Online Appendix: Determinants of Healthcare Provider Networks: Risk Selection vs. Administrative Costs

Natalia Serna

Appendix 1 Current risk adjustment system

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

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

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

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

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

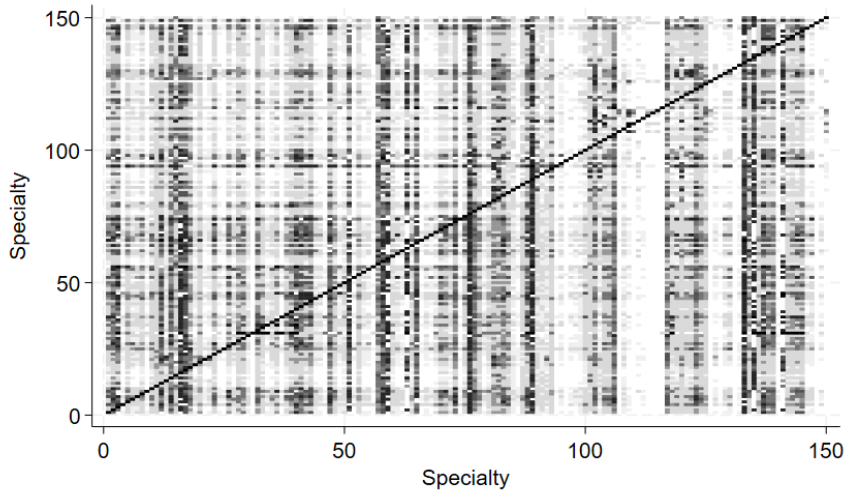
Age group	Sex	Multiplier
Less than 1	—	3.0000
1-4	—	0.9633
5-14	—	0.3365
15-18	M	0.3207
15-18	F	0.5068
19-44	M	0.5707
19-44	F	1.0588
45-49	—	1.0473
50-54	—	1.3358
55-59	—	1.6329
60-64	—	2.1015
65-69	—	2.6141
70-74	—	3.1369
More than 74	—	3.9419

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

Appendix 2 Service categories

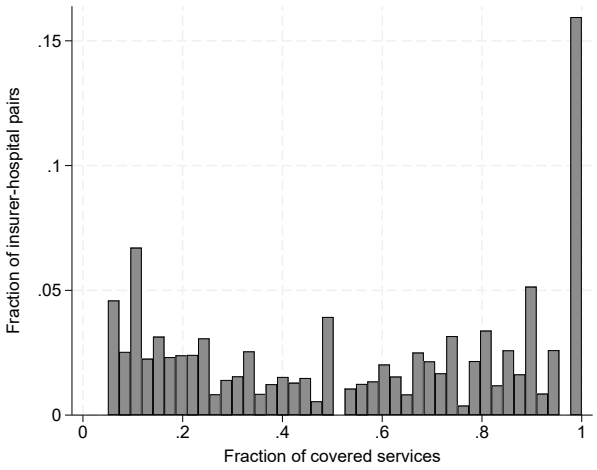
The service-level provider network data reports 150 unique specialities over which insurers and providers bargain. Some of these specialties are highly correlated in the sense that insurers tend to include them together at a particular provider. Appendix Figure 1 presents a heatmap of the fraction of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

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



Appendix Figure 1: Heatmap of specialty pairs network inclusions

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



Appendix Figure 2: Service inclusions within hospital

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

Appendix Table 3: List of services

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

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

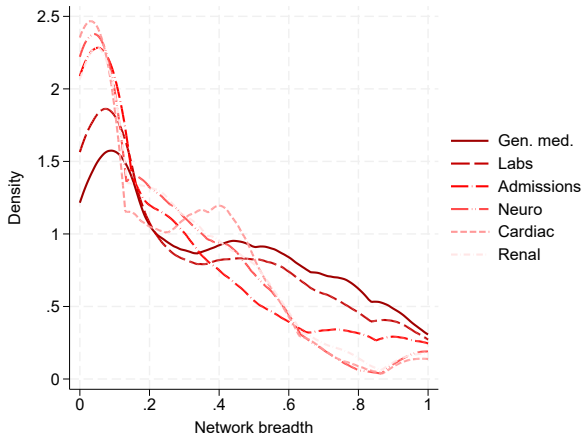
Appendix Table 4: Service coverage at hospitals

	Cardiac care			Renal care			Hospital admissions		
Insurer	Valle del Lili	Santa Fe	Pablo Tobón	Valle del Lili	Santa Fe	Pablo Tobón	Valle del Lili	Santa Fe	Pablo Tobón
Insurer A	1	0	0	1	0	0	1	1	1
Insurer B	1	0	1	1	0	1	1	1	1
Insurer C	1	0	0	0	0	0	1	1	1
Insurer D	1	1	1	1	1	1	1	1	1
Insurer E	1	1	0	1	1	0	1	1	0
Insurer F	0	0	1	0	0	1	0	0	1
Insurer G	1	1	1	1	1	1	1	1	1
Insurer H	1	1	0	1	1	0	1	1	0
Insurer I	1	0	1	1	0	1	1	1	1
Insurer J	1	1	1	1	1	1	1	1	1
Insurer K	0	1	0	0	1	0	1	1	1
Insurer L	1	1	1	1	1	1	1	1	1
Insurer M	0	0	0	0	0	0	1	1	1
Insurer N	1	1	1	1	1	1	1	1	1

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

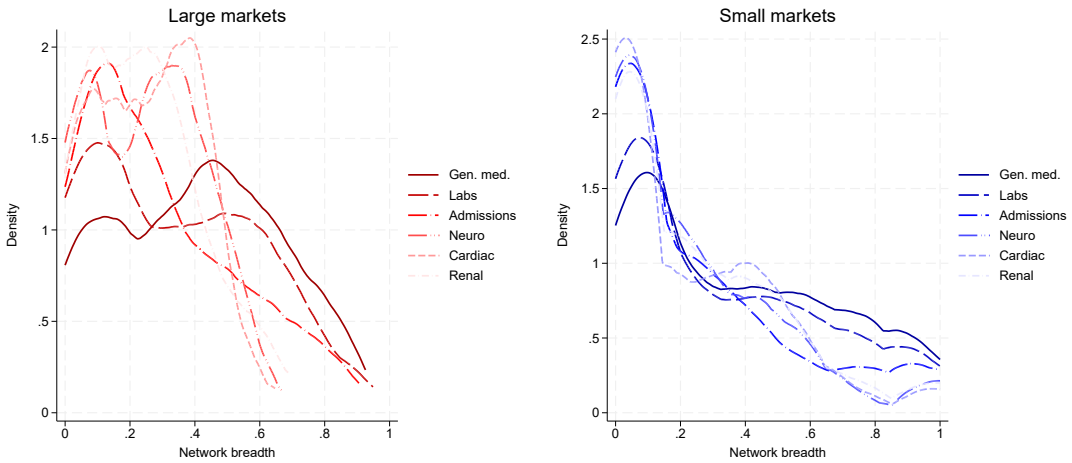
Appendix 3 Additional descriptives in full sample

This appendix presents additional descriptive evidence in the full sample of enrollees. Appendix Figure 3 shows the distribution of network breadth during 2011 for general medicine, laboratory testing, hospital admissions, neurological care, cardiac care, and renal care. Appendix Figure 4 presents these distributions conditional on the 5 largest markets and the rest of markets. These two figures illustrate the types of services that are under-covered across all markets. Appendix Table 5 shows OLS regressions of an indicator for switching out of an insurer on that insurer's service network breadth conditional on patients with different health conditions. The estimation uses the full sample of individuals and findings provide evidence of adverse selection in service network breadth.



Appendix Figure 3: Distribution of network breadth per service

Note: Figure presents kernel density estimates for the distribution of network breadth conditional on six services: general medicine, laboratory testing, hospital admissions, neurological care, cardiac care, and renal care. An observation to construct these distribution is a combination of insurer, market, and year.



