

The second term on the right side 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 insurers in the *observed* equilibrium since myopic, healthy new enrollees disproportionately choose narrow network insurers with lower implied out-of-pocket costs.[8]

Given the distribution of the preference shock, the probability that consumer i in market m enrolls with insurer j is:

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

Hospital identity in demand. Substantial research in the US suggests that patients care strongly about whether their preferred hospital is included in the network (e.g., Shepard, 2022; Ho and Lee, 2017). Hospital preference may depend on distance to the patient or on hospital quality, both of which may vary within insurers and markets in ways that my measure of network breadth does not seem to capture.

[8] My specification for out-of-pocket costs at the level of consumer types and insurers is equivalent to aggregating service-level out-of-pocket costs with weights given by the claim probabilities.

There is however a clear relation between the patient's hospital valuation for a given service and service network breadth. Take one market, and consider a simple model of hospital choice where individual i's indirect utility from choosing hospital h for service k in the network of insurer j is:

$$u_{ijkh} = \xi_{kh} + \nu_{ijkh}$$

Here ξ_{kh} captures hospital h's average quality in service k and ν_{ijkh} is a preference shock distributed T1EV. Following McFadden (1996), individual i's value for insurer j's network of hospitals (G_{jk}) in this model is:

$$w_{ijk} = \log\Big(\sum_{h \in G_{jk}} \exp(\xi_{kh})\Big)$$

Let $|G_k|$ be the total number of hospitals that provide service k and $|G_{jk}|$ the number of hospitals in insurer j's network. The measure of network value derived from a hospital choice model relates to my measure of network breadth as follows:

$$\begin{aligned} w_{ijk} &= \log\Big(\sum_{h \in G_{jk}} \exp(\xi_{kh})\Big) \geq \log\Big(\frac{1}{|G_k|}\sum_{h \in G_{jk}} \exp(\xi_{kh})\Big) \geq \frac{1}{|G_k|}\sum_{h \in G_{jk}} \log(\exp(\xi_{kh})) \\ &= \frac{1}{|G_k|}\sum_{h \in G_{jk}} \xi_{kh} = \frac{|G_{jk}|}{|G_k|}\sum_{h \in G_{jk}} \frac{1}{|G_{jk}|}\xi_{kh} = \overline{\xi}_{jk} H_{jk} \end{aligned} \tag{3}$$

where the second inequality follows from Jensen's inequality and $\overline{\xi}_{jk} = |G_{jk}|^{-1} \sum_{h \in G_{jk}} \xi_{kh}$ is the average quality of the hospitals in insurer j's network. I parameterize $\overline{\xi}_{jk}$ in the demand model as $\beta x_j q_{\theta k}$. Equation (3) is indicative of a measurement error problem in a non-linear model. More explicitly, preferences for service network breadth will be biased towards zero relative to preferences for service network value where consumers care about specific hospitals being included in the network. The problem is thus one of identification of preferences for service network breadth, which I turn to next.

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 cardiac care in two different markets, but one of these markets has a higher prevalence of cardiovascular conditions, then we should observe higher insurer demand in the market where people are relatively sicker.

There are however several threats to identification. Network breadth may be correlated with unobserved insurer quality or unobserved consumer characteristics, such as their valuation for specific hospitals described earlier. Network breadth could also reflect how good the insurer is in processing health claims. These types of unobserved insurer characteristics potentially do not vary across markets conditional on the consumer type. The inclusion of insurer-by-market fixed effects in the utility function therefore allows me to identify preferences for network breadth off of exogenous variation in market demographics.

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

4.2 Insurer Average Costs per Enrollee

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

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

where K_m is the number of services available in market m, A_k is the government's reference price for service k (explained in more detail in appendix 6), and λ_θ, η_m, and δ_j are consumer type, market, and insurer fixed effects, respectively.

The coefficient τ_0 captures whether insurers bargain higher or lower prices than the reference price with the average hospital in their network. $\tau_1 q_{\theta k}$ represents the elasticity of average costs with respect to network breadth for service k. τ_2 captures the average degree of complementarity between pairs of services. If $\tau_2 < 0$, then insurers have economies of scope across services, thus greater coverage for service $l \neq k$ makes it more attractive to provide higher coverage for service k. 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 8).

Identification. The parameters of equation (4) are identified from variation in average costs within consumer types and across insurers that are identical except for their service network breadth. My source of identification does not rely on different consumers implying different costs for similar insurers as in Tebaldi (2017) but, *conditional* on the composition of enrollee pools, for different service coverage levels to imply different costs to the insurer. In this case, variation in network breadth across insurers is exogenous conditional on the rich set of fixed effects. However, one worry is that consumers may select into insurers based on their unobservables. One way to check this is to test whether estimates are robust to a more granular definition of consumer type. I conduct robustness checks of this style in appendix table 15 using patient-level data.

4.3 Competition in Network Breadth

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

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

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

Insurers take into account the future profits associated with each enrollee since, after making their first enrollment choice, individuals do not switch. Insurers in the model therefore maximize the net present value of their profits as they do in reality. $N_{\theta m}$ is the fixed market size of consumers type θ. In the expression for future profits, $\rho_{\theta m}$ represents the probability that a type-θ consumer drops out of the contributory system. This probability is (assumed) exogenous to the choice of network breadth as it is mostly governed by the event of being unemployed. $\mathcal{P}(\theta'|\theta)$ is the transition probability from type θ in period t to type θ' in period $t+1$. Future profits at year t are discounted by a factor of ζ^t. I set ζ equal to 0.95 and forward simulate this profit function for 100 periods.[9]

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. This is an administrative cost associated with the inclusion of an additional hospital to the network. The network formation cost is non-linear in network breadth and heterogeneous across insurers with $\xi_{jkm} = \xi_j + \vartheta_{jkm}$. In this specification, ξ_j represents the observed insurer-specific cost component and ϑ_{jkm} represents the idiosyncratic cost shock that is observed by insurance companies but unobserved to the econometrician. The multiplicative structure of the unobserved cost is needed to obtain a first-order condition that is linear in ϑ_{jkm}.

Profit maximization involves a set of $J \times K$ first-order conditions (FOC) in each market, which assuming an interior solution in network breadth, is given by:

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

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

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

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

$$\frac{\partial \pi_{ijm}}{\partial H_{jkm}} = \underbrace{\overbrace{R_{\theta m}\frac{\partial s_{ijm}}{\partial H_{jkm}}}^{\substack{\text{Selection}\\ \text{incentives}}} + \overbrace{R_{\theta m}\frac{\partial s_{ijm}}{\partial AC_{\theta jm}}\frac{\partial AC_{\theta jm}}{\partial H_{jkm}}}^{\text{Cost incentives}}}_{\text{Marginal revenue}} \tag{6}$$
$$- \underbrace{(1-r_i)\Big(\overbrace{AC_{\theta jm}\frac{\partial s_{ijm}}{\partial H_{jkm}} + s_{ijm}\frac{\partial AC_{\theta jm}}{\partial H_{jkm}}}^{\text{Selection incentives}} + \overbrace{AC_{\theta jm}\frac{\partial s_{ijm}}{\partial AC_{\theta jm}}\frac{\partial AC_{\theta jm}}{\partial H_{jkm}}}^{\text{Cost incentives}}\Big)}_{\text{Marginal cost}}$$

Equation (6) shows the effect of selection and marginal cost incentives 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' marginal 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 marginal consumer (selection effect). Cost incentives have opposite effects on revenues and costs. Expanding 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 if relatively sicker consumers disenroll due to higher out-of-pocket payments, then the marginal consumer is now cheaper.

