

# Appendices for Non-Price Competition and Risk Selection Through Hospital Networks

## Appendix 1 Current risk adjustment system

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

Department/city	Transfer
National (pesos)	505,627.2
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	$\times 1.06$

Appendix 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

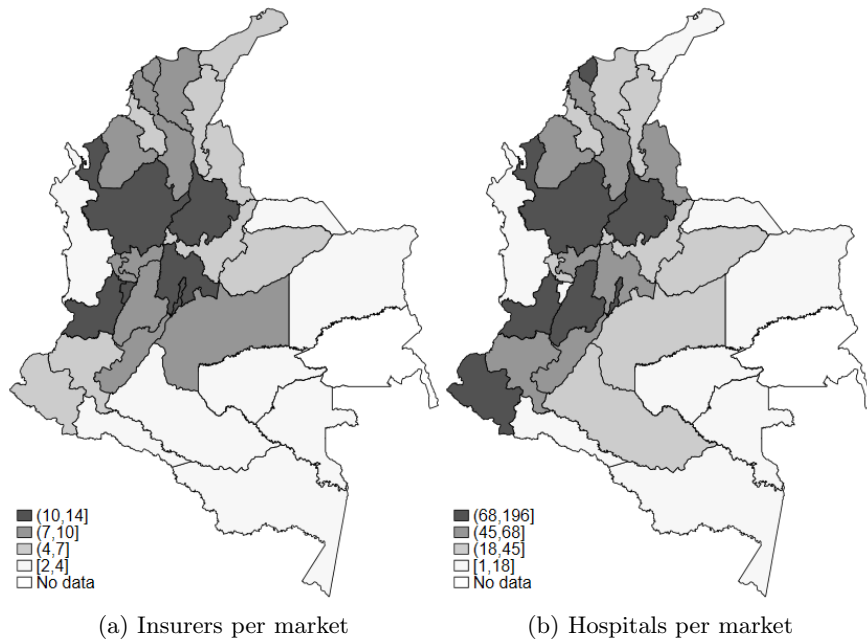
## Appendix 2 Service categories

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

## Appendix 3 Description of Colombian health care market

Appendix Table 3: National market shares in 2011

Insurer	Market share
EPS013	21.4
EPS016	15.2
EPS037	11.1
EPS002	9.3
EPS017	7.2
EPS010	7.1
EPS005	4.5
EPS018	4.4
EPS003	4.0
EPS008	3.7
EPS023	3.1
EPS009	1.8
EPS001	1.6
EPS012	1.6



Appendix Figure 1: Number of insurers and hospitals per market

Appendix Table 4: Service coverage per insurer at Fundación Valle del Lili in Cali in 2011

	Insurer EPS0-													
	01	02	03	05	08	09	10	12	13	16	17	18	23	37
Adult intermediate care	1	0	0	1	1	0	0	1	0	0	0	1	0	1
Anesthesia	1	0	0	1	1	0	1	0	0	0	0	1	0	1
Cardiac vessel implant	1	0	0	0	1	0	0	0	0	0	0	1	0	0
Cardiology	1	0	0	0	1	0	1	1	0	0	0	1	0	1
Cardiovascular diagnostics	0	1	0	1	1	0	1	1	0	0	0	0	0	1
Cardiovascular surgery	1	1	0	1	1	0	0	1	0	0	0	1	0	1
Dialysis	0	1	0	0	1	0	1	0	0	0	0	0	0	1
Family medicine	0	0	0	0	1	0	1	0	0	0	0	1	0	1
Gastroenterology	0	1	0	0	1	0	1	1	0	0	0	1	0	1
Gastrointestinal surgery	0	1	0	1	1	0	0	0	0	0	0	1	0	1
General adult admission	1	0	0	1	1	0	0	1	0	1	0	1	0	1
General medicine	0	0	0	0	1	0	1	0	0	0	0	1	0	1
General pediatric admission	1	0	0	1	1	0	0	1	0	0	0	1	0	1
General surgery	0	1	0	1	1	0	0	1	0	0	0	1	0	1
ICU	1	0	0	1	1	0	0	1	0	0	0	1	0	1
Infectology	1	0	0	0	1	0	1	1	0	0	0	1	0	1
Internal medicine	1	0	0	0	1	0	1	1	0	0	0	1	0	1
Nephrology	1	0	0	0	1	0	1	1	0	0	0	1	0	1
NICU	1	0	0	0	1	0	0	1	0	0	0	1	0	1
Nuclear medicine	0	1	0	1	1	0	0	0	0	0	0	1	0	1
Oncological surgery	1	0	0	0	1	0	0	1	0	0	0	1	0	1
Oncology	1	0	0	1	1	0	1	1	0	0	0	1	0	1
Ophthalmological surgery	0	1	0	1	1	0	0	1	0	0	0	1	0	1
Ophthalmology	0	0	0	0	1	0	1	1	0	0	0	1	0	1
Pathology	1	0	0	0	1	0	1	0	0	0	0	1	0	1
Radiotherapy	1	0	0	1	1	0	1	1	0	0	0	0	0	1
Ultrasound	0	0	0	1	1	0	0	0	0	0	0	1	0	0

*Note:* Table presents service coverage per insurer at the largest hospital in Cali during 2011 for a sample of services. Data comes from the National Health Superintendency.

Appendix Table 5: Service coverage per insurer at Fundación Santa Fe in Bogotá in 2011

	Insurer EPS0-													
	01	02	03	05	08	09	10	12	13	16	17	18	23	37
Adult intermediate care	0	0	0	1	1	0	1	1	0	0	1	0	0	1
Anesthesia	0	0	0	0	1	0	1	0	0	0	0	0	0	1
Cardiac vessel implant	0	0	0	0	1	0	0	0	0	0	0	0	0	1
Cardiology	0	0	0	1	1	0	1	1	0	0	0	1	0	1
Cardiovascular diagnostics	0	0	0	1	1	0	1	1	0	0	1	0	0	1
Cardiovascular surgery	0	0	0	1	1	0	1	1	0	0	0	0	0	1
Dialysis	0	0	0	1	0	0	0	0	0	0	1	0	0	0
Family medicine	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Gastroenterology	0	0	0	1	0	0	1	1	0	0	1	1	0	1
Gastrointestinal surgery	0	0	0	0	1	0	0	0	0	0	0	0	0	0
General adult admission	0	0	0	1	1	0	1	1	0	1	1	0	0	1
General medicine	0	0	0	0	0	0	1	0	0	0	1	1	0	1
General pediatric admission	0	0	0	1	1	0	1	1	0	0	0	0	0	1
General surgery	0	0	0	1	1	0	1	1	0	0	1	1	0	1
ICU	0	0	0	1	1	0	1	1	0	0	1	0	0	1
Infectology	0	0	0	1	0	0	1	1	0	0	0	0	0	1
Internal medicine	0	0	0	1	0	0	1	1	0	0	1	1	0	1
Nephrology	0	0	0	1	0	0	1	1	0	0	1	0	0	1
NICU	0	0	0	0	1	0	1	1	0	0	1	0	0	1
Nuclear medicine	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Oncological surgery	0	0	0	1	1	0	1	1	0	0	0	0	0	1
Oncology	0	0	0	1	1	0	1	1	0	0	1	0	0	1
Ophthalmological surgery	0	0	0	1	1	0	1	1	0	0	0	0	0	1
Ophthalmology	0	0	0	0	0	0	1	1	0	0	1	0	0	1
Pathology	0	0	0	0	0	0	1	0	0	0	0	0	0	1
Radiotherapy	0	0	0	1	0	0	1	0	0	0	1	0	0	1
Ultrasound	0	0	0	0	1	0	0	0	0	0	0	0	0	0

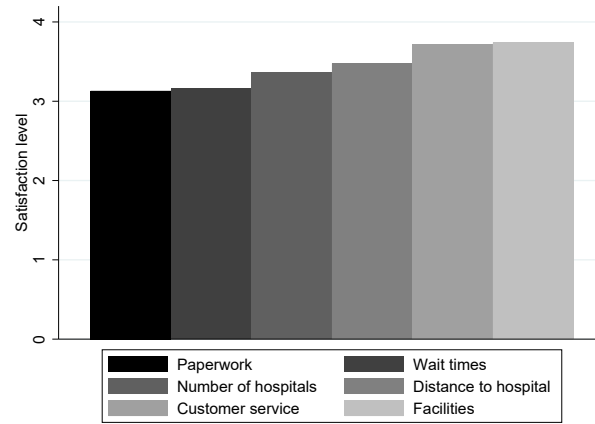
*Note:* Table presents service coverage per insurer at the largest hospital in Bogotá during 2011 for a sample of services. Data comes from the National Health Superintendency.

