Non-Price Competition and Risk Selection Through Hospital Networks

Natalia Serna*

Draft - Do not cite

Click here for appendix

Abstract

In this paper I study non-price competition in health insurance through hospital networks and the resulting risk selection incentives. Using a model of insurer demand, average costs, and competition in networks, I measure the impact of risk adjustment and premiums on hospital network breadth and welfare, with data from Colombia as a case study. Every aspect of the Colombian national insurance plan is regulated by the government, except for hospital networks which insurers can form separately by service. I find that insurers risk select by narrowing their networks in services that sick, costly patients demand the most, conditional on risk adjustment. Eliminating risk adjustment reduces average network breadth by 2.7% and consumer welfare by 7.4%. Improving the risk adjustment formula to compensate for diagnoses, incentivizes insurers to broaden their networks in services like hospital admissions by 3.2% and increases consumer welfare by 3.5%. A zero-premium policy exacerbates the underprovision of insurance coverage. Deregulating premiums results in almost complete networks but the welfare of healthy consumers falls substantially.

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

JEL codes: I11, I13, I18, L13.

*Department of Economics, University of Wisconsin-Madison. E-mail: nserna@wisc.edu. I am deeply grateful to the Colombian Ministry of Health and Alvaro Riascos at Quantil for providing the data for this research. I want to thank Alan Sorensen, Ken Hendricks, Dan Quint, JF Houde, Naoki Aizawa, Lorenzo Magnolfi, Christopher Sullivan, and Corina Mommaerts for their invaluable mentorship and thoughtful advice. I also thank participants to the IO Workshop at UW-Madison, Midwest Economics Association Conference, and Quantil Seminar.

1 Introduction

Risk selection in health insurance is the attempt of insurers to enroll profitable, usually healthy individuals as opposed to unprofitable patients, usually with chronic diseases. When they have discretion over the insurance plan design, private insurers typically engage in risk selection by offering a menu of insurance contracts that may differ in their cost-sharing rules and premiums, and then having individuals self-select into the contracts. In public health insurance systems, governments are aware of these incentives, so along with plan design they make risk-adjusted payments to insurers with the purpose of managing risk selection. In some cases, the government's risk adjustment formula is rich enough to attenuate selection incentives. But, when plan characteristics are regulated and the risk adjustment formula controls for only a few risk factors, selection incentives can be exacerbated and insurers will try to minimize costs and select risks using the elements of the insurance plan that they have discretion over, which can unravel the market (Kong et al., 2022).

Existing literature has focused on how risk selection arises from insurer competition in prices while holding other elements of the insurance plan fixed, even if these other elements –such as hospital networks or cost-sharing– are an insurer's choice variables as well (e.g, Ho and Lee, 2017; Dafny et al., 2015). Though it is possible to identify selection through price variation in the presence of other endogenous product characteristics (Crawford, 2012), less is known about insurer competition in variables other than premiums.[1] In this paper I study non-price competition in health insurance through hospital networks, and I show how risk selection emerges from this type of competition using as case study the Colombian health care system.

In Colombia, private insurers provide one national health insurance plan with near universal coverage. The government sets the premium for the national plan to zero, and standardizes covered services and prescription medications, coinsurance rates, copays, and maximum out-of-pocket amounts, across insurers, hospitals, and services. In addition to the strict plan regulation, insurers receive ex-ante (before health claims are made, at the beginning of every year) capitated risk-adjusted payments from the government that control only for the enrollee's sex, age, and location. Insurers also receive ex-post (after health claims are made, at the end of every year) payments that control for only a few comorbidities. These risk adjustment systems have been insufficient to capture the variance in health care costs, so even after transfers are made incentives for selection are still present (Riascos and Camelo, 2017; Riascos, 2013). The Colombian government argues that

[1] Some papers getting at the endogeneity of both hospital networks and prices include Ghili (2020); Ho and Lee (2019); Liebman (2018)

the strict regulation of the national insurance plan's characteristics incentivizes insurers to compete on quality. I claim, instead, that this type of regulation incentivizes competition in risk selection.

The only element of the public health insurance plan that is unregulated is hospital networks, which insurers can form separately by service. In Colombia, insurers have discretion over which services to cover at which hospitals, so they can use these service-specific hospital networks as a mechanism to select risks and minimize costs. By offering a narrow network in services that sick patients demand the most, insurers can effectively avoid enrollment from costly patients usually with chronic conditions, who need health care the most.[2] This is similar to Park et al. (2017) who find that insurers in Medicare Advantage in the United States engage in service-level selection by placing services that sick individuals need in higher cost-sharing tiers.

Insurers' discretion over service-specific hospital networks and universal coverage of the national health insurance plan, means that the relevant margin for risk selection in Colombia is not whether an individual decides to enroll at all, but which insurer she enrolls. Plan regulation or coarse risk adjustment in this context can distort insurers' incentives to invest in having a complete network, which impacts individuals' decisions over carriers and the distribution of health risk across insurers. The purpose of this paper, therefore, is to show that insurers effectively use their networks to engage in selection, and to quantify the effect of alternative risk adjustment formulae and premium deregulation on service-level hospital network breadth, risk, and total health care costs.

Quantifying the effect of risk adjustment and premiums on network breadth, so far, has been unexplored in the literature. Although the study of provider networks as a risk selection mechanism is not unprecedented. Shepard (2022) in the context of the Massachusetts Health Exchange, shows that sick individuals' strong preferences for expensive providers, incentivizes insurers to drop these providers from their networks. I build on his intuition to show that in the presence of zero premiums and coarse risk adjustment those exclusion incentives are exacerbated.

To do so, I model insurers' market-level network coverage decisions per service. I assume insurers engage in a simultaneous-move game where they maximize total current profits plus future discounted profits, by choosing once the vector of service-level network coverage, conditional on their competitors' choices. My solution concept is a full commitment Nash equilibrium. This approach resembles the first stage of the game in Liebman (2018), where insurers commit to network size before entering price negotiations with hospitals. I reinterpret network size as the fraction of hospitals

[2] If health insurers are unable to charge different prices for different quality, then they will find it optimal to set lower quality levels in equilibrium relative to a situation where they can price discriminate and they will focus on minimizing costs by insuring the healthiest consumers.

in a market that can provide a particular service that are covered by the insurer. This continuous measure in the unit interval considerably simplifies computation of insurers' equilibrium network breadth choices relative to existing methods in the literature (Prager and Tilipman, 2020; Ghili, 2020; Liebman, 2018; Ho and Lee, 2019), but its limitation is that it ignores hospital identity. If consumers care about hospital identity, collapsing networks to a one-dimensional object per service could be problematic. Nonetheless I show that, conditional on the service, there is no evidence of differential sorting of consumers into hospitals. My results are also robust to specifications in which I drop markets with known star hospitals, where identity matters.

My supply model is similar to that in Shepard (2022) as well, who models insurers' binary decision to include a high-priced hospital in their network and the resulting incremental costs per enrollee due to adverse selection. My solution concept can be thought of as mixed strategy equilibrium of the binary decision game, where insurers include a hospital in their network with a probability equal to my network breadth measure. Unlike the author, insurers in my context compete in network breadth for several service categories and, in that sense, the object of choice is multidimensional. When setting up their networks, I also allow insurers to consider the future profits associated to each enrollee due to consumer inertia, which has not been addressed before. Because switching rates are nearly zero in the Colombian market, insurers take into account the future profits from each active new enrollee as they transition to future diagnoses while being locked into their initial carrier, but commit to a one-time network choice.

My model consists of three steps. First, I model new enrollees' active discrete choice of insurance carrier. Their indirect utility is a function of insurers' network breadth per service and average out-of-pocket payments. These out-of-pocket payments are exogenous because the fraction due to copayments and required contributions to the health care system is more than 4 times the fraction due to coinsurance payments, where service prices are endogenous. The variation in choice sets across markets for individuals in the same cost-sharing tier identifies the consumer's disutility for out-of-pocket payments. Identification of the marginal utility for network breadth relies on exogenous variation in demographics and diagnoses across markets within insurer, once the endogenous variation in network breadth is averaged out by insurer fixed effects. I allow for sufficient observed heterogeneity in both parameters to capture adverse selection on the coverage level. This means that risk selection in my model occurs only on observable, un-reimbursed (or poorly reimbursed) consumer characteristics like diagnoses.

Second, I model insurers' average cost as a flexible function of network breadth and reference service prices. This function can be thought of as a reduced-form approximation of an equilibrium where insurers and hospitals bargain over service prices and then consumers make claims for those services. I use the list of references prices produced by the government in 2005 for each service in the national plan to capture average cost differences across consumer types. Insurers and hospitals use these reference prices in their bilateral negotiations, so they are a strong predictor of the cost of a consumer type. My average cost model allows insurers to exhibit economies of scope across services, which helps explain why I observe carriers that offer a broad network in one service, doing so for other services as well.

While my model of average cost can be derived from a more involved hospital choice model, it imposes fewer functional form assumptions and flexibly captures the relation between insurer costs and my object of interest, namely, the network breadth per service. For example, allowing insurers to exhibit complementarities across service categories would not be possible with a model of discrete hospital choice without relying on non-parametric identification (Compiani, 2019), but these complementarities or scope economies are necessary to explain insurers' network breadth choices in Colombia.

Finally, I model insurers' network formation cost as a service-specific administrative cost associated with inclusion of an additional hospital to the network. Preference heterogeneity for network breadth and cost heterogeneity across insurers, guarantee the existence of an asymmetric equilibrium in network breadth per service. Without sufficient heterogeneity in preferences and costs, my model would predict that all insurers choose narrow networks in all services, because with zero premiums and regulated cost-sharing, insurers have no incentives to invest in complete networks. Despite the strong regulation of plan characteristics, I show that there is a trade-off associated with provision of a narrow network: a narrow network carrier attracts fewer sick enrollees who are unprofitable, but also fewer healthy enrollees who are profitable under the current risk adjustment mechanisms, which reduces its total future profits.

I estimate the model on a novel administrative dataset that encompasses all enrollees to the contributory health care system in Colombia during 2010 and 2011, which represents nearly 50% of the population in the country (25 million individuals) and their associated health claims (650 million). I focus on individuals with insurance coverage over the entire sample period (9 million) and their claims (250 million). This dataset contains enrollee demographic characteristics such as sex, age, and municipality of residence, as well as enrollment spell length and insurer. For every claim the data reports provider, service, price, hospital length-of-stay, and contract type with the insurer. As is usual in the literature on hospital networks (Gowrisankaran et al., 2015; Capps et al., 2003), I use the claims-level data to recover each insurer's network of hospitals in each of the service categories provided in the national insurance plan, and conduct additional robustness checks on my network measure.

I find that consumers have a strong preference for broader networks and lower out-of-pocket payments, but this preference decreases with age and sickness. This is consistent with the intuition that older and sicker individuals have had more interaction with the health care system and likely concentrate their care in a few providers. It is also consistent with these individuals being less responsive to price than their healthy and young counterparts. Put together, these estimates imply that, conditional on sex and age, individuals with chronic conditions have a significantly higher willingness-to-pay for network breadth than healthy individuals.

Insurers' average cost is hump-shaped with respect to network breadth. At low coverage levels, the average cost is increasing in network breadth, but at high coverage insurers enjoy economies of scope. For example, for a narrow network carrier increasing network breadth for primary care by 10% raises the average cost of healthy females aged 19-44 by $8,236 pesos, while for a broad network carrier the increase equals $6,893 thousand pesos.

I also find that the network formation cost is strictly convex in the level of network breadth. The average network formation cost is at least 80% of long-run variable profits for the majority of insurers, which matches the ratio of administrative expenses to accounting profits from insurers' income statements. A decomposition of profit changes following an insurer unilaterally increasing network breadth shows that adverse selection, or the change in the composition of consumer types in demand, explains on average 46% of the variation in profits; while the direct effect of networks on average costs and network formation costs explains the remaining 54%.

In view of the growing regulation and literature regarding network adequacy in countries like the US (Mattocks et al., 2021; Prager and Tilipman, 2020; Ghili, 2020; Haeder et al., 2015) as well as concerns about access to health care in Colombia, I use my model to quantify how hospital networks respond to changes in the regulatory environment, and how these network changes translate into health care costs. The extent to which hospital networks respond to regulation that only affects insurers is informative of the degree of risk selection in the market. The findings of my paper speak to the trade-off between providing better access to care through broader networks, and containing health care costs.

In the first counterfactual, I eliminate the risk adjustment systems and reimburse insurers with a fixed per capita rate equal to the average cost per enrollee in the population, while holding government spending fixed. Eliminating compensations for health risk factors should exacerbate risk selection, which incentivizes insurers to narrow their networks in services that costly patients require. My findings show that without any type of risk adjustment, insurer competition becomes a race to the bottom in terms of network breadth. Insurers drop coverage of relatively expensive services such as hospital admissions by 19.6%, and significantly narrow the networks for relatively cheap services like consultations and laboratory by 4.1%. In the short-run, eliminating risk adjustment reduces consumer welfare by 7.4% due to worsened access to and quality of care.

In my second counterfactual analysis, I improve the government's risk adjustment formula by reimbursing diagnoses ex-ante, a dimension currently un-reimbursed. If diagnoses help better predict an individual's health care costs, including them in the risk adjustment formula should decrease risk selection incentives and promote broader networks. In this exercise, I find that average network breadth increases across insurers. Average coverage for relatively expensive services such as hospital admissions increases 3.2% relative to the observed scenario, while average network breadth for commonly used services like consultations increases 1.5%. Broader network coverage also raises consumer welfare for patients with chronic diseases by 3.2%. This is because access to care improves in services these patients disproportionately claim relative to healthy individuals.

In the last counterfactual exercise, I allow insurers to compete simultaneously over premiums and network breadth. I assume insurers can discriminate premiums along the dimensions of enrollee sex, age group, and income level, similar to the way they set premiums in their supplementary health insurance plans. Moreover, I assume that counterfactual premiums replace the consumers' required contributions to the health care system, and that insurers receive premium revenues directly, which implies government spending is zero in counterfactual. I find that premiums are U-shaped with respect to the enrollee's age, higher for males than for females, and weakly higher for higher income individuals. Deregulating premiums incentivizes insurers to broaden their networks 46.0% on average across the board of services. Put differently, my findings show that a zero-premium policy exacerbates the underprovision of insurance coverage. The network expansion that results after premium deregulation increases consumer welfare by 58.8% for individuals with chronic conditions, who are more likely to make health claims. However, welfare decreases 87.5% for those without diseases, who have low willingness-to-pay for network breadth and high demand elasticity with respect to out-of-pocket expenses.

While competition in hospital networks at the service-level is unique to the Colombian market, it can be reinterpreted as a game where insurers choose the probability of covering a specialized hospital. This alternative interpretation better fits the way other health care systems are designed such as the United States, where inclusion of a hospital to the network implies coverage of all services the hospital has to provide. The findings of my paper are relevant for Colombia where one the main reasons for dissatisfaction with an insurance company is narrow hospital networks (Ministerio de Salud y Protección Social, 2015). Findings are also relevant for public health insurance systems, where premiums and cost-sharing are standardized, that rely on regulated competition for insurance provision. One such market is Medicaid Managed Care in the United States, which as of 2016 covered over 73 million individuals in that country (Layton et al., 2018). My paper provides recommendations for how to improve the risk adjustment formula or regulate premiums to increase network breadth, decrease the concentration of risky patients in a few insurance carriers, and decrease the likelihood of insurer bankruptcy, which is a pressing issue in Colombia.

My paper contributes to several branches of the literature. In terms of the design of risk adjustment systems and premiums, there is extensive evidence of the effect of premiums on enrollment (Einav et al., 2019; Finkelstein et al., 2019; Tebaldi, 2017; Decarolis, 2015), of the impact of risk adjustment on selection effort (Brown et al., 2014; McWilliams et al., 2012; Nicholson et al., 2004), and of the relation between risk adjustment and premiums, mostly in the context of Medicare Advantage (Cabral et al., 2018; McGuire et al., 2013; Pauly and Herring, 2007). But, quantifying the effect of risk adjustment and premiums on hospital network breadth, risk, and costs is unique.

Related to the literature on risk selection in health insurance, I show that insurers can use service-level hospital networks to select risks. Selection mechanisms are usually unobservable to researchers, which is why this paper makes an important contribution to the understanding of insurer behavior in a tightly regulated market. Park et al. (2017) present evidence of selection through cost-sharing at the service level in Medicare Advantage, while I present evidence of selection through hospital networks. Shepard (2022) shows that adverse selection is possible through inclusion of one expensive hospital in the network. I explore the effect on adverse selection of including many hospitals and services in the network.

Other studies have identified different selection mechanisms. For example, Aizawa and Kim (2018) show that Medicare Advantage insurers advertise more in places where there are healthier, more profitable, newly eligible enrollees. And in a working paper, McNamara et al. (2021) study insurers' selective entry into markets with demographic characteristics that are highly reimbursed by CMS' risk adjustment formula. In addition to identifying service-level hospital networks as a risk selection mechanism, my paper differs from these in that I model insurers' strategic interactions in their network choices. I develop a tractable model of insurer competition in hospital networks that can be used to make predictions about market structure in health insurance systems where insurers compete mainly in non-price characteristics.

2 Background

The Colombian health care system was established in 1993 with Law 100. The system is divided into a "contributory" and a "subsidized" regime. The first covers formal employees and independent workers who are able to pay their monthly contribution to the system (nearly 51% of the population). The second covers individuals who are poor enough to qualify and are unable to contribute (nearly the remaining 49%). The national health care system has almost universal coverage with variation in the number of uninsured across departments due to difficult geographical access. Universal coverage means that risk selection does not happen on the individual's decision of whether to enroll or not but on the decision of which insurer to enroll.