Model discussion. My 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 insurers in equilibrium.

Identification. Rewriting the FOC as

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

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 service network breadths. 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 $E[\vartheta_{jkm}|H_{jkm}] < 0$.[10] Identification of the network formation cost shock thus relies on insurer fixed effects given by ξ_j, which capture the endogenous variation in marginal variable profits across insurers. I estimate the FOC via OLS since only 1 percent of observations correspond to corner solutions in H_{jkm} in my sample of markets.[11]

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 5 show that insurer demand is decreasing in out-of-pocket costs and increasing in network breadth. A 10 thousand pesos increase in out-of-pocket costs reduces insurer demand by 24 percent, corresponding to an

[10] 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.

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

average elasticity of -0.26.[12] A ten percentage point increase in network breadth across all services increases the choice probability by roughly 23 percent.[13] These results suggest not only that there is selection on network breadth but also that consumers prefer broad service networks overall.

Interactions between consumer and insurer characteristics matter for enrollment decisions. Males and females have similar sensitivity to out-of-pocket costs, but males have a stronger taste for network breadth potentially reflecting their higher likelihood of developing chronic conditions. Sensitivity to out-of-pocket costs is also decreasing in the consumer's income level. Patients aged 65 or older are both more likely to enroll in broad-network insurers and more sensitive to out-of-pocket costs compared to younger patients. One explanation for this is that old individuals need more expensive care. Given that old consumers tend to have higher out-of-pocket costs, the findings also imply that the average demand elasticity for patients aged 65 or older (-0.40) is larger than that of patients aged 19-24 (-0.12).

Findings show that individuals with cancer and renal disease have stronger preferences for broader networks than their healthy peers. However, the preference for network breadth is similar between consumers with diabetes, pulmonary disease, and those without diagnoses. Individuals with chronic conditions are all significantly less responsive to out-of-pocket costs than healthy ones. Interactions between diagnosis indicators and out-of-pocket costs overcompensate interactions between diagnoses and network breadth. Because healthy individuals have substantially lower out-of-pocket costs, the implied average elasticity for this type of consumer (-0.21) is lower than for patients with chronic diseases (-0.50). Appendix 5.2 presents some measures of in-sample model fit.

With my estimates of the preference for network breadth and out-of-pocket costs, I calculate patient willingness-to-pay (wtp) for an additional percentage point of net-

[12] 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.

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

Table 5: Insurer demand

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

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

Table 6: Average willingness-to-pay per service and diagnosis

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

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

work breadth in each service as $\frac{1}{-\alpha_i}\frac{\partial s_{ijm}}{\partial H_{jkm}}$. Differences in *wtp* across consumer types will be suggestive of patient sorting based on network breadth. Consumers with relatively high *wtp* for a particular service will tend to enroll with insurers that have broad networks for that service.

Table 6 presents the average *wtp* for some services among patients with chronic diseases, normalizing healthy individuals to 1. Findings show that patients with chronic conditions have a significantly higher *wtp* for network breadth than individuals without diagnoses. For example, patients with renal disease are willing to pay 27 times more than a healthy individual for an additional hospital in the network for renal care services.[14] This variation in *wtp* implies that, in principle, insurers can avoid unprofitable patients by offering narrow networks in the services they require, and that for some services insurers can find it profitable to offer broad networks.

Robustness checks. In the appendix I conduct several robustness checks to provide encouraging evidence that unobserved insurer quality does not pose a threat to identification. Appendix table 12 presents a demand function that includes an indicator of star hospital coverage. Appendix table 14 shows a version of demand where consumers are uncertain about their diagnoses. Finally, because requiring that

[14] 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 27 times more than a healthy individual for a reduction of approximately 10 minutes per visit in travel time to the nearest hospital that offers renal care services.

new enrollees know their diagnoses before enrolling can create mechanical bias, in appendix table 13 I identify new enrollees' diagnoses using only the information from claims made during January 2011.

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 7 shows the results, and appendix figure 9 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. Insurer coverage decisions are hence 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 4.8 percent per enrollee.[15] The magnitude of scope economies is larger than the direct effect of network breadth on average costs. The estimate for τ_1 indicates that a 1 percent increase in network breadth raises average costs by 2.2 percent per enrollee.[16]

Table 7 also shows that there is substantial average cost heterogeneity across insurers. Some insurers, such as EPS008, EPS009, and EPS018, have average costs per enrollee that are between 6 and 20 percent higher than the average cost of EPS037. Other insurers, such as EPS002, EPS005, and EPS013 have average costs that are between 24 and 13 percent lower than the reference insurer.

Appendix table 15 shows that my average cost model is robust to more granular definitions of consumer type. Appendix table 16 also shows that results are robust to explicitly modelling hospital quality with inclusion of a star hospital indicator. These exercises provide suggestive evidence of no relevant unobserved cost heterogeneity within consumer types.

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

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

Table 7: Insurer average costs per enrollee

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

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

Model-based evidence of adverse selection. With my demand and average costs estimates I can test for the model-implied adverse selection by looking at the correlation between insurers' marginal cost and patients' wtp, along the lines of Einav et al. (2010). The marginal cost of service k is $\frac{\partial(AC(H_{jkm})s_{ijm}(H_{jkm}))}{\partial H_{jkm}}$, and the wtp 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 10. This finding indicates that the endogenously selected patients with the highest wtp 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, for which I use insurers' FOCs. Demand and average cost estimates allow me to compute marginal variable profits in the left-hand side of equation (5), while dropout and transition probabilities are calculated off-line non-parametrically from the data. Appendices 8 and 9 present summary statistics of these probabilities and of marginal variable profits, respectively. The fact that MVPs are positive for all insurer-service combinations suggests a role for network formation costs in explaining the profit maximizing choices of network breadth.

Table 8: Model of insurer network formation costs

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

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

Table 8 presents the results of equation (7) for the log of marginal variable profits. I find that network formation costs are convex in network breadth and substantially heterogeneous across insurers. The unobserved cost component explains nearly 65 percent of the variation in MVPs. The estimates suggest that there are two types of insurers in this health system (see appendix table 20): those with fixed costs below and above 30 percent of total variable profits.

Magnitude of adverse selection and cost incentives. The heterogeneity in costs across insurers suggests that the decision to offer narrow (broad) service networks need not respond to selection incentives coming from demand but to services being associated with higher (lower) fixed and average costs. To see how important each of these components are for determining network breadth, I conduct a partial equilibrium exercise where I allow an insurer to unilaterally deviate and increase network breadth for general medicine by 1 percent, while holding its rivals fixed.