Appendix Table 6: Service coverage per insurer at Hospital Pablo Tobón in Medellín in 2011

	Insurer EPS0-													
	01	02	03	05	08	09	10	12	13	16	17	18	23	37
Adult intermediate care	0	0	0	0	0	1	0	0	0	0	0	1	0	1
Anesthesia	0	1	0	0	0	1	0	0	0	0	0	1	0	1
Cardiology	0	0	0	0	0	1	0	0	0	0	0	1	0	1
Cardiovascular diagnostics	0	1	0	0	0	1	0	0	0	0	0	0	0	1
Cardiovascular surgery	0	1	0	0	0	0	0	0	0	0	0	1	0	1
Dialysis	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Gastroenterology	0	1	0	0	0	1	0	0	0	0	0	1	0	1
Gastrointestinal surgery	0	1	0	0	0	0	0	0	0	0	0	0	0	1
General adult admission	0	1	0	0	0	0	0	0	0	1	0	1	0	1
General medicine	0	1	0	0	0	1	0	0	0	0	0	1	0	1
General pediatric admission	0	1	0	0	0	1	0	0	0	0	0	1	0	1
General surgery	0	1	0	0	0	1	0	0	0	0	0	1	0	1
ICU	0	0	0	0	0	1	0	0	0	0	0	1	0	1
Internal medicine	0	0	0	0	0	1	0	0	0	0	0	1	0	1
Nephrology	0	0	0	0	0	1	0	0	0	0	0	0	0	1
NICU	0	1	0	0	0	0	0	0	0	0	0	1	0	1
Nuclear medicine	0	1	0	0	0	0	0	0	0	0	0	0	0	1
Oncological surgery	0	1	0	0	0	1	0	0	0	0	0	0	0	1
Oncology	0	1	0	0	0	1	0	0	0	0	0	1	0	1
Ophthalmological surgery	0	1	0	0	0	1	0	0	0	0	0	1	0	1
Ophthalmology	0	0	0	0	0	1	0	0	0	0	0	1	0	1
Radiotherapy	0	0	0	0	0	1	0	0	0	0	0	0	0	1
Ultrasound	0	0	0	0	0	0	0	0	0	0	0	0	0	1

*Note:* Table presents service coverage per insurer at the largest hospital in Medellín during 2011 for a sample of services. Data comes from the National Health Superintendency.

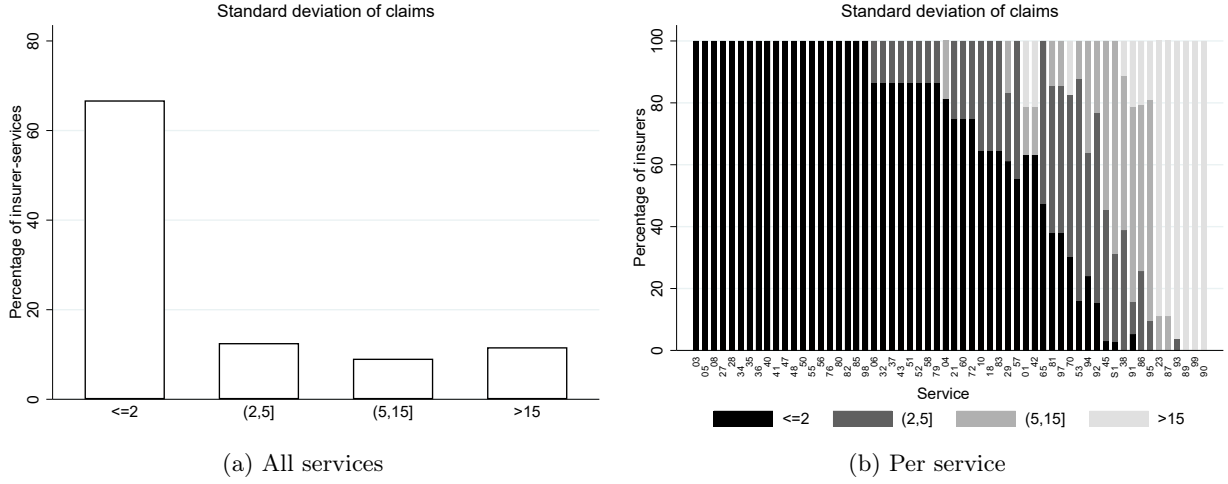
## Appendix 4 Variation in provider quality within service



Appendix Figure 2: Average satisfaction levels

*Note:* This figure shows the average enrollee satisfaction level with six measures of insurer quality: amount of paperwork, number of hospitals in-network, customer service, wait times for doctor appointments, distance to the nearest hospital, and insurer facilities. The data for this graph are the pooled enrollee satisfaction surveys from 2013 to 2016 conducted by Ministry of Health to individuals in the contributory regime.

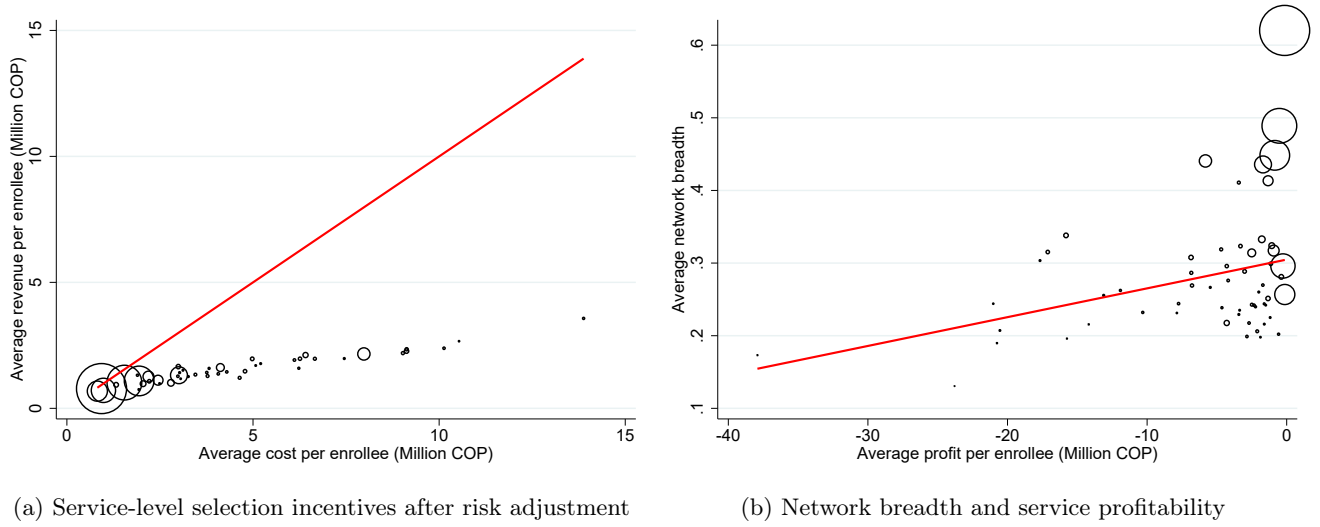




Appendix Figure 3: Standard deviation of claims across providers

*Note:* Panel (a) of this figure shows distribution of the standard deviation of number of claims across insurer-service combinations. Panel (b) shows the distribution of the standard deviation of number of claims across insurers separately for every service in the x-axis.

## Appendix 5 Robustness checks on correlates of network breadth



Appendix Figure 4: Service-level selection incentives and profitability

*Note:* This figure uses data from all enrollees to the contributory system without constraining enrollment to be continuous. Panel (a) presents a scatter plot of average revenue and average cost per enrollee among those who make claims for each service. Individual costs are annualized by dividing by the number of days the individual is enrolled to the system and multiplying by 360. Panel (b) presents a scatter plot of profits per enrollee among those that make claims for each service, where individual costs are annualized as before, and average network breadth per service.

Appendix Table 7: Network breadth and service profitability

	Network breadth		
	(1)	(2)	(3)
Profit per enrollee	0.002*** (0.0002)	0.002*** (0.0002)	0.001*** (0.0002)
Market	—	Y	Y
Insurer	—	Y	Y
Service	—	—	Y
<i>N</i>	11,377	11,377	11,377
<i>R</i> <sup>2</sup>	0.005	0.32	0.40

*Note:* Table presents results of an OLS regression of network breadth per insurer, service, and market, on the average profit of enrollees who make a claim for the service. The profit per enrollee is calculated as ex-ante plus ex-post government transfers, plus revenues from copayments and coinsurance rates, minus total health care costs. Column (1) has no fixed effects, column (2) includes market and insurer fixed effects, and column (3) includes market, insurer, and service fixed effects. Robust standard errors in parenthesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Appendix Table 8: Robustness on regression specification: Network breadth and health care cost per service during 2011

	$1\{\text{service cost}^{2011} > 0\}$	$\log(\text{service cost}^{2011})$
<u>Panel A: Stayers</u>		
$H_{jmk}$	0.14*** (0.01)	0.09*** (0.01)
$N$	622,256	
$R^2$	0.31	
<u>Panel B: New enrollees</u>		
$H_{jmk}$	0.26*** (0.02)	0.16*** (0.02)
$N$	226,400	
$R^2$	0.27	
<u>Panel C: Full sample</u>		
$H_{jmk}$	0.18*** (0.01)	0.08*** (0.01)
$N$	14,831,006	
$R^2$	0.26	

*Note:* Table reports coefficients from a two-part model of health care cost per service during 2011. The first stage is a logistic regression for the probability of having non-zero cost. The second-stage is a log-linear regression of cost per service conditional on having non-zero cost. The independent variable is the 2011 network breadth per service. Panel A conditions on the sample of stayers. Panel B conditions on the sample of new enrollees. Panel C uses the full sample without conditioning enrollment to be continuous. In the last panel, costs are annualized by dividing by the number of days enrolled and multiplying by 360. All the models control for consumer demographics and diagnoses, and include market, service, and insurer fixed effects. Robust standard errors in parenthesis. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