Private insurers in Colombia's contributory system provide the national benefits package to enrollees who contribute a proportion of their monthly income.[3] The national plan covers a comprehensive list of more than 7,000 procedures and 673 medications as of 2010. The government sets premiums for the national plan to zero and sets cost-sharing rules as functions of the enrollee's income level, but they are homogenous across insurers and providers.[4,5] Enrollment of formal workers and independent workers is compulsory, so dropping out of the system while still receiving a monthly income can lead to monetary sanctions by the National Department of Taxes and Customs.

Hospital networks are the only dimension of insurer differentiation. Insurers can form these networks separately for each of the services offered in the national health insurance plan. Although the government does stipulate a set of network adequacy rules to guarantee appropriate access to health services, these rules are very coarse.[6] The rules recommend that insurers estimate demand

[3] Contributions equal 12% of the monthly income for independent workers and 8% for formally employed individuals with an additional 4% paid by the employer.

[4] 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%, the copay equals 2,100 pesos, and the maximum expenditure amount in a year equals 57.5% times the MMW. For those with incomes between 2 and 5 times the MMW, the coinsurance rate is 17.3%, the copay is 8,000 pesos, and the maximum expenditure is 230% times the MMW. Finally, for people with incomes above 5 times the MMW, the coinsurance rate equals 23%, the copay 20,900 pesos, and the maximum expenditure amount is 460% times the MMW.

[5] The average exchange rate during 2011 was $1,847 COP/USD.

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

for health services by risk group in each market; analyze hospital supply and installed capacity; and decide which hospitals can meet their demand for primary care, urgent care, oncology, and treatment of certain chronic diseases. After deciding on their network, insurers and hospitals engage in bilateral negotiations over service prices and type of contract (capitation or fee-for-service).

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

Department/city	Transfer
National	\$505,627.2
Multiplier a_k	
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

Private insurers 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 on a patient's previous diagnoses. For year t, the base transfer is calculated using the claims data from all insurers from year $t-2$ and it is roughly equal to the present value of the average health care cost per enrollee. For each risk pool defined by a combination of sex, age group, and municipality, the government calculates a risk adjustment multiplier. Table 1 shows the national base transfer and its value for some special municipalities and table 2 shows the risk group multipliers for 2011. Because of the coarsely defined risk pools, the current ex-ante risk adjustment formula behaves poorly when trying to predict health care costs (Riascos et al., 2017a,b, 2014).

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

Age group	Sex	Multiplier a_g
<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
>74	—	3.9419

The High-Cost Account, on the other hand, compensates insurers for enrollees 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. Because this list of diseases is not exhaustive, selection incentives are still present after ex-post risk adjustment (Riascos, 2013; Riascos and Camelo, 2017). Regulation of plan characteristics and imperfect risk adjustment, incentivizes private insurers in Colombia's contributory health care system to select risks through the discretionary elements of the national insurance plan, namely the service-level hospital networks. The next section describes the dimensions of insurer heterogeneity and the extent of risk selection through hospital networks.

3 Data and descriptive evidence

The data for this paper are the cross sections of all enrollees to the contributory system during 2010 and 2011 (25 million), and their claims through the national insurance plan (650 million). I focus on the sample of individuals with continuous enrollment spells or no gaps in enrollment per year (9 million) and their associated claims (270 million). Of the continuously enrolled, 2/3 are *current enrollees* or individuals who are enrolled throughout 2010 and 2011. The remaining 1/3 are *new enrollees* or individuals who enroll for the first time in 2011. Current enrollees can be further characterized by their switching status. I observe only 0.06% of current enrollees switch their insurance carrier from 2010 to 2011, which evidences the extent of consumer inertia in this market. Because there is near universal coverage, new enrollees to the contributory system can be individuals who move from the subsidized system after they get a job, or those who for some reason were uncovered at any particular moment in time and then enroll the health care system.

For every claim, the data reports date of provision, service description, service price, provider, insurer, hospital length-of-stay, ICD-10 diagnosis code, and contract type under which the insurer reimbursed the provider. For every enrollee, the data has basic demographic characteristics like sex, age, municipality of residence, and enrollment spell length in the year. I assign each individual the average income in its municipality, sex, and age group, using aggregate income data from enrollees to the contributory system, made available by the Colombian Ministry of Health. With the demographic information, I recover the per capita ex-ante risk-adjusted transfer received by each insurer. I also have data on the total High-Cost Account transfers that each insurer received in every year. Providers that are reimbursed by an insurer in the claims data are considered to be in-network, so I do not observe out-of-network claims. This is not an issue for my analysis given that out-of-network claims are infrequent, they are usually not reimbursed by the insurer, and thus the enrollee has to pay completely out-of-pocket for out-of-network care.

During my sample period there are 23 private insurers in Colombia's contributory health care system. I focus on the 14 largest insurers that account for over 97% of enrollees. Insurers compete separately in every market, which I define as a Colombian state (similar to an MSA in the US). There are 32 markets in my data. The Colombian insurance market is highly concentrated, with the top 3 companies covering over 50% of individuals (see appendix table 1). Panel A of appendix figure 1 shows that all insurers have presence in the central region of the country but peripheral markets, characterized by difficult geographical access, have fewer insurers, the smallest of which has a duopoly of carriers.

Table 3 presents average demographic characteristics of new enrollees and current enrollees by switching status. Switch-ins and new enrollees are on average younger and have lower prevalence of chronic conditions than stayers during 2011. New enrollees are disproportionately male, while switch-ins and stayers are disproportionately female. Across the three samples, EPS037 has the oldest population of enrollees with an average age of 38 years among stayers, 42 years among switch-ins, and 43 years among new enrollees. This insurer's population of enrollees is relatively sicker than it's competitors', with 49% of stayers having a chronic disease.

Current and new enrollees also differ in their total health care cost. Figure 1 shows the 2011 average health care cost by age (in the solid line) and its associated 1st and 99th percentiles (in the shaded areas). The solid black line represents current enrollees and the solid red line corresponds

Table 3: Demographic characteristics of current and new enrollees

	Stayers 2011			Switch-ins 2011			New enrollees 2011		
Insurer	Age	Male	Sick	Age	Male	Sick	Age	Male	Sick
EPS013	37.3	44.3	32.6	38.4	39.3	37.6	30.4	58.7	7.2
EPS016	37.6	43.1	33.3	32.9	48.4	28.9	32.1	55.3	6.5
EPS037	52.9	39.6	48.9	42.3	37.2	29.4	42.7	50.6	9.5
EPS002	36.6	44.1	35.8	31.8	45.4	26.7	31.3	59.0	8.8
EPS017	35.2	43.4	30.1	32.7	44.4	23.1	31.9	60.0	9.5
EPS010	37.9	43.3	33.1	29.5	44.0	21.6	33.9	56.3	6.7
EPS018	38.1	44.6	25.9	31.2	40.6	16.4	30.5	56.2	7.4
EPS005	45.4	40.9	19.6	38.8	37.2	21.6	34.9	55.9	5.0
EPS003	38.7	44.0	32.1	37.4	34.6	19.8	33.9	56.2	5.9
EPS008	37.6	42.4	25.3	34.6	41.0	20.8	32.9	57.7	8.8
EPS023	35.0	45.2	27.0	30.3	43.2	18.2	29.3	60.8	6.4
EPS009	38.7	42.5	32.7	33.5	47.2	22.6	32.8	56.7	7.7
EPS012	40.5	43.7	39.3	33.6	39.4	33.3	31.7	58.0	9.7
EPS001	41.7	44.3	27.9	32.1	40.4	11.9	36.6	50.8	4.8

Note: Average age, percentage of males, and percentage of enrollees with chronic diseases by insurer in the sample of stayers, switchers, and new enrollees during 2011.

to new enrollees. In both samples, the mean and variance of total health care costs are increasing with age. Current enrollees are more expensive and exhibit a higher dispersion in costs relative to new enrollees, who are considerably cheaper conditional on age.

Since the government's ex-ante risk adjustment formula controls for age group, the rising trend in average cost suggests that insurers have incentives to engage in selection against old individuals with high risk scores. The rising trend in variance suggests there is scope to select individuals in the upper tail of the distribution who are more likely to be overcompensated by the formula (Brown et al., 2014). Selection based on age is possible in health care plans that cover the entire population, but less so in markets where insurance is provided to a specific age group like Medicare Advantage in the United States. Conditional on risk adjustment, selection incentives in Medicare Advantage arise from differences in diagnoses rather than differences in age. In Colombia, however, selection on both age and diagnoses is possible with the current risk adjustment mechanisms.

The coarse nature of the risk adjustment formula generates heterogeneous profits per enrollee. Table 4 presents the mean, 1st and 99th percentiles of profits per enrollee in the sample of new enrollees and current enrollees by switching status. The profit is calculated as the government's ex-ante and ex-post transfers, plus revenues from copayments and coinsurance rates, minus total health care costs.[7]

[7] Strictly speaking, insurer revenues should additionally include government transfers that compensate for administrative costs, but I'm only focusing on the risk-adjusted part of revenues in the surplus calculation.

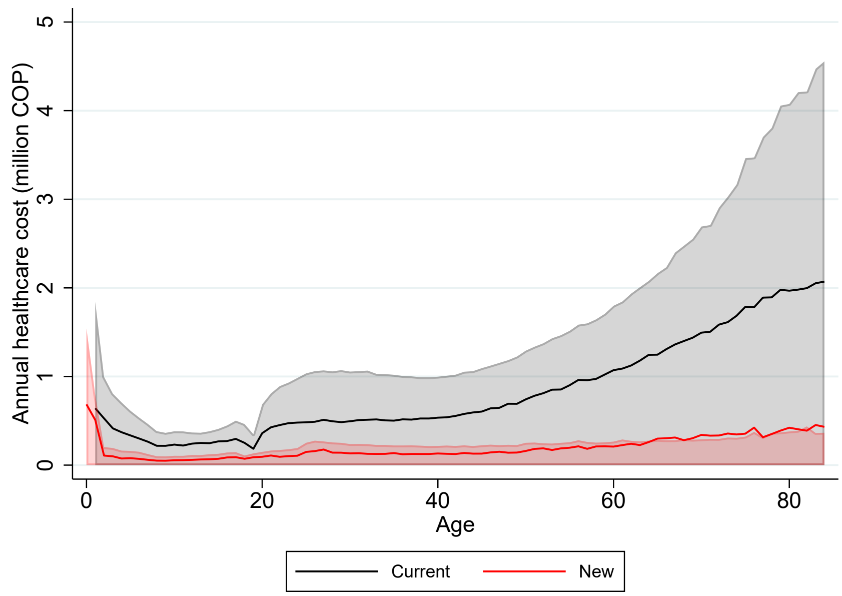


Figure 1: Average health care cost by age

Note: Average health care cost during 2011 by age in the solid line and its associated 1st and 99th percentiles in the shaded area. The black line corresponds to the sample of current enrollees and the red line to the sample of new enrollees.

If the risk adjustment formulas were able to completely eliminate risk selection incentives, we should see similar variance in the distribution of profits per enrollee across insurers, but this is not the case. For EPS037, the risk adjustment formulas highly overcompensate health care costs for its cheapest stayers, resulting in the highest 99th percentile across insurers. The formulas also severely undercompensate costs for this insurer's most expensive stayers, resulting in losses of $17 million pesos per enrollee in the 1st percentile of the distribution.

New enrollees' average profit is significantly higher than the average profit of switch-ins and stayers. New enrollees' distribution of profits is also less skewed to the left. Switch-ins are more profitable than stayers, which evidences positive selection into carriers. So, if insurers engage in risk selection, we should observe high selection efforts for new enrollees and switch-ins who are more profitable than stayers, in general.

Given that almost every aspect of the national insurance plan is regulated by the government except for service-level hospital networks, insurers can use these networks as a risk selection mechanism. Colombia's provider landscape is characterized by a large number of individual or stand-alone doctors and a relatively small number of hospitals and clinics. In 2011, there were around 11,200 hospitals and small clinics, which comprised nearly 1/6 of all providers in the country. I focus on the

Table 4: Distribution of profit per enrollee by switching status

Insurer	Stayers 2011			Switch-ins 2011			Newly enrolled 2011		
	Mean	P1	P99	Mean	P1	P99	Mean	P1	P99
EPS013	0.08	-5.30	1.86	-0.03	-4.24	1.68	0.34	-1.00	1.99
EPS016	0.08	-6.25	1.90	0.08	-4.80	1.64	0.38	-1.27	1.99
EPS037	0.13	-17.16	2.15	0.49	-1.51	1.99	0.66	-1.34	3.12
EPS002	0.09	-5.62	1.82	0.24	-1.89	1.47	0.35	-1.12	1.99
EPS017	0.04	-6.23	1.83	-0.33	-26.33	1.29	0.31	-1.60	1.99
EPS010	0.10	-5.69	1.87	0.10	-3.82	0.82	0.39	-1.03	2.11
EPS005	0.14	-8.38	1.99	0.12	-7.47	1.67	0.39	-1.62	2.11
EPS018	0.04	-6.23	1.83	0.14	-2.57	1.07	0.28	-1.63	1.68
EPS003	0.07	-5.93	1.88	0.13	-7.64	1.63	0.42	-0.82	2.11
EPS008	0.08	-6.54	1.88	0.02	-5.97	1.59	0.30	-2.40	2.02
EPS023	0.10	-4.51	1.68	0.18	-1.86	1.94	0.33	-0.82	1.68
EPS009	-0.36	-15.25	1.87	0.21	-2.67	2.11	0.26	-3.32	1.99
EPS001	0.15	-7.76	2.06	0.35	-1.14	1.40	0.49	-0.73	2.11
EPS012	0.08	-6.35	1.80	0.30	-1.48	1.72	0.36	-1.00	1.90
Total	0.08	-7.76	1.99	0.13	-3.70	1.63	0.44	-1.14	2.11

Note: Mean, 1st and 99th percentiles of profit per enrollee for each insurer in the sample of current enrollees by switching status and new enrollees. Stayers are individuals whose choice of insurer doesn't change across years. Switchers are individuals whose choice of insurer changes from 2010 to 2011. Surplus is calculated as the risk-adjusted transfers (ex-ante and ex-post), plus revenues from copays and coinsurance rates, minus total health care cost. Surplus is measured in millions of 2011 Colombian pesos.

sample of hospitals and clinics ("hospitals" for short) that provide inpatient, surgical, urgent care, and diagnostic services, which are in the upper tail of the distribution of health care costs, where variance is high and risk selection incentives are more important. For this sample of hospitals I am able to observe claims for each service category that they have contracted with insurers, while for small providers or stand alone doctors it is more likely that there are zero claims for certain services even though they are part of the network. Because I recover service-specific hospital networks from observed claims, to avoid this measurement problem I drop small providers. My results are robust to alternative network definitions.

I obtain the list of 1,663 hospitals in 2011 and 1,453 in 2010 from the Special Registry of Health Care Providers by the Ministry of Health.[8] I match hospitals in my claims data to the registry and end up with a 97% match rate in 2010 and an 87% match rate in 2011. The matched sample of hospitals represents 13% of individual health care costs, 10% of individual claims, 32% of total health care costs, and 27% of total claims in the in the contributory system. Panel B of appendix figure 1 shows the total number of hospitals per market that results from this sample selection criteria. The largest market has 196 hospitals and the market with the capital city of Bogotá has 86 hospitals.

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

Unlike countries such as the United States, where links between insurers and hospitals involve contracts over all services in the insurance plan, insurers in Colombia have discretion over which services to cover at which hospitals. This set of services is in part determined by differences in hospital specialty and available capacity, but service coverage also depends on the type of consumers that insurance companies want to risk select upon. Appendix table 2 explores insurer generosity in the services covered with their hospital network across markets. The table presents several statistics of the percentage of offered services. Although by law insurers have to guarantee access to all services and procedures in the national benefits package, there is variation in service coverage across insurers and markets because I am only focusing on a subset of all providers in the country. For every claim in my data, I observe the 6-digit service code from the national insurance plan, which I assign to one of 58 service categories ("service" for short) describing surgical and non-surgical procedures.[9] The complete list of service categories is provided in appendix 1. Large insurers, such as EPS013 and EPS037, cover more than 80% of services in the average market both during 2010 and 2011. Other insurers like EPS012 and EPS023, cover only 63% and 52% of services on average during 2011, respectively.

If insurers use their hospital and service networks to select risks, then differences in risk selection efforts should be reflected in differences in network breadth, defined as the fraction of all hospitals in a market that can provide the service that are covered by the insurer. Table 5 shows that there is significant heterogeneity in network breadth across insurers and markets. EPS013 and EPS037 have relatively broad networks in almost all markets, covering an average of 86% and 61% of hospitals during 2011, respectively. Smaller insurers, like EPS008 and EPS023, tend to cover between 43% and 57% of hospitals in the average market during the same year. For the majority of insurance companies, network breadth exhibits small increases from 2010 to 2011. For 5 out of the 32 markets there are no changes in the absolute number of hospitals covered by any insurer.