Table 9: Decomposition of short-run profits

Insurer	Demand	Total avg. cost	Fixed cost
EPS001	0.17	0.18	0.74
EPS002	0.28	0.31	0.86
EPS003	0.29	0.31	0.92
EPS005	0.22	0.24	0.86
EPS008	0.17	0.18	0.72
EPS010	0.37	0.38	0.95
EPS013	0.48	0.50	0.95
EPS016	0.40	0.42	0.85
EPS017	0.26	0.29	0.95
EPS018	0.15	0.16	0.68
EPS023	0.28	0.31	0.91
EPS037	0.20	0.22	0.80

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

Table 9 presents the percentage change in short-run demand, total average costs, and network formation costs from this exercise. I find that changes in demand explain 48 percent of the variation in insurers' total variable profits, while average cost incentives explain the remaining 52 percent. Network formation costs also account for 75 percent of the variation in insurers' total costs. The decomposition exercise shows that adverse selection and cost incentives are equally important for determining the choices of network breadth.

6 The Importance of Insurers' Cost Structure

Policy debate in the US on guaranteeing high-quality coverage for less profitable patients and addressing the proliferation of narrow-network plans, has lead to the implementation of several network adequacy rules. These rules go from requiring minimum hospital-to-enrollee ratios to making coverage of essential community providers mandatory.[17] Whether these coverage rules expand access to care at the expense of higher costs to the insurer is not well understood. In this section I use my model estimates to assess whether the market can generate broad-network insurers absent network adequacy standards.

My model provides two explanations for why insurers can choose to offer broad networks even with adverse selection. The first is that consumers on average prefer to have broad networks. Although willingness-to-pay for service network breadth is lower for healthy individuals relative to those with chronic diseases, it is not zero. The second is that insurers are sufficiently heterogeneous in their average costs and network formation costs. If some insurers enjoy economies of scope or have network formation costs that decline in network size, these insurers may have incentives to offer broad networks. The decomposition exercise presented in section 5.3 suggested that demand and cost incentives are equally important for determining network breadth. The finding of substantial adverse selection also suggests that demand can mute the effects of cost incentives on the decision to offer broad networks.

In this section I compute new market equilibria in service network breadth making costs homogeneous across insurers, which gets at the importance of cost heterogeneity for producing broad hospital networks. I start by eliminating average cost heterogeneity imposing the median insurer fixed effect ($\bar{\delta}$) to all insurers. I then remove heterogeneity in network formation costs in addition to average costs by assigning the median insurer-specific fixed cost component ($\bar{\xi}$) to every insurer. For tractability, I conduct these counterfactuals with data from the largest market, Bogotá, where 29

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

percent of all continuously enrolled individuals reside and where all private insurers compete.

One concern in the counterfactual analyses is that the model may admit multiple equilibria. For instance, my measure of scope economies can make it such that every firm choosing complete networks or no coverage at all are both feasible equilibria. While a direct proof of uniqueness is challenging, in appendix 10 I compute 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 network breadth and confirm that they all converge to the same equilibrium.

Table 10: Networks, costs, and welfare under homogeneous costs

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

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

Table 10 presents results of first imposing homogeneous average costs and then homogeneous network formation costs in columns (1) and (2), respectively. The main takeaway from these results is that, absent fixed cost heterogeneity, network breadth collapses. Column (1) shows that average cost homogeneity has very little impact on service network breadth. Instead, column (2) shows that if insurers had homogeneous network formation costs, mean network breadth would decrease 7.6 percent relative to the observed scenario. Insurers' total average cost increases 2.6 percent because they can no longer take advantage of scope economies. Consumer surplus for individuals with and without diagnoses increase by a moderate amount, which suggests that welfare losses due to lower network coverage are slightly overcompensated by welfare gains from lower out-of-pocket costs.

Panel B of column (2) shows that the reduction in network breadth is larger for services that mostly sick individuals tend to claim. Network breadth for general medicine –a highly claimed service among healthy enrollees– decreases roughly 7 percent, while network breadth for renal care –a highly claimed service among patients with renal disease– falls around 10 percent relative to the observed scenario. These results are robust to different ways of imposing cost homogeneity, for example, results are qualitatively similar when using the average rather than the median insurer fixed effect as seen in appendix table 21.

The role of universal insurance. The finding that cost heterogeneity explains why insurers choose broad networks might depend on the fact that Colombia's health insurance system has near universal coverage. From a demand perspective, insurers play a zero-sum game and the market is always covered. Insurers that end up providing coverage for less profitable individuals might choose broad networks to take advantage of scope economies in equilibrium. If consumers could instead choose uninsurance, risk selection incentives may be more salient than cost incentives and thus the equilibrium one where insurers choose narrow networks.

To explore this argument, I conduct a counterfactual exercise where I remove individuals who have *wtp* in the bottom 10 percent of the distribution, holding insurers' cost structure fixed. Results in appendix table 22 suggest that universal insurance also plays a role in determining broad service networks. The table shows that in a situation where low-wtp individuals can drop out of the market, selection incentives are stronger than cost incentives, and insurers reduce average service network breadth by around 11 percent.

7 The Effect of Premiums on Network Breadth

One outstanding question in my setting is whether the strict regulation of premiums –and hence forcing insurers to compete on other dimensions of the health insurance plan– effectively generates competition in risk selection, which leads to narrow-network insurers. This is an interesting economic question since little is known about the interaction between the price and the non-price elements of health insurance contracts.

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 can price discriminate based on the enrollee's sex, age group, and income group in each market; however, they cannot undercut each other by charging premiums below zero. Insurers receive premiums in addition to the government's risk-adjusted transfers, and consumers pay premiums in addition to their tax contributions to the system.

The fact that premiums are zero in the observed scenario poses two important challenges: the first is having to predict how demand will respond to premiums; the second is predicting how insurers' average cost will change with premiums, given evidence on the relation between these two variables in markets with adverse selection (Mahoney and Weyl, 2017; Starc, 2014). I explain my approach for addressing these issues next.

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:

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

Equation (8) implicitly assumes that the sensitivity of demand to premiums is the same as to coinsurance payments and copays. I also conduct analyses allowing the sensitivity to premiums to be greater than to out-of-pocket costs, following previous findings that consumers are more responsive to premiums than to other measures of cost-sharing (e.g., Abaluck and Gruber, 2011).[18]

To predict how insurers' average costs change with premiums, I use the fact that insurers will price-in consumers' health risk that is not already accounted for by the risk adjustment formula. This suggests that I can estimate the sensitivity of average costs to premiums by including the difference between average costs and risk-adjusted transfers on the right-hand side of equation (4), holding the rest of parameters fixed. Appendix table 23 presents my estimate of this predicted average cost sensitivity.[19]

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

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

where $P_m = \{\{P_{\theta_1 jm}\}_{\theta_1(i)=\theta}\}_{j=1}^{|\mathcal{J}_m|}$. Insurers simultaneously choose premiums and ser-

[18] 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.