Appendix Table 9: Robustness on regression specification: Selection on baseline costs

	$1\{\text{service cost}^{2010} > 0\}$	$\log(\text{service cost}^{2010})$
$H_{j'mk}^{2010} - H_{j'mk}^{2011}$	0.34*** (0.02)	0.24*** (0.02)
Switch	-0.36*** (0.08)	-0.14** (0.06)
Switch $\times (H_{j'mk}^{2010} - H_{j'mk}^{2011})$	-1.74** (0.79)	0.56 (0.58)
$N$	14,457,009	
$R^2$	0.51	

*Note:* Specification uses a random sample of 250,000 current enrollees. Table presents results of a two-part model of total costs per service. Independent variables include a switching indicator and the difference between network breadth in 2010 and network breadth in 2011 for the insurer chosen in 2011 ( $j'$ ). The first stage is a logistic regression for the probability of having non-zero cost. The second stage is a log-linear regression of total costs per service conditional on non-zero costs. Model includes demographics and diagnoses indicators, as well as insurer, service, and market fixed effects. Robust standard errors in parenthesis. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

Appendix Table 10: Selection on current enrollees' baseline costs

	$\log(\text{total cost}_{ijmt}^{2010} + 1)$ (1)	any claim $_{ijmt}^{2010}$ (2)
$(H_{jmk}^{2010} - H_{j'mk}^{2011})$	0.004*	-0.0001
	0.002	0.0002
Switch	-0.09***	-0.007***
	0.02	0.001
Switch $\times (H_{jmk}^{2010} - H_{j'mk}^{2011})$	-0.08	-0.01
	0.06	0.006
Demog+Diag	Y	Y
Market	Y	Y
Service	Y	Y
Insurer	Y	Y
$N$	14,457,009	14,457,009
$R^2$	0.50	0.51

*Note:* Specifications use a random sample of 250,000 current enrollees.  $j$  denotes the choice of insurer in 2010.  $j'$  denotes the choice of insurer in 2011. Column (1) presents results of an OLS regression of the logarithm of 2010 total service-specific costs on a switching indicator and the difference between network breadth in 2010 for the insurer chosen in 2010 ( $j$ ) and network breadth in 2011 for the insurer chosen in 2011 ( $j'$ ). 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. Robust standard errors in parenthesis. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

Appendix Table 11: Selection on baseline costs conditional on switchers

	$\log(\text{total cost}_{ijmt}^{2010} + 1)$ (1)	any claim $_{ijmt}^{2010}$ (2)
$(H_{j'mk}^{2010} - H_{j'mk}^{2011})$	-0.09	-0.01
	0.08	0.008
Demog+Diag	Y	Y
Market	Y	Y
Service	Y	Y
Insurer	Y	Y
$N$	8,870	8,870
$R^2$	0.51	0.51

*Note:* Results in columns (1) and (2) condition on the sample of switchers.  $j'$  denotes the choice of insurer in 2011. Column (1) presents results of an OLS regression of the logarithm of 2010 total service-specific costs on the difference between network breadth in 2010 and network breadth in 2011, for the insurer chosen in 2011 ( $j'$ ). 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. Robust standard errors in parenthesis. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

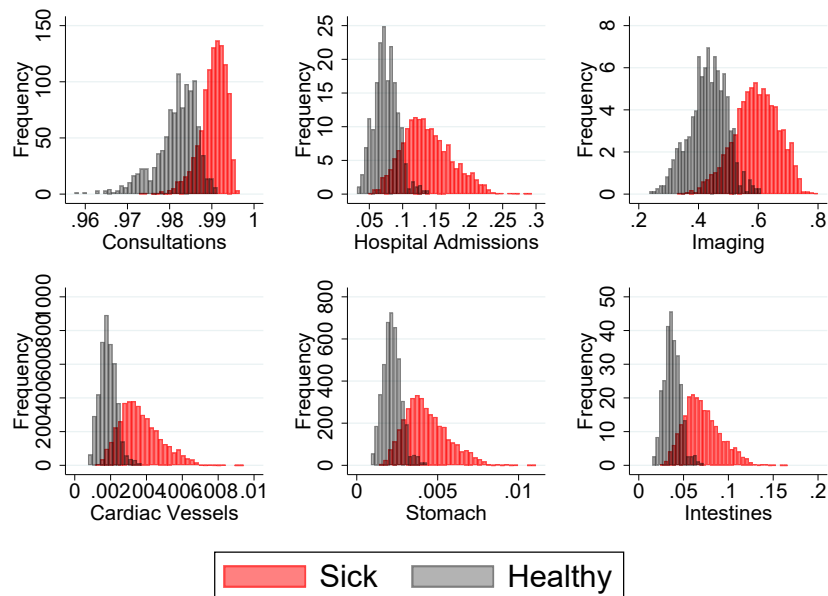
Appendix Table 12: New enrollees' risk scores and network breadth levels

	$\log(\text{risk transfer}^{2011})$
$H_{jk}^{2011}$	0.003 0.005
Demog+Diag	—
Market	Y
Service	—
Insurer	Y
$N$	2,653,415
$R^2$	0.06

*Note:* Table presents results of an OLS regression of the logarithm of new enrollees' risk-adjusted transfers on the insurer's 2011 total network breadth level. Specification includes market and insurer fixed effects. Robust standard errors in parenthesis. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

## Appendix 6 Additional demand results

### 6.1 Service claim probability



Appendix Figure 5: Distribution of service claim probability

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

## 6.2 Demand with expectation over diagnoses

I estimate an alternative demand model where consumers take expectations over future diagnoses when choosing their carrier. The utility function is given below:

$$u_{ijk} = \beta_i^D \sum_{l \in L} \gamma_{\theta l} \sum_m \gamma_{\theta l m k} H_{j m k} - \alpha_i \sum_{l \in L} \gamma_{\theta l} c_{\theta l y j k} + \delta_j + \varepsilon_{ijk}$$

where  $\gamma_{\theta l}$  is the probability that consumer of sex and age  $\theta$  has diagnosis  $l$ , calculated non-parametrically from the data.

Appendix Table 13: Insurer demand with expectation over diagnoses

Variable		Coefficient	Std. Error
Network		2.33***	0.01
OOP spending	(million COP)	-3.61***	0.19
Interactions			
Network	Demographics		
	Male	0.34***	0.01
	Age	-0.02***	0.00
	Location		
	Normal	0.05***	0.01
	Special	0.72***	0.03
	Urban	(ref)	(ref)
OOP spending	Demographics		
	Male	-0.87***	0.09
	Age	-0.01**	0.00
	Location		
	Normal	4.63***	0.10
	Special	2.12***	0.40
	Urban	(ref)	(ref)
<i>N</i>		5,800,610	
<i>N</i> enrollees		500,000	
Pseudo- $R^2$		0.17	

*Note:* This table presents results of an insurer choice model with uncertainty over diagnosis. Includes insurer fixed effects. Robust standard errors reported. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

## 6.3 Demand with random coefficients

The model with unobserved consumer heterogeneity that is common across alternatives is as follows:

$$u_{ijk} = \beta_i^D \sum_m \gamma_{\theta l m k} H_{j m k} - \alpha_i c_{\theta l y j k} + \phi_j + \varepsilon_{ijk}$$

with

$$\begin{aligned} \beta_i^D &= x_i' \beta^D + \sigma_\beta \nu_i \\ \alpha_i &= x_i' \alpha + \sigma_\alpha \nu_i \\ \nu_i &\sim N(0, 1) \end{aligned}$$

Appendix Table 14: Insurer demand with random coefficients

Variable		Coefficient	Std. Error
Network		2.37***	0.01
$\sigma$ Network		0.00	0.02
OOP spending		-6.50***	0.16
$\sigma$ OOP spending		1.03***	0.17
Interactions			
Network	Demographics		
	Male	0.30***	0.01
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	-0.34***	0.02
	Cardiovascular	-0.33***	0.02
	Diabetes	-0.44***	0.05
	Renal	-0.58***	0.08
	Other	-0.52***	0.02
	$\geq 2$ diseases	-0.62***	0.02
	Healthy	(ref)	(ref)
	Location		
	Normal	0.05***	0.01
	Special	0.74***	0.03
	Urban	(ref)	(ref)
	OOP spending		
	Demographics		
	Male	0.04	0.09
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	4.98***	0.19
	Cardiovascular	5.51***	0.15
	Diabetes	5.65***	0.40
	Renal	6.35***	0.17
	Other	5.20***	0.18
	$\geq 2$ diseases	5.78***	0.14
	Healthy	(ref)	(ref)
	Location		
	Normal	1.34***	0.10
	Special	0.52*	0.30
	Urban	(ref)	(ref)
$N$		5,800,610	
$N$ enrollees		500,000	

*Note:* This table presents results of the insurer choice model with random coefficients on network breadth and average out-of-pocket payments denoted by  $\sigma$  and distributed standard normal. Includes insurer fixed effects. Robust standard errors reported. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.