Network breadth as a continuous measure in the unit interval is my primary object of interest in the rest of this paper. Enrollee satisfaction surveys conducted by the Colombian Ministry of Health show that network size is one of the top three reasons of dissatisfaction with an insurance company, together with long waiting times for scheduling doctor appointments, and the amount of paperwork needed to claim services, procedures, and medications (see appendix figure 2). By collapsing networks to a one-dimensional object for each service, I am effectively ignoring the hospital's identity

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

Table 5: Distribution of number of in-network providers across departments

	2010				2011			
Insurer	Mean	SD	P25	P75	Mean	SD	P25	P75
EPS013	88.7	11.5	81.0	98.3	85.6	12.5	72.9	95.2
EPS016	83.5	9.8	76.7	92.9	84.7	11.6	79.5	92.3
EPS037	76.7	22.1	70.3	93.5	61.0	25.2	38.0	80.5
EPS002	75.2	11.4	67.7	82.4	78.3	14.4	70.9	88.4
EPS017	50.2	22.3	34.1	70.8	43.5	20.6	32.1	52.2
EPS010	54.6	22.6	33.3	72.3	55.2	21.5	37.7	71.7
EPS005	61.3	22.0	51.2	75.9	58.9	22.0	42.5	71.7
EPS018	44.2	31.4	17.8	74.1	38.9	28.5	11.2	64.7
EPS003	69.4	22.3	56.1	88.9	69.4	21.5	57.8	87.0
EPS008	46.8	25.2	19.9	66.1	43.8	30.0	16.0	71.2
EPS023	53.4	24.8	41.2	71.2	57.1	24.4	47.8	79.0
EPS009	40.7	37.2	13.1	84.3	35.0	38.8	4.7	84.6
EPS001	48.8	15.7	35.6	61.0	44.0	12.4	39.0	50.0
EPS012	50.7	26.6	23.5	76.7	51.4	39.3	6.0	75.5

Note: Mean, standard deviation, and 25th and 75th percentiles of the percentage of in-network hospitals for each insurer across markets during 2010 and 2011.

from the network. This approach would be problematic if consumers care about hospital identity conditional on the service.

One way to check for differential sorting on hospital identity is to compute the standard deviation of claims across hospitals conditional on the service category and the insurer. A relatively high standard deviation would be suggestive of patients sorting differentially across providers. While this strategy does not allow me to distinguish the reason for sorting –whether it is the hospital's identity, distance to the hospital, or adverse selection on hospital choice– it does allow me to check for the presence of sorting conditional on the service. Panel (a) of appendix figure 3 shows that for over 65% of insurer-service combinations the standard deviation of claims across hospitals is at most 2, while for 10% of insurer-services the standard deviation is greater than 15 claims.

Panel (b) of appendix figure 3 shows the standard deviation of claims separately for each service category in the x-axis. Service categories are arranged in increasing order with respect to the total number of claims. For more than half of services, the standard deviation of claims is at most 5. These correspond to highly complex services like procedures in brain, skull, cardiac vessels, and stomach. As we move up in number of claims towards less complex service categories such as primary care, laboratory, and imaging, the standard deviation of claims increases suggesting a role for hospital identity as a determinant of insurer choice. For these less complex services, however, it seems reasonable to assume that consumers care more about distance to the hospital than they do hospital identity. In that case, my measure of network breadth would be reflective of the probability that a hospital is located near the consumer. Because the main capital cities in the country have known star hospitals,[10] another way to deal with hospital identity is to exclude these markets from my estimation. I provide robustness checks of this style later in the document.

3.1 Selection and network breadth

The descriptive statistics so far show that there is substantial variation in network breadth and service coverage across insurers and markets. The high concentration of patients with chronic diseases and new enrollees, as well as differences in the distribution of profits per enrollee across insurers, are suggestive of differences in selection efforts. A natural question that arises is why is there heterogeneity in network breadth if insurers in this health care system can not charge higher prices for higher coverage or can not apply different cost-sharing rules to different types of consumers. In absence of price competition, it seems natural to predict that all insurers would choose narrow networks across the board of services.

In the reduced-form analysis that follows I show evidence of trade-offs to the provision of a narrow and a broad network that are consistent with an asymmetric equilibrium in network breadth in absence of price competition. These trade-offs will evidence the scope for risk selection at the service level and patients' response to insurers' network breadth choices. These descriptive exercises will only be suggestive of the presence of selection in the Colombian market but are not meant to represent causal relationships, as I won't be able to disentangle supply-side risk selection from demand-side adverse selection without a model of how consumers choose carriers and how carriers choose coverage after taking expectations of demand.

I start by estimating the correlation between network breadth and insurer costs using the following equation:

$$\log(y_{ijmk}^{2011} + 1) = \beta_0 + \beta_1 H_{jmk}^{2011} + \mathbf{d}_i \beta_2 + \eta_m + \delta_j + \gamma_k + \varepsilon_{ijk}$$

where y_{ijmk} is the 2011 annual health care cost in service category m of individual i enrolled to insurer j in market k, H_{jmk}^{2011} is insurer j's network breadth in market k for service m during 2011, $\mathbf{d}_i$ is a vector of consumer demographics and diagnoses, η_m is a service fixed effect, δ_j is an insurer fixed effect, and γ_k is a market fixed effect.[11]

[10] For example, Fundación Santa Fe in Bogotá, Fundación Valle del Lili in Cali, Hospital Pablo Tobón in Medellín, and Hospital Metropolitano in Barranquilla.

[11] Results of this regression are robust to alternative transformations of the dependent variable such as the inverse

Table 6 shows that greater network breadth for a particular service is associated with higher annual health care costs in that service category, both for the sample of stayers and new enrollees during 2011. This positive correlation holds after controlling for consumer observable characteristics that are accounted for in the government's risk adjustment mechanisms. More generally, results suggest that broad network carriers have higher costs than narrow network carriers, which is consistent with the literature on network formation in health insurance that finds that insurers with broad networks have lower bargaining leverage with hospitals (Ho and Lee, 2019).

Table 6: Network breadth, utilization, and costs

	log(*total service cost*)	
	Stayers	New
H_{jmk} / H_{jk}	0.06***	0.02***
	(0.003)	(0.001)
Demog + Diag	Y	Y
Market	Y	Y
Service	Y	Y
N	14,487,530	14,496,056
R^2	0.44	0.21

Note: OLS regression of the logarithm of 2011 health care cost per service in column (1) and 2011 total cost in column (2), on insurer network breath during 2011. Includes demographic controls and market fixed effects. Column (1) additionally includes service fixed effects. Robust standard errors in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

Next, I estimate the correlation between consumer choice and network breadth using the following linear regression at the insurer-market level:

$$s^m_{jk} = \beta_0 + \beta_1 H^m_{jk} + \gamma_k + \varepsilon^m_{jk}$$

Here s^m_{jk} is either the share of patients with renal disease, cancer, arthritis, or share of childbirths, and H^m_{jk} is either network coverage for dialysis, chemotherapy and radiotherapy, procedures in bones and facial joints, and delivery, respectively. Results in table 7 show a positive correlation between network breadth for a particular service and the share of patients with health conditions whose treatment requires that service. This correlation is consistent with positive selection at the service level. Put together results show that insurers that offer broad networks have higher demand overall, but specially higher demand from individuals with chronic conditions who are

hyperbolic sine. They are also robust to alternative estimation methods such as two-part models, with a logit for the probability of having non-zero cost and a log-linear regression of costs conditional on having non-zero cost.

usually costlier, conditional on risk adjustment. Alternatively, insurers that offer narrow networks have lower demand but also tend to attract individuals who are usually healthier and less costly, conditional on risk adjustment.

Table 7: Service-specific network breadth and share of chronic patients

Share of patients with	Stayers	New
(1) *Renal disease*		
H_{jk} Dialysis	0.005	0.001***
	(0.004)	(0.0004)
(2) *Cancer*		
H_{jk} Therapy	0.06***	0.007*
	(0.01)	(0.004)
(3) *Arthritis*		
H_{jk} Procedures in bones	0.02***	0.002
	(0.006)	(0.001)
(4) *Childbirth*		
H_{jk} Delivery	0.01**	0.006***
	(0.005)	(0.001)
N	241	241

Note: OLS regressions of the share of patients with renal disease, cancer, arthritis, and childbirth, on the percentage of in-network providers of a particular service. All models include market fixed effects. Sample is constrained to insurers with at least 200 enrollees in each market. ***p<0.01, **p<0.05, *p<0.1.

Because these set of results use current market shares and costs as dependent variables, they conflate the effects of selection and moral hazard in health insurance demand. From these regressions it is difficult to assess how much of the increase in costs is coming from (i) individuals consuming more health care because of better network access (moral hazard), (ii) individuals selecting plans that have better coverage in services they anticipate needing (adverse selection into moral hazard), or (iii) insurers designing their coverage to maximize market share and profits (risk selection).

To separate (i) and (ii) from (iii), table 8 presents results of a regression on the sample of current enrollees where the dependent variable is the number of service-specific claims in 2011, conditional on 2010 patient diagnoses and demographics. The regression specification given by:

$$y_{ijk}^{m,2011} = \beta_0 + \beta_1 H_{jk}^{m,2011} + \mathbf{d}_i^{2010}\beta_2 + \gamma_k + \varepsilon_{ijk},$$

captures the extent of selection into moral hazard in the lines of Einav et al. (2013). A positive correlation between service-specific network breadth and number of claims for that service in $t+1$ would be suggestive of patients enrolling carriers with more generous coverage for services that they anticipate needing given their health conditions in t. More specifically, $y_{ijk}^{m,2011}$ is the number of claims of individual i for service m in 2011, $H_{jk}^{m,2011}$ is insurer j's network breadth in service m during 2011, $\mathbf{d}_i^{2010}$ is a vector of demographics and diagnoses in 2010, and γ_k is a market fixed effect.

Table 8 shows that the probability of childbirth in 2011 among women who were in childbearing age during 2010 is increasing in the coverage for delivery services. An increase of one percentage point in the proportion of in-network hospitals for delivery is associated to a 0.02% increase in the likelihood of childbirth. The number of dialysis claims, antirheumatic drug claims, and chemotherapy claims are also positively correlated with the percentage of covered hospitals for dialysis treatment, for procedures in bones and joints, and for chemotherapy, respectively. In particular, a percentage point increase in coverage of hospitals for dialysis is related to a 0.10% increase in the number of dialysis claims in 2011 among the sample of patients that were diagnosed with renal disease during 2010.

The fact that even for very standard medical treatment guidelines, as for renal disease, rheumatoid arthritis, and cancer, there is a positive correlation between demand and network breadth, means that service-level selection is possible in this health care system. An effective strategy to avoid enrollment from costly current enrollees would be to provide lower coverage in services these patients are most likely to claim. These results, however, are also suggestive of consumer moral hazard. To isolate the effect of selection from moral hazard, (ii) and (iii) from (i) above, I explore whether insurers' network breadth choices are correlated with their enrollees baseline costs and risk scores next.

I estimate a regression in the spirit of Brown et al. (2014) to compare baseline costs of switchers into insurers that reduce their network breadth over time to baseline costs of stayers in insurers that expand their network breadth. This analysis allows me to isolate the effect of selection from moral hazard by focusing on baseline costs rather than current costs as outcome. The regression specification is as follows:

$$\begin{aligned} y_{imk}^{2010} = & \beta_0 + \beta_1 (H_{j'mk}^{2010} - H_{j'mk}^{2011}) + \beta_2 Switch_{ik} + \beta_3 Switch_{ik} \times (H_{j'mk}^{2010} - H_{j'mk}^{2011}) \\ & + \mathbf{d}_i \beta_4 + \lambda_m + \delta_{j'} + \eta_k + \varepsilon_{imk} \end{aligned} \tag{1}$$

y_{imk}^{2010} is either the logarithm of total health care cost of individual i in service m during 2010 or an indicator for having non-zero claims for service m in 2010, $Switch_{ik}$ is an indicator for whether the consumer switched carriers from one year to the other, j' denotes the insurer chosen in 2011, $H_{j'mk}^{2010}$ is the 2010 network breadth of insurer j' and $H_{j'mk}^{2011}$ is the 2011 network breadth of insurer

Table 8: Service-specific network breadth and types of claims

	Stayers 2011
(1) $1\{Childbirth\}$	
H_{jk} Delivery	0.02***
	(0.001)
N	1,085,206
(2) $ihs(dialysis\ claims)$	
H_{jk} Dialysis	0.10***
	(0.01)
N	83,765
(3) $ihs(antirheumatic\ drugs)$	
H_{jk} bones and joints	0.004*
	(0.003)
N	102,602
(4) $ihs(chemotherapy\ claims)$	
H_{jk} Therapy	0.005
	(0.004)
N	439,176
Demog	Y
Market	Y

Note: OLS regressions of the probability of childbirth, and the inverse hyperbolic sine of dialysis claims, antirheumatic drug claims, and chemotherapy claims during 2011, on the percentage of service-specific in-network providers in 2011 conditional on diagnoses received during 2010. All models include market, age group, and sex fixed effects. Robust standard errors in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

j'. Additionally, $\mathbf{d}_i$ is a vector of demographics and diagnoses, λ_m is a service fixed effect, $\delta_{j'}$ is an insurer fixed effect, and η_k is a market fixed effect. The coefficient of interest is β_3.[12]

Column (1) of table 9 shows that individuals who switch into carriers that reduce their network coverage over time tend to be less costly in that service category than individuals who don't switch. A percentage point decrease in network breadth during 2011 relative to 2010 is associated to a 23% reduction in baseline costs. This could be either because the individual knows which services she will need and switches into a carrier with good coverage for that service, or because the insurer selectively narrows a service network to attract individuals with lower baseline cost in that service. Results in column (2) for the probability of making a claim in each service category are consistent with these interpretations. A person who switches into a carrier that reduces its network breadth for a particular service, tends to be 2% less likely to make a claim in that service category relative to stayers.[13] However, these regressions alone do not allow me to separate the different sources of

[12]Results in column (1) of table 9 are robust to an inverse hyperbolic sine transformation of the dependent variable. They are also robust to estimating baseline costs using a two-part model, with a first-stage logit for the probability of having non-zero cost, and a log-linear regression for the second stage conditional on having non-zero cost.

[13]In equation (1) the change in network breadth is defined conditional on the insurer that was chosen in 2011 (j'). An alternative definition compares network breadth between the insurer chosen in 2010 (j) and the one chosen in

selection, for which I will need a model of insurer demand and competition in network breadth.

Table 9: Selection on baseline costs and risk

	$\log(total\ cost_{ijmt}^{2010}+1)$ (1)	$any\ claim_{ijmt}^{2010}$ (2)	$\log(Risk\ transfer_{new}^{2011})$ (3)
$H_{j'mk}^{2010}-H_{j'mk}^{2011}$	0.004*	-0.0001	-0.17***
	0.002	0.0002	0.008
Switch	-0.09***	-0.007***	—
	0.02	0.001	
Switch × $(H_{j'mk}^{2010}-H_{j'mk}^{2011})$	-0.23***	-0.02***	—
	0.08	0.007	
Demog+Diag	Y	Y	—
Market	Y	Y	Y
Service	Y	Y	—
Insurer	Y	Y	Y
N	14,457,009	14,457,009	2,653,415
R^2	0.50	0.51	0.06

Note: Columns (1) and (2) use a random sample of 250,000 current enrollees. Column (1) presents results of an OLS regression of the logarithm of 2010 total service-specific costs on a switching indicator and the difference in network breadth between 2010 and 2011 for the 2011 choice of insurer. Column (2) shows results of an OLS regression for an indicator of non-zero service-specific claims on the same variables as before. Both columns include demographics and diagnoses indicators, as well as insurer, service, and market fixed effects. Column (3) presents results of an OLS regression of the logarithm of new enrollees' risk-adjusted transfer on the difference in network breadth between 2010 and 2011, and market and insurer fixed effects. Robust standard errors in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

The estimates from the previous equation may have low statistical power since they are identified from the 0.06% of current enrollees that switch their insurer in 2011. Another way to show that selection is meaningful in this market by leveraging consumer baseline characteristics is to regress individuals' risk scores on insurers' 2011 network breadth choices. The risk score is given by the per capita risk-adjusted transfer from the government that varies across sex, age group, and municipality combinations, and that is known to insurance companies before networks are formed. I estimate the following equation:

$$\log(Risk\ transfer_{ijk}^{2011}) = \beta_0 + \beta_1(H_{j'mk}^{2010} - H_{j'mk}^{2011}) + \delta_j + \eta_k + \varepsilon_{ijk}$$

where H_{jk}^{2010} and H_{jk}^{2011} are insurer j's total network breadth (across all services) in market k during 2010 and 2011, respectively, δ_j is an insurer fixed effect and η_k a market fixed effect. Column (3) of table 9 presents the results. Insurers that narrow their overall hospital network breadth tend to attract new enrollees with lower ex-ante risk scores compared to insurers that expand their network

2011 (j'). Results of this alternative specification are provided in appendix table 3 and are consistent with results in table 9. Moreover, it is possible to estimate equation (1) only on the sample of switchers. I present results using this sample restriction in appendix table 4.

breadth over time. If after making her first carrier choice, the enrollee is locked into this insurer due to inertia, then subsequent health care consumption will be greater the broader the networks are because of moral hazard. In section 4, I explicitly model how the observed selection effects can arise from both demand and supply, with consumers sorting into insurers based on their diagnoses and insurers choosing network breadth per service to maximize profits.

3.2 Consumer inertia

I turn now to exploring the magnitude of consumer inertia in insurer choice. In addition to preference heterogeneity for network breadth and out-of-pocket costs, inertia can help explain why insurers make different choices of network breadth if they are not better compensated for having a broad network, given that premiums are zero and other plan characteristics are standardized across insurers. If the dynamic effects introduced by inertia are meaningful, this could prevent the unraveling of insurers that offer broad networks to maximize initial market share, but face subsequently higher uninsurable costs (Polyakova, 2016).