[19] Appendix table 23 does not include consumer type fixed effects as these are collinear with the risk-adjusted transfers.

vice network breadth to maximize profits given by:

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

The solution concept is a Nash equilibrium. 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

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

and

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

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

Table 11 presents the average premium (in millions of pesos) conditional on consumer demographics and insurer. I set the coefficient on premiums to 0.5, 1.0, and 1.5 times the average component of the coefficient for out-of-pocket costs in columns (1), (2), and (3), respectively. A comparison across columns shows that premiums are lower the higher is consumer sensitivity to premiums, as expected. I find that premiums are higher for males than for females, and hump-shaped in age. Insurers therefore cross-subsidize older, relatively sicker individuals by charging higher prices to younger, relatively healthier individuals. Findings show that premiums are increasing with income, which suggests that premium deregulation is a progressive policy. In the bottom panel of the table I find substantial heterogeneity in average premiums across insurers. In equilibrium, premiums are negatively correlated with insurer market share as seen in appendix figure 12.

Table 11: Average equilibrium premiums

Variable	Low (1)	Medium (2)	High (3)
Sex			
Female	0.53	0.27	0.20
Male	0.91	0.61	0.53
Age group			
19-34	0.97	0.56	0.44
35-49	1.13	0.75	0.64
50-64	0.87	0.52	0.43
65 or older	0.12	0.03	0.04
Income group			
Low	0.05	0.01	0.02
Medium	0.78	0.47	0.39
High	—	—	—
Insurer			
EPS001	0.89	0.59	0.52
EPS002	0.55	0.28	0.18
EPS003	0.56	0.29	0.19
EPS005	0.50	0.22	0.13
EPS008	1.18	0.89	0.89
EPS010	0.78	0.50	0.43
EPS013	0.52	0.26	0.16
EPS016	0.73	0.47	0.41
EPS017	0.53	0.24	0.15
EPS018	0.86	0.61	0.56
EPS023	0.52	0.25	0.16
EPS037	0.76	0.45	0.36

Note: Table presents average premiums (in millions of pesos) conditional on demographic characteristics and insurer. Columns (1), (2), and (3) set α for premiums equal to 0.5, 1.0, and 1.5 times the average component of α_i for out-of-pocket costs, respectively.

Panel A of table 12 presents the percentage change in mean network breadth, short-run average cost per enrollee, and insurer total average costs in the premium deregulation counterfactuals. The panel also presents the implied average elasticity with respect to premiums.[20] I set the coefficient on premiums to be 0.5, 1.0, and 1.5 times the average component of the coefficient for out-of-pocket costs in each column, respectively.

Table 12: Networks, costs, and welfare under premium deregulation

Variable		Low (1)	Medium (2)	High (3)
A. Overall	Mean network breadth	34.99	22.72	15.99
	Avg. cost per enrollee	-0.71	-0.14	1.29
	Total avg. cost	-2.69	-4.65	-7.12
	Premium elasticity	-0.80	-0.97	-1.22
B. Service network breadth	Otorhinolaryngologic care	37.18	24.13	16.98
	Cardiac care	33.11	21.41	14.98
	Gastroenterologic care	34.03	22.02	15.43
	Renal care	35.04	22.66	15.85
	Gynecologic care	35.12	22.72	15.92
	Orthopedic care	33.55	21.70	15.18
	Imaging	29.82	19.50	13.87
	General medicine	30.23	20.22	14.77
	Laboratory	30.20	19.83	14.19
	Hospital admission	30.04	19.45	13.64

Note: Panel A presents the percentage change in mean network breadth, insurer total average costs, short-run average cost per enrollee, and long-run consumer welfare for the healthy and sick in the premium deregulation scenario. Panel B presents the percentage change in mean network breadth by service category. The average component of α for premiums is set to 0.5, 1.0, and 1.5 times the average component of α_i for out-of-pocket costs in columns (1), (2), and (3), respectively.

Deregulating premiums incentivizes insurers to offer much broader networks. This implies a large pass-through of hospital coverage to premiums, which is similar to findings in Cabral et al. (2018). Mean network breadth increases 35 percent in column (1) and 16 percent in column (3). The model thus implies that the price and the non-price elements of insurance contracts in markets with universal coverage 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 network breadth. The increase in coverage is relatively homogeneous across services as seen in panel B of the table, consistent with premiums accounting for the residual variation

[20]The elasticity with respect to premiums is calculated as $\frac{\partial s_{ijm}}{\partial P_{\theta_1 jm}} \frac{P_{\theta_1 jm}}{s_{ijm}}$ and averaged across individuals.

in average costs after risk-adjusted transfers. Column (3) shows that mean network breadth for general medicine and cardiac care services both increase by around 15 percent.

8 Conclusions

Private insurers in health care markets respond to different incentives when crafting their insurance contracts. In this paper I show that risk selection and cost incentives are the main drivers of insurers' decision to offer hospital network breadth. Risk selection leads insurers to offer narrow networks, while fixed cost heterogeneity has some insurers choosing broad networks despite selection incentives. I use a structural model of insurer competition in hospital network breadth to decompose the relative importance of these incentives in counterfactuals. The empirical setting is Colombia, where the government regulates premiums and cost-sharing, and allows insurers to choose which and how many hospitals to cover for each health service.

I find that service network breadth collapses in markets with universal coverage when insurers have homogeneous fixed costs. This highlights the importance of cost incentives above risk selection incentives in the decision to offer broad networks. Mean network breadth in this exercise falls 7.6 percent, with reductions being larger in services that sick individuals require the most. Finally, I show that one way to further promote broad networks is to foster insurer competition on premiums. I find that deregulating premiums incentivizes insurers to increase mean network breadth between 16 and 35 percent, depending on how sensitive consumers are to premiums.

The findings of this paper speak to the increasing use of network adequacy rules in markets where narrow-network plans have proliferated. When health systems have universal coverage, insurer competition alone can generate broad hospital networks, hence maintaining healthy levels of competition is crucial to improve access to health care.

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.

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

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.

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.

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

Mahoney, N. and Weyl, G. (2017). Imperfect Competition in Selection Markets. *The Review of Economics and Statistics*, 99(4):637–651.

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.

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

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

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

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

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

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

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

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.