## 6.4 Control function approach

To account for potential endogeneity of average out-of-pocket costs, I estimate demand using a control function approach. The instrument is the government's reference price per service created in 2005. In the first step I estimate the following linear regression:

$$c_{\theta lyjk} = \beta_0 + \beta_1(1 - r_y) \sum_m \gamma_{\theta lmk} A_m + \lambda_\theta + \lambda_l + \delta_j + \eta_k + \mu_{\theta lyjk}$$

where  $r_y$  is the coinsurance rate of an individual with income level  $y$  and  $A_m$  is the government's reference price for service  $m$ . From this regression I obtain  $\hat{\mu}_{\theta lyjk}$  and standardize it at the market level, by subtracting the mean and dividing by the standard deviation. Then estimate the following utility function using a conditional logit:

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

where  $\mu_{\theta lyjk}^z$  is the standardized control function and  $x_i$  is a vector of demographics including sex, age, and diagnosis indicators. First stage results are presented in appendix table 15 and demand estimation is presented in table 16

Appendix Table 15: Control function for out-of-pocket costs

	OOP spending	
	Coef	SE
$(1 - r_y) \sum_m \gamma_{\theta lmk} A_m$	3.87***	0.01
Male	0.02***	0.00
Age<1	(ref)	(ref)
Age 1-4	-0.04***	0.01
Age 5-14	-0.03***	0.01
Age 15-18	-0.03***	0.01
Age 19-44	-0.05***	0.01
Age 45-49	-0.04***	0.01
Age 50-54	-0.04***	0.01
Age 55-59	-0.03***	0.01
Age 60-64	-0.03***	0.01
Age 65-69	-0.02***	0.01
Age 70-74	-0.03***	0.01
Age>=75	-0.04***	0.01
Cancer	(ref)	(ref)
Cardio	0.01***	0.00
Diabetes	0.02***	0.00
Renal	0.23***	0.01
Other	0.02***	0.00
>=2 diagnoses	0.13***	0.00
Healthy	-0.02***	0.00
$N$	5,800,610	
$R^2$	0.20	

*Note:* First stage regression of average out-of-pocket costs on service reference prices. Includes insurer fixed effects. Robust standard errors in parenthesis. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Appendix Table 16: Insurer demand with control function approach

Variable	Coefficient	Std. Error
Network	2.37***	0.01
OOP spending	-7.89***	0.55
$\hat{\mu}_{\theta lyjk}^z$	0.05	0.04
Interactions		
Network		
Demographics		
Male	0.30***	0.01
Age	-0.01***	0.00
Diagnoses		
Cancer	-0.33***	0.02
Cardiovascular	-0.33***	0.01
Diabetes	-0.44***	0.04
Renal	-0.58***	0.08
Other	-0.53***	0.02
>=2 diseases	-0.62***	0.02
Healthy	(ref)	(ref)
Location		
Normal	0.05***	0.01
Special	0.74***	0.04
Urban	(ref)	(ref)
OOP spending		
Demographics		
Male	1.54***	0.24
Age	-0.03***	0.01
Diagnoses		
Cancer	3.80***	0.51
Cardiovascular	3.42***	0.48
Diabetes	2.23**	0.94
Renal	3.64***	0.67
Other	3.12***	0.50
>=2 diseases	3.08***	0.47
Healthy	(ref)	(ref)
Location		
Normal	5.07***	0.42
Special	-2.52	2.06
Urban	(ref)	(ref)
$\hat{\mu}_{\theta lyjk}^z$		
Demographics		
Male	-0.12***	0.02
Age	0.00***	0.00
Diagnoses		
Cancer	0.10**	0.04
Cardiovascular	0.19***	0.03
Diabetes	0.27***	0.07
Renal	0.18***	0.04
Other	0.18***	0.04
>=2 diseases	0.20***	0.03
Healthy	(ref)	(ref)
Location		
Normal	-0.25***	0.02
Special	0.25**	0.12
Urban	(ref)	(ref)
N	5,800,610	
N enrollees	500,000	
Pseudo- $R^2$	0.17	

Note: This table presents results of the insurer choice model estimated with a control function approach to instrument for average out-of-pocket costs. Includes insurer fixed effects. Robust standard errors reported. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

## 6.5 Other robustness checks

Appendix Table 17: Insurer demand in small markets

Variable		Coefficient	Std. Error
Network		1.14***	0.02
OOP spending	(million COP)	-1.65***	0.26
Interactions			
Network	Demographics		
	Male	0.25***	0.01
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	-0.03*	0.02
	Cardiovascular	-0.13***	0.01
	Diabetes	-0.23***	0.03
	Renal	-0.41***	0.07
	Other	-0.26***	0.02
	>=2 diseases	-0.26***	0.01
	Healthy	(ref)	(ref)
	Location		
	Normal	0.37***	0.02
	Special	0.79***	0.03
	Urban	(ref)	(ref)
OOP spending	Demographics		
	Male	0.02	0.06
	Age	0.00***	0.00
	Diagnoses		
	Cancer	1.76***	0.17
	Cardiovascular	2.23***	0.11
	Diabetes	1.55***	0.29
	Renal	2.38***	0.14
	Other	2.18***	0.12
	>=2 diseases	2.12***	0.11
	Healthy	(ref)	(ref)
	Location		
	Normal	-0.40*	0.24
	Special	-0.43	0.26
	Urban	(ref)	(ref)
$N$		4,391,032	
$N$ enrollees		500,000	
Pseudo- $R^2$		0.19	

*Note:* This table presents results of the insurer choice model estimated on the subsample of markets that excludes the four main capital states: Antioquia, Atlántico, Bogotá, and Valle del Cauca. Includes insurer fixed effects. Robust standard errors reported.

\*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

Appendix Table 18: Insurer demand with additional quality measures

Variable		Coefficient	Std. Error
Network		2.32***	0.01
OOP spending	(million COP)	-5.84***	0.22
Avg. quality		0.003***	0.00
Avg. wait time		-0.01***	0.00
Interactions			
Network	Demographics		
	Male	0.29***	0.01
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	-0.33***	0.02
	Cardiovascular	-0.33***	0.01
	Diabetes	-0.46***	0.04
	Renal	-0.60***	0.08
	Other	-0.52***	0.02
	>=2 diseases	-0.65***	0.02
	Healthy	(ref)	(ref)
	Location		
	Normal	0.19***	0.01
	Special	0.66***	0.03
	Urban	(ref)	(ref)
	Demographics		
	Male	0.11	0.10
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	4.57***	0.24
	Cardiovascular	5.26***	0.19
	Diabetes	5.34***	0.32
	Renal	5.71***	0.21
	Other	4.80***	0.22
	>=2 diseases	5.24***	0.19
	Healthy	(ref)	(ref)
	Location		
	Normal	1.08***	0.11
	Special	0.67	0.45
	Urban	(ref)	(ref)
OOP spending			
Demographics			
Male		0.11	0.10
Age		-0.01***	0.00
Diagnoses			
Cancer		4.57***	0.24
Cardiovascular		5.26***	0.19
Diabetes		5.34***	0.32
Renal		5.71***	0.21
Other		4.80***	0.22
>=2 diseases		5.24***	0.19
Healthy		(ref)	(ref)
Location			
Normal		1.08***	0.11
Special		0.67	0.45
Urban		(ref)	(ref)
$N$		5,358,649	
$N$ enrollees		489,442	
Pseudo- $R^2$		0.17	

*Note:* This table presents results of the insurer choice model including additional insurer quality measures: the average quality from a likert scale and average wait time for an appointment with the primary care doctor or specialist. Both variables are measured at the insurer-market level and obtained from enrollee-level survey data conducted by the Colombian Ministry of Health during 2013 to 2016. Includes insurer fixed effects. Robust standard errors reported. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Appendix Table 19: Insurer demand with star hospital coverage indicator

Variable		Coefficient	Std. Error
Network		2.08***	0.01
OOP spending		-6.84***	0.24
Star hospital		1.27***	0.02
Interactions			
Network	Demographics		
	Male	0.29***	0.01
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	-0.30***	0.02
	Cardiovascular	-0.32***	0.02
	Diabetes	-0.45***	0.04
	Renal	-0.56***	0.08
	Other	-0.51***	0.02
	>=2 diseases	-0.64***	0.02
	Healthy	(ref)	(ref)
	Location		
	Normal	0.14***	0.01
	Special	0.93***	0.04
	Urban	(ref)	(ref)
	Demographics		
	Male	0.24**	0.11
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	5.29***	0.23
	Cardiovascular	5.88***	0.20
	Diabetes	5.66***	0.36
	Renal	6.38***	0.23
	Other	5.42***	0.22
	>=2 diseases	5.79***	0.19
	Healthy	(ref)	(ref)
	Location		
	Normal	1.43***	0.13
	Special	1.08***	0.42
	Urban	(ref)	(ref)
OOP spending			
Demographics			
Male		0.24**	0.11
Age		-0.01***	0.00
Diagnoses			
Cancer		5.29***	0.23
Cardiovascular		5.88***	0.20
Diabetes		5.66***	0.36
Renal		6.38***	0.23
Other		5.42***	0.22
>=2 diseases		5.79***	0.19
Healthy		(ref)	(ref)
Location			
Normal		1.43***	0.13
Special		1.08***	0.42
Urban		(ref)	(ref)
$N$		5,800,610	
$N$ enrollees		500,000	
Pseudo- $R^2$		0.17	