Among the sample of current enrollees, I observe only 0.06% of individuals switching their carrier in 2011. Even conditional on age and number of additional diagnoses received in 2011 as seen in appendix figure 4, I find negligible switching probabilities, suggestive of substantial inertia. These probabilities, nonetheless, conflate the effect of switching costs/inattention and preference persistence induced by the variation in networks. To isolate the effect of switching costs and inattention, I follow Polyakova (2016)'s strategy and compare 2011 choices of current and new enrollees in markets where no insurer changed its network breadth over time. There are 5 out of the 32 markets that satisfy this condition.[14] Appendix table 5 presents the insurer market share in each of the samples. Current enrollees who face non-negative switching costs and are potentially inattentive and new enrollees who face zero switching costs and are attentive, make substantially different choices in the same year even after controlling for supply-induced changes in preferences. For example, EPS013 is more attractive to new enrollees than to current enrollees in 2011, while the opposite is true for EPS005.

[14] While this strategy separates the part of inertia due to consumer heterogeneity from the part due to state dependence as in Heckman (1991), it does not allow me to say how much of the effect comes from pecuniary costs associated to switching –like the cost of driving to the insurer's office or the cost of filling out the enrollment form– or from the consumer being inattentive about insurer characteristics.

3.3 Dynamics and networks

Consumer inertia can generate two types of dynamics in a health care market. The first type is known as insurers' invest-harvest incentives, which have been studied extensively in the literature (Klemperer, 1987, 1995; Farrell and Klemperer, 2007; Ho et al., 2017). If insurers can not commit to network breadth, an optimal strategy would be to offer broad networks in period t to attract more consumers and achieve a target market share (invest stage), and then narrow the networks in period $t+1$ to minimize per-enrollee costs once consumers are locked-in (harvest stage). I abstract from this type of dynamics because my short panel does not allow me to see the harvest incentives at play, given that networks are formed at the beginning of every year. Therefore, I assume insurers fully commit to their initial choices of network breadth, where I can rationalize risk selection through the investment stage.

The second way in which dynamics can arise in my context is from shocks to consumers' health status that induce switching. If the probability of switching is higher for groups of consumers with worsened health states, then the evolution of diagnoses may incentivize consumers to switch away from carriers that offer narrow networks in services they will need. This requires that the consumer reoptimizes her carrier choice every period. However, the descriptive evidence of the previous subsection showed that switching rates are very small conditional on age and diagnoses, so I can abstract from this source of dynamics as well, and focus only on new enrollees' decisions assuming they make a one-time myopic choice and are locked-in there after. New enrollees either experience infinite inertia or ignore the decision to reoptimize altogether after making their first choice. By definition, for this sample there is no initial conditions problem (Heckman, 1991).

If my data were reflective of a steady state, I would be able to model insurer competition in network breadth as a static game. However, in appendix 6 I show that several of the sufficient conditions for a steady state do not hold in my data. Thus, with a model of demand for insurance carriers and a model of insurer competition in network breadth that captures future profits, I can compute the effect of counterfactual policies such as alternative risk adjustment formulae or premium deregulation, that would affect not just the level of network breadth but the steady state as a whole.

My approach is unlike Ho et al. (2017), who focus only on comparing price levels for different degrees of consumer inattention in the context of Medicare Part D, conditional on being in steady state. It is also unlike Shepard (2022) who, in the context of the Massachusetts Health Insurance

Exchange, assumes that the share of inertial enrollees is fixed and that future markups are constant, in a counterfactual where insurers can drop hospitals from their networks but premiums and prices are not allowed to adjust dynamically. In my case, the premiums' channel is unavailable to insurers and average out-of-pocket expenses depend not only on service prices but of coinsurance rates and copayments which are exogenous to network choice. This means that the first order conditions that describe the insurers' profit maximization problem fully characterize the commitment equilibrium.

4 Econometric Model

In this section, I describe my model of demand for insurance carriers and competition in service-specific network breadth that allows me to disentangle demand- from supply-side selection. For carrier demand, I take a discrete choice approach and model new consumers' decisions as a function of network breadth per service and average out-of-pocket payments. For competition in network breadth, I assume insurers maximize the sum of current and future discounted profits by choosing once their vector of network coverage conditional on their competitors' choices. This one-time choice characterizes a full commitment Nash equilibrium in networks. To specify the profit function, I model insurers' average costs per consumer type as a function of network breadth, allowing average costs to exhibit economies of scope across pairs of services.

Existence of an asymmetric equilibrium in service-specific network coverage is obtained from preference heterogeneity for networks and out-of-pocket costs, as well as from insurer heterogeneity in average costs and network formation costs. While my model of insurer average costs can be derived from a model of consumer hospital choice and price bargaining as in Ho and Lee (2019), it can be understood as a reduced-form approximation to a bargaining equilibrium between insurers and hospitals. A model of hospital choice and price bargaining for each service category becomes intractable even with a small set of hospitals and services. My model of insurer competition coupled with the strict regulation of the Colombian health care market, thus, allows me predict market equilibria under counterfactual conditions in a tractable way.

4.1 Insurer demand

Start with insurer demand. Assume a new enrollee i of type (θ, l) with income level y living in market k, has diagnosis l. Conditional on diagnosis, with probability $\gamma_{\theta lmk}$ she will need each of the $m = \{1, ..., M\}$ service categories ("service" for short). An individual's type is given by the combination of sex, age category (<1, 1-4, 5-14, 15-18, 19-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, ≥75), and diagnosis $l \in L$, where $L =${cancer only, cardiovascular disease only, diabetes only, renal disease only, other disease only, two or more comorbidities, and no diseases}. The diagnoses in this list are both exhaustive and mutually exclusive. Sex and age group combinations are captured by θ. I assume the individual knows her diagnosis before making enrollment choices. This could be either because of medical family history or because, while being uncovered, she went to the doctor and received a diagnosis. My estimators on the preference for network breadth and out-of-pocket expenses are robust to a specification of demand where I allow for uncertainty in diagnosis (presented in appendix 7). My preferred demand specification is one where new enrollees know their diagnosis because risk selection occurs on observable, un-reimbursed (or poorly reimbursed) consumer characteristics such as diagnosis.

The consumer observes the insurer's index of coverage for each service m in market k, H_{jmk}, weights each service by the probability of claiming it $\gamma_{\theta lmk}$, conditional on the diagnosis, and then makes a one-time myopic choice of carrier. Denote by u_{ijk} the indirect utility of a new enrollee i in market k with diagnosis l for insurer j, which takes the following form:

$$u_{ijk} = \beta_i^D \sum_m \gamma_{\theta lmk} H_{jmk} - \alpha_i c_{\theta lyjk} + \phi_j + \varepsilon_{ijk} \tag{2}$$

where,

$$\begin{pmatrix} \beta_i^D \\ \alpha_i \end{pmatrix} = \begin{pmatrix} \beta^D \\ \alpha \end{pmatrix} \mathbf{X}_i$$

$\mathbf{X}_i$ is a vector of consumer demographics including sex, age, indicators for the diagnoses in L, and an intercept. $c_{\theta lyj}$ is the average out-of-pocket payment of a type-(θ, l) consumer with income level y at insurer j. ϕ_j is an insurer fixed effect capturing insurer unobserved quality. $\beta_i^D \gamma_{\theta lmk}$ is the average marginal utility for an additional percentage point of coverage for service m, α_i is the marginal disutility associated to average out-of-pocket payments, and ε_{ijk} is an unobserved shock to preferences assumed to be distributed T1EV.

I allow preferences for network breadth to vary across demographic characteristics and diagnoses to capture the extent of service-specific adverse selection documented in table 8. The average out-of-pocket payment captures differences in prices and utilization across insurer-hospital pairs and generates sorting of consumers into carriers based on average service prices. This sorting is needed to rationalize the existence of narrow network carriers in a full commitment equilibrium, where they exhibit lower costs compared to broad network carriers as documented in table 6. Average out-of-pocket payments are indexed to the enrollee's income level as the government's cost-sharing rules dictate so. In addition to the expenses incurred while going to the doctor, employed individuals also have to contribute 4% of their monthly income to the health care system. These contributions are included in my measure of average out-of-pocket payments.

Let $\mathcal{J}_k$ be the set of insurers available in market k and assume individuals have no outside option, since health insurance coverage in Colombia is almost universal. Let $H_{jk} = \{H_{jmk}\}_{m=1}^{M}$ and $H_k = \{H_{jk}\}_{j=1}^{\#\mathcal{J}_k}$; consumer i in market k enrolls insurer j with probability:

$$s_{ijk}(H_k) = \frac{\exp\left(\beta_i^D \sum_m \gamma_{\theta lmk} H_{jmk} - \alpha_i c_{\theta lyjk} + \phi_j\right)}{\sum_{g \in \mathcal{J}_k} \exp\left(\beta_i^D \sum_m \gamma_{\theta lmk} H_{gmk} - \alpha_i c_{\theta lyjk} + \phi_g\right)}$$

Identification. Even though network breadth is an endogenous equilibrium object, the demand specification aggregates H_{jmk} across services to eliminate the endogenous variation in this dimension. To account for the endogenous variation across insurers, note that my demand function includes insurer fixed effects ϕ_j, which help capture unobserved insurer-level characteristics that may be correlated with H_{jmk}. This means that β_i^D can be identified from variation in the demographic composition across markets within an insurer, that generates exogenous variation in $\gamma_{\theta lmk}$. However, if network breadth is correlated with additional unobserved quality measures that vary within insurers and are not captured by the fixed effects, then estimates of consumer preferences for coverage can be biased upwards. In appendix table 8, I provide a robustness check on demand that includes additional quality measures obtained from the Ministry's enrollee satisfaction surveys.

While service prices are endogenous because they are the result of bilateral bargaining between insurers and hospitals, the fact that coinsurance rates, copayments, and required contributions to the system are regulated by the government and fixed across insurers, hospitals, and services, alleviates the endogeneity problem of out-of-pocket expenses. The ratio of total copayments plus contributions, to total out-of-pocket expenses equals 82% on average across individuals, which suggests that service prices account at most for the remaining 18% of variation in average out-of-pocket payments.[15] Hence, two types of variation identify α_i. First, I leverage variation in cost-sharing rules across patients in the same market. Second, I leverage the variation in choice sets

[15]The extent to which negotiated prices respond to risk adjustment and premium deregulation through changes in the service-specific hospital networks is an interesting avenue for future research.

across markets, which generates variation in cost-sharing.

The insurer fixed effect ϕ_j is identified from variation in insurer market shares across markets. Finally, $\gamma_{\theta lmk}$ is the average prediction per consumer type, service, and market, of a logistic regression estimated at the individual level given by:

$$1\{Claim_{imk}\} = \psi_m + \psi_\theta + \psi_l + \psi_k + \psi_{imk} \tag{3}$$

The dependent variable is an indicator for whether consumer i living in market k made a claim in service category m. On the right side, ψ_m, ψ_θ, ψ_l, and ψ_k are service, sex and age group, diagnosis, and market fixed effects, respectively. ψ_{imk} is a mean zero shock to the claim probability that is independent of network breadth conditional on consumer observable characteristics. Even though new consumers make myopic decisions of insurance carrier, I assume that their expectations on the type of services they will need are correct on average. I estimate this regression on data from both current and new enrollees in 2010 and 2011. Appendix figure 7 presents the distribution of the resulting γ separately for healthy and sick individuals and for a few service categories including consultations, hospital admissions, imaging, and procedures in cardiac vessels, stomach, and intestines.

4.2 Insurer average costs

Following Shepard (2022), I assume the realized annual health care cost of consumer i of type (θ, l) under the observed service-specific networks, $C^{obs}_{ijk}(H_{jk})$, equals the cost of a type-(θ, l) consumer plus a random shock: $C^{obs}_{ijk}(H_{jmk}) = AC_{\theta ljk}(H_{jmk}) + \omega_{ijk}$. If sex, age, and diagnoses are observable or predictable by the insurer before the cost shock is realized, then ω_{ijk} is orthogonal to $AC_{\theta ljk}(H_{jmk})$ conditional on (θ, l) and I can recover this cost directly from the data by taking the average of observed costs across individuals of type (θ, l): $AC_{\theta ljk}(H_{jmk}) = \frac{1}{N_{\theta ljk}} \sum_{\theta(i)=\theta, l(i)=l} C_{ijk}(H_{jmk})$. By the law of large numbers, this average will equal the insurer's expected costs per consumer type. I then model the logarithm of average costs as a function of network breadth, as follows:

$$\begin{aligned}\log(AC_{\theta ljk}(H_{jk})) = \beta^S_0 \Big(\sum_m \gamma_{\theta lmk} A_m\Big) + \beta^S_1 \Big(\sum_m \gamma_{\theta lmk} H_{jmk}\Big) + \frac{1}{2M_k} \beta^S_2 \sum_m \sum_{n \neq m} \gamma_{\theta lmk} \gamma_{\theta lnk} H_{jmk} H_{jnk} \\ + \lambda_{\theta l} + \delta_j + \eta_k \end{aligned} \tag{4}$$

where M_k is the number of service categories available in market k, that is, services that existing hospitals in the market can provide. A_m is the average reference price for service m explained in more detail below, $\lambda_{\theta l}$ is a consumer type (θ, l) fixed effect, δ_j is an insurer fixed effect, and η_k is a market fixed effect. In appendix 8, I show that this average cost function has a direct relation to a model where consumers choose a hospital to receive service m. In my case, the average cost function need only capture adverse selection through hospital networks, so a complete model of consumer choice of hospital is not needed. My specification imposes fewer assumptions than a discrete hospital choice model, but is flexible enough to allow for cost variation across consumer types and across network breadth levels, so it can be understood as a reduced-form approximation to an equilibrium where insurers and hospitals engage in bilateral negotiation over service prices.

β_1^S captures differences in average costs that are driven by differences in network breadth alone. This coefficient represents the elasticity of average costs with respect to insurer j's network breadth. β_2^S captures the average degree of complementarity between pairs of services. If $\beta_2^S < 0$, then insurer j exhibits economies of scope across services, so greater coverage for service $n \neq m$ makes it more attractive to the insurer to provide higher coverage for service m. If instead $\beta_2^S \geq 0$, then insurer j's coverage decisions across services are at least independent. I include this measure of scope economies to rationalize the fact that I observe insurers that have a broad network in one service, offering broad networks in other services as well. EPS013, EPS016, and EPS037 tend to have generous coverage across all services and markets, while EPS008, EPS009, and EPS012 tend to have narrower networks across the board of services. My specification with two-way interactions between services is the simplest way to capture economies of scope. Any evidence of economies of scope between pairs of services will likely hold for higher-order interactions.

The first two terms in the right-hand side of equation 4 are multiplied by $\gamma_{\theta lmk}$ to capture the fact that increasing network breadth for one service does not increase the average cost of all consumer types by the same magnitude. The effect on average costs will depend on how likely it is that the consumer makes a claim in that service category, which is known by insurance companies before making coverage decisions. For example, increasing network breadth for c-sections is likely going to increase the average cost of women in childbearing ages but not the average cost of men.

$\lambda_{\theta l}$ measures the average degree of selection by type-(θ, l) consumers and provides a test for the type of selection occurring in this market, in the lines of Einav et al. (2010). In the presence of adverse selection, insurer j's average costs would be increasing with consumer type and thus with preferences for network breadth. In this case, the competitive equilibrium in network breadth would be below the efficient coverage level. On the contrary, if the market exhibits advantageous selection, average costs and consumer types would be negatively correlated, such that the competitive equilibrium in network coverage would be above the efficient level. If the Colombian health insurance market is characterized by adverse selection, then my counterfactual analyses will shed light on whether alternative risk adjustment formulae or competition in premiums can bring network breadth closer to efficiency.

Identification. The parameters of equation (4) can be identified from the variation in average costs within a consumer type across insurers that are identical except for their network coverage per service. My source of identification does not rely on different consumers implying different costs for similar insurers as in Tebaldi (2017) but, *conditional* on the composition of enrollee pools, for different coverage levels per service to imply different costs to the insurer. While network breadth is endogenous, the average cost specification aggregates this measure across services and includes a rich set of fixed effects to capture most of the endogenous variation in network breadth across and within insurers. I estimate the average cost function using OLS, which generates an error term $\nu_{\theta ljk}$ that is assumed not structural nor observable to insurers. The estimating equation is:

$$\log(AC_{\theta ljk}(H_{jk})) = f(A_m, H_{jmk}, \gamma_{\theta lmk}; \beta) + \nu_{\theta ljk}$$

where $f(A_m, H_{jmk}, \gamma_{\theta lmk}; \beta)$ equals the right-hand side of equation (4).

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. These lists were determined 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).[16] 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). I use the reference prices as a measure of average claim cost for service m in the insurers' average cost function. This means that β_0^S will adjust up or down depending on whether insurers bargain markups or markdowns with hospitals for the average service.

