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

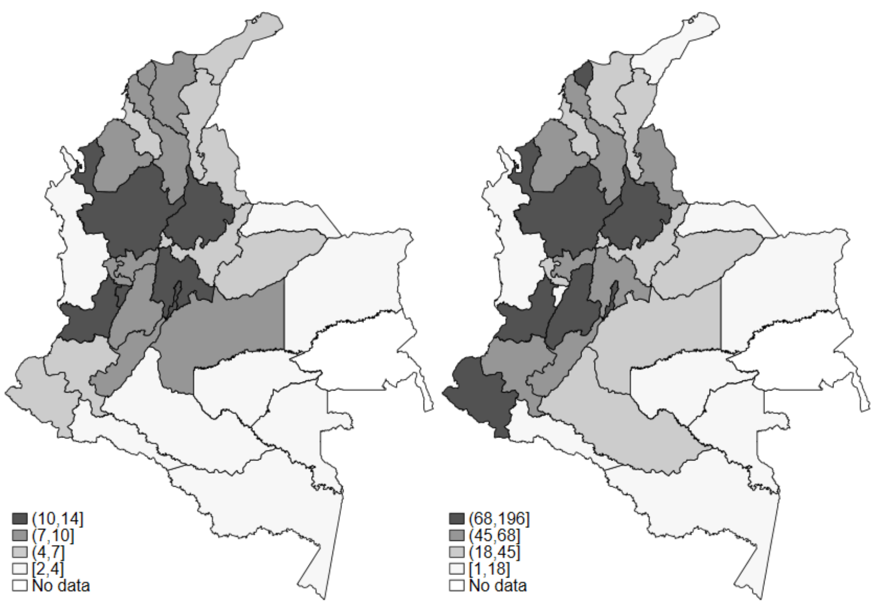
Appendix 1 Service categories

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 2 Description of Colombian health care market

Appendix Table 1: 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



(a) Insurers per market (b) Hospitals per market

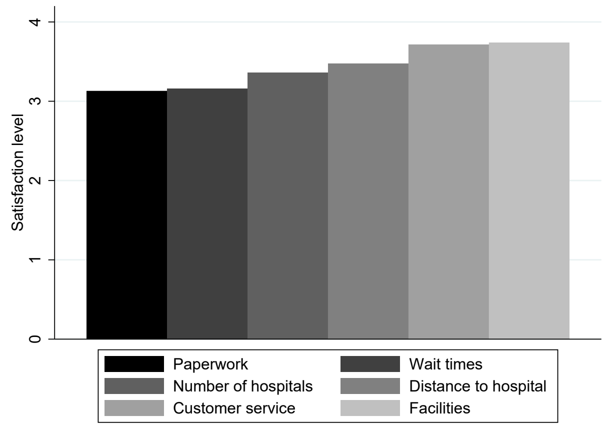
Appendix Figure 1: Number of players

Appendix Table 2: Distribution of percentage of covered services across departments in selected hospitals