Appendix Figure 4: Distribution of service network breadth per market

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

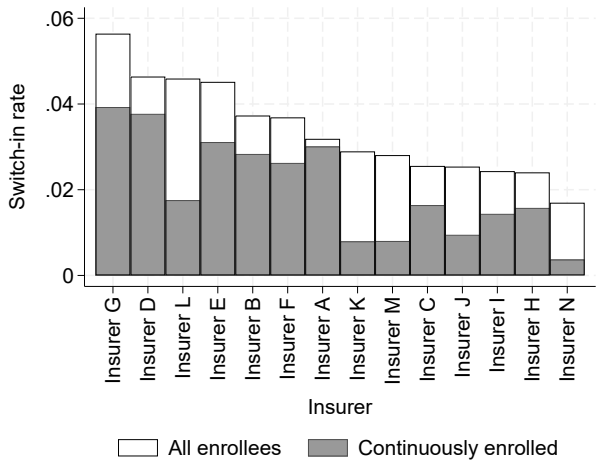
Appendix Table 5: Determinants of switching

Sample:	Healthy	Cancer	Diabetes	Cardio
Service:	General medicine	Therapy	Laboratory	Cardiac care
	(1)	(2)	(3)	(4)
Service network breadth	0.0721	-0.0484	-0.0134	-0.0249
	(0.00106)	(0.00234)	(0.00251)	(0.00188)
Demographics (male, age)	Yes	Yes	Yes	Yes
Enrollment spell length	Yes	Yes	Yes	Yes
Market FE	Yes	Yes	Yes	Yes
N	10703097	770951	346017	1718751
R^2	0.145	0.0740	0.0790	0.0808

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

Appendix 4 Descriptives in subsample

This appendix replicates the descriptive evidence in the main text on the sample for model estimation which comprises individuals who have continuous enrollment spells. Appendix Figure 5 presents the fraction of consumers who switch into each insurer in 2011 relative to 2010 among the full sample and the sample of continuously enrolled.



Appendix Figure 5: Insurer switch-in rates in the continuously enrolled

Note: Figure presents the number of enrollees that switch into each insurer in 2011 relative to 2010 divided by the total number of enrollees at each insurer in 2011. The white bars use the full sample of enrollees. The gray bars use the sample of continuously enrolled.

Appendix Table 6 presents summary statistics of the continuously enrolled, new enrollees, and the random sample of new enrollees used for demand model estimation. New enrollees are relatively healthier and have lower healthcare spending than the continuously enrolled. However, Appendix Table 7, shows that insurers' population of enrollees is in steady state since new enrollees and disenrollees have similar observable characteristics. Appendix Figures 6 and 7 replicate the evidence on selection incentives at the service level and positive correlation between service profitability and network breadth among the continuously enrolled. Finally, Appendix Table 8 shows results of OLS regressions of an indicator for switching out of an insurer on that insurer's network breadth conditional on the continuously enrolled with certain health conditions.

Appendix Table 6: Summary statistics of new and continuous enrollees in 2011

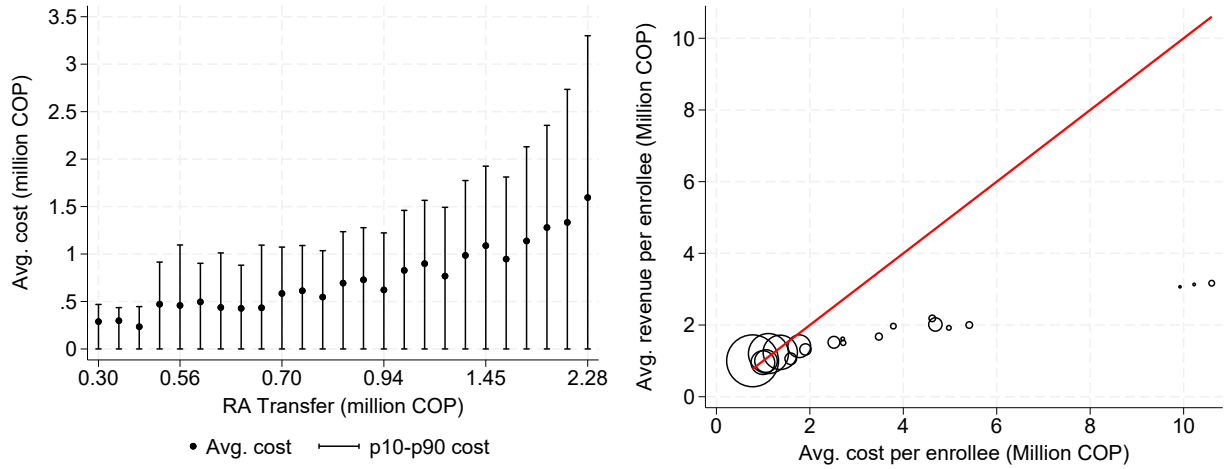
	(1) Continuous		(2) New		(3) Demand sample	
	mean	sd	mean	sd	mean	sd
Male	0.447	0.497	0.508	0.500	0.507	0.500
Age	45.526	15.960	41.791	15.343	41.849	15.354
Cancer	0.088	0.283	0.058	0.233	0.057	0.232
Diabetes	0.042	0.200	0.018	0.135	0.019	0.136
Cardiovascular disease	0.194	0.396	0.099	0.299	0.099	0.299
Pulmonary disease	0.021	0.142	0.010	0.099	0.010	0.100
Renal disease	0.018	0.133	0.007	0.081	0.007	0.082
Other disease	0.071	0.257	0.039	0.193	0.039	0.193
Total healthcare cost†	0.591	3.623	0.358	3.708	0.360	2.219
Risk-adjusted transfer†	0.761	0.477	0.666	0.415	0.668	0.416
Income†	1.280	0.431	1.238	0.402	1.256	0.385
Observations	7675021		3101064		500000	

Note: Table presents mean and standard deviation in parenthesis of demographic and health characteristics of the continuously enrolled in column (1), new enrollees in column (2), and the sample of new enrollees for demand model estimation in column (3) during 2011. (†) measured in millions of COP.

Appendix Table 7: Summary statistics of new enrollees and disenrollees

	(1) 2011 New enrollees		(2) 2010 Disenrollees	
	mean	sd	mean	sd
Male	0.508	0.500	0.510	0.500
Age	41.79	15.34	41.91	15.44
Cancer	0.058	0.233	0.037	0.190
Diabetes	0.018	0.135	0.014	0.116
Cardiovascular disease	0.099	0.299	0.070	0.255
Pulmonary disease	0.010	0.099	0.008	0.090
Renal disease	0.007	0.081	0.005	0.069
Other disease	0.039	0.193	0.029	0.168
Total healthcare cost†	0.358	3.708	0.342	2.540
Risk-adjusted transfer†	0.666	0.415	0.640	0.405
Income†	1.238	0.402	1.230	0.405
Observations	3101064		2973001	

Note: Table presents mean and standard deviation in parenthesis of demographic and health characteristics of 2011 new enrollees in column (1) and of individuals who disenroll after 2010 in column (2). Both columns consider individuals who are continuously enrolled. (†) measured in millions of COP. Monetary measures for disenrollees are adjusted for inflation and expressed in pesos of 2011.



(a) Cost by risk-adjusted transfer

(b) Service-level selection incentives

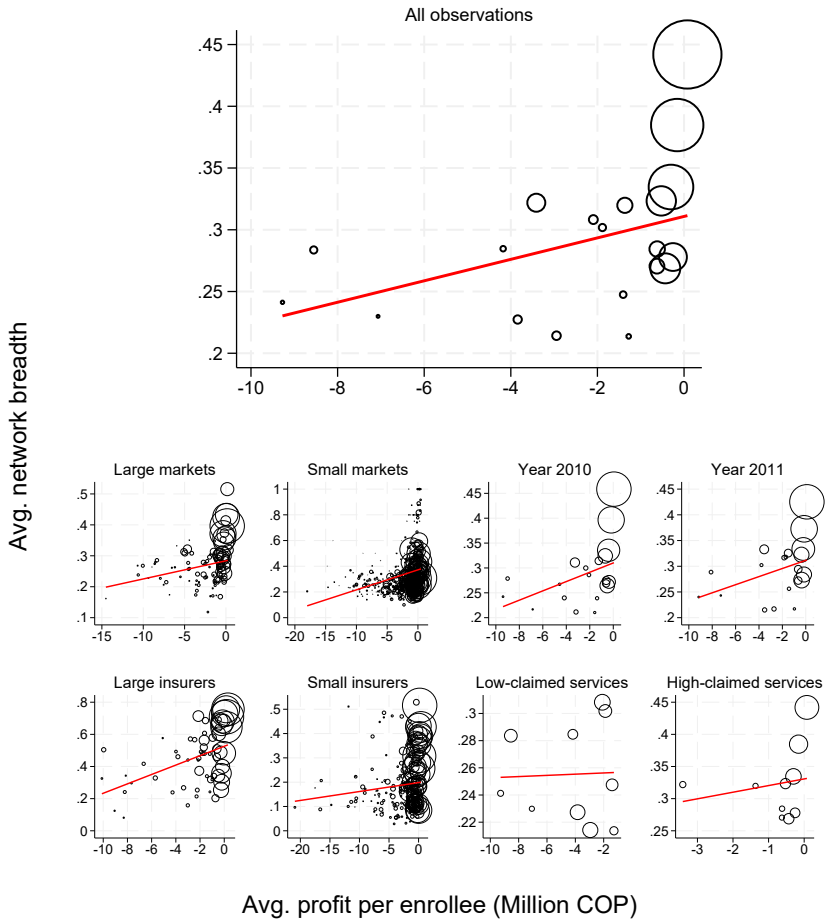
Appendix Figure 6: Costs and selection incentives in the continuously enrolled