Because the reference prices are exogenous by nature, I can hold them fixed in my counterfactual exercises. Therefore, the only endogenous object in the average cost is network breadth, which I explicitly model in the next subsection. Figure 2 shows the average total claim cost for service m calculated directly from the claims data in black and the price instrument in red. The correlation

[16] The lists are known as the Compulsory Traffic Accident Insurance Tariffs Manual and the Social Security Institute Tariffs Manual.

between these two measures equals 0.77.[17]

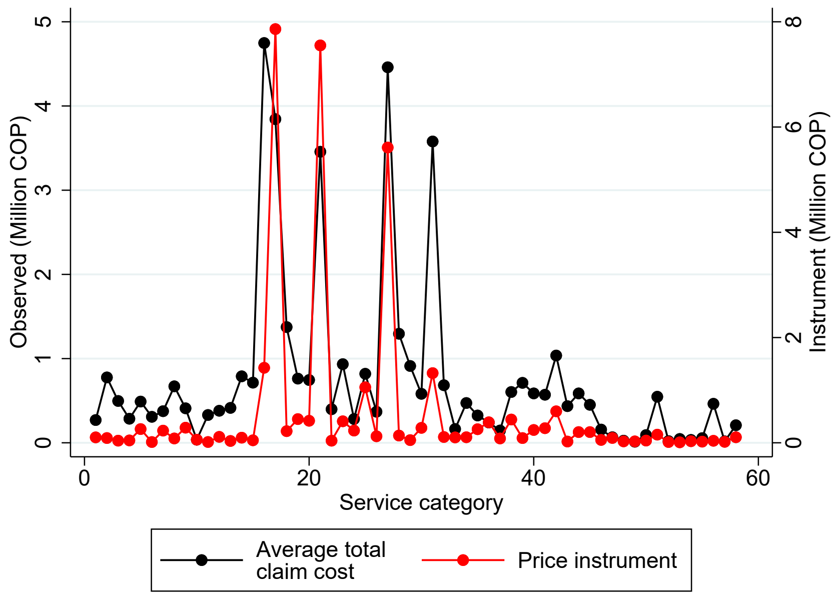


Figure 2: Average claims cost and reference prices

4.3 Competition in network coverage

Insurers compete separately in every market choosing their degree of service-specific network breadth after taking expectations of demand and costs, with Nash equilibrium as solution concept. Let $\pi_{ijk}(H_k, \theta, l)$ be insurer j's short-run per-enrollee profit in market k, which depends on j's network breadth and its competitors', H_k, as well as on the enrollee's type (θ, l). The short-run per-enrollee profit is given by:

$$\pi_{ijk}(H_k, \theta, l) = (R_{\theta k} - (1 - r_y)AC_{\theta ljk}(H_{jk}))s_{ijk}(H_k)$$

where $R_{\theta k}$ is the per-capita revenue including the ex-ante and ex-post risk-adjusted transfers from the government and average copayments, $AC_{\theta ljk}$ is the average cost of a type-(θ, l) consumer net of patients' coinsurance payments with r_y denoting the income-indexed coinsurance rate, and s_{ijk} is consumer i's choice probability for insurer j in market k.

I focus on a full commitment Nash equilibrium in which insurers choose networks once. Insurers maximize total profits given by the sum of short-run and long-run discounted profits per enrollee

[17] Let p_m^R be the reference price for service m and f the average inflation rate from 2005 to 2011, then $A_m = p_m^R \times (1 + f)^6$.

minus the cost of network formation:

$$\Pi_{jk}(H_k) = \sum_{\theta,l}\Big(\underbrace{\pi_{ijk}(H_k,\theta,l)N_{\theta lk}}_{short-run\ profit} + \underbrace{\sum_{s=t+1}^{T}\beta^s\sum_{\theta',l'}(1-\rho_{\theta'l'k})\mathcal{P}(l'|\theta,l)\pi_{ijk}(H_k,\theta',l')N_{\theta'l'k}\Big)}_{long-run\ profit}$$
$$\underbrace{-\sum_{m}\Big(\omega H_{jmk}+\xi_{jmk}\Big)H_{jmk}}_{network\ formation\ cost}$$

$N_{\theta lk}$ represents the market size of consumers type (θ, l), which is fixed over time, so there are no dynamics introduced by changes in population. In the expression for long-run profits, I assume that the probability of switching across carriers is zero as shown in section 3.2. $\rho_{\theta lk}$ represents the probability that type (θ, l) drops out of the contributory system. This probability is (assumed) exogenous to the choice of network breadth as it is governed by the event of falling in unemployment or by mortality. $\mathcal{P}(l'|\theta, l)$ is the transition probability from type (θ, l) to diagnosis l' (the transition across θ is deterministic). Future profits at year t are discounted by a factor of β^t.[18]

In addition to its indirect effect on insurer profits through expected costs and demand, I assume network breadth involves a direct cost to the insurer, which can be interpreted as an administrative cost associated with inclusion of an additional hospital to the network, where:

$$\xi_{jmk} = \xi_j + \xi_k + \xi_m + \Delta\xi_{jmk}$$

This network formation cost is non-linear in network breadth, ω capturing the convexity of the cost function. The network formation cost is also heterogeneous across insurers, markets, and services, with ξ_j, ξ_k, and ξ_m representing the insurer-, market-, and service-specific cost components, respectively.

I assume $\Delta\xi_{jmk}$ is a network formation cost shock that is *iid* across insurers, services, and markets, as well as over time. This cost shock is observed by insurers but unobserved to the econometrician, and it is mean independent of insurers' network formation cost shifters. By making unobserved costs have a multiplicative effect on network breadth, I am implicitly assuming that cost shocks can affect network breadth across all services and markets. An example of this can be that the insurer has an outstanding managerial or bargaining team that makes it less costly to offer

[18] In the formulation of insurer profits, I use θ to denote sex-age combinations as opposed to sex-age groups, to avoid introducing more notation, but to be consistent between transition probabilities and periods over which future profits are calculated (years).

broad networks across the board of services.

In presence of adverse selection, the trade-off associated to providing broader network coverage for a given service is that it increases both demand and costs. By doing so, not only does the insurer attract more consumers overall, but disproportionately attracts those with a high likelihood of claiming the service. A competitors' choice of service-specific network breadth also enters insurer j's profits through demand and costs. Insurer j's markups per service are a function of its network breadth choices across other service categories and of its competitors' network breadth choices in the same market. Competitors' choices also affect insurer j's total average cost through their effect on the composition of insurer j's enrollee types, given that there is no outside option.

My model of competition in service-specific network breadth can be understood as expanding on the first stage of the game in Liebman (2018), where insurers commit to network size before negotiating prices with hospitals and the solution concept is Markov-perfect equilibrium. I reinterpret network size as a continuous measure in the unit interval per service and focus on a simultaneous-move game with Nash equilibrium as solution concept. I do not model price negotiations as my focus is on the network formation stage.

My model of insurer competition also extends and complements the work in Shepard (2022), who models the binary decision of an insurer to include or exclude a star hospital from its network in the context of the Massachusetts Health Exchange. In my case, I allow for insurer heterogeneity in network breadth across different services and model insurer competition in network breadth accounting for future profits, although I abstract from the hospital's identity in the insurers' maximization problem.

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

$$\sum_i \left(\frac{\partial \pi_{ijk}}{\partial H_{jmk}} N_{\theta lk} + \sum_{s=t+1}^{T} \beta^t \sum_{\theta', l'} (1 - \rho_{\theta lk}) \mathcal{P}(l' | \theta, l) \frac{\partial \pi'_{ijk}}{\partial H_{jmk}} N_{\theta' l' k} \right) = 2\omega H_{jmk} + \xi_{jmk} \tag{5}$$

The left-hand side of equation (5) represents the marginal variable profit MVP_{jmk} and the right-hand side is the marginal cost of network formation.

Identification. Rewriting the FOC as:

$$MVP_{jmk}(H_{jmk}) = 2\omega H_{jmk} + \xi_j + \xi_k + \xi_m + \Delta\xi_{jmk}, \quad \forall \ H_{jmk} \in (0,1) \tag{6}$$

makes explicit the endogeneity between H_{jmk} and $\Delta\xi_{jmk}$. Insurance companies observe $\Delta\xi_{jmk}$ before or simultaneously to their network breadth choices. 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 $E[\Delta\xi_{jmk}|H_{jmk}] < 0, \ \forall \ H_{jmk} \in (0,1)$ and OLS estimation of (6) would result in ω that is biased towards the null.[19] Identification of network formation cost parameters, thus, relies on instrumental variables Z_{jmk} that are correlated with network breadth but not with the cost shock, and that are correlated with marginal variable profits only through network breadth, such that $E[\Delta\xi_{jmk}Z_{jmk}] = 0, \ \ \forall \ H_{jmk} \in (0,1)$.

The instrument set is populated as follows. First, I include the set of insurer, market, and service fixed effects in equation (6). Second, because I use data from 2011 in estimating the model and $\Delta\xi_{jmk}$ is assumed *iid* over time, I use the service-specific network breadth in 2010. Third, I include the average probability that a female, a person aged 19-44, and an individual without diseases make a claim for service m in market k. These probabilities are calculated as the average prediction of equation (3) across consumers that share the demographic traits above. Finally, I include the interaction between 2010 network breadth and the average service claim probability of a person aged 19-44.

The moment conditions at an interior solution given by $E[\Delta\xi_{jmk}Z_{jmk}] = 0$, can only rationalize the observed equilibrium in markets where no insurer chooses a corner solution in any of the services. In other words, the model is rejected by the data in markets where insurers choose service-level network breadth equal to zero or one. Thus, to estimate the parameters of the network formation cost as described by the FOC, I restrict my sample to the four largest markets in the country (Antioquia 05, Atlántico 08, Bogotá 11, and Valle de Cauca 76) that cover 60% of the population in the contributory regime, and to the top 10 insurers in these markets that cover 87% of enrollees (EPS001, EPS002, EPS003, EPS005, EPS010, EPS013, EPS016, EPS017, EPS018, EPS037). In this final sample, all insurers choose an interior solution in $H_{jmk}, \ \forall \ m, k$. In dropping the smallest insurers I am assuming that only competition among the largest carriers determines their network breadth choices, and that small insurers move only after the top companies make their choices.[20]

[19] 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.

[20] More generally, without dropping observations at the corners, the parameters of the network formation cost are partially identified. Note that the FOC at $H_{jmk} = 0$ is $MVP_{jmk} - \xi_j - \xi_k - \xi_m - \Delta\xi_{jmk} < 0$, and at $H_{jmk} = 1$, is $MVP_{jmk} - \tilde{\omega} - \xi_j - \xi_k - \xi_m - \Delta\xi_{jmk} > 0$. Estimation can be pursued using these moment inequalities following Pakes et al. (2015). But, to rationalize the corner at zero, the network formation cost should additionally include a fixed cost term. While this approach would utilize the entirety of markets for estimation, it is computationally costlier. Moreover, set identification of ω and the extensive set of fixed effects is not necessarily guaranteed.

5 Estimation

5.1 Insurer demand

I proceed with the estimation of demand for insurance carriers using a conditional logit with insurer fixed effects. For computational simplicity, I estimate equation (2) in a random sample of 500,000 new enrollees. Results are reported in table 10. I find that demand for insurers is decreasing in average out-of-pocket payments, suggestive of patient moral hazard; and increasing in network breadth, suggestive of positive selection into health insurance. In particular, an increase of one million pesos in the average out-of-pocket spending reduces demand for carriers by 6.64 percentage points. An increase of one percentage point in network coverage increases the probability of choosing an insurer by 2.37 percentage points.

I find that interactions between demographics and diagnoses with carrier characteristics also matter for enrollment decisions. Males are as sensitive to out-of-pocket expenses as females but have a stronger taste for network breadth. Older patients are less likely to enroll broad network carriers and are more sensitive to out-of-pocket expenses, while younger patients have a higher preference for network coverage and their insurance demand is more inelastic. These findings are consistent with the idea that older patients have had more interaction with the health care system, which makes them more likely to concentrate their care in a few providers.

Findings also show that patients with chronic conditions do not necessarily have stronger preferences for network breadth than their healthy peers, but they are significantly less responsive to out-of-pocket expenses. On the one hand, the interaction between network breadth and indicators for each of the chronic diseases are all negative and significant compared to individuals without diagnoses. On the other hand, interactions between average out-of-pocket spending and diagnosis indicators are all positive and significant. In fact, for individuals with renal disease, the effective demand elasticity with respect to the out-of-pocket expenditure is nearly zero. Finally, relative to people living in urban areas, I find that those who live in special peripheral municipalities characterized by fewer providers and more difficult access to care, have stronger preferences for coverage and a relatively inelastic demand with respect to out-of-pocket expenses.

Estimators on the preference for network breadth and out-of-pocket expenses are robust to an alternative demand specification where consumers are uncertain about their diagnoses and maximize utility over expected diagnoses and service utilization (see appendix table 6). In appendix table 7, I also provide results of a robustness check dropping markets 05, 08, 11, and 76, with the main

Table 10: Insurer demand

Insurer choice		Coefficient	Std. Error
Network		2.37***	0.01
OOP spending		-6.64***	0.21
Interactions			
Network	Demographics		
	Male	0.30***	0.01
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	-0.34***	0.02
	Cardiovascular	-0.33***	0.01
	Diabetes	-0.44***	0.04
	Renal	-0.61***	0.08
	Other	-0.53***	0.02
	>=2 diseases	-0.64***	0.02
	Healthy	(ref)	(ref)
	Location		
	Normal	0.05***	0.01
	Special	0.73***	0.04
	Urban	(ref)	(ref)
OOP spending	Demographics		
	Male	0.05	0.09
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	5.36***	0.22
	Cardiovascular	5.89***	0.18
	Diabetes	5.77***	0.31
	Renal	6.27***	0.22
	Other	5.56***	0.20
	>=2 diseases	5.89***	0.18
	Healthy	(ref)	(ref)
	Location		
	Normal	1.15***	0.11
	Special	0.67	0.43
	Urban	(ref)	(ref)
N		5,800,610	
N enrollees		500,000	
Pseudo-R^2		0.17	

Note: This table reports results of the insurer choice model estimated on a random sample of 500,000 new enrollees. Network coverage per service is weighted by the probability of claiming the service. Includes insurer fixed effects. Robust standard errors reported. ***p<0.01, **p<0.05, *p<0.1.

Table 11: Average willingness-to-pay per consumer type

Characteristic	Willingness-to-pay
Diagnosis	
Cancer	0.07
Cardiovascular	0.63
Diabetes	0.48
Renal	0.08
Other	0.37
>=2 diseases	0.57
Healthy	0.01
Sex	
Female	0.27
Male	0.35
Age group	
<1	0.04
1-4	0.13
5-14	0.30
15-18	0.31
19-44	0.23
45-49	0.54
50-54	0.19
55-59	0.77
60-64	0.73
65-69	0.17
70-74	0.08
>=75	0.06

Note: This table presents the average (across services, insurers, and markets) willingness-to-pay for an additional percentage point of network breadth conditional on consumer characteristics. Willingness-to-pay is calculated as $\frac{1}{\alpha_i}\frac{\partial s_{ijk}}{\partial H_{jmk}}\frac{H_{jmk}}{s_{ijk}} = \frac{1}{\alpha_i}\beta_i^D\gamma_{\theta lmk}(1-s_{ijk})H_{jmk}$ and is measured in millions of COP.

capital cities where hospital identity might be an issue. Results there are consistent with my main finding that consumers care about broader networks and lower out-of-pocket expenditures, and that willingness-to-pay for an additional percentage of coverage is greater for individuals with chronic conditions than for healthy ones.

Because my measure of network coverage might be correlated with other insurer-market level quality measures that may induce bias and are not captured by the insurer fixed effect, I estimate demand with additional quality controls as a robustness check. Using enrollee satisfaction survey data from the Ministry of Health, I obtain a measure of average waiting times for a doctor appointment through the insurer and average insurer quality as measured by a likert scale. Results of this exercise are provided in appendix table 8 and show that my estimators are robust to the inclusion of additional quality measures that vary within insurer.

In appendix table 9, I re-estimate demand calculating network breadth on the sample of 316 largest hospitals in the country and on the entire sample of providers, to see whether my hospital sample selection criteria matters for final results. The coefficients of interest associated to preference for network breadth and out-of-pocket spending are robust to these alternative network measures as well. Finally, in appendix table 10, I conduct a robustness check estimating demand on the sample of adults aged 19 or older, to get at the issue that children themselves may not be making informed insurance carrier choices but parents in their behalf. Including children could potentially lead to overestimation of the preference for network breadth given that I do not observe households nor head of household. My results are, nonetheless, robust to limiting the sample to only adults.

Using the estimates for network breadth preference and disutility of average out-of-pocket payments, I recover the patients' willingness-to-pay for an additional percentage of network coverage per service. Willingness-to-pay is calculated as $\frac{1}{\alpha_i}\frac{\partial s_{ijk}}{\partial H_{jmk}}\frac{H_{jmk}}{s_{ijk}}$ and is measured in millions of pesos. Differences in willingness-to-pay across consumer types will be suggestive of adverse selection from the demand side. Consumers with higher willingness-to-pay for a particular service category will tend to sort into carriers with high coverage for that service.

Table 11 presents the average willingness-to-pay across services, insurers, and markets, conditional on consumer observable characteristics. I find that patients with chronic conditions have a higher willingness-to-pay than individuals without diseases, overall. Within the group of patients with diseases, there is heterogeneity in willingness-to-pay across conditions. For instance, patients with cardiovascular disease are willing to pay \$0.56 million pesos more for an additional percentage point of coverage than patients with cancer. This is because individuals with cardiovascular conditions require a broader array of services than patients with cancer, for whom care is mostly concentrated in services like chemotherapy and radiotherapy. Findings also show that willingness-to-pay is \$80 thousand pesos higher for males than for females, the first of which also have a relatively higher prevalence of long-term diseases. Moreover, average willingness-to-pay is non-monotonic with respect to age. Consumers aged 55-59 have the highest willingness-to-pay equal to \$0.77 million pesos, an increase of \$0.23 million pesos from individuals aged 45-49.