*Note:* This table presents results of the insurer choice model including a measure of star hospital coverage equal to  $\sum_m \gamma_{lmk} Star_{jmk}$ , where  $Star_{jk}$  is an indicator for insurer  $j$  covering a star hospital in market  $k$  for service  $m$ . Includes insurer fixed effects. Robust standard errors reported. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Appendix Table 20: Insurer demand with alternative network measures

		(1) Largest hospitals		(2) Only hospitals		(3) All providers	
Variable		Coef.	SE	Coef.	SE	Coef.	SE
Network		2.32***	0.01	2.37***	0.01	2.54***	0.01
OOP spending	(million COP)	-5.33***	0.21	-6.64***	0.21	-4.90***	0.16
Interactions							
Network	Demographics						
	Male	0.31***	0.01	0.30***	0.01	0.14***	0.01
	Age	-0.01***	0.00	-0.01***	0.00	0.00***	0.00
	Diagnoses						
	Cancer	-0.29***	0.03	-0.34***	0.02	-0.71***	0.02
	Cardiovascular	-0.39***	0.02	-0.33***	0.01	-0.50***	0.02
	Diabetes	-0.46***	0.05	-0.44***	0.04	-0.65***	0.06
	Renal	-0.63***	0.09	-0.61***	0.08	-0.64***	0.10
	2 diseases	-0.56***	0.02	-0.53***	0.02	-0.69***	0.03
	>=3 diseases	-0.71***	0.02	-0.64***	0.02	-0.93***	0.02
	Healthy	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
	Location						
	Normal	-0.01	0.01	0.05***	0.01	-0.25***	0.01
	Special	0.81***	0.05	0.73***	0.04	-0.59***	0.02
	Urban	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
OOP spending	Demographics						
	Male	0.33*	0.11	0.05	0.09	-0.02	0.06
	Age	-0.01***	0.00	-0.01***	0.00	-0.01***	0.00
	Diagnoses						
	Cancer	4.11***	0.23	5.36***	0.22	4.41***	0.15
	Cardiovascular	4.67***	0.18	5.89***	0.18	4.57***	0.15
	Diabetes	4.41***	0.37	5.77***	0.31	4.81***	0.25
	Renal	5.11***	0.19	6.27***	0.22	4.82***	0.17
	2 diseases	4.19***	0.21	5.56***	0.20	4.32***	0.15
	>=3 diseases	4.56***	0.18	5.89***	0.18	4.61***	0.15
	Healthy	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
	Location						
	Normal	1.12***	0.11	1.15***	0.11	0.74***	0.08
	Special	1.15***	0.28	0.67	0.43	0.92***	0.16
	Urban	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
N		5,800,610		5,800,610		5,800,610	
N enrollees		500,000		500,000		500,000	
Pseudo-R2		0.17		0.17		0.12	

*Note:* This table presents results of the insurer choice model under alternative specifications of network breadth. Column (1) reports coefficients and standard errors of a model where network breadth is constructed from a sample of the largest hospitals in each market. Large hospitals are defined as having number of beds above the 70th percentile of the distribution of beds in each market. There are 314 hospitals under this definition. Column (2) shows coefficients and standard errors of a model where network breadth is constructed from the sample of providers that report provision of inpatient services through the Ministry's Registry. There are 1,444 providers under this definition. Column (2) presents coefficients and standard errors of a model where network breadth is constructed from the sample of all institutional providers of which there are 16,609. All models include insurer fixed effects. Robust standard errors reported. \*\*\*p<0.01, \*\*p<0.05, \*p<0.1.

Appendix Table 21: Insurer demand on adults aged 19 or older

Variable		Coef.	SE
Network		2.28***	0.01
OOP spending (million COP)		-5.41***	0.19
Interactions			
Network	Demographics		
	Male	0.30***	0.01
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	-0.31***	0.02
	Cardiovascular	-0.34***	0.01
	Diabetes	-0.36***	0.03
	Renal	-0.63***	0.07
	Other	-0.48***	0.02
	>=2 diseases	-0.62***	0.01
	Healthy	(ref)	(ref)
	Location		
	Normal	0.01	0.01
	Special	0.54***	0.03
	Urban	(ref)	(ref)
OOP spending	Demographics		
	Male	-0.08	0.07
	Age	-0.01***	0.00
	Diagnoses		
	Cancer	3.44***	0.22
	Cardiovascular	4.46***	0.17
	Diabetes	4.47***	0.29
	Renal	5.03***	0.18
	Other	4.30***	0.16
	>=2 diseases	4.43***	0.14
	Healthy	(ref)	(ref)
	Location		
	Normal	1.36***	0.11
	Special	1.40***	0.16
	Urban	(ref)	(ref)
N		5,849,583	
N enrollees		500,000	
Pseudo- $R^2$		0.17	

*Note:* This table presents results of the insurer choice model estimated on a sample of adults aged > 18. All models include insurer fixed effects. Robust standard errors reported. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

## Appendix 7 Note on primitives of average cost function

Suppose there is a second stage in the demand model where consumers choose a hospital to receive service  $m$ . The cost of consumer  $i$  enrolled to insurer  $j$  can be written as:

$$c_{ij}(H_j) = \sum_m \gamma_{\theta(i)l(i)m} \sum_{h \in H_{jm}} p_{jhm} s_{ihm}(H_{jm})$$

where  $p_{jhm}$  is the negotiated price for service  $m$  between insurer  $j$  and hospital  $h$ ,  $s_{ihm}$  is the probability that consumer  $i$  chooses hospital  $h$  for service  $m$ , and  $\gamma$  is the probability that consumer  $i$  of type  $(\theta, l)$  makes a claim for service  $m$ .

With this specification of individual costs, the insurer profit function is:

$$\pi_j = \sum_i (R_{ij} - c_{ij}(H_j)) s_{ij}(H_j)$$

where  $s_{ij}$  is the probability that consumer  $i$  enrolls insurer  $j$ . We can rewrite the previous equation as:

$$\pi_j = (R_j - AC_j(H_j)) D_j$$

where  $R_j = \sum_i R_{ij}$ ,  $D_j = \sum_i s_{ij}$ , and

$$AC_j(H_j) = \frac{1}{D_j} \sum_i c_{ij}(H_j) s_{ij}(H_j)$$

Suppose  $s_{ihm}$  and  $s_{ij}$  are obtained from a discrete choice model with preference shocks that are distributed T1EV. Also, for simplicity, assume there are two hospitals and two insurers. Then, the average cost is:

$$AC_j(H_j) = \sum_i \left( \sum_m \gamma_{\theta(i)l(i)m} \underbrace{\sum_{h \in H_{jm}} p_{jhm} \frac{\exp(\delta_{ihm})}{1 + \exp(\delta_{ihm})} \left( \frac{1 + \exp(\eta(H_j))}{\exp(\eta(H_j))} \right)}_{A_m} \right) \frac{\exp(\eta_i(H_j))}{1 + \exp(\eta_i(H_j))}$$

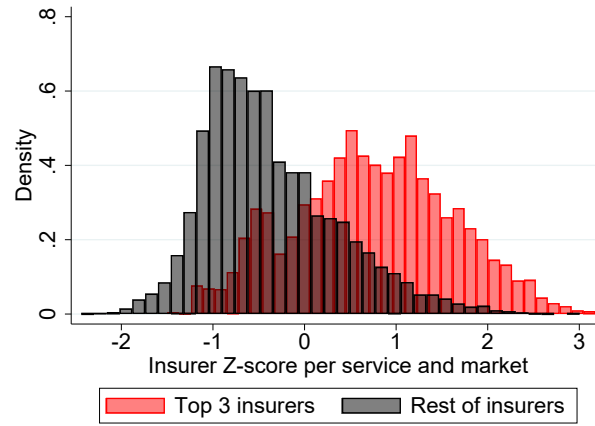
where  $\eta(H_j)$  is the average utility of choosing insurer  $j$  and  $\delta_{ihm}$  is the average utility of choosing hospital  $h$  for service  $m$ . For one consumer, taking logs of the equation above yields:

$$\log(AC_{ij}(H_j)) = \log \left( \sum_m \gamma_{\theta(i)l(i)m} A_m \right) + \underbrace{\eta_i(H_j)}_{\sum_m \gamma_{\theta(i)l(i)m} H_{jm}} - \log(1 + \exp(\eta_i(H_j)))$$

which shows that my average cost function can be obtained from a more involved model of hospital choice under certain conditions.