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

Appendix Table 8: Determinants of switching in the continuously enrolled

Sample:	Healthy	Cancer	Diabetes	Cardio
Service:	General medicine	Therapy	Laboratory	Cardiac care
	(1)	(2)	(3)	(4)
Service network breadth	-0.0010	-0.00048	-0.00024	-0.00093
	(0.00013)	(0.00022)	(0.00016)	(0.00015)
Demographics (male, age)	Yes	Yes	Yes	Yes
Enrollment spell length	Yes	Yes	Yes	Yes
Market FE	Yes	Yes	Yes	Yes
N	2783571	422246	243125	1147976
R^2	0.00027	0.00027	0.00011	0.00025

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



Appendix Figure 7: Network breadth and service profitability in the continuously enrolled

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

Appendix 5 Micro-foundation

Substantial research in the U.S. finds that patients' enrollment decisions depend on whether a specific health plan covers their preferred healthcare providers (e.g., **??**). Following **?**, this literature uses random utility models that allow for preference heterogeneity across providers to derive patients' expected utility for their insurer's network. This expected utility is fed into models of insurer and provider Nash bargaining with the goal of endogenizing negotiated prices (**??**). The structural unobservable of the Nash-in-Nash surplus function is typically the provider's marginal cost, but off-equilibrium prices in the event that the insurer and the provider disagree are assumed to be fixed at their equilibrium values (**?**).

When trying to endogenize insurers' decision over which providers to include in its network in addition to negotiated prices, the assumption that disagreement payoff prices are fixed can be too strong. Nash-in-Nash proves to be infeasible in this case because both the provider's marginal cost and the off-equilibrium price are unobserved, resulting in a system with more unknowns than equations.[1] **?** and **?** provide some solutions to this problem. The first group of authors apply their model to a setting with a monopolist insurer and the second group of authors showcase their model through simulations.

My model provides an alternative solution by redefining the problem of *which* providers to include in the network and at what *price* as a problem of *how* many providers to include and at what *cost*. Below I show how my insurer cost functions can be derived from standard models of provider choice and Nash-in-Nash bargaining.

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

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

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

[1] **?** discuss some of the limitations of allowing for endogenous networks in Nash-in-Nash in the context of television markets.

in service k, G_{jk}, is:

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

which simplifies to:

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

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

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

where $\phi_j = \sum_k \phi_{jk}$. This shows that insurer demand can be modelled as a function of $\sum_k \overline{\xi}_k H_{jk}$ and insurer fixed effects ϕ_j. It also shows that a demand function defined over network breadth is an exact representation of markets where providers are homogeneous *conditional on the service.*

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

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

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

providers that deliver service k. Then:

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

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

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

The log Nash surplus function is:

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

where β is the bargaining power of the insurer, and t_h^j and t_j^h are the insurer and provider disagreement payoffs, respectively. The insurer disagreement payoff is defined as the profit it would enjoy if it excludes provider h from the network, while reimbursing the rest of providers at their equilibrium prices. Provider disagreement payoffs are defined analogously.

The FOC of the log Nash surplus function with respect to the negotiated price is:

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

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

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

where variables with over-lines denote the value for the average provider in service k.

This derivation shows that an average cost function that is quadratic in network breadth is a correct simplification of certain bargaining environments under the assumption that insurer costs do not depend on the disagreement payoffs.

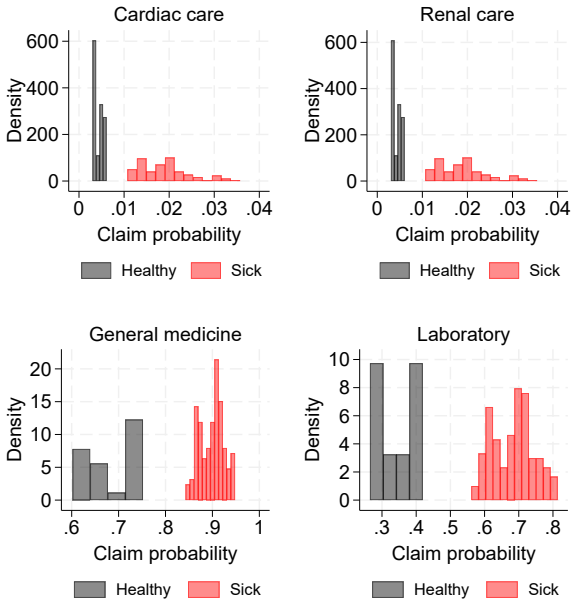
Appendix 6 Model inputs

6.1 Service claim probabilities

I estimate the claim probabilities using the following logistic regression:

$$\text{logit}(\text{any claims})_{ik} = \psi_k + \psi_\theta + \epsilon_{ik} \tag{1}$$

The dependent variable is an indicator for whether patient i makes a claim for service k. On the right side, ψ_k and ψ_θ are service and consumer type fixed effects, respectively. ϵ_{ikm} is a mean zero shock to the claim probability that is independent of network breadth conditional on consumer observables. I assume that new enrollees' expectations over the services they will need are correct on average and that these expectations do not depend on the insurer they enroll with. I estimate equation (1) on data from both current and new enrollees in 2010 and 2011. Appendix Figure 8 presents the resulting distribution of $q_{\theta k}$ for a few services such as cardiac care, renal care, general medicine, and laboratory testing.



Appendix Figure 8: Distribution of service claim probability

Note: Figure presents the distribution of the probability of making a claim in a few specific services, separately for individuals with (sick) and without (healthy) diagnoses. Services reported in the figure include cardiac care, renal care, general medicine, and laboratory testing.

6.2 Support for model assumptions

This subsection presents supporting evidence for the assumptions of my model. Appendix Table 9 shows that negotiated service prices and total service healthcare costs are higher the broader is the network. Hence, consumers who enroll broad-network insurers face a trade-off between higher out-of-pocket payments and greater coverage. Appendix Table 10 shows that service claim probabilities are uncorrelated with service network breadth conditional on the consumer type, providing suggestive evidence for using demographic variation in claim probabilities to identify the parameter on network breadth in the demand model. Appendix Table 11 illustrates that the probability of receiving a chronic disease diagnosis in 2011 conditional on not having one in 2010 is uncorrelated with network breadth. Thus, transition probabilities in the insurer profit function are orthogonal to network breadth. Appendix Figure 9 displays each insurer's average network breadth in each market. This figure evidences that insurers that offer broad networks in one market, tend to do so in other markets as well, supporting my specification of network formation costs being insurer-market specific.

Appendix Table 9: Correlation between negotiated prices and network breadth

	Log negotiated price (1)	Log total cost (2)
Service network breadth	0.423 (0.154)	3.793 (0.195)
Fraction male	Yes	Yes
Fraction of enrollees with diseases	Yes	Yes
Log average income	Yes	Yes
Number of enrollees	Yes	Yes
Market-service-year FE	Yes	Yes
N	7014	7014
R^2	0.424	0.597

Note: Table presents an OLS regression using as outcomes the log of negotiated service prices in column (1) and the log of total healthcare costs per service in column (2). An observation is an insurer-service-market-year. Negotiated prices are calculated as the average price across providers weighted by number of claims. Specifications control for the fraction of males; fraction of enrollees with cancer, diabetes, cardiovascular disease, pulmonary disease, renal disease, and other other diseases; log of average income; and number of enrollees. Specifications include market-service-year fixed effects. Robust standard errors in parenthesis.