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

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

Online Appendix

Appendix 1 Current risk adjustment system

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

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

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

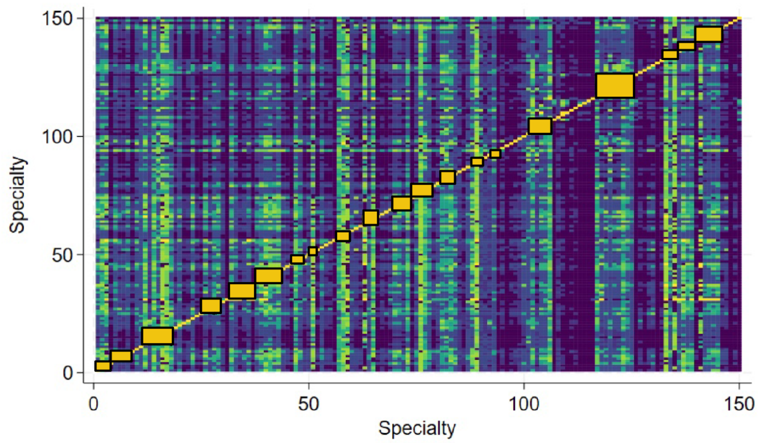
Appendix Table 2: Risk Adjustment Factors in the Contributory System during 2011

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

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

Appendix 2 Service categories

The service-level hospital network data reports 150 unique specialities over which insurers and hospitals can bargain. Some of these specialties are highly correlated in the sense that insurers tend to include them together at a particular hospital. Appendix figure 1 presents a heatmap of the fraction of insurer-hospital pairs including the specialty in the horizontal axis, that also include the specialty in the vertical axis, with light colors representing a higher fraction. The heatmap shows that (i) there are very common specialties such as general medicine and internal medicine characterized by vertical light-colored lines, and (ii) some specialties are correlated along the diagonal. The yellow rectangles along the diagonal show how the different specialties are grouped into the final 20 service categories. Excluded specialties with very few insurer-hospital pairs are, for example, ambulance or helicopter transportation and certain types of transplants. The 20 resulting service categories can be mapped to the claims data based on the 6-digit service code reported for each claim. Appendix table 3 provides a description of this final list of services and appendix table provides a data excerpt for three hospitals and three services.



Appendix Figure 1: Heatmap of specialty pairs network inclusions

Note: Figure presents a heatmap of the fraction of insurer-hospital pairs in the network data that include the specialty in the horizontal axis and the specialty in the vertical axis. Lighter colors represent higher fractions. Yellow rectangles along the diagonal represent relatively correlated groups of specialties that constitute a service category in the final network data.

Appendix Table 3: List of services

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

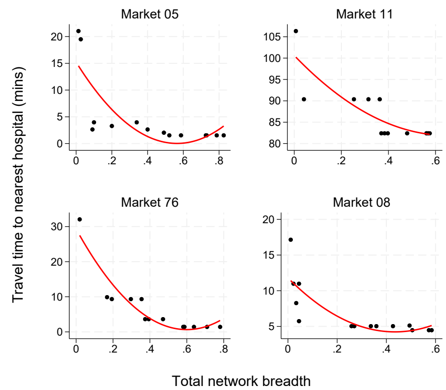
Note: Table presents the final list of 20 services and their description.

Appendix Table 4: Service coverage at hospitals

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

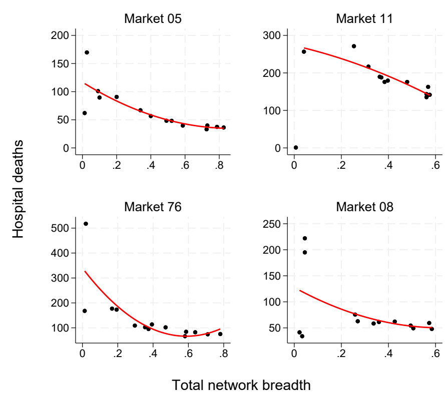
Note: Table presents service coverage per insurer at three hospitals in the country and for three services. Data comes from the National Health Superintendency.

Appendix 3 Correlates of Network Breadth



Appendix Figure 2: Average network breadth and travel times

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



Appendix Figure 3: Average network breadth and hospital deaths

Note: Figure presents a scatter plot of network breadth and total hospital deaths. The red line represents a quadratic fit.

Appendix 4 Descriptives in subsample

Appendix Table 5: Summary statistics in estimation sample

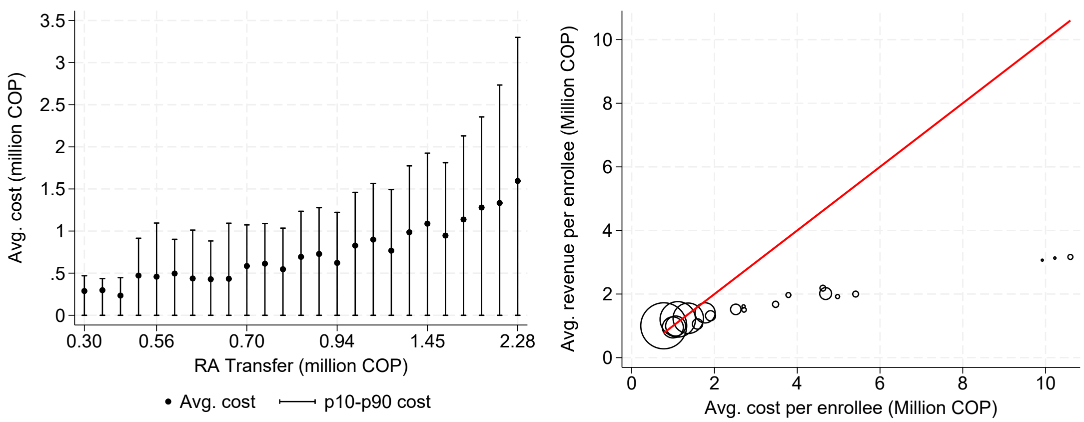
	mean	sd
Demographics		
Male	50.85	(49.99)
Age	41.70	(15.29)
Low income (%)	28.52	(45.15)
Rural (%)	24.73	(43.14)
Diagnoses (%)		
Cancer	5.55	(22.90)
Diabetes	1.77	(13.17)
Cardiovascular disease	9.55	(29.40)
Long-term pulmonary disease	0.99	(9.92)
Renal disease	0.64	(7.99)
Other disease	3.77	(19.04)
Health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

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

Appendix Table 6: Distribution of profit per enrollee in subsamples

Insurer	All enrollees			Continuously enrolled			New enrollees		
	mean	p1	p99	mean	p1	p99	mean	p1	p99
EPS001	0.37	-5.26	2.15	0.37	-6.66	2.19	0.49	-2.49	2.20
EPS002	0.32	-2.14	1.74	0.30	-3.39	1.97	0.34	-2.03	1.98
EPS003	0.35	-2.27	1.89	0.34	-3.23	2.04	0.42	-1.92	2.07
EPS005	0.44	-2.65	2.07	0.49	-3.06	2.13	0.47	-2.06	2.13
EPS008	0.27	-4.26	1.82	0.24	-5.74	2.01	0.31	-3.73	1.95
EPS009	0.07	-7.98	1.75	-0.07	-11.80	1.96	0.17	-6.53	1.97
EPS010	0.27	-3.46	1.75	0.28	-4.82	2.02	0.35	-2.75	1.91
EPS012	0.24	-3.35	1.74	0.20	-5.23	1.83	0.27	-3.35	1.75
EPS013	0.31	-2.17	1.81	0.29	-3.11	1.99	0.33	-2.09	2.00
EPS016	0.31	-2.95	1.88	0.32	-3.93	2.05	0.39	-2.36	2.07
EPS017	0.30	-2.77	1.75	0.27	-4.16	1.98	0.33	-2.56	1.92
EPS018	0.22	-3.73	1.65	0.16	-5.53	1.88	0.23	-3.65	1.74
EPS023	0.34	-1.86	1.72	0.32	-2.65	1.82	0.36	-1.79	1.75
EPS037	0.42	-7.92	2.20	0.48	-8.77	2.20	0.65	-3.10	2.20
Total	0.32	-3.28	2.02	0.33	-4.75	2.07	0.40	-2.58	2.11

Note: Table shows mean, and 1st and 99th percentiles of profit per enrollee for each insurer in the full sample in column (1), in the sample of continuously enrolled in column (2), and in the sample of new enrollees in column (3). Profits per enrollee are calculated as ex-ante and ex-post risk-adjusted transfers minus total health care cost.