5.2 Insurer expected costs

I estimate equation (4) for the logarithm of insurers' average cost per consumer type-(θ, l) and market in the sample of new *and* current enrollees, conditional on observed choices in 2010 and 2011. Table 12 shows the results and appendix figure 8 presents the estimated consumer type fixed effects and their corresponding 95% confidence intervals. Average costs are increasing in network breadth and decreasing in the correlation between network coverage for different pairs of services. This suggests that insurer coverage decisions are characterized by economies of scope. A 1% increase in network breadth for service m decreases the average cost of providing service $n \neq m$ by 0.51% per enrollee.[21] The effect of scope economies is smaller in magnitude than the direct effect of network breadth. My estimates show that a 1% increase in network breadth, raises average costs by 1.93%.[22] Findings also show that average costs decrease with service m's reference price, which is due to cheaper services, like consultations and laboratory, having a higher likelihood of being claimed.

Table 12: Insurer expected costs

Variable	Coefficient	Std. Error
Network	0.30***	0.04
Scope economies	-5.27***	0.84
Avg. ref. price	-0.71***	0.19
Insurer		
EPS001	-0.05	0.04
EPS002	-0.48***	0.03
EPS003	-0.19***	0.02
EPS005	0.03	0.02
EPS008	0.15**	0.07
EPS009	0.11	0.07
EPS010	-0.19***	0.03
EPS012	-0.24***	0.04
EPS013	-0.07***	0.02
EPS016	-0.21***	0.02
EPS017	-0.27***	0.04
EPS018	-0.16***	0.04
EPS023	-0.26***	0.03
EPS037	(ref)	(ref)
N	27,747	
R^2	0.42	

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

Figure 3 better depicts the magnitude of scope economies by service. The figure plots the predicted average cost against average network breadth in different service categories including: procedures in cardiac vessels, stomach, intestines; imaging; consultations; laboratory; nuclear medicine; and hospital admissions. In general, average costs are hump-shaped with respect to network breadth.

[21] This effect is calculated as the average of $100 \times \frac{1}{2M_k}\hat{\beta}_2^S \sum_{n\neq m} \gamma_{\theta lmk}\gamma_{\theta lnk} H_{jnk}$

[22] This effect is calculated as the average of $100 \times \hat{\beta}_1^S \gamma_{\theta ljmk}$

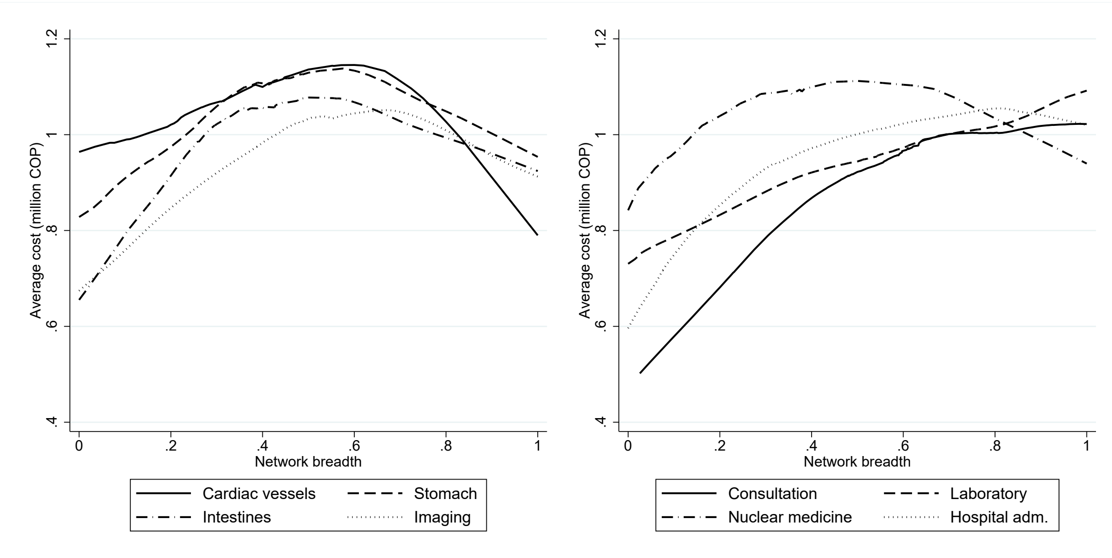
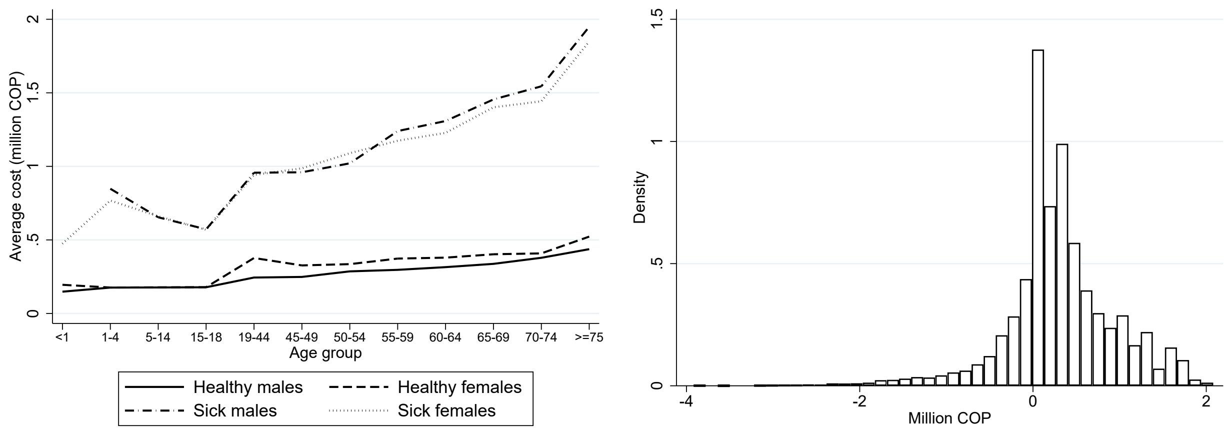


Figure 3: Expected cost function

In the left side panel of the figure, going from 20% to 30% coverage of procedures in stomach, raises the average cost per enrollee by \$0.16 million pesos; while going from 70% to 80% network breadth, reduces the average cost per enrollee by \$0.27 million pesos. In the right side panel, average cost reductions associated to coverage of nuclear medicine occur only at levels of network breadth exceeding 70%. While for consultations, laboratory, hospital admissions, and procedures in cardiac vessels the average cost per enrollee is increasing at all values of the average network breadth.

Scope economies also vary across consumer types. Appendix table 11 shows the change in the average cost of healthy females aged 19-44 and the average cost of females aged 19-44 with cancer, following a 10% increase in network breadth for the service in each row, separately for narrow and broad network carriers. I define a broad network carrier for service m as an insurer with average network breadth across all services $n \neq m$ above 80%, and a narrow network carrier as the opposite.

Following a 10% increase in network breadth for consultations, results in the appendix show that the average cost of healthy adult females enrolled to a narrow network carrier increases \$1,343 pesos more than for those enrolled to a broad network carrier. There is no evidence of scope economies among females with chronic diseases. For example, a 10% increase in network breadth for laboratory services, raises the average cost of adult females with cancer at a narrow network carrier by \$14,711 pesos and at a broad network carrier by \$17,086 pesos. Appendix table 12 presents similar comparisons across broad and narrow network carriers in the case of males aged 19-44 with diabetes and males aged 19-44 without diseases.



(a) Predicted average cost

(b) Predicted profit per consumer type

Figure 4: Predicted average cost and surplus per consumer type

Panel (a) of figure 4 shows that, in the sample of new enrollees, the predicted average cost per consumer type is U-shaped with respect to the enrollee's age, higher for healthy females than for healthy males after the age of 19, and higher for patients with chronic illnesses than for healthy enrollees, overall. These predictions are consistent with the descriptive evidence of section 3. Between the ages of 19-44, healthy females are 55% more expensive than healthy males, because of the costs associated to childbirth. Among enrollees with chronic conditions aged 19-44, males are 2% more expensive than females, and this difference in costs increases with age.

Differences in average costs generate significant variation in profits per consumer type as seen in panel (b) of the figure 4. Profits per new enrollee are calculated as the average risk-adjusted transfers (ex-ante and ex-post) minus the average cost. Since the ex-ante risk adjustment formula controls for only sex, age category, and municipality of residence, most of the variation in profits per consumer type comes from differences in cost across diagnoses and network breadth.[23] The fact that there are types of consumers for which the average profit is negative, such as females aged 19 or older with cancer, males aged 60 or older with renal disease, or males aged 19-44 with diabetes, means that insurers have incentives to risk select and choose their service specific hospital networks to avoid these enrollees.

[23]Profit variation per enrollee can also be partly explained by differences in sex, age, and location if the risk adjustment formula imperfectly compensates insurers in these dimensions.

5.3 Competition in network breadth

With the insurer demand and average cost estimations, I compute the left-hand side of equation (5) denoting total marginal variable profits. Appendix 10 presents summary statistics of this variable as well as of dropout probabilities and transition probabilities needed to recover marginal revenues and marginal costs. The average marginal variable profit per service and market ranges between $248 million pesos for EPS005 and $1,302 million pesos for EPS016. There is significant dispersion in this variable within insurers as its standard deviation exceeds $900 millions for the majority of carriers. The fact that marginal variable profits are positive for all insurer-service-market combinations, suggests a role for network formation costs in explaining profit maximizing choices of network breadth.

Table 13 presents the results of a 2-step GMM estimation for equation (6) on the subsample of four largest markets and 10 largest insurers in these markets, where the observed equilibrium is an interior solution.[24] The specification includes insurer, market, and service fixed effects, but only the first two are reported for exposition. I find that network formation costs are strictly convex in the level of network breadth. The coefficient associated to this variable is positive and significant. The first-stage F-statistic for the endogenous variable, network breadth, is greater than the 1% critical value. Appendix table 16 presents the first-stage regression, where the sign of the coefficients associated with each instrument are as expected.

Convexity of the average cost and the network formation cost with concavity of demand, guarantee that the insurer's profit function is concave in the level of network breadth. While concavity would be violated in presence of scope economies as those included in the insurer average cost function, the magnitude of these scope economies is small enough relative to the direct effect of network breadth on average costs, so as to not pose concerns about existence of an equilibrium in counterfactual. Existence of this counterfactual equilibrium also relies on the continuity and differentiability of the profit function at an interior solution.

Using the estimators for the FOC, I calculate the average total network formation cost across markets per insurer in column (1) of appendix table 17. Column (2) reports the network formation cost as a percentage of total variable profits. For 7 out of the 10 insurers, network formation costs represent at least 80% of an insurer's total variable profits. These percentages almost match the ratio of administrative expenses to accounting profits calculated from insurers' balance sheets and

[24] Appendix figure 9 reports the distribution of network breadth in this subsample.

income statements.[25]

Table 13: Model of insurer network formation costs

MVP_{jmk}	Coefficient	Std. Error
Network	3,139***	335
Insurer FEs		
EPS001	-37	189
EPS002	-9	171
EPS003	-25	184
EPS005	-490***	173
EPS010	410**	181
EPS013	-300*	168
EPS016	-235	173
EPS017	180	182
EPS018	18	197
EPS037	(ref)	(ref)
Market FEs		
Market 05	(ref)	(ref)
Market 08	-354***	110
Market 11	381***	108
Market 76	-77	106
First stage F-stat	929.7	
N	2,262	
R^2	0.63	

Note: This table presents a 2-step GMM estimation of equation (6) on the subsample of markets 05, 08, 11, 76, and the subsample of the 10 largest insurers in these markets. Excluded instruments are described in section 4.3. Robust standard errors and first-stage F-statistic for the endogenous variable, network breadth, are reported. ***p<0.01, **p<0.05, *p<0.1.

Changes in network breadth generate variation in profits that can be decomposed into its portions explained by variations in demand, average costs, and network formation costs. How much of the variation in profits comes from changes in demand and changes in the composition of enrollee types per insurer evidence the extent of adverse selection from the demand-side. The change in networks induced by changes in the regulatory environment that affect insurers but not consumers, instead, evidence risk selection from supply.

To decompose short-run profit changes following an increase in network coverage, I conduct a counterfactual exercise where I allow each insurer to unilaterally increase its network breadth for service m by 10%, while holding its competitors' choices fixed. In this exercise, I do not impose the FOC, so predictions of network formation costs are in absence of the cost shock, or $(\hat{\omega}H_{jmk} + \hat{\xi}_j + \hat{\xi}_k + \hat{\xi}_m)H_{jmk}$. This decomposition can be understood as a partial equilibrium exercise, where

[25] See https://docs.supersalud.gov.co/PortalWeb/SupervisionRiesgos/EstadisticasEPSRegimenContributivo/RC%20Estados%20financieros%20Dic%202011-CT2011.pdf

each insurer engages in a one-shot deviation. If my model is able to rationalize the choices of network breadth observed in the data, then the decomposition exercise should show that there are no profitable deviations, or that changes in total costs overcompensate changes in revenues when these changes are positive.

Table 14 presents the average percentage change in short-run demand ($\%\Delta s_{ijk}$), total revenues ($\%\Delta R_{\theta t} s_{ijk}$), total average costs ($\%\Delta AC_{\theta ljk} s_{ijk}$), average cost per enrollee ($\%\Delta AC_{\theta ljk}$), and network formation costs ($\%\Delta F_{jk}$) across insurers and markets, following a 10% increase in network breadth for the service in the row. First of all, I find that for every service the increase in total average costs plus the increase in network formation costs is greater than the increase in total revenues, so there are no profitable one-shot deviations.

Second, because changes in network breadth are weighted by $\gamma_{\theta lmk}$, I find that a 10% increase in coverage of consultations generates the largest variations in demand and costs relative to other services. For this service, demand increases 12.70% relative to observed levels, while total average costs and network formation costs increase 14.20% and 7.18%, respectively. In the case of hospital admissions, I find that insurers that increase coverage by 10%, experience a 0.69% increase in demand, a 0.74% increase in total average costs and a 0.47% increase in network formation costs.

For the services in this table, the change in demand explains on average 46% of the change in insurer total costs (total average costs plus network formation costs), while the change in networks, which directly affects average costs per enrollee and network formation costs, explains the remaining 54%. This means that the effect of demand-side selection generated from insurers' network breadth choices is substantial. Moreover, because transfers from the government are fixed, the change in total revenues equals the change demand.

Table 14: Decomposition of short-run profit changes after network breadth increase

Service	$\%\Delta s_{ijk}$	$\%\Delta R_{\theta t} s_{ijk}$	$\%\Delta AC_{\theta ljk} s_{ijk}$	$\%\Delta AC_{\theta ljk}$	$\%\Delta F_{jk}$
Cardiac vessels	0.01	0.01	0.01	0.00	0.08
Stomach	0.01	0.01	0.02	0.00	0.10
Intestines	0.27	0.27	0.29	0.02	0.20
Imaging	3.91	3.91	4.25	0.32	2.21
Consultations	12.70	12.70	14.20	1.31	7.18
Laboratory	5.71	5.71	6.24	0.49	3.12
Nuclear medicine	0.09	0.09	0.10	0.01	0.10
Hospital admissions	0.69	0.69	0.74	0.05	0.47

Note: This table shows the average percentage change in demand ($\%\Delta s_{ijk}$), total revenues ($\%\Delta R_{\theta k} s_{ijk}$), average costs per enrollee ($\%\Delta AC_{\theta ljk}$), total average costs ($\%\Delta AC_{\theta ljk} s_{ijk}$), and network formation costs ($\%\Delta F_{jk}$), following an insurer unilaterally increasing network breadth for the service in the row by 10%, while holding its competitors' choices fixed and assuming $\Delta\xi_{jmk} = 0$.

In appendix table 18, I show that most of the variation in demand and costs following a 10% increase in coverage of consultations comes from individuals with chronic conditions other than renal disease, particularly from those with cardiovascular diseases. Broader networks for consultations increases the demand from individuals with diseases and changes the composition of enrollee types within each insurer. The latter effect explains why insurer's average costs per enrollee increase 1.01% relative to observed levels for people with conditions other than renal disease. In the case of hospital admissions, I find that a 10% increase in network breadth attracts more patients with renal disease, followed by patients with other diseases, and lastly by healthy consumers. Raising network breadth for imaging and laboratory services generates similar effects in demand from sick and healthy individuals.

6 The effect of risk adjustment on network breadth

I use my model estimates to conduct two types of counterfactual exercises that help understand the effect of risk adjustment on service-level hospital network breadth, risk, and costs. In view of the growing network adequacy rules in countries like the US, analyzing how hospital networks respond to changes in the regulatory environment is important. While incentivizing insurers to broaden their networks might seem desirable to improve access to care, broader networks are also associated to higher health care costs. Quantifying the extent to which hospital networks respond to risk adjustment and how much of the network changes translate into higher health care costs can help policymakers in the design of public health insurance systems. This matters in particular for public health systems that are privately provided where there is no competition in prices, such as Medicaid Managed Care in the US.

In the first counterfactual exercise, I eliminate the observed risk adjustment systems and impose a uniform rate across all consumer types that is equal to the national average cost per enrollee. In the second exercise, I improve the government's risk adjustment formula by reimbursing diagnoses ex-ante. In both scenarios, I hold government spending, service prices, transition probabilities, and dropout probabilities fixed. While this would be problematic if negotiated prices change with counterfactual risk adjustment, these prices do not enter my model directly. Moreover, dropout probabilities are mostly governed by the event of becoming unemployed rather than by the individual choosing not to enroll due to changes in the network. The counterfactual scenarios do not impact the cost of enrollment in the contributory health care system, which is a fixed percentage of the consumer's monthly income, so there is no reason to believe that the individual will be more likely to dropout of the system in counterfactual. Finally, I assume that the probability of transitioning to new diagnoses depends mostly on the natural disease and age progression rather than on the type of care received in the network.