Appendix Table 10: Correlation between claim probabilities and network breadth

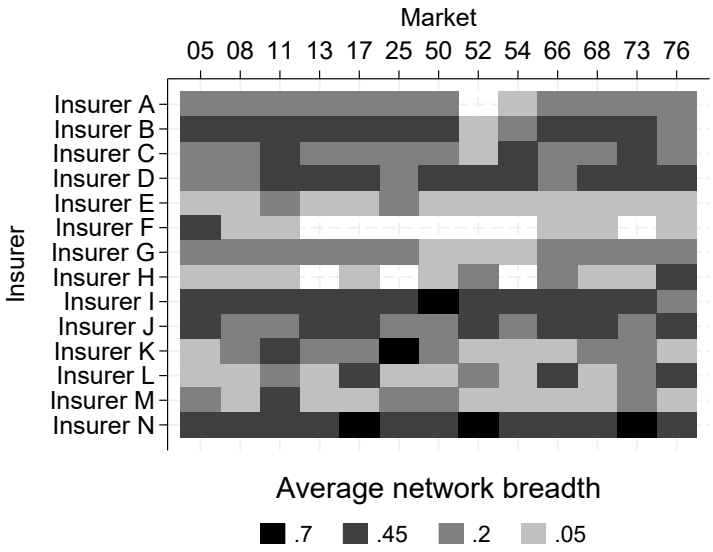
	Claim probability
Service network breadth	0.00033
	(0.0006)
Constant	0.0957
	(0.0002)
Consumer type FE	Yes
Service FE	Yes
Market FE	Yes
N	10000000
R^2	0.945

Note: Table presents an OLS regression of the claim probability on service network breadth. An observation is an individual-service. Estimation uses the subsample of new enrollees for model estimation. Specification includes consumer type, service, and market fixed effects. Standard errors are clustered at the consumer type level.

Appendix Table 11: Correlation between diagnosis probability and network breadth

	Any diagnosis in 2011
Claim-weighted service network breadth	0.00022
	(0.00072)
Claim probabilities	Yes
Market FE	Yes
N	2783735
R^2	0.997

Note: Table presents an OLS regression of an indicator of being diagnosed with any chronic condition in 2011 on service network breadth weighted by the claim probabilities. Estimation uses the subsample of continuously enrolled individuals who did not have a diagnosis in 2010. Specification controls for the service claim probabilities and includes market fixed effects. Standard errors are clustered at the insurer and market level.



Appendix Figure 9: Average network breadth across markets

Note: Figure presents the average network breadth (averaged across services) for each insurer in the row and each market in the column. Darker colors indicate higher network breadth.

6.3 In-sample demand model fit and additional results

This appendix shows the observed and model-predicted national market shares and market shares in the three largest markets for every insurer.

Appendix Table 12: National market shares

Insurer	Observed	Predicted
EPS001	2.219	2.228
EPS002	8.654	8.583
EPS003	4.441	4.444
EPS005	5.655	5.667
EPS008	5.039	5.019
EPS009	1.995	2.010
EPS010	7.973	7.992
EPS012	1.620	1.640
EPS013	14.878	14.887
EPS016	16.702	16.687
EPS017	7.572	7.548
EPS018	4.280	4.323
EPS023	2.961	2.964
EPS037	16.011	16.008

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

Appendix Table 13: Market shares in three largest markets

	Market 05		Market 11		Market 76	
Insurer	Obs	Pred	Obs	Pred	Obs	Pred
EPS001	0.778	3.295	4.298	2.194	1.183	1.117
EPS002	5.155	6.209	9.547	11.051	2.952	3.463
EPS003	3.070	3.572	8.152	7.541	.875	1.670
EPS005	1.420	3.797	11.346	7.427	2.634	4.505
EPS010	26.999	16.899	3.155	5.966	4.451	4.331
EPS013	11.335	17.231	9.206	12.304	7.424	7.491
EPS016	25.154	21.606	3.691	6.924	27.331	28.424
EPS023	2.300	2.637	6.541	6.731	1.885	1.084
EPS037	14.210	14.851	12.273	7.534	16.888	16.114

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

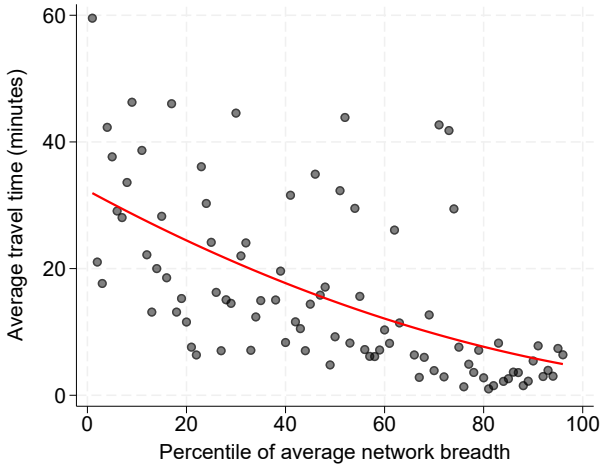
6.4 Alternative specifications

This appendix presents alternative demand specifications to provide encouraging evidence of my identification arguments and modelling choices.

Appendix Table 14: Average willingness-to-pay per service and age group

Age group	Cardiac care	Renal care	Imaging	General medicine	Laboratory	Hospital admissions
Age 19-24	1.111	1.111	1.395	1.790	1.507	1.150
Age 25-29	0.885	0.885	1.051	1.217	1.108	0.909
Age 30-34	0.841	0.841	1.022	1.223	1.087	0.867
Age 35-39	0.519	0.519	0.643	0.794	0.689	0.537
Age 40-44	0.952	0.952	1.151	1.396	1.224	0.980
Age 45-49	0.810	0.809	0.958	1.139	1.012	0.831
Age 50-54	0.845	0.845	1.005	1.185	1.061	0.869
Age 55-59	1.101	1.101	1.252	1.424	1.305	1.124
Age 60-64	1.002	1.002	1.091	1.177	1.118	1.017
Age 65-69	0.912	0.912	0.983	1.054	1.005	0.924
Age 70-74	0.920	0.920	0.974	1.026	0.991	0.929
Age 75+	1.000	1.000	1.000	1.000	1.000	1.000

Note: Table presents average willingness-to-pay for a percentage point increase in network breadth for the service in the column. Willingness-to-pay is calculated as $\frac{1}{-\alpha_i}\frac{\partial s_{ijm}}{\partial H_{jkm}}$, averaged across all consumers with the age group in the row, and normalized to 1 for individuals aged 75 or older.



Appendix Figure 10: Correlation between network breadth and travel times

Note: Figure presents the a scatter plot of average travel time (in minutes) for every percentile of average network breadth (averaged across services). Travel times are calculated between the municipality centroid and the nearest in-network provider for every insurer. The red line corresponds to a quadratic fit.

Appendix Table 15: Insurer demand with star hospital indicator

	Network breadth		OOP spend		Star hospital	
	coef	se	coef	se	coef	se
Mean	3.441	(0.021)	-1.586	(0.117)	0.037	(0.005)
Interactions Male	0.543	(0.010)	0.121	(0.066)		
Cancer	-0.601	(0.013)	-0.012	(0.093)		
Cardiovascular	-0.902	(0.011)	-0.209	(0.076)		
Diabetes	-0.465	(0.023)	0.007	(0.105)		
Other disease	-0.785	(0.015)	0.457	(0.068)		
Pulmonary	-0.611	(0.031)	0.836	(0.095)		
Renal	0.039	(0.037)	0.860	(0.069)		
Age 19-24	0.055	(0.020)	0.577	(0.158)		
Age 25-29	-0.576	(0.019)	0.328	(0.120)		
Age 30-34	-0.560	(0.019)	0.346	(0.130)		
Age 35-39	-0.456	(0.020)	-0.306	(0.214)		
Age 40-44	-0.356	(0.020)	0.456	(0.160)		
Age 45-49	-0.384	(0.019)	0.236	(0.141)		
Age 50-54	-0.324	(0.020)	0.319	(0.134)		
Age 55-59	-0.246	(0.021)	0.476	(0.122)		
Age 60-64	-0.147	(0.023)	0.162	(0.121)		
N			5200890			
Pseudo-R^2			0.112			

Note: Table presents the insurer demand model including a measure of star hospital coverage equal to $\sum_k q_{\theta k} Star_{jkm}$, where $Star_{jkm}$ is an indicator for whether insurer j covers a star hospital in market m for service k. Star hospitals are the top 50 academic medical centers in the country. Specification includes insurer fixed effects. Robust standard errors in parenthesis.