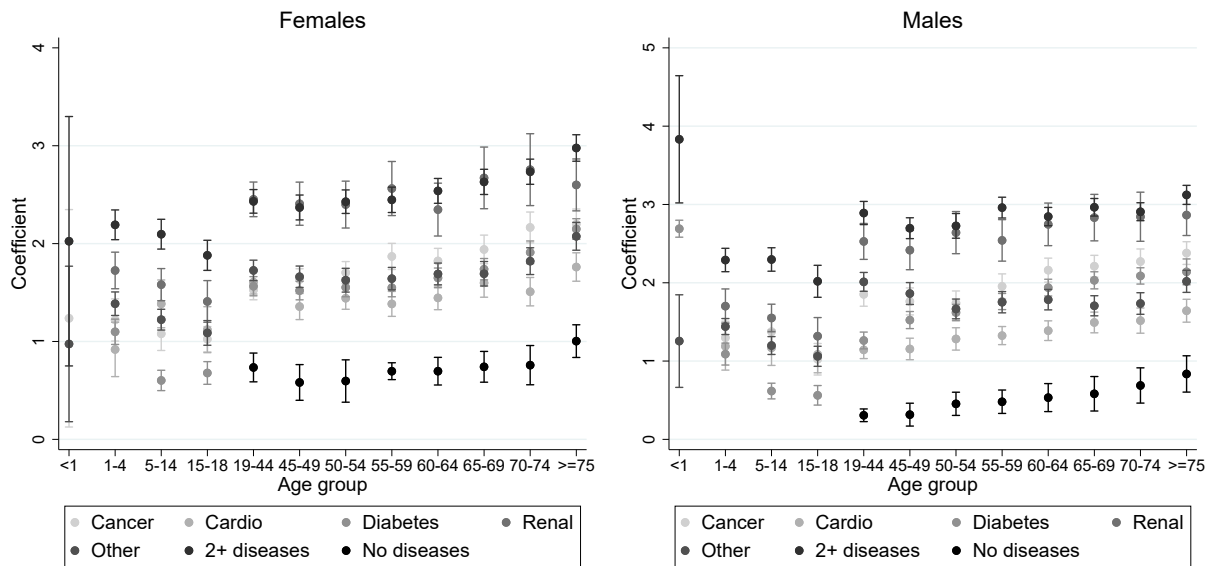


## Appendix 8 Additional average cost results



Appendix Figure 6: Standardized network breadth per service and market

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



Appendix Figure 7: Consumer type fixed effects

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

Appendix Table 22: Average cost over more granular consumer types

Variable	Coefficient	Std. Error
Network	0.22***	0.01
Scope economies	-4.24***	0.35
Avg. ref. price	5.94***	0.03
<u>Insurer</u>		
EPS001	-0.11***	0.01
EPS002	-0.28***	0.01
EPS003	0.01	0.01
EPS005	0.30***	0.01
EPS008	0.16***	0.01
EPS009	0.15***	0.01
EPS010	-0.01	0.01
EPS012	-0.01	0.01
EPS013	0.15***	0.01
EPS016	0.04***	0.01
EPS017	-0.14***	0.01
EPS018	0.11***	0.01
EPS023	-0.05***	0.01
EPS037	(ref)	(ref)
$N$	653,902	
$R^2$	0.25	

*Note:* OLS regression of logarithm of average costs per insurer, market, and consumer type on network breadth, economies of scope, and service reference price. Consumer type is defined as combinations of sex, age, and diagnosis from a list of 30 exhaustive and mutually exclusive diagnoses in appendix table 31. Includes insurer, market, and consumer type fixed effects. Robust standard errors reported.

Appendix Table 23: Predicted change in females' average cost by type of carrier and diagnosis

Service	F, 19-44, Healthy		F, 19-44, Cancer	
	Narrow	Broad	Narrow	Broad
Cardiac vessels	455	1,588	787	6,784
Stomach	457	1,592	799	6,802
Intestines	659	1,784	1,638	7,559
Imaging	3,399	4,071	9,841	14,105
Consultations	8,236	6,893	17,764	18,175
Laboratory	5,413	5,847	14,711	17,086
Nuclear medicine	1,296	2,724	3,189	8,816
Hospital admissions	1,755	3,035	4,949	9,967

*Note:* This table shows the average change in the average cost of a healthy female aged 19-44 and a female aged 19-44 with cancer following a 10% increase in network breadth for the service in the row, separately for broad and narrow network carriers. Broad network carriers are defined as insurers with average network breadth across all other services above 70% and narrow network carriers as the complement. Units are in Colombian pesos.

Appendix Table 24: Predicted change in males' average cost by type of carrier and diagnosis

Service	M, 19-44, Healthy		M, 19-44, Diabetes	
	Narrow	Broad	Narrow	Broad
Cardiac vessels	210	757	772	3,365
Stomach	211	759	779	3,374
Intestines	317	860	1,180	3,782
Imaging	1,902	2,213	6,302	8,209
Consultations	5,305	4,252	13,752	12,302
Laboratory	3,248	3,639	9,951	11,182
Nuclear medicine	804	1,733	2,331	5,312
Hospital admissions	1,047	1,904	3,277	5,969

*Note:* This table shows the average change in the average cost of a healthy male aged 19-44 and a male aged 19-44 with diabetes following a 10% increase in network breadth for the service in the row, separately for broad and narrow network carriers. Broad network carriers are defined as insurers with average network breadth across all other services above 70% and narrow network carriers as the complement. Units are in Colombian pesos.

## Appendix 9 Dropout and transition probabilities

To estimate the marginal cost of network formation in the third step of my model, I first need to compute the probability that consumer type  $(\theta, l)$  drops out of the contributory system and the probability that consumer type  $(\theta, l)$  in period  $t$  transitions into diagnosis  $l'$  in period  $t + 1$ . Both of these probabilities weight future per-enrollee profits in the insurer's total profit function.

I use the data from *all* enrollees to the contributory system in 2010 and 2011, regardless of their enrollment spell length, to compute dropout probabilities. For each consumer type  $(\theta, l)$ , I calculate the probability that she drops out of the system non-parametrically as the number of individuals of type  $(\theta, l)$  observed only in 2010 but not 2011, divided by the total number of type  $(\theta, l)$  individuals in 2010. Table 25 presents the mean and standard deviation of the dropout probability overall, and conditional on health status, sex, and age. Healthy individuals are on average 10 percentage points more likely to dropout of the system compared to sick patients, and consumers aged less than 44 are on average 3.8 percentage points more likely to dropout compared to individuals aged 45 or older. The table also shows that males are more likely to dropout of the contributory system relative to females, with the difference in means equal to 3.6 percentage points.

I use a non-parametric approach to compute transition probabilities as well, using data from continuously enrolled new *and* current enrollees in 2010 and 2011. Given that the transition from  $\theta$  to  $\theta'$  is deterministic, I only need to compute transition probabilities across diagnoses. The probability that type  $(\theta, l)$  transitions into  $(\theta', l')$  equals the number of type  $(\theta, l)$  in 2010 that end up with diagnosis  $l'$  in 2011, divided by the number of type  $(\theta, l)$  individuals in 2010. Table 26 presents the mean and standard deviation in parenthesis of transition probabilities from having cancer, cardiovascular disease, diabetes, renal disease, other diseases, 2 or more diseases, and no diseases in period  $t$  to having each of these 7 diagnoses in period  $t + 1$ . Because my list of diagnoses is mutually exclusive, the table shows that the probability of transitioning from a particular diagnosis in  $t$  to a single different diagnosis in  $t + 1$  is zero, but the probability for the transition into the same diagnosis or to an added diagnosis ( $\geq 2$  diseases) in  $t + 1$  is non-zero. For patients without diseases in period  $t$ , remaining healthy in period  $t + 1$  has the higher likelihood, followed by receiving a diagnosis for cardiovascular disease and other diseases like long-term pulmonary disease. The fact that the diagnosis list is mutually exclusive simplifies the computation of future profits and future marginal variable profits per enrollee that are needed to recover the marginal cost of network formation. I move to the estimation of this marginal cost next.

Appendix Table 25: Summary statistics of dropout probability

	Mean	SD
Overall	0.111	0.114
Sick	0.094	0.110
Healthy	0.201	0.092
Age>44	0.094	0.100
Age≤44	0.135	0.129
Male	0.129	0.124
Female	0.093	0.101

Appendix Table 26: Summary statistics of transition probabilities across diagnoses

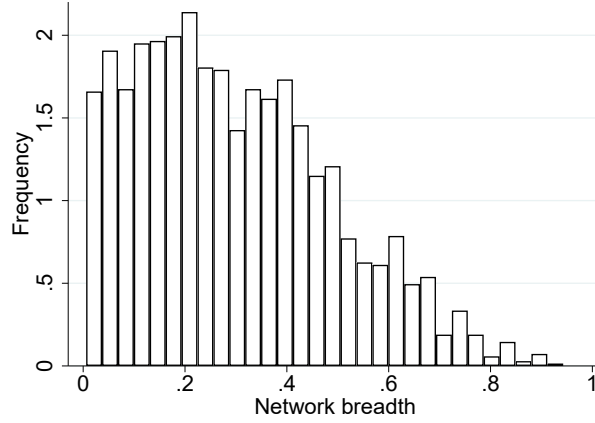
Diagnosis $t/t+1$	Cancer	Cardio	Diabetes	Renal	Other	$\geq 2$ disea.	No disea.
Cancer	0.789 (0.136)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.211 (0.136)	0.000 (0.000)
Cardio	0.000 (0.000)	0.774 (0.148)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.226 (0.148)	0.000 (0.000)
Diabetes	0.000 (0.000)	0.000 (0.000)	0.654 (0.209)	0.000 (0.000)	0.000 (0.000)	0.346 (0.209)	0.000 (0.000)
Renal	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.643 (0.167)	0.000 (0.000)	0.357 (0.167)	0.000 (0.000)
Other	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.711 (0.171)	0.289 (0.171)	0.000 (0.000)
$\geq 2$ diseases	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	1.000 (0.000)	0.000 (0.000)
No diseases	0.034 (0.038)	0.087 (0.086)	0.007 (0.015)	0.004 (0.016)	0.044 (0.028)	0.054 (0.084)	0.770 (0.148)

*Note:* This table shows the mean and standard deviation in parenthesis of transition probabilities from cancer, cardiovascular disease, diabetes, renal disease, other diseases, 2 or more diseases, and no diseases in period  $t$  to each of these 7 diagnoses in period  $t + 1$ . Summary statistics are calculated across sex-age combinations in each cell.