(a) Health care cost by risk-adjusted transfer

(b) Service-level selection incentives

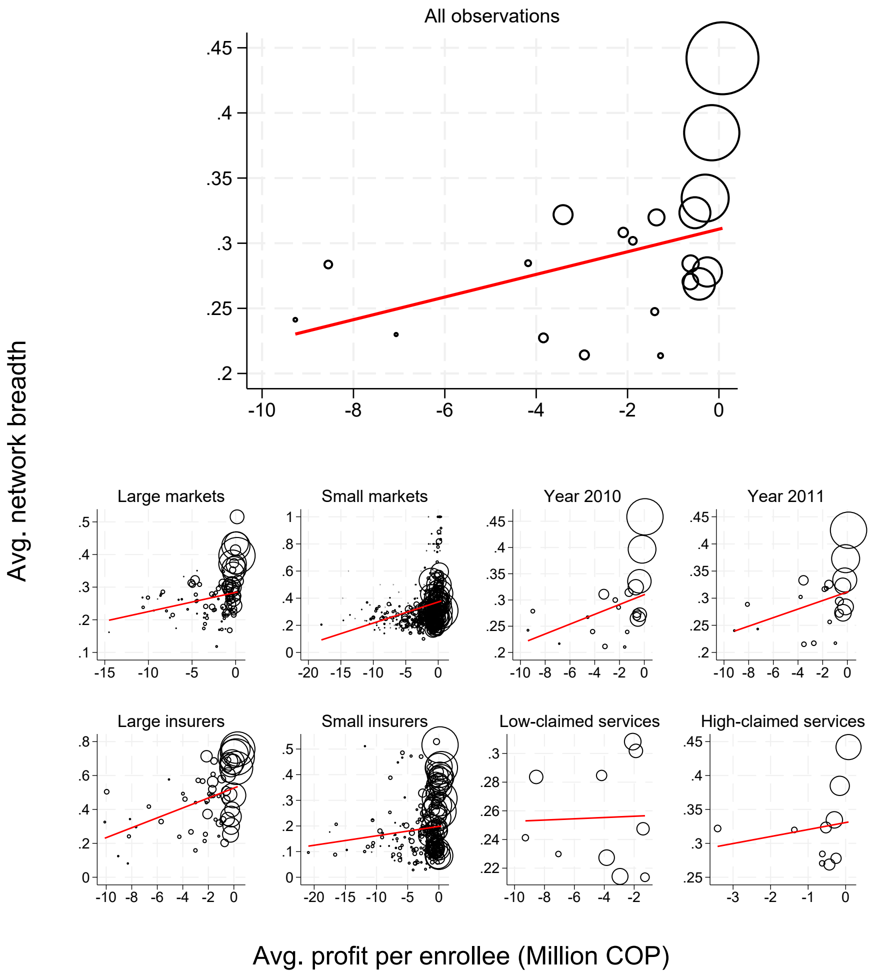
Appendix Figure 4: Costs and selection incentives in the continuously enrolled

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

Appendix Table 7: Determinants of switching in the continuously enrolled

	(1)	(2)	(3)	(4)
Sample:	Healthy	Cancer	Diabetes	Cardio
Service:	General medicine	Therapy	Therapy	Cardiac care
Network breadth	6.94	-4.72	-4.08	-4.24
	(0.16)	(0.34)	(0.38)	(0.26)
Controls				
Demographic	x	x	x	x
Days enrolled	x	x	x	x
Market FE	x	x	x	x
N	5,384,234	533,647	283,591	1,358,599

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



Appendix Figure 5: Network breadth and service profitability in the continuously enrolled

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

Appendix Table 8: Switchers' baseline costs

Sample:	(1) Healthy	(2) Cancer	(3) Diabetes	(4) Cardio
Switch	0.75	-0.37	-1.19	-0.47
	(0.11)	(0.25)	(0.53)	(0.20)
Network breadth	-0.27	-0.16	0.26	0.10
	(0.01)	(0.03)	(0.03)	(0.01)
Switch x Network breadth	-2.53	0.03	2.11	0.25
	(0.27)	(0.63)	(1.36)	(0.49)
Controls				
Demographic	x	x	x	x
Days enrolled	x	x	x	x
Insurer FE	x	x	x	x
Market FE	x	x	x	x
N	2,783,579	422,585	243,127	1,150,890

Note: Table presents OLS regression of total health care costs in 2010 on a switching indicator, mean service network breadth for the 2011 insurer, and their interaction. Sample is restricted to individuals with continuous enrollment spells. Column (1) uses the sub-sample of individuals without diagnoses in 2010. Column (2) uses the sub-sample of individuals with cancer in 2010. Column (3) uses the sub-sample of individuals with diabetes in 2010. Column (4) uses the sub-sample of individuals with cardiovascular disease in 2010. All specifications control for enrollees' demographic characteristics, days enrolled, insurer fixed effects, and market fixed effects. Robust standard errors in parenthesis.

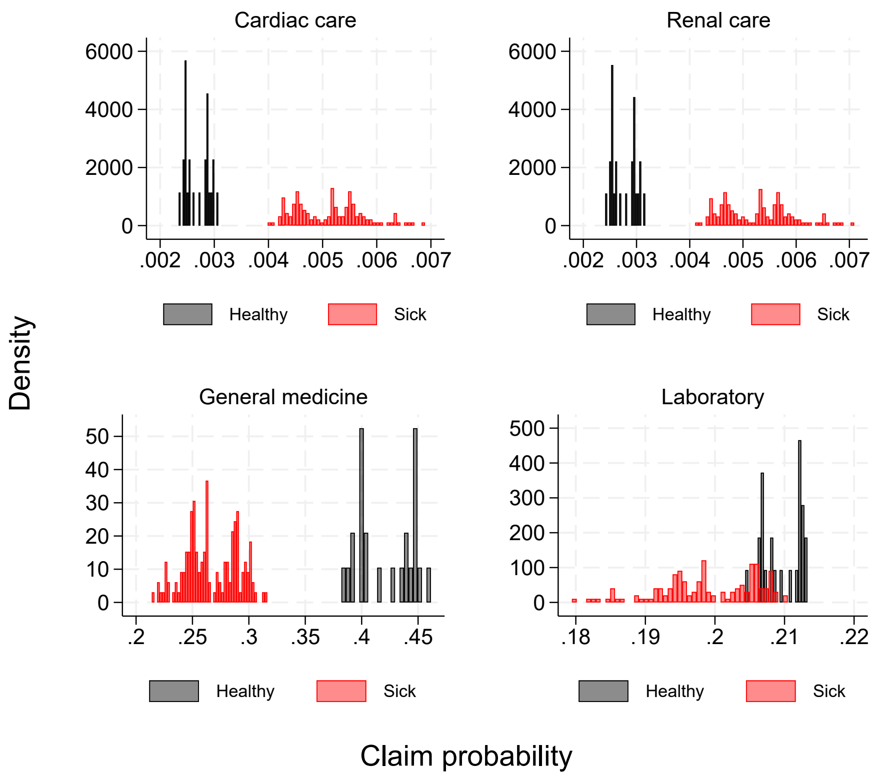
Appendix Table 9: Determinants of new enrollee risk in the continuously enrolled