Appendix Table 16: Insurer demand with diagnosis received in January

Variable	Network Breadth		OOP spend	
	coef	se	coef	se
Mean	2.725	(0.021)	-1.048	(0.096)
Interactions				
Male	0.638	(0.010)	-0.013	(0.068)
Cancer	-0.512	(0.031)	-0.164	(0.234)
Cardiovascular	-0.542	(0.019)	-0.605	(0.168)
Diabetes	-0.645	(0.042)	0.338	(0.227)
Other disease	-0.674	(0.033)	0.233	(0.155)
Pulmonary	-0.933	(0.065)	0.830	(0.203)
Renal	-0.519	(0.083)	0.810	(0.087)
Age 19-24	0.411	(0.020)	0.489	(0.148)
Age 25-29	-0.234	(0.019)	0.232	(0.117)
Age 30-34	-0.237	(0.019)	0.375	(0.126)
Age 35-39	-0.163	(0.020)	-0.394	(0.236)
Age 40-44	-0.094	(0.020)	0.547	(0.152)
Age 45-49	-0.176	(0.020)	0.430	(0.140)
Age 50-54	-0.195	(0.020)	0.420	(0.134)
Age 55-59	-0.158	(0.021)	0.337	(0.118)
Age 60-64	-0.090	(0.023)	0.191	(0.130)
N		5200890		
Pseudo-R^2		0.107		

Note: Table presents the insurer demand model obtaining individuals' diagnoses from claims made in January 2011. Specification includes insurer fixed effects. Robust standard errors in parenthesis.

Appendix Table 17: Insurer demand with network breadth weighted by number of beds

Variable	Network Breadth		OOP spend	
	coef	se	coef	se
Mean	4.325	(0.026)	-1.263	(0.111)
Interactions				
Male	0.846	(0.014)	-0.066	(0.061)
Cancer	-0.636	(0.019)	0.018	(0.080)
Cardiovascular	-1.243	(0.016)	-0.117	(0.068)
Diabetes	-0.609	(0.032)	-0.013	(0.102)
Other disease	-1.073	(0.021)	0.247	(0.065)
Pulmonary	-0.905	(0.041)	0.669	(0.097)
Renal	-0.069	(0.057)	0.481	(0.065)
Age 19-24	0.126	(0.029)	1.034	(0.161)
Age 25-29	-0.828	(0.026)	0.612	(0.113)
Age 30-34	-0.807	(0.027)	0.767	(0.116)
Age 35-39	-0.669	(0.028)	0.154	(0.185)
Age 40-44	-0.528	(0.028)	0.878	(0.140)
Age 45-49	-0.551	(0.027)	0.641	(0.130)
Age 50-54	-0.472	(0.028)	0.597	(0.126)
Age 55-59	-0.351	(0.030)	0.641	(0.120)
Age 60-64	-0.215	(0.032)	0.272	(0.120)
N	5200890			
Pseudo-R^2	0.116			

Note: Table presents the insurer demand model with service network breadth weighted by each provider's number of beds. The denominator in network breadth is the total number of beds in the market and the numerator is the number of beds included in the insurer's network. Specification includes insurer fixed effects. Robust standard errors in parenthesis.

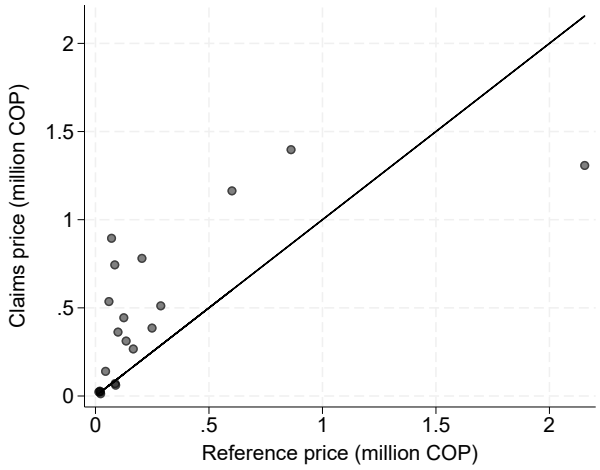
Appendix Table 18: Insurer demand with expectations over diagnoses

Variable	Network Breadth		OOP spend	
	coef	se	coef	se
Mean	2.381	(0.018)	-3.865	(0.134)
Interactions				
Male	0.688	(0.009)	-1.114	(0.135)
Age 19-24	1.038	(0.018)	8.139	(0.310)
Age 25-29	0.267	(0.016)	-0.175	(0.213)
Age 30-34	0.232	(0.017)	-1.721	(0.291)
Age 35-39	0.281	(0.018)	-2.919	(0.297)
Age 40-44	0.325	(0.017)	1.076	(0.351)
Age 45-49	0.160	(0.018)	0.290	(0.262)
Age 50-54	0.058	(0.018)	-0.370	(0.247)
Age 55-59	0.085	(0.018)	1.344	(0.235)
Age 60-64	0.031	(0.020)	0.197	(0.229)
N	5200890			
Pseudo-R^2	0.116			

Note: Table presents the insurer demand model assuming consumers form expectations over diagnoses and services. Let $\tilde{\theta}$ be a combination of sex and age group. Network breadth is calculated as $\sum_d \sum_k \gamma_{d\tilde{\theta}} q_{\theta k} H_{jkm}$ and out-of-pocket costs are calculated as $\sum_d \gamma_{d\tilde{\theta}} c_{\theta jm}$, where $\gamma_{d\tilde{\theta}}$ is the probability of having diagnosis d conditional on $\tilde{\theta}$. Specification includes insurer fixed effects. Robust standard errors in parenthesis.

Appendix 7 Service reference prices

In 2005, the Colombian government published a list of reference prices for all the services included in the national health insurance plan. These prices were created by a group of government officials and medical experts with the purpose of reimbursing healthcare providers in the event of terrorist attacks, natural disasters, and car accidents (Decree 2423 of 1996). Although they were not meant to guide price negotiations between insurers and providers, there is evidence that insurers use these reference prices as starting points in their negotiations with providers (**?**). Appendix Figure 11 shows that references prices are highly correlated with negotiated prices from the claims data and that negotiated prices tend to be higher than the reference price.



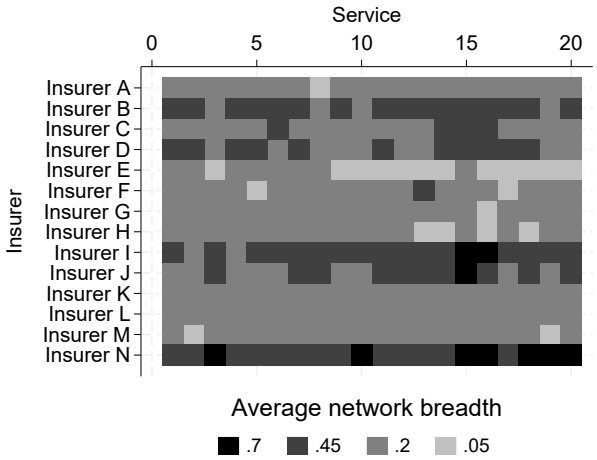
Appendix Figure 11: Negotiated prices and reference prices