Appendix Table 27: Summary statistics of marginal variable profits per insurer

Insurer	MVP
EPS001	287 (1,202)
EPS002	859 (3,231)
EPS003	419 (1,636)
EPS005	248 (981)
EPS010	930 (3,203)
EPS013	757 (2,472)
EPS016	1,302 (4,258)
EPS017	663 (3,591)
EPS018	597 (2,312)
EPS037	1,138 (3,675)

*Note:* Mean and standard deviation in parenthesis of marginal variable profits in the left-hand side of equation (4). Measured in millions of Colombian pesos per service per market.



Appendix Figure 8: Distribution of network breadth for FOC

*Note:* This figure shows the distribution of network breadth in the fourlargest markets and top 10 insurance companies, where no one chooses a corner solution.

Appendix Table 28: First stage regression of network breadth

$H_{jmk}$	Coefficient	Std. Error
$H_{jmk}^{t-1}$	0.76***	0.01
$\bar{\gamma}_{female,m,k}$	33.94***	8.45
$\bar{\gamma}_{healthy,m,k}$	14.31***	4.20
$\bar{\gamma}_{age\ 19-44,m,k}$	-55.33***	13.63
$H_{jmk}^{t-1} \times \bar{\gamma}_{age\ 19-44,m,k}$	0.16***	0.05
<u>Insurer FEs</u>		
EPS001	-0.02***	0.01
EPS002	0.01*	0.01
EPS003	-0.04***	0.01
EPS005	0.02**	0.01
EPS010	-0.01	0.01
EPS013	0.00	0.01
EPS016	0.14***	0.01
EPS017	0.00	0.01
EPS018	-0.01	0.01
EPS037	(ref)	(ref)
<u>Market FEs</u>		
Market 05	(ref)	(ref)
Market 08	0.00	0.01
Market 11	0.00	0.00
Market 76	-0.02***	0.01
N	2,262	
F-stat	929.67	

*Note:* This table presents the first stage of the GMM estimation of equation (5).  $H_{jmk}^{t-1}$  is the network breadth in 2010.  $\bar{\gamma}_{i,m,k}$  is the average probability that a consumer with characteristic  $i$  makes a claim for service  $m$  in market  $k$ . The specification includes insurer, market, and service fixed effects. Robust standard errors and first-stage F-statistic reported. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Appendix Table 29: Predicted average total network formation cost per market

Insurer	(1) Total	(2) %
EPS001	1,381	14
EPS002	16,810	134
EPS003	6,917	85
EPS005	-2,576	-105
EPS010	11,721	91
EPS013	13,084	98
EPS016	23,851	83
EPS017	13,208	66
EPS018	10,545	90
EPS037	19,541	92

*Note:* Column (1) presents the predicted average total cost of network formation in millions of pesos across markets and column (2) presents this cost as a percentage of total variable profits.

Appendix Table 30: Decomposition of profit changes after network breadth increase by diagnosis

Service	Healthy		Renal		Other disease	
	$\% \Delta s_{ijk}$	$\% \Delta AC_{\theta ljk}$	$\% \Delta s_{ijk}$	$\% \Delta AC_{\theta ljk}$	$\% \Delta s_{ijk}$	$\% \Delta AC_{\theta ljk}$
Cardiac vessels	0.01	0.00	0.03	0.00	0.02	0.00
Stomach	0.01	0.00	0.03	0.00	0.03	0.00
Intestines	0.25	0.02	0.60	0.02	0.49	0.02
Imaging	3.98	0.29	6.08	0.31	5.62	0.33
Consultations	13.17	1.24	12.35	0.95	13.07	1.01
Laboratory	5.45	0.43	7.19	0.41	6.97	0.43
Nuclear medicine	0.07	0.01	0.18	0.01	0.15	0.01
Hospital admissions	0.67	0.05	1.56	0.07	1.27	0.06

*Note:* This table shows the average percentage change in demand ( $\% \Delta s_{ijk}$ ) and average costs per enrollee ( $\% \Delta AC_{\theta ljk}$ ) for healthy individuals, patients with renal disease, and patients with other chronic conditions, after a 10% unilateral increase in network breadth for the service in the row by insurer  $j$ , while holding its competitors' choices fixed.

## Appendix 10 Concavity of the profit function

The second partial derivative of the insurers' profit function with respect to network breadth, all else equal, is:

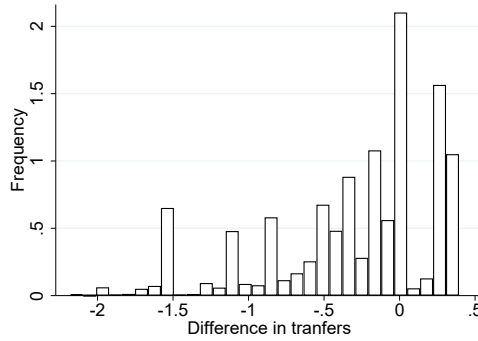
$$\begin{aligned} \frac{\partial^2 \Pi_{jk}}{\partial H_{jmk}^2} = & (R_{\theta k} - (1 - r_i)AC_{\theta ljk})(\beta_i^D \gamma_{\theta lmk})^2 s_{ijk}(1 - s_{ijk})(1 - 2s_{ijk}) \\ & - 2(1 - r_i)\beta_i^D \gamma_{\theta lmk} s_{ijk}(1 - s_{ijk})AC_{\theta ljk} \left( \beta_1^S \gamma_{\theta lmk} + \frac{1}{2M_k} \beta_2^S \sum_m \sum_{n \neq m} \gamma_{\theta lmk} \gamma_{\theta lnk} H_{jnk} \right) \\ & - (1 - r_i)s_{ijk}AC_{\theta ljk} \left( \beta_1^S \gamma_{\theta lmk} + \frac{1}{2M_k} \beta_2^S \sum_m \sum_{n \neq m} \gamma_{\theta lmk} \gamma_{\theta lnk} H_{jnk} \right)^2 - 2\omega \end{aligned}$$

Given the estimated  $\beta_1^S$  and  $\beta_2^S$ , we have that:

$$\begin{aligned} \left( \beta_1^S \gamma_{\theta lmk} + \frac{1}{2M_k} \beta_2^S \sum_m \sum_{n \neq m} \gamma_{\theta lmk} \gamma_{\theta lnk} H_{jnk} \right) & > 0 \quad \forall H_{jmk} \in (0, 1) \\ AC_{\theta ljk} & > 0 \quad \forall H_{jmk} \in (0, 1) \end{aligned}$$

thus, the second and third lines on the right-hand side are negative.

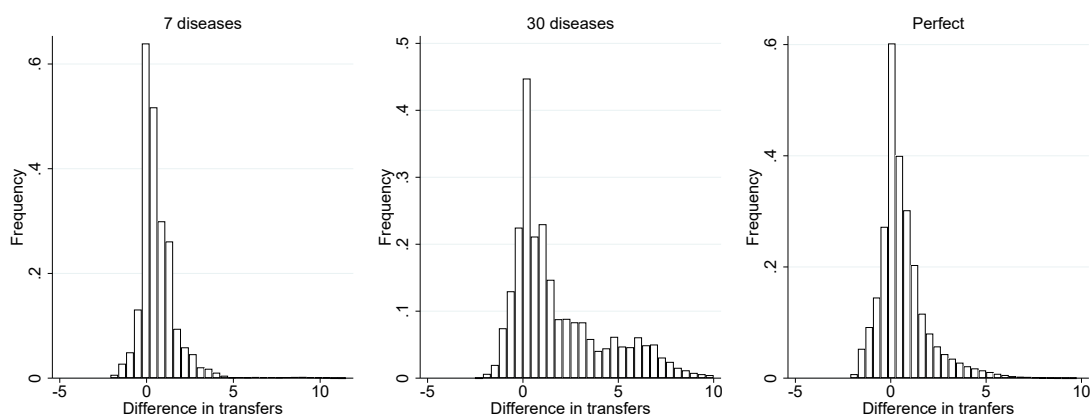
## Appendix 11 Additional counterfactual results



Appendix Figure 9: Distribution of counterfactual minus observed transfer under no risk adjustment

*Note:* This figure shows the distribution of the difference between the observed risk-adjusted transfer and the counterfactual under no risk adjustment across consumer types.