	(1)	(2)	(3)	(4)
Sample:	Healthy	Cancer	Diabetes	Cardio
Service:	General medicine	Therapy	Therapy	Cardiac care
Network breadth	-17.23	74.20	153.53	77.69
	(1.92)	(5.07)	(13.32)	(8.79)
Controls				
Demographic	x	x	x	x
Days enrolled	x	x	x	x
Market FE	x	x	x	x
N	2,538,800	179,088	57,341	307,388

Note: Table presents OLS regression of new enrollees' risk-adjusted transfer in 2011 on service network breadth for the 2011 insurer. Sample is restricted to individuals with continuous enrollment spells. Column (1) uses the sub-sample of individuals without diagnoses and network breadth for general medicine. Column (2) uses the sub-sample of individuals with cancer and network breadth for chemotherapy. Column (3) uses the sub-sample of individuals with diabetes and network breadth for laboratory. Column (4) uses the sub-sample of individuals with cardiovascular disease and network breadth for cardiac care services. All specifications control for enrollees' demographic characteristics, days enrolled, and market fixed effects. Robust standard errors in parenthesis.

Appendix 5 Additional demand results

5.1 Distribution of claim probabilities



Appendix Figure 6: Distribution of service claim probability

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

5.2 In-sample demand model fit

Appendix Table 10: National market shares

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

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

Appendix Table 11: Market shares in three largest markets

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

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

5.3 Robustness checks

Appendix Table 12: Insurer demand with star hospital indicator

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

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

Appendix Table 13: Insurer demand with diagnosis in January

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

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

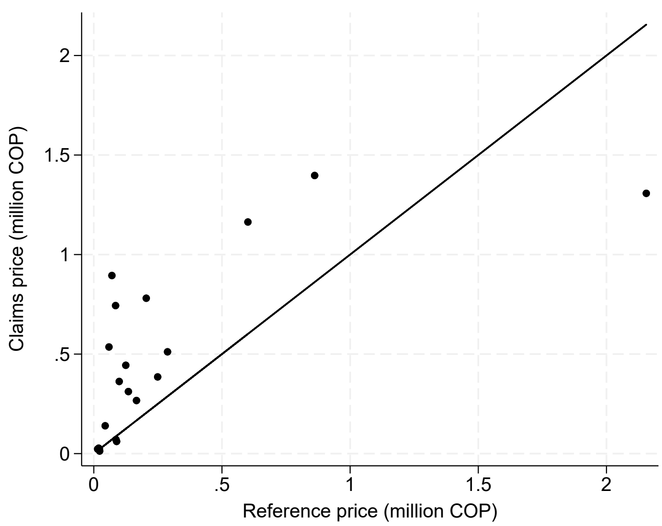
Appendix Table 14: Insurer demand with expectations over diagnoses

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

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

Appendix 6 Service reference prices

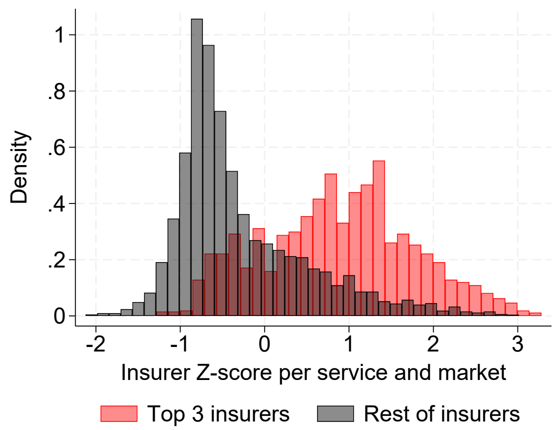
In 2005, the Colombian government published a list of reference prices for all the services included in the national health insurance plan. The list was created by a group of government officials and medical experts with the purpose of reimbursing hospitals in the event of terrorist attacks, natural disasters, and car accidents (Decree 2423 of 1996). 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 with hospitals (Ruiz et al., 2008). Appendix figure 7 shows that references prices are highly correlated with negotiated prices from the claims data.



Appendix Figure 7: Negotiated prices and reference prices

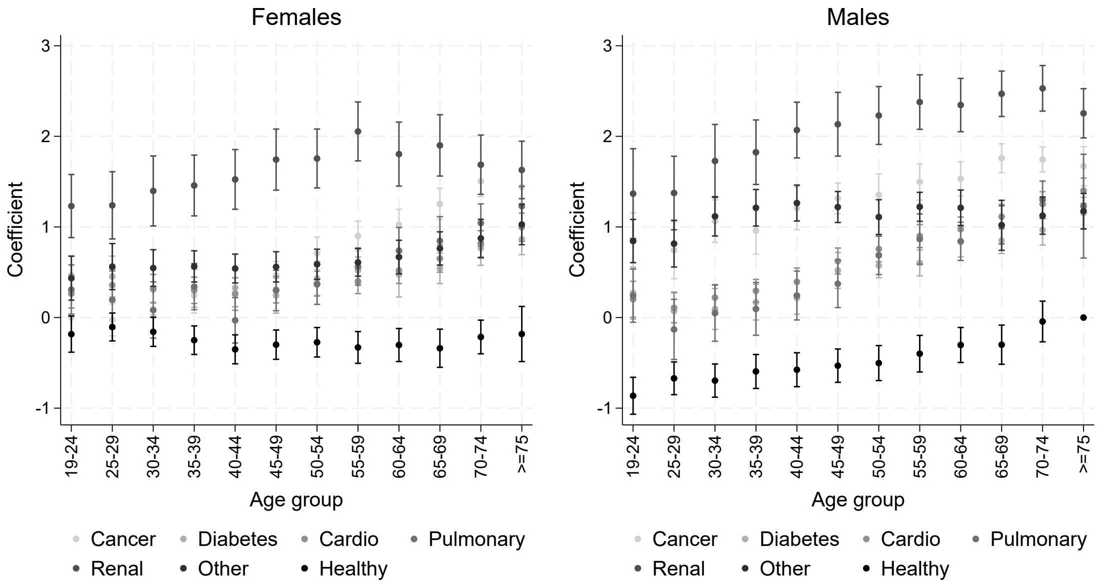
Note: Figure presents a scatter plot of average negotiated price obtained from the claims data and average reference price per service. The black line is a 45 degree line.