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

Appendix 8 Additional average cost results

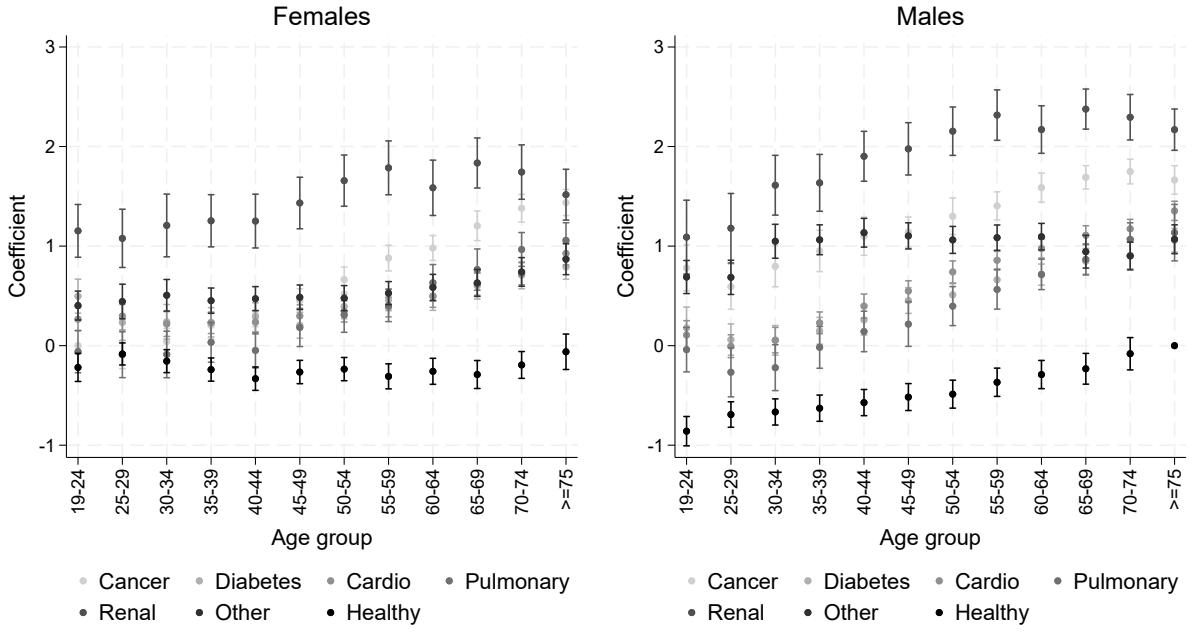
This appendix presents additional descriptive evidence and robustness checks to support my identification and functional form assumption for the average cost model. Appendix Figure 12 shows that insurers' network breadth decisions are correlated across services. Appendix Figure 13 shows the consumer type fixed effects in my main specification. Appendix Figure 14 presents the average cost model fit. Appendix Table 19 shows evidence that insurers that

cover several services within a provider receive price discounts, suggestive of economies of scope.



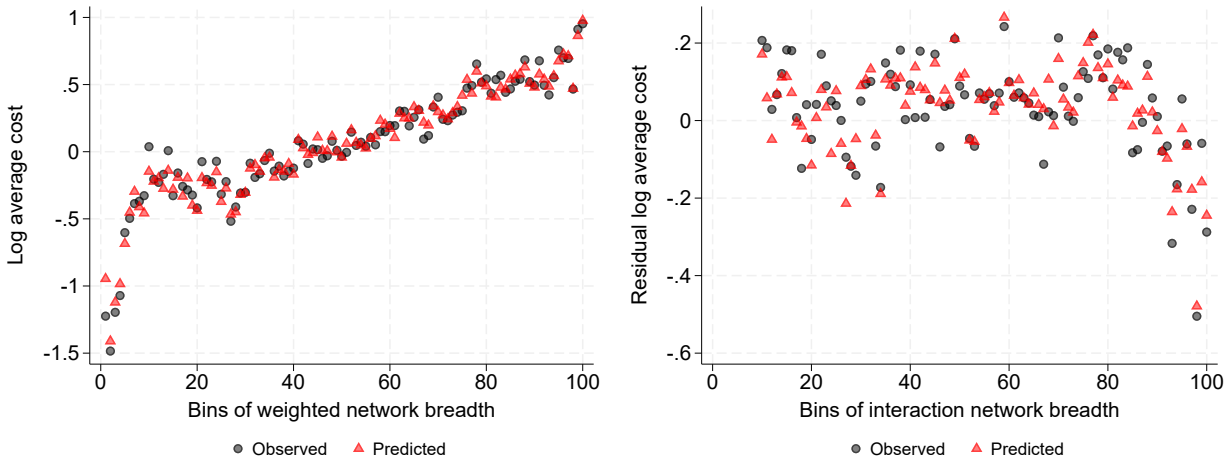
Appendix Figure 12: Average network breadth per insurer and service

Note: Figure presents the average service network breadth (averaged across markets) for each insurer in the row and each service in the column. Darker colors represent higher network breadth.



Appendix Figure 13: Consumer type fixed effects

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



(a) Service network breadth

(b) Service network breadth interactions

Appendix Figure 14: Average cost model fit

Note: Panel A presents a scatter plot of observed (black circles) and predicted (red triangles) log average cost per enrollee by percentiles of weighted service network breadth ($\sum_k q_{\theta k} H_{jkm}$). Panel B presents a scatter plot of observed (black circles) and predicted (red triangles) residual log average cost per enrollee after controlling for service network breadth by percentiles of scope economies ($\sum_k^{K_m} \sum_{l \neq k}^{K_m} q_{\theta k} q_{\theta l} H_{jkm} H_{jlm}$).

Appendix Table 19: Price mechanism for economies of scope

	Log negotiated price	
	(1)	(2)
Provider in network for other services	-0.091	
	(0.002)	
Avg. network breadth for other services		-1.536
		(0.027)
Demographics (male, diagnoses, income)	Yes	Yes
Number of enrollees	Yes	Yes
Market-Service-Year FE	Yes	Yes
N	77312	80129
R^2	0.655	0.660

Note: Table presents OLS regression of the log of the negotiated price for service k between insurer j and provider h on the fraction of other services $l \neq k$ for which the provider is in network in column (1) and on average network breadth across all other services $l \neq k$ in column (2). An observation in this table is a combination of insurer, provider, service, market, year. Specifications control for the fraction of males; fraction of enrollees with cancer, diabetes, cardiovascular disease, pulmonary disease, renal disease, and other other diseases; log of average income; and number of enrollees for the insurer. Specifications include market-service-year fixed effects. Estimation is conditional on the provider being in-network for service k. The reduction in the number of observations in column (1) is due to providers that are in-network for service k only and for which the fraction is indeterminate because it is divided by zero. Robust standard errors in parenthesis.

Appendix Tables 20, 21, and 22 present robustness on my average cost model using patient-level data, including an indicator for star hospital coverage, and weighting network breadth by each provider's number of beds.

Appendix Table 20: Patient-level estimates of average cost

	Log total healthcare cost+1	
	coef	se
Service network breadth	5.591	(0.024)
Scope economies	-54.07	(0.265)
Reference price	-2.169	(0.062)
Insurer FE		
Insurer A	-0.645	(0.395)
Insurer B	0.762	(0.292)
Insurer C	0.954	(0.459)
Insurer D	1.803	(0.259)
Insurer E	2.588	(0.298)
Insurer F	0.976	(0.442)
Insurer G	1.968	(0.268)
Insurer H	0.248	(0.396)
Insurer I	0.650	(0.250)
Insurer J	0.827	(0.254)
Insurer K	1.214	(0.306)
Insurer L	0.106	(0.298)
Insurer M	1.897	(0.426)
Constant	-7.664	(0.167)
Individual FE	Yes	
Market FE	Yes	
N	8003976	
R^2	0.640	

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

Appendix Table 21: Average cost with star hospital indicator

	Log average cost per enrollee	
Variable	coef	se
Service network breadth	0.273	(0.047)
Scope economies	-1.147	(0.579)
Star hospital	0.041	(0.009)
Reference price	—	—
Insurer FE		
Insurer A	0.154	(0.038)
Insurer B	-0.093	(0.022)
Insurer C	0.017	(0.025)
Insurer D	-0.152	(0.027)
Insurer E	0.273	(0.047)
Insurer F	0.010	(0.061)
Insurer G	0.071	(0.034)
Insurer H	0.036	(0.040)
Insurer I	-0.008	(0.018)
Insurer J	0.107	(0.021)
Insurer K	-0.132	(0.033)
Insurer L	0.113	(0.035)
Insurer M	-0.050	(0.032)
Constant	-0.330	(0.053)
Consumer type FE	Yes	
Market FE	Yes	
N	18369	
R^2	0.612	

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

Appendix Table 22: Average cost with network breadth weighted by number of beds

	Log average cost per enrollee	
Variable	coef	se
Service network breadth	0.307	(0.065)
Scope economies	-1.272	(0.683)
Reference price	2.818	(0.570)
Insurer FE		
Insurer A	0.093	(0.037)
Insurer B	-0.166	(0.021)
Insurer C	-0.050	(0.023)
Insurer D	-0.242	(0.025)
Insurer E	0.195	(0.046)
Insurer F	0.001	(0.057)
Insurer G	0.015	(0.034)
Insurer H	0.012	(0.040)
Insurer I	-0.028	(0.017)
Insurer J	0.091	(0.021)
Insurer K	-0.212	(0.032)
Insurer L	0.133	(0.035)
Insurer M	-0.122	(0.031)
Constant	-1.393	(0.104)
Consumer type FE	Yes	
Market FE	Yes	
N	18369	
R^2	0.610	

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

Appendix 9 Additional network administrative cost results

9.1 Dropout and transition probabilities

To estimate the marginal cost of network formation in the third step of my model, I first compute the probability that consumer type θ drops out of the contributory system and the probability that consumer type θ in period t transitions into θ' in period $t+1$. I use data from *all* enrollees to the contributory system in 2010 and 2011, regardless of their enrollment spell length, to compute dropout probabilities. For each consumer type θ, I calculate the probability that they drop out of the system non-parametrically as the number of individuals of type θ observed only in 2010 but not 2011, divided by the total number of type θ individuals in 2010. Appendix Table 23 presents the mean and standard deviation of the dropout probability conditional on diagnoses, sex, and age.