Appendix Figure 10: Distribution of counterfactual minus observed transfer under improved risk adjustment

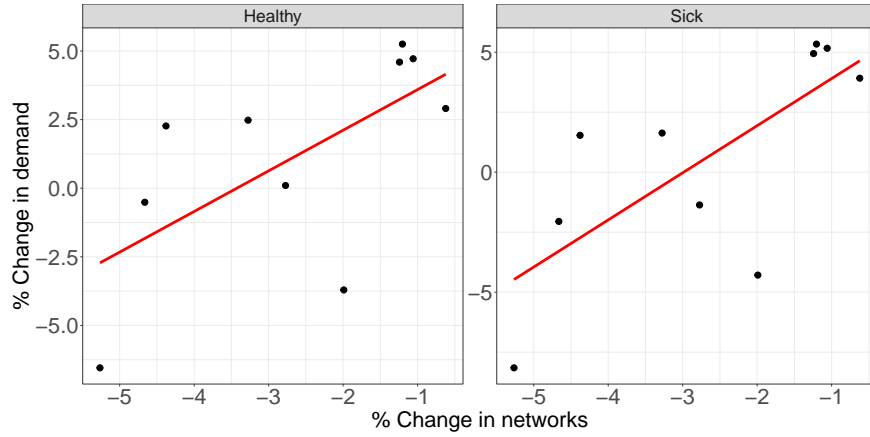
*Note:* This figure shows the distribution of the difference between the observed risk-adjusted transfer and the counterfactual transfer under a improved risk adjustment formula across consumer types. In the left panel, the counterfactual transfer compensates insurers for the list of 7 diseases included in the model in addition to sex, age group, and municipality. In the middle panel, the counterfactual transfer compensates insurers for a list of 30 exhaustive and mutually exclusive disease categories (presented in appendix table 31), in addition to sex, age group, and municipality. The right panel corresponds to “perfect” risk adjustment, where the transfer equals the individual’s average cost.

Appendix Table 31: Disease categories

---

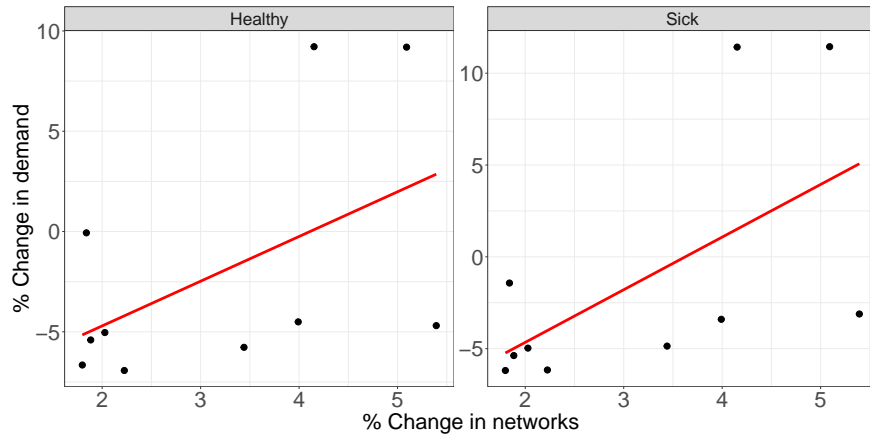
Arthritis
Arthrosis
Asthma
Autoimmune disease
Breast cancer
Cancer in digestive organs
Cancer in female genitalia
Cancer in male genitalia
Cancer therapy
Invasive cervical cancer
Local cervical cancer
Diabetes
Epilepsy
Genetic anomalies
HIV-AIDS
Hypertension
Cancer in respiratory organs
Lymphatic cancer
Melanoma or skin cancer
Other types of cancer
Other types of cardiovascular disease
Long-term pulmonary disease
Renal disease
Chronic kidney disease
End-stage renal disease
Long-term renal disease
Transplant
Tuberculosis
More than 2 diseases
No diseases

---



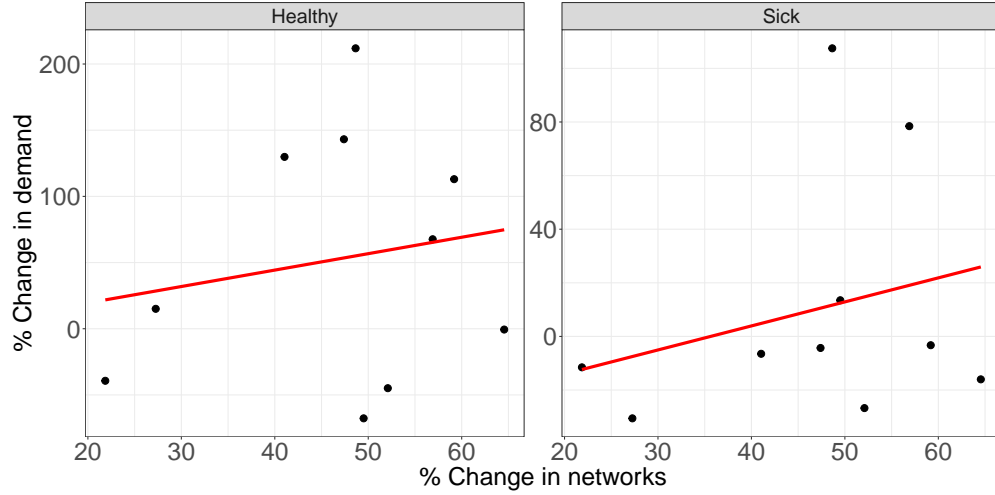
Appendix Figure 11: Correlation between network changes and changes in demand under no risk adjustment

*Note:* This figure shows the correlation between changes in network breadth and changes in demand at the insurer level in the counterfactual of no risk adjustment relative to the observed scenario. Every black dot is an insurer and the red line corresponds to a linear fit. The left panel focuses on changes in demand from healthy individuals and the right panel focuses on changes in demand from sick individuals.



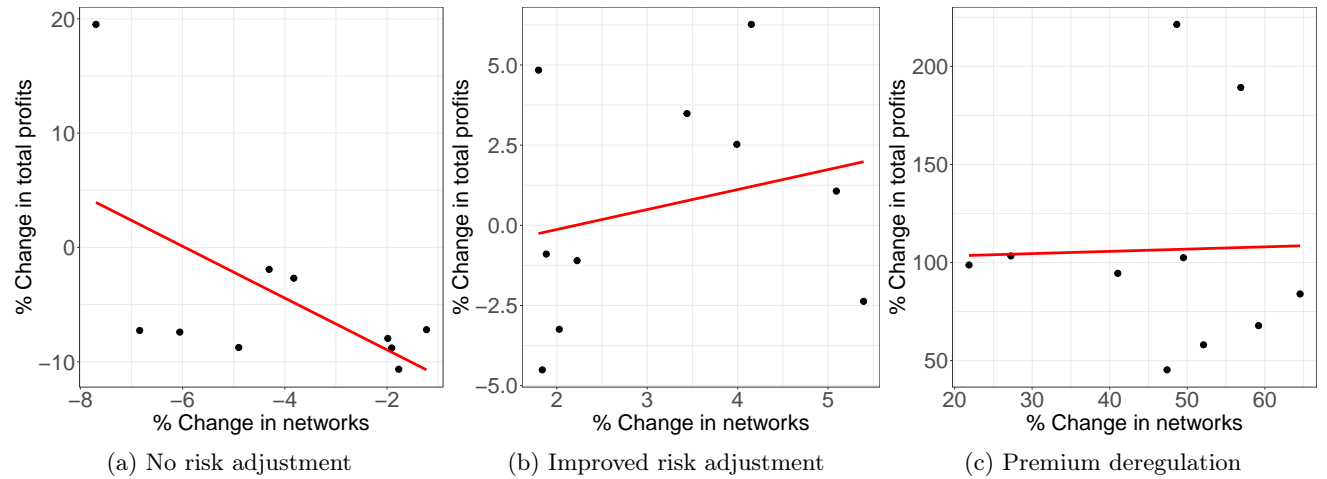
Appendix Figure 12: Correlation between network changes and changes in demand under improved adjustment

*Note:* This figure shows the correlation between changes in network breadth and changes in demand at the insurer level in the counterfactual of improved risk adjustment with compensations for 30 disease categories relative to the observed scenario. Every black dot is an insurer and the red line corresponds to a linear fit. The left panel focuses on changes in demand from healthy individuals and the right panel focuses on changes in demand from sick individuals.



Appendix Figure 13: Correlation between network changes and changes in demand under premium deregulation

*Note:* This figure shows the correlation between changes in network breadth and changes in demand at the insurer level in the counterfactual of premium setting relative to the observed scenario. Every black dot is an insurer and the red line corresponds to a linear fit. The left panel focuses on changes in demand from healthy individuals and the right panel focuses on changes in demand from sick individuals.



Appendix Figure 14: Correlation between profit and network changes in counterfactual

*Note:* This figure shows the correlation between changes in network breadth and changes in total profits at the insurer level in the counterfactual of no risk adjustment in panel (a), improved risk adjustment with 30 disease categories in panel (b), and premium setting in panel (c), all relative to the observed scenario. Every black dot is an insurer and the red line corresponds to a linear fit.