In counterfactual, I also assume that choice sets and utility and cost parameters remain fixed, so other than through changes in networks and demand, insurers will not incur additional costs on top of their average cost per enrollee and total network formation costs. In other words, I assume insurers do not exit markets nor enter new ones, but they can drop coverage of specific services altogether as long as coverage is non-zero for at least one service in each market where the insurer is present. For computational simplicity, I conduct my counterfactual analyses in the largest market in the country (Bogotá). This market represents 28.7% of all continuously enrolled individuals and has presence of all 14 insurers.

6.1 No risk adjustment

I start by describing the effect of eliminating the risk adjustment systems. In this counterfactual scenario, the per capita revenue to the insurer equals the national base transfer from table 1:

$$R^{cf}_{\theta lk} = UPC_{National}, \quad \forall(\theta, l, k)$$

The counterfactual transfer eliminates variation across sex and age groups, but is designed so that government spending and, thus, insurer revenues are the same as in the observed scenario. Failure to compensate for individuals' health risk should exacerbate risk selection, which would incentivize insurers to narrow their networks in services that baseline costlier patients require. Figure 5 shows the distribution of the difference between the counterfactual transfer without risk adjustment and the observed transfer per consumer type. Insurers receive lower payments for old individuals with chronic conditions and higher payments for young, healthy consumers, relative to the observed scenario. For example, for males aged 19-44 with cancer in Bogotá, insurers receive $25 thousand pesos less than in the observed risk adjustment system where the transfer equals $625 thousand pesos. For healthy males in the same age bracket and market, insurers receive $236 thousand pesos more than the observed transfer of $347 thousand pesos.

Using the equation describing the FOC condition of the insurer's profit maximization problem, I conduct an iterative procedure until convergence up to a tolerance level of 10^{-5} in the vector

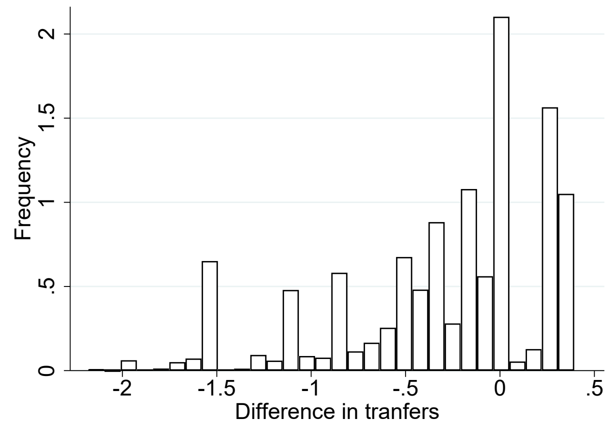


Figure 5: Distribution of counterfactual minus observed transfer under no risk adjustment

of service-specific network breadth. Because estimation of the FOC uses data from the 10 largest insurance companies in the 4 largest markets, a fair comparison of counterfactual results to the observed equilibrium requires using my model to predict network breadth under the observed risk adjustment when these 10 insurers instead of 14 compete in the market.

Panel A of table 15 presents the percentage change in counterfactual relative to the observed scenario, of average network breadth, insurer total average costs, short-run average costs per enrollee, and short-run consumer welfare separately for healthy and sick individuals.[26] My findings show that under no risk adjustment, average network breadth falls 2.7%. The reduction in coverage explains the 1.4% decrease in insurer total average costs relative to the observed risk adjustment. This corresponds to a 0.8% reduction in average costs per enrollee. The fall in network breadth means insurers do not enjoy the cost savings from scope economies, but the direct effect of lower network breadth on average costs overcompensates the effect on scope economies.

Eliminating the risk adjustment systems results in a 7.4% decrease in short-run consumer welfare for individuals without diseases. Consumer welfare falls by a larger magnitude, equal to 7.8%, for individuals with chronic conditions who use health care services more often. This is because with lower network breadth, access to and quality of care worsen in counterfactual, which differentially impacts those with a higher likelihood of making claims. Column (1) of appendix table 19 presents a more granular decomposition of welfare changes across subgroups of consumers.

[26] Insurer total average costs are calculated as:
$\sum_{ij}\left(AC_{\theta ljk}s_{ij}+\sum_{s=t+1}^{T}\beta^{s}\sum_{\theta',l'}(1-\rho_{\theta l})\mathcal{P}(l'|\theta,l)AC_{\theta' l' jk}s'_{ij}\right)$,
Short-run average cost per enrollee is $(1/N)\sum_{ij}AC_{\theta lj}$,
and short-run consumer welfare is $\sum_i|\alpha_i|^{-1}\log(\sum_j exp(\beta_i^D\sum_m\gamma_{\theta lm}H_{jm}-\alpha_i c_{\theta jy})+\delta_j)$

Panel B of table 15 shows that the reduction in average network breadth happens across the board of services. I collapse the 58 service categories into 15 broader groups that describe medical procedures in certain parts of the body. When they are not compensated for their enrollees' health risk, insurers drop coverage of relatively expensive services like hospital admissions by 19.6% and procedures in skull and spine by 3.5%. For less expensive services, the reduction in average network breadth is smaller but still sizable. For instance, average coverage of procedures in joints and bones falls by only 2.1%. Figure 6 shows the correlation between changes in service-level network breadth and the difference between the probability that a sick consumer and a healthy consumer make a claim for that service. When this difference is small, insurers tend to broaden their networks for that particular service to attract healthy individuals. When this difference is large, the correlation is negative, suggesting that insurers drop coverage of services that sick individuals are differentially more likely to claim.

Table 15: Counterfactual changes in networks, costs, and welfare under no risk adjustment

Variable	%Δ in CF
Panel A. Overall	
Avg. Network	-2.7
Total avg. cost	-1.4
Avg. cost per enrollee	-0.8
Consumer welfare (healthy)	-7.4
Consumer welfare (sick)	-7.8
Panel B. Avg. network per service	
Skull, spine, nerves, glands	-3.5
Eyes, ears, nose, mouth	-4.2
Pharynx, lungs	-1.3
Heart and cardiac vessels	-0.4
Lymph nodes, bone marrow	4.5
Esophagus, stomach and intestines	1.3
Liver, biliary tract	-4.8
Abdominal wall	-1.3
Urinary system	-1.7
Reproductive system	-2.3
Bones and facial joints	-2.3
Joints, bones, muscles, tendons	-2.1
Skin	-7.6
Imaging, lab, consultation	-4.1
Hospital admission	-19.6

Note: Panel A of this table presents the percentage change in counterfactual under no risk adjustment relative to predictions at observed risk adjustment, of average network breadth across insurers, insurer total average costs, short-run average cost per enrollee, and short-run consumer welfare for the healthy and sick. Panel B presents the percentage change in counterfactual of average network breadth per service category, where I collapse the 58 original categories into 15 broader groups. The counterfactual is calculated with data from Bogotá only and the 10 largest insurers.

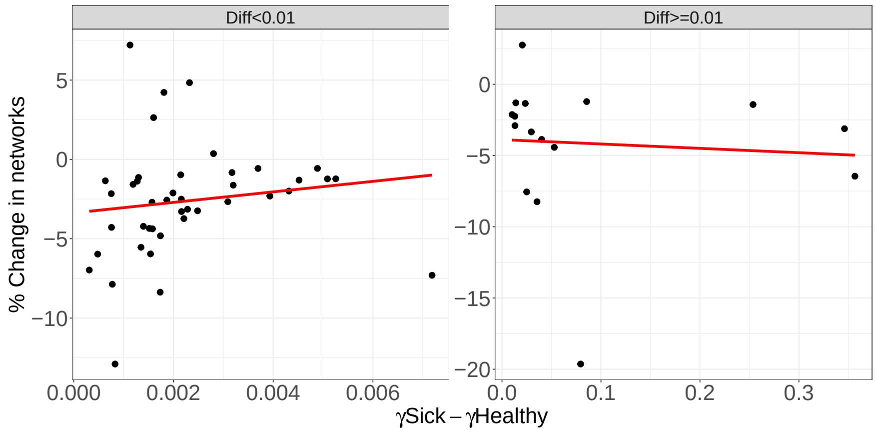


Figure 6: Network changes per service and likelihood of making claims under no risk adjustment

Consistent with adverse selection, insurers for which average network breadth falls by a greater percentage in counterfactual see a decrease in total short-run demand from healthy and sick individuals as seen in appendix figure 10. The figure shows the correlation between the change in average network breadth and the change in total demand across insurers. Under no risk adjustment, healthy individuals substitute away mostly from EPS005 into EPS001, EPS010, and EPS037, given that EPS005 has the largest reduction in average network breadth equal to 5.3%. The pattern is similar for demand from consumers with chronic diseases. This suggests that health risk becomes more concentrated relative to the observed scenario, since the baseline largest insurers end up receiving most of the switchers with diseases.

Overall, panel A of appendix figure 13 shows that in a situation where all insurers choose narrow networks, those that reduce their network breadth by a larger percentage have increases in total long-run profits, as opposed to those that drop network coverage by a small amount. The figure presents the correlation between the change in average total profits and the change in average network breadth, across insurers. The red line corresponding to a linear fit, shows that this correlation is negative, which implies that the reduction in costs from lower network breadth overcompensates the reduction in demand.

6.2 Improved risk adjustment

I now move to the opposite exercise where I improve the current risk adjustment system by compensating for the list of 7 diagnoses ex-ante. If allowing for variation in per-capita transfers across diagnoses better captures the individuals' health risk, then risk selection incentives should decrease and network breadth should increase. I assume the counterfactual risk-adjusted transfer is given by the annualized average cost per consumer type (θ, l). More formally, this is:

$$R^{cf}_{\theta lk} = a_k \times 360 \times \frac{\sum_{i \in \theta l} X_i}{\sum_{i \in \theta l} d_i}$$

where X_i is the total health care cost of individual i of type (θ, l) living in market k, d_i is number of days enrolled to the contributory system during a year, and a_k is the market multiplier from table 1. When the prediction of the annualized health care cost equals zero for a consumer type, I replace it for the value in the observed risk adjustment system which conditions on sex, age group, and location. Unlike the observed risk adjustment system, the counterfactual payment allows for variation within θ and k, while keeping government spending fixed. Figure 7 presents the distribution of the difference between the counterfactual payment and the observed risk adjusted payment per consumer type. In Bogotá, insurers receive \$1.8 million pesos more for males aged 19-44 with cancer compared to the observed transfer, but receive only \$0.4 million pesos more for females with cancer aged 19-44.

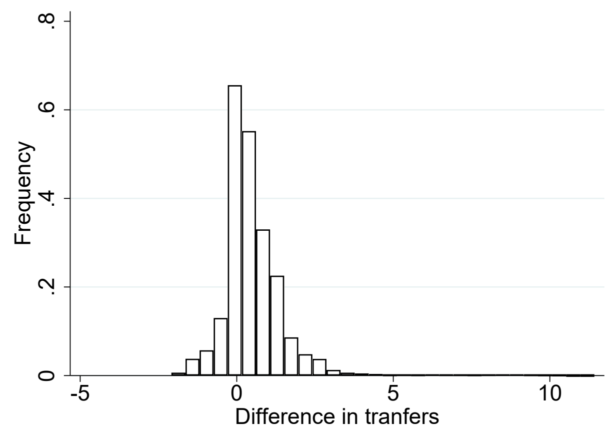


Figure 7: Distribution of counterfactual minus observed transfer under improved risk adjustment

Panel A of table 16 shows my model's predictions regarding average network breadth, insurers' total cost, average cost per enrollee, and consumer welfare under the improved risk adjustment formula. I find that average network breadth increases 1.1% relative to the observed scenario. This increase in coverage happens across the board of services as seen in panel B of the table, but disproportionately so for services that mostly sick patients tend to claim, which is consistent with weakened selection incentives. For instance, average network breadth for hospital admissions increases 3.2%, while for primary care or consultations it increases only 1.5%.

In fact, figure 8 shows that changes in network breadth are positively correlated with the differ-

Table 16: Counterfactual changes in networks, costs, and welfare under improved risk adjustment

Variable	%Δ in CF
Panel A. Overall	
Avg. Network	1.1
Total avg. cost	0.6
Avg. cost per enrollee	0.3
Consumer welfare (healthy)	3.5
Consumer welfare (sick)	3.2
Panel B. Avg. network per service	
Skull, spine, nerves, glands	0.7
Eyes, ears, nose, mouth	1.3
Pharynx, lungs	0.3
Heart and cardiac vessels	0.8
Lymph nodes, bone marrow	0.9
Esophagus, stomach and intestines	1.1
Liver, biliary tract	1.3
Abdominal wall	0.5
Urinary system	0.7
Reproductive system	0.9
Bones and facial joints	0.4
Joints, bones, muscles, tendons	0.8
Skin	4.4
Imaging, lab, consultation	1.5
Hospital admission	3.2

Note: Panel A of this table presents the percentage change in counterfactual under improved risk adjustment relative to predictions at observed risk adjustment, of average network breadth across insurers, insurer total average costs, short-run average cost per enrollee, and short-run consumer welfare for the healthy and sick. Panel B presents the percentage change in counterfactual of average network breadth per service category, where I collapse the 58 original categories into 15 broader groups. The counterfactual is calculated with data from Bogotá only and the 10 largest insurers.

ence between the probability that a sick consumer and a healthy one make a claim for that service, when this difference is large. But there is a negative correlation when the difference in claim probabilities is small, which indicates that compensating insurers ex-ante for chronic diseases improves access to care in services that patients with those diseases are significantly more likely to claim relative to healthy consumers.

Because enrollees prefer broader networks on average, insurers that expand their network breadth by a larger amount in counterfactual have a higher demand than those that narrow their networks. Appendix figure 11 shows that the correlation between changes in networks and changes in demand from healthy and sick individuals is positive.

The demand increase translates into weakly higher profits as seen in appendix figure 13. Every black dot corresponds to the average profit change and average network breadth change for each

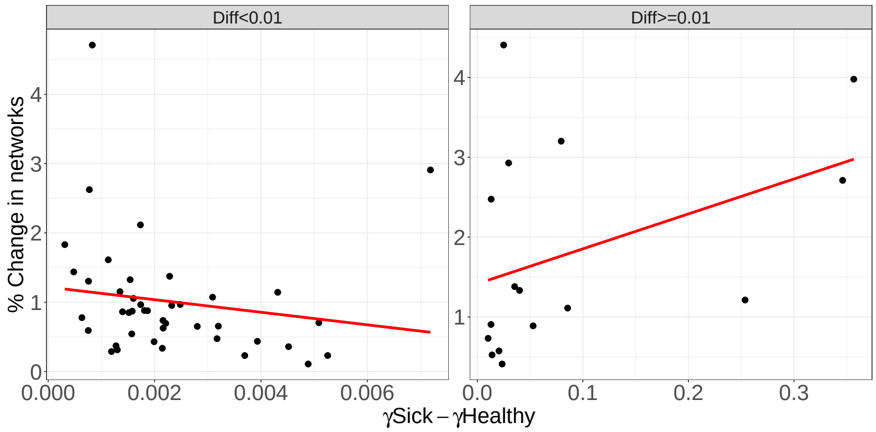


Figure 8: Network changes per service and likelihood of making claims under improved risk adjustment

insurer, and the red line corresponds to a linear fit. The positive correlation is consistent with the direct effect of expanding the networks on total variable costs and network formation costs being smaller than the effect of scope economies and the effect on demand. Panel A of table 16 also shows that short-run average costs per enrollee increase only 0.3% in counterfactual. But total demand for insurers EPS005 and EPS017, for example, increase 2.8% and 2.2%, respectively.

With the improved ex-ante risk adjustment formula and the resulting broader networks in services that risky health types need the most, I find that consumer welfare for patients with any chronic condition increases 3.2% relative to the observed scenario as seen in panel A of table 16. This welfare increase is similar to the one for consumers without diseases, which equals 3.5%. Column (2) of appendix table 19 presents a more detailed decomposition of welfare changes for subgroups of consumers characterized by sex, age groups, and diagnoses. Findings there show that the greatest effect of improved risk adjustment on welfare happens for patients with renal disease, who are traditionally the most expensive consumer type in Colombia. For renal patients welfare increases 3.7%, followed by patients with cancer and other chronic diseases.

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

7 The effect of premiums on network breadth

In this section, I study the effect of deregulating premiums on market outcomes. I assume insurers compete on prices and are allowed to price discriminate based on the enrollee's income level, sex, and age group, similar to their pricing rules in supplementary health insurance plans. Insurers engage in Bertrand competition in premiums separately by market. In the observed scenario, the monthly contribution to the health care system equals 12% of the enrollee's monthly income. 1/3 of this contribution is paid by the enrollee and the remaining 2/3 by her employer.[27] Under counterfactual premiums, I assume the same split of the cost of enrollment. Premiums replace the government's risk-adjusted transfers and insurers receive premium revenues directly. So, government spending is zero in counterfactual.

More formally, in the observed scenario, total average out-of-pocket expenses equal average coinsurance payments, plus average copayments and average contributions to the health care system as seen below.

$$c_{\theta lyjk} = \text{Coins}_{\theta lyjk} + \text{Copay}_{\theta lyjk} + \text{Contribution}_y$$

Allowing insurers to compete in premiums means that the counterfactual measure of total average out-of-pocket spending is:

$$\tilde{c}_{\theta lyjk} = \text{Coins}_{\theta lyjk} + \text{Copay}_{\theta lyjk} + (1/3) * \tilde{P}_{\theta yjk}$$

where $\tilde{P}_{\theta yjk}$ is insurer j's total premium in market k for consumer type θ with income level y. Because premiums are paid instead of the required contributions to the health care system in the observed scenario, I assume that dropout probabilities remain fixed in counterfactual conditional on network breadth. To the extent that premiums are not significantly greater than observed contributions to the system, conditional on network coverage, an individual is not necessarily more likely to disenroll and remain uncovered in counterfactual. This provides a natural test of the assumption of fixed dropout probabilities, where the null hypothesis is equality of the conditional distribution of premiums and contributions, conditional on networks.