To compute transition probabilities I use a non-parametric approach as well, using data from all the continuously enrolled in 2010 and 2011. The probability that type θ transitions into θ' equals the number of type θ in 2010 that end up with diagnosis l' in 2011, divided by the number of type θ individuals in 2010. Appendix Table 24 presents the mean and standard deviation in parenthesis of transition probabilities from having cancer, diabetes, cardiovascular disease, pulmonary disease, renal disease, other diseases, and no diseases in period t to having each of these 7 health conditions in period $t+1$.

9.2 Marginal variable profits

This appendix presents additional descriptive evidence for marginal variable profits and measures of out-of-sample model fit.

Appendix Table 23: Dropout probabilities

	mean	sd
Female	0.084	(0.131)
Male	0.106	(0.165)
Age 19-24	0.120	(0.177)
Age 25-29	0.087	(0.134)
Age 30-34	0.081	(0.135)
Age 35-39	0.085	(0.141)
Age 40-44	0.085	(0.146)
Age 45-49	0.085	(0.149)
Age 50-54	0.089	(0.153)
Age 55-59	0.091	(0.158)
Age 60-64	0.092	(0.158)
Age 65-69	0.096	(0.159)
Age 70-74	0.104	(0.160)
Age 75+	0.124	(0.164)
Cancer	0.048	(0.024)
Diabetes	0.027	(0.008)
Cardiovascular	0.028	(0.009)
Pulmonary	0.040	(0.015)
Renal	0.044	(0.018)
Other disease	0.026	(0.011)
Healthy	0.450	(0.073)

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

Appendix Table 24: Transition probabilities

Diagnosis	Stat	Cancer	Cardio	Diabetes	Renal	Pulmonary	Other	Healthy
Cancer	mean	0.316	0.017	0.139	0.014	0.007	0.047	0.460
	sd	0.067	0.014	0.090	0.013	0.006	0.019	0.176
Diabetes	mean	0.030	0.557	0.170	0.009	0.013	0.021	0.200
	sd	0.026	0.078	0.100	0.010	0.011	0.010	0.140
Cardio	mean	0.043	0.028	0.554	0.014	0.011	0.034	0.316
	sd	0.036	0.018	0.205	0.012	0.010	0.009	0.224
Pulmonary	mean	0.055	0.019	0.191	0.234	0.007	0.078	0.416
	sd	0.046	0.014	0.089	0.152	0.006	0.034	0.231
Renal	mean	0.044	0.036	0.214	0.012	0.371	0.058	0.265
	sd	0.035	0.030	0.132	0.013	0.062	0.031	0.154
Other	mean	0.056	0.016	0.156	0.023	0.008	0.343	0.398
	sd	0.040	0.013	0.106	0.020	0.004	0.058	0.095
Healthy	mean	0.055	0.012	0.108	0.014	0.004	0.045	0.762
	sd	0.042	0.008	0.068	0.014	0.003	0.021	0.109

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

Appendix Table 25: Summary statistics of marginal variable profits

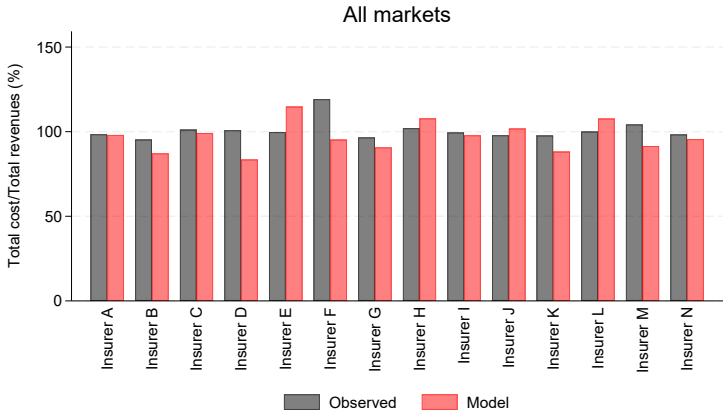
Insurer	mean	sd
Insurer A	6697.998	28615
Insurer B	34998.53	156924.5
Insurer C	16191.7	86058.47
Insurer D	23004.19	111649.9
Insurer E	93551.45	277196.6
Insurer F	11046.99	30985.02
Insurer G	27080.69	97283.26
Insurer H	41008.3	88315.88
Insurer I	31207.44	137559.4
Insurer J	25403.59	95181.3
Insurer K	57247.54	275614.2
Insurer L	14026.04	51929.83
Insurer M	40024.9	150600.8
Insurer N	27492.63	99826.83

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

Appendix Table 26: First-stage regression for network formation cost estimation

	Service network breadth	
	coef	se
Claim probability, healthy	0.310	(0.0194)
Insurer FE		
Insurer A	-0.322	(0.013)
Insurer B	-0.103	(0.016)
Insurer C	-0.255	(0.014)
Insurer D	-0.199	(0.014)
Insurer E	-0.249	(0.020)
Insurer F	-0.301	(0.026)
Insurer G	-0.330	(0.014)
Insurer H	-0.031	(0.029)
Insurer I	-0.074	(0.016)
Insurer J	-0.178	(0.015)
Insurer K	-0.198	(0.020)
Insurer L	-0.034	(0.026)
Insurer M	-0.303	(0.018)
Constant	0.524	(0.015)
Market FE	Yes	
F-statistic	249.47	
N	2280	

Note: Table presents first stage regression for my preferred specification of the network administrative cost. The outcome variable is service network breadth and the covariates are insurer fixed effects, market fixed effects, and the instrument which is the average service claim probability among healthy consumers. The table reports the F-statistic associated with the instrument. Robust standard errors in parenthesis.



Appendix Figure 15: Out-of-sample model fit

Note: Figure presents the model-predicted ratio of total costs (average costs plus network administrative costs) to total revenues and the observed ratio from insurers' public income statements submitted to the National Health Superintendency. Because my model is estimated on the sample of new enrollees with continuous enrollment and public income statements correspond to all enrollees, I scale up estimated insurer revenues and costs by multiplying by the total number of enrollees in the country and dividing by the number of new enrollees.