Appendix 7 Additional average cost results



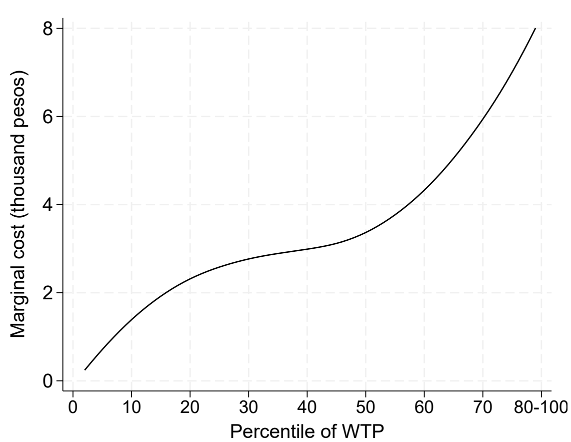
Appendix Figure 8: Standardized network breadth per service and market

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



Appendix Figure 9: Consumer type fixed effects

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



Appendix Figure 10: Marginal cost and willingness-to-pay for network breadth

Note: Figure presents average marginal cost per service by percentile of willingness-to-pay for service network breadth.

Appendix Table 15: Patient-level estimates of average cost

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

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

Appendix Table 16: Average cost with star hospitals

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

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

Appendix 8 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 17 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 18 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 17: Dropout probability

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

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

Appendix Table 18: Transition probabilities

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

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

Appendix 9 Additional network formation results

Appendix Table 19: Summary statistics of marginal variable profits

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

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

Appendix Table 20: Fixed cost as percentage of variable profit

Insurer	% of variable profits
EPS018	23.01
EPS023	25.46
EPS001	26.22
EPS037	26.41
EPS010	32.02
EPS002	32.11
EPS017	32.54
EPS008	34.95
EPS003	35.15
EPS005	35.98
EPS013	39.12
EPS016	39.46

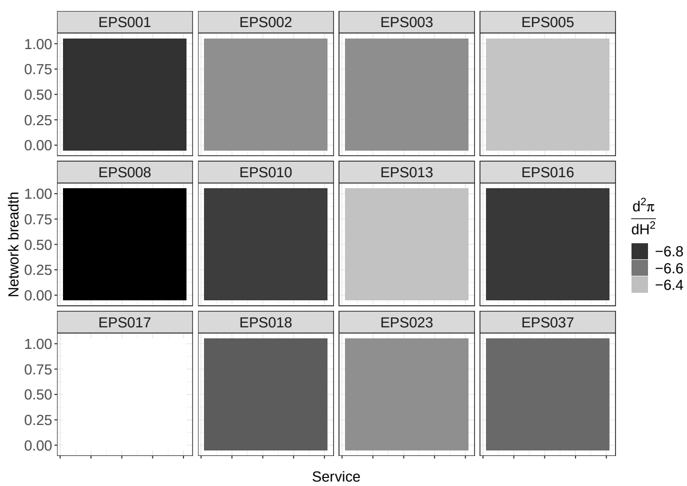
Note: Table shows the estimated network formation cost as percentage of total variable profits for each insurer averaged across markets.

Appendix 10 Concavity of the profit function

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

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

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

Appendix 11 Additional counterfactual results

Appendix Table 21: Networks, costs, and welfare under homogeneous costs

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

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

Appendix Table 22: Networks, costs, and welfare removing low-*wtp* enrollees

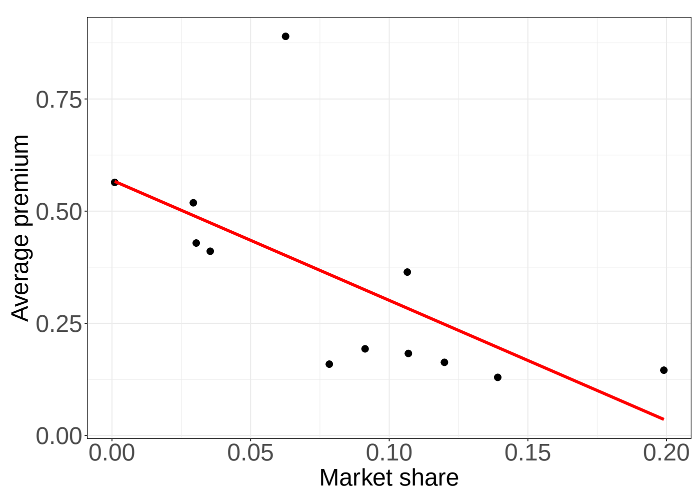
	Variable	Outside option
A. Overall	Mean network breadth	-10.99
	Avg. cost per enrollee	6.04
	Total avg. cost	-6.76
	Consumer surplus (sick)	-15.72
	Consumer surplus (healthy)	-3.31
B. Service network breadth	Otorhinolaryngologic care	-12.35
	Cardiac care	-12.07
	Gastroenterologic care	-11.80
	Renal care	-12.68
	Gynecologic care	-11.16
	Orthopedic care	-12.04
	Imaging	-7.84
	General medicine	-8.33
	Laboratory	-8.27
	Hospital admission	-9.34

Note: Panel A presents the percentage change in mean network breadth, insurer total average costs, short-run average cost per enrollee, and long-run consumer welfare for sick and healthy individuals, in the scenario where I drop individuals with average willingness-to-pay for service network breadth in the bottom 10 percent of the distribution. Panel B presents the percentage change in mean network breadth by service category.

Appendix Table 23: Impact of premiums on average costs

Variable	Coefficient	Std. Error
Avg. cost − RA transfer	0.17	(0.01)
Network breadth	0.52	(0.08)
Scope economies	-74.18	(44.44)
Reference price	73.99	(1.32)
Insurer		
EPS001	0.13	(0.05)
EPS002	0.05	(0.04)
EPS003	0.10	(0.04)
EPS005	0.05	(0.04)
EPS008	0.27	(0.05)
EPS009	0.32	(0.04)
EPS010	0.17	(0.06)
EPS012	0.06	(0.05)
EPS013	0.07	(0.03)
EPS016	0.18	(0.03)
EPS017	0.04	(0.04)
EPS018	-0.04	(0.06)
EPS023	0.08	(0.04)
EPS037	(ref)	(ref)
N	8,662	
R^2	0.67	

Note: OLS regression of logarithm of average costs per on the difference between average costs and risk-adjusted transfers, network breadth, economies of scope, and service reference prices. Includes insurer and market fixed effects. Robust standard errors in parenthesis.



Appendix Figure 12: Premiums and insurer market share

Note: Figure presents a scatter plot of average premiums and insurer market share. The red line is a linear fit.