[27] Self-employed individuals pay the full amount of their contribution to the health care system. In counterfactual, I am implicitly assuming no one is self-employed.

Let $P_{\theta(i)y(i)jk} = (1/3) * \tilde{P}_{\theta(i)y(i)jk}$, individual i's choice probability for insurer j in market k is:

$$s_{ijk}(H_k) = \frac{\exp\left(\beta_i^D \sum_m \gamma_{\theta lmk} H_{jmk} - \alpha_i \tilde{c}_{\theta lyjk} + \phi_j\right)}{\sum_{g\in\mathcal{J}_k} \exp\left(\beta_i^D \sum_m \gamma_{\theta lmk} H_{gmk} - \alpha_i \tilde{c}_{\theta lyjk} + \phi_g\right)} \tag{7}$$

Equation (7) implicitly assumes that the marginal disutility of premiums is the same as the marginal disutility of contributions, coinsurance payments, and copays, which is equal to α_i. This assumption is based on the fact that consumers should not care whether increases in out-of-pocket costs come from premiums or cost-sharing if they are making rational choices (Abaluck and Gruber, 2011).

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

$$\pi_{ijk}(H_k, P_k, \theta, l) = (\tilde{P}_{\theta yjk} - (1 - r_y) AC_{\theta lj}(H_{jk})) s_{ijk}(H_k, P_k)$$

The insurer now has two choice variables per market: the network breadth per service and the premium per consumer type. Notice that the average cost function is the same as in the observed scenario, that is, I do not model the relation between premiums and average costs. While some papers use that relation to test for adverse selection (Tebaldi, 2017; Einav et al., 2010), in my case it is more difficult to argue what the functional form of average cost is with respect to premiums than to assume that the average cost function does not change in counterfactual. Moreover, because I allow insurers to discriminate premiums along income level, sex, and age group, selection in this market can now happen both through premiums and service-level network breadth.

Insurers simultaneously choose premiums and service-level network breadth in period $t = 0$ and commit to this choice thereafter. The solution concept in counterfactual is a full commitment Nash equilibrium. Insurers simultaneously choose premiums and networks to maximize:

$$\Pi_{jk}(H_k, P_k) = \sum_{\theta,l} \left(\pi_{ijk}(H_k, P_k, \theta, l) N_{\theta lk} + \sum_{s=t+1}^{T} \beta^s \sum_{\theta',l'} (1 - \rho_{\theta lk}) \mathcal{P}(l'|\theta, l) \pi_{ijk}(H_k, P_k, \theta', l') N_{\theta' l' k}\right)$$
$$- \sum_m \left(\omega_0 H_{jmk} + \xi_{jmk}\right) H_{jmk}$$

The FOC with respect to premiums is:

$$\frac{\partial \Pi_{jk}}{\partial P_{\theta yjk}} = \sum_i \left(\frac{\partial \pi_{ijk}}{\partial P_{\theta yjk}} N_{\theta lk} + \sum_{s=t+1}^{T} \beta^t \sum_{\theta', l'} (1-\rho_{\theta lk}) \mathcal{P}(l'|\theta, l) \frac{\partial \pi'_{ijk}}{\partial P_{\theta yjk}} N_{\theta' l' k} \right) = 0$$

where,

$$\frac{\partial \pi_{ijk}}{\partial P_{\theta yjk}} = \Omega \Big(\tilde{P}_{\theta yjk} - (1-r_y) AC_{\theta ljk} \Big) + s_{ijk}(H_k, P_k)$$

and,

$$\Omega(j,g) = \begin{cases} -s_{ijk}(1-s_{ijk})\alpha_i & \text{if} \quad j = g \\ s_{ijk} s_{igk} \alpha_i & \text{if} \quad j \neq g \end{cases}$$

Note that the short-run premium level can be solved for from the premium FOC. This determines a function of the form $\tilde{P}_{\theta yjk} = f(\tilde{P}_{\theta yjk})$, that can be used to obtain equilibrium premiums in a fixed point manner.

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

$$\begin{aligned} \frac{\partial \Pi_{jk}}{\partial H_{jmk}} = & \sum_i \left(\frac{\partial \pi_{ijk}}{\partial H_{jmk}} N_{\theta lk} + \sum_{s=t+1}^{T} \beta^t \sum_{\theta', l'} (1-\rho_{\theta lk}) \mathcal{P}(l'|\theta, l) \frac{\partial \pi'_{ijk}}{\partial H_{jmk}} N_{\theta' l' k} \right) \\ & - \Big(2\omega_0 H_{jmk} + \xi_{jmk} \Big) = 0 \end{aligned}$$

Figure 9 presents the distribution of observed contributions to the health care system, total premiums, and premium pass-through to consumers, which equals 1/3 of total premiums. First of all, with deregulation, I find that insurers will charge non-zero premiums. These premiums are significantly greater than observed contributions to the system, but have a higher dispersion across consumer types. On the one hand, the average contribution equals \$50.8 thousand pesos and its standard deviation across consumer types is \$17.7 thousand pesos. On the other hand, the average consumer premium equals \$351.7 thousand pesos and its standard deviation is \$158.5 thousand pesos.

Table 17 presents the average total premium across demographics and insurers in millions of pesos. I find that average premiums are \$100 thousand pesos higher for males than female, which is consistent with men being more likely to develop chronic diseases compared to women. The breakdown across age groups shows that average total premiums are generally U-shaped with respect

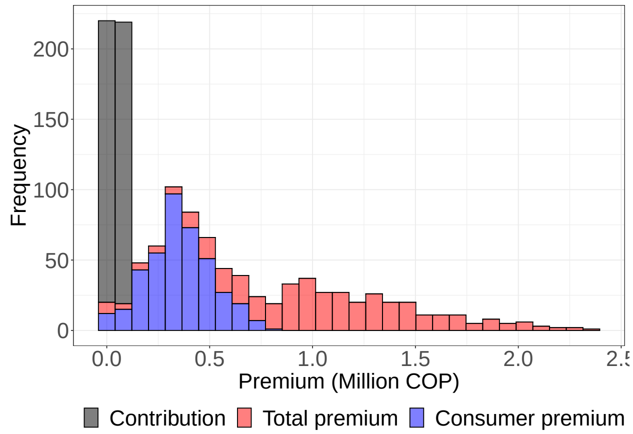


Figure 9: Distribution of counterfactual premiums

to age. For children aged 5-14, the average premium equals \$1.26 million pesos, which decreases by nearly half for individuals aged 45-49. Finally, the average premium for individuals aged 75 or older is \$1.14 million pesos. My findings show that counterfactual premiums weakly increase with the enrollee's income level. People earning between 2 and 5 times the monthly minimum wage pay premiums that are \$10 thousand pesos higher than individuals who make less that 2 times the minimum monthly wage. Across insurers, I find a positive correlation between premiums and insurer size. Relatively large insurers such as EPS037 have higher premiums than smaller insurers such as EPS002 and EPS003.

Counterfactual premiums are higher than observed contributions to the health care system, because insurers significantly expand their networks relative to the observed scenario. This means that there is a large pass-through of insurance coverage to premiums as noted in Cabral et al. (2018). Panel A of table 18 presents the percentage change in counterfactual of average network breadth, insurer total average costs, short-run average cost per enrollee, insurer total revenue, and short-run consumer welfare for healthy and sick individuals. Allowing insurers to compete in premiums, increases average network breadth by 46.0%. This increase in coverage happens across the board of services as seen in panel B of the table. While relatively expensive services such as hospital admissions or procedures in skull and spine see smaller increases in average network breadth compared to relatively cheap services such as procedures in skin or abdominal wall, the effect is substantial.

Given that average network breadth expands considerably, and that the direct effect of network breadth on average costs is greater than the effect of scope economies, I find that the short-run per-enrollee average cost increases 9.1% relative to observed predictions. However, weighting by

Table 17: Average premium per consumer characteristics and insurer

Variable	Avg. premium
Sex	
Female	1.01
Male	1.10
Age group	
<1	—
1-4	—
5-14	1.26
15-18	1.34
19-44	1.53
45-49	0.63
50-54	0.51
55-59	1.05
60-64	0.98
65-69	1.00
70-74	1.11
>=75	1.14
Income group	
$<$ 2 x MMW	1.05
$[2, 5]$ x MMW	1.06
$>$ 5 x MMW	—
Insurer	
EPS001	1.46
EPS002	0.66
EPS003	0.93
EPS005	1.82
EPS010	1.00
EPS013	0.70
EPS016	0.62
EPS017	1.17
EPS018	0.97
EPS037	1.22

Note: This table presents the counterfactual average total premium per sex, age category, income group, and insurer. Measured in million COP.

demand in the long-run, total insurer average costs fall 0.5%. The effect on insurers' total costs is explained entirely by risk selection. If average cost per enrollee increases, only large increases in demand from relatively cheap individuals, such as those without diseases, can explain the fall in insurer's total costs in the long-run.

Appendix figure 12 in fact shows that under premium deregulation, demand from healthy individuals sees large changes across insurers compared to demand from sick patients, conditional on network breadth changes. Since on average all insurers expand their networks, such that there is less network heterogeneity across carriers, risk selection in this counterfactual happens mostly on the premium level. This translates into a low correlation between changes in network breadth and

Table 18: Changes in networks, costs, and welfare under premium deregulation

Variable	%Δ in CF
Panel A. Overall	
Avg. Network	46.0
Total avg. cost	-0.5
Avg. cost per enrollee	9.1
Total revenue	108.9
Consumer welfare (healthy)	-87.5
Consumer welfare (sick)	58.8
Panel B. Avg. network per service	
Skull, spine, nerves, glands	34.3
Eyes, ears, nose, mouth	54.2
Pharynx, lungs	60.6
Heart and cardiac vessels	85.8
Lymph nodes, bone marrow	30.7
Esophagus, stomach and intestines	26.6
Liver, biliary tract	38.3
Abdominal wall	54.4
Urinary system	43.7
Reproductive system	41.3
Bones and facial joints	23.7
Joints, bones, muscles, tendons	60.4
Skin	272.6
Imaging, lab, consultation	32.7
Hospital admission	26.2

Note: Panel A of this table presents the percentage change in counterfactual under premium deregulation relative to predictions at observed risk adjustment, of average network breadth across insurers, insurer total average costs, short-run average cost per enrollee, insurer total revenue, and short-run consumer welfare for the healthy and sick. Panel B presents the percentage change in counterfactual of average network breadth per service category, where I collapse the 58 original categories into 15 broader groups. The counterfactual is calculated with data from Bogotá only and the 10 largest insurers.

changes in profits as seen in panel C of appendix figure 13.

Panel A of table 18 also shows that insurer total revenues increase 84.0% relative to the observed scenario. But, this effect is completely at the expense of consumers without diseases, for whom welfare falls 87.5% in counterfactual. For healthy individuals, who have a low willingness-to-pay for network breadth and a high sensitivity to out-of-pocket expenses, welfare losses due to out-of-pocket premium payments overcompensate the welfare gains from greater network breadth in every service. Instead, for consumers with chronic diseases who have a high willingness-to-pay for coverage and who are relatively inelastic to out-of-pocket expenditures, welfare increases 58.8% relative to the observed scenario. Column (3) of appendix table 19 shows a decomposition of welfare changes across sex, age groups, and the 7 disease categories. Findings there show that most welfare gains for patients with chronic conditions are accrued by those with renal disease, followed by patients with 2 or more diseases, and by patients with cardiovascular disease.

8 Conclusions

Risk selection is a main concern in public health insurance systems with regulated competition, where governments make risk adjusted transfers to private insurers. In this paper I show that in a tightly regulated market in which insurers have no discretion over the design of the health insurance plan, they can engage in risk selection at the service level using their hospital networks. The existing literature has focused on the study of risk selection arising from insurer competition in premiums, but less explored is the effect of risk selection on non-price competition. I model insurer competition in service-level hospital networks using as case study the Colombian health care system, where the government sets premiums to zero and compensates private insurers with a coarse risk adjustment formula.

In Colombia, insurers have discretion over which services to cover at which hospitals, so hospital networks are service specific. This is similar to allowing insurers to include specialized hospitals in their networks as in the United States. I find that insurers engage in risk selection in this setting by narrowing their networks in services that costly patients, usually with chronic conditions, tend to demand the most. This strategy either disincentivizes enrollment from these costly individuals or selects those with lower baseline health care costs conditional on risk adjustment.

The increasing popularity of network adequacy rules in countries like the United States and the policy debate surrounding access to care in Colombia, raises the question of how to incentivize insurers to expand their service-level networks, while containing health care costs. I use my model to answer this question by measuring the effect of risk adjustment and premium deregulation on service-level hospital network breadth while holding government spending fixed.

I find that eliminating risk adjustment makes insurer competition a race to the bottom in terms of network breadth. This implies lower quality of and access to care, which is why consumer welfare falls 7.4% relative to the observed scenario. Improving the current risk adjustment formula by compensating for sex, age, location, and a list of 7 exhaustive and mutually exclusive diagnoses all ex-ante, results in service-level networks that are broader on average, particularly for services that individuals with chronic diseases are differentially more likely to claim. With the improved formula, welfare of patients with chronic diseases increases 3.2%. Despite the positive welfare gains at no additional cost to the government, implementing this improved risk adjustment is difficult because of information frictions between hospitals, insurers, and the regulator. Finally, allowing insurers to compete in premiums results in nearly complete hospital networks for every insurer and service. This network expansion comes at the expense of higher out-of-pocket costs, which significantly reduces the welfare of consumer without diseases but increases the welfare of those with chronic conditions who are less elastic with respect to out-of-pocket expenses.

The findings of this paper provide evidence on the trade-off between better access to care through broader hospital networks and lower health care costs. In quantifying the extent to which networks respond to risk adjustment and premiums, the findings here help policymakers in the design of public health systems with private provision of health insurance. Policy implications, nonetheless, extend beyond these types markets, to ones where private insurers actively engage in risk selection using non-price elements of their health insurance plans.

References

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

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

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

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

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

Compiani, G. (2019). Market counterfactuals and the specification of multi-product demand: A nonparametric approach. *Forthcoming, Quantitative Economics*.

Crawford, G. (2012). Endogenous Product Choice: A Progress Report. *International Journal of Industrial Organization*, 30(3):315–320.

Dafny, L., Hendel, I., and Wilson, N. (2015). Narrow Networks on the Health Insurance Exchanges: What do they Look Like and How do they Affect Pricing? A Case Study of Texas. *American Economic Review*, 105(5):110–14.

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

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

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

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

Farrell, J. and Klemperer, P. (2007). *Coordination and Lock-In: Competition with Switching Costs and Network Effects*. The Handbook of Industrial Organization.

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.

Ghili, S. (2020). Network Formation and Bargaining in Vertical Markets: The Case of Narrow Networks in Health Insurance.

Gowrisankaran, G., Nevo, A., and Town, R. (2015). Mergers When Prices Are Negotiated: Evidence from the Hospital Industry. *American Economic Review*, 105(1):172–203.

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

Heckman, J. (1991). Identifying the Hand of the Past: Distinguishing State Dependence from Heterogeneity. *American Economic Review*, 81(2):75–79.

Ho, K., Hogan, J., and Scott-Morton, F. (2017). The impact of consumer inattention on insurer pricing in the Medicare Part D program. *The Review of Economic Studies*, 48(4):877–905.

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

Ho, K. and Lee, R. (2019). Equilibrium provider networks: Bargaining and exclusion in health care markets. *American Economic Review*, 109(2):473–522.

Klemperer, P. (1987). Markets with Consumer Switching Costs. *Quarterly Journal of Economics*, 102(2):375–394.

Klemperer, P. (1995). Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade. *The Review of Economic Studies*, 62(4):515–539.

Layton, T., Ndikumana, A., and Shepard, M. (2018). *Health Plan Payment in Medicaid Managed Care: A Hybrid Model of Regulated Competition*. Elsevier.

Liebman, E. (2018). Bargaining in Markets with Exclusion: An Analysis of Health Insurance Networks.

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

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

McNamara, C., Serna, N., and Sorensen, A. (2021). Targeted Advertising in Medicare Advantage.

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.

Ministerio de Salud y Protección Social (2015). Estudio de evaluación de los servicios de las EPS por parte de los usuarios, en el régimen contributivo y subsidiado.

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.

Polyakova, M. (2016). Regulation of Insurance with Adverse Selection and Switching Costs: Evidence from Medicare Part D. *American Economic Journal: Applied Economics*, 8(3):165–195.

Prager, E. and Tilipman, N. (2020). Regulating Out-of-Network Hospital Payments: Disagreement Payoffs, Negotiated Prices, and Access.

Riascos, A. (2013). Complementary compensating mechanisms of ex ante risk adjustment in colombian competitive health insurance market. *Revista Desarrollo Y Sociedad*, 71(1):165–191.

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. and Camelo, S. (2017). A note on risk-sharing mechanisms for the colombian health insurance system. *Documentos CEDE*, (30):1–14.

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

Riascos, A., Serna, N., and Guerrero, R. (2017b). Capital requirements of health insurers under different risk-adjusted capitation payments. *Documentos CEDE*, (6).

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

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

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