Appendix Table 27: Model of network administrative costs with networks weighted by number of beds

	Log marginal variable profit	
Variable	coef	se
Service network breadth	20.56	(1.076)
<u>Insurer FE</u>		
Insurer A	4.236	(0.448)
Insurer B	1.110	(0.332)
Insurer C	3.494	(0.408)
Insurer D	2.319	(0.344)
Insurer E	4.106	(0.663)
Insurer F	5.511	(0.727)
Insurer G	6.355	(0.499)
Insurer H	-0.353	(0.858)
Insurer I	1.394	(0.317)
Insurer J	4.493	(0.396)
Insurer K	3.065	(0.421)
Insurer L	-0.602	(0.474)
Insurer M	4.815	(0.559)
Constant	-5.216	(0.838)
Market FE	Yes	
First-stage F-stat	344.78	
N	2280	

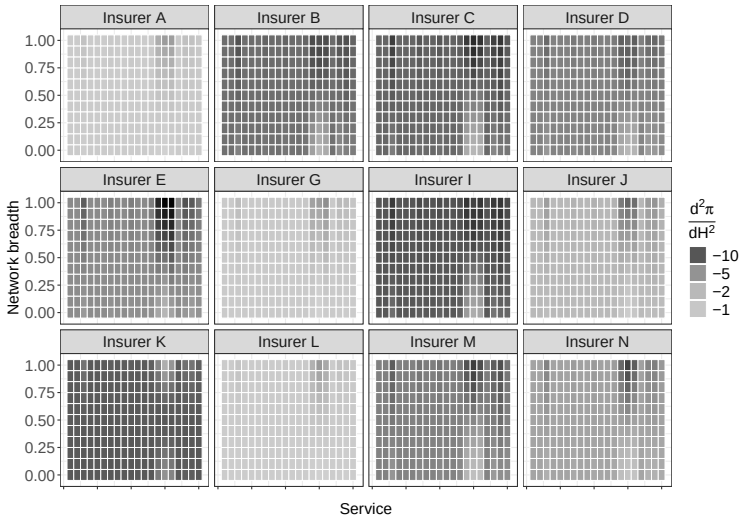
Note: Table presents 2SLS regression of the log of marginal variable profit on service network breadth, insurer fixed effects, and market fixed effects. Service network breadth is calculated using each provider's number of beds as weights. An observation is a combination of insurer, service, and market. The instrument for service network breadth is the average service claim probability among healthy consumers. The table reports the F-statistic for the first stage regression. Robust standard errors in parenthesis.

Appendix 10 Concavity of the profit function

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

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

To check whether this derivative is negative at all values of network breadth, I conduct a partial equilibrium exercise where each insurer is allowed to deviate and set $H_{jkm} = \{0, 0.1, 0.2, 0.3, ..., 1\}$ for each service k, holding its rivals' choices fixed at observed levels. I compute this exercise with data from Bogotá as in my counterfactuals. Appendix Figure 16 presents the results. Each panel corresponds to the deviating insurer, and displays the value of the second partial derivative for each service in the horizontal axis and for each value of network breadth in the vertical axis. Results show that the second partial derivative of the short-run profit function is negative for all insurers and services.



Appendix Figure 16: Second partial derivative of short-run profit function

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

Appendix 11 Additional counterfactual results

Appendix Table 28: Insurer demand

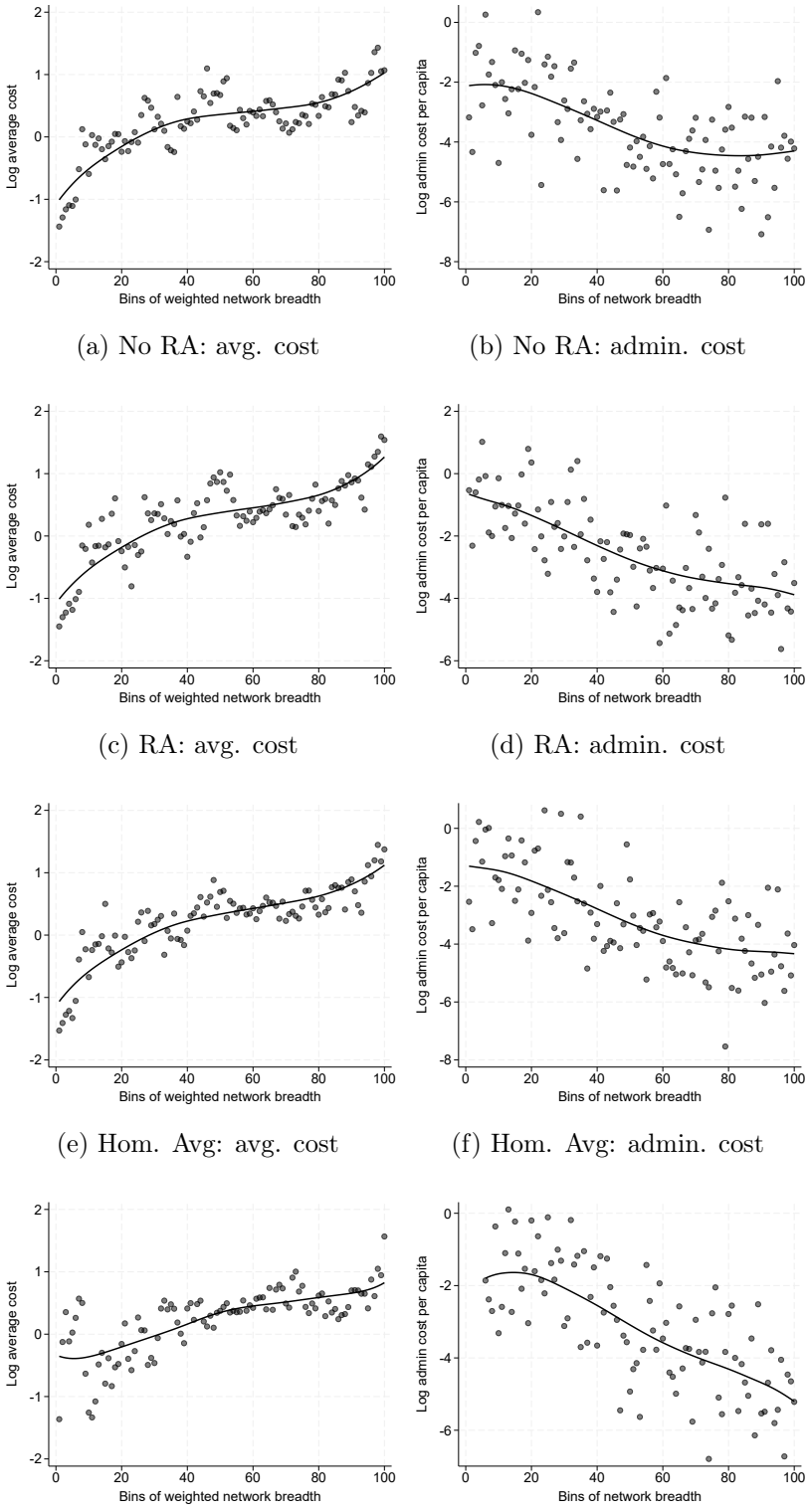
Diagnosis list
Healthy
Cardiovascular disease
Other Disease
Cervical Cancer
Breast Cancer
Other Renal Disease
Other Cancer
Chronic Kidney Disease
Diabetes
Skin Cancer
Lymphoma
Stomach Cancer
HIV-AIDS
Lung Cancer

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

Appendix Table 29: Networks, costs, and welfare under homogeneous costs

Variable	(1) Avg cost	(2) Admin cost
Panel A. Overall		
Average network breadth	-0.252	-1.497
Average cost per enrollee	-19.08	-18.77
Total average cost	-10.63	-8.642
Consumer surplus (without diagnoses)	-0.488	2.442
Consumer surplus (with diagnoses)	0.003	3.126
Panel B. Service network breadth		
Otorhinolaryngologic care	-0.252	-2.434
Cardiac care	-0.280	-2.329
Gastroenterologic care	-0.273	-2.349
Renal care	-0.296	-2.465
Gynecologic care	-0.281	-1.990
Orthopedic care	-0.281	-2.354
Imaging	-0.136	-0.431
General medicine	-0.048	-1.469
Laboratory	-0.106	-0.925
Hospital admission	-0.234	-1.516

Note: Panel A presents the percentage change relative to the observed scenario in average network breadth, insurer total average costs, short-run average cost per enrollee, and long-run consumer welfare for individuals with and without diagnoses, in the scenario with homogeneous average costs in column (1), and the scenario with homogeneous average and network formation costs in column (2). Insurer fixed effects in average costs and network formation costs are set to the average fixed effect. Panel B presents the percentage change relative to the observed scenario in average network breadth by service category.



(a) No RA: avg. cost (b) No RA: admin. cost

(c) RA: avg. cost (d) RA: admin. cost

(e) Hom. Avg: avg. cost (f) Hom. Avg: admin. cost

(g) Hom. admin: avg. cost (h) Hom. admin: admin. cost

Appendix Figure 17: Functional form stability in counterfactuals

Note: Figure presents the predicted log average cost per enrollee and the predicted log network administrative cost per enrollee by percentile of service network breadth. Each row corresponds to a counterfactual: without risk adjustment, with improved risk adjustment, and imposing homogeneity in average and network administrative costs.