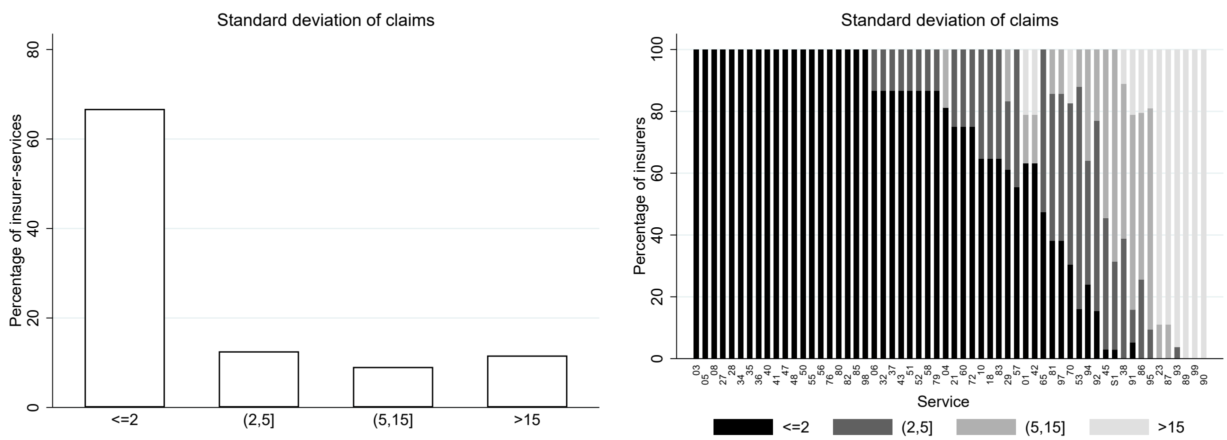
	2010				2011			
Insurer	Mean	SD	P25	P75	Mean	SD	P25	P75
EPS013	92.1	9.8	87.3	100.0	93.5	11.4	91.2	100.0
EPS016	93.4	7.8	89.7	100.0	94.3	12.0	96.6	100.0
EPS037	81.5	25.4	73.9	100.0	81.1	24.2	67.2	100.0
EPS002	83.7	24.2	82.8	100.0	83.9	22.6	69.0	100.0
EPS017	64.2	26.7	46.3	87.9	66.0	30.9	34.5	93.1
EPS010	52.3	38.4	13.8	94.8	50.9	40.6	10.5	98.3
EPS005	82.7	16.5	72.2	96.6	85.9	14.9	78.4	96.6
EPS018	53.9	32.6	24.1	84.5	52.6	35.6	22.4	82.8
EPS003	77.7	23.7	66.7	93.1	77.2	25.1	77.6	94.8
EPS008	57.2	34.0	37.9	86.2	61.4	31.5	44.8	86.2
EPS023	51.4	34.2	22.4	87.9	51.9	33.9	25.9	84.5
EPS009	43.2	33.1	24.1	66.7	51.5	34.1	15.5	75.9
EPS001	77.0	20.7	54.4	98.3	77.6	20.2	55.2	96.6
EPS012	61.5	39.7	20.7	100.0	62.6	39.9	20.7	100.0

Note: Mean, standard deviation, and 25th and 75th percentiles of the percentage of services offered by each insurer across markets through the sample of providers in my data. A service is defined by its first 2 digits from the procedure codes covered in the national health insurance plan.

Appendix 3 Variation in provider quality within service



Appendix Figure 2: Average satisfaction levels



(a) All services (b) Per service

Appendix Figure 3: Standard deviation of claims across providers

Appendix 4 Robustness checks on correlates of network breadth

Appendix Table 3: Robustness on regression specification: Network breadth and health care cost per service

	(1) $ihs(service\ cost)$	(2) Two-part model	
		$\mathbf{1}\{service\ cost > 0\}$	$\log(service\ cost)$
Panel A: Stayers			
H^{2011}_{jmk}	0.03***	0.14***	0.09***
	(0.003)	(0.01)	(0.01)
N	14,500,000	622,256	
R^2	0.44	0.31	
Panel B: New enrollees			
H^{2011}_{jmk}	0.02***	0.26***	0.16***
	(0.002)	(0.02)	(0.02)
N	14,500,000	226,400	
R^2	0.21	0.27	
Demog + Diag	Y	Y	
Market	Y	Y	
Service	Y	Y	
Insurer	Y	Y	

Note: Column (1) uses the inverse hyperbolic sine transformation of 2011 total cost per service as dependent variable. Column (2) models 2011 total cost per service a two-stage process. The first stage is a logistic regression for the probability of having non-zero cost. The second-stage is a log-linear regression of service cost conditional on having non-zero cost. The independent variable in both columns is the 2011 network breadth per service. Panel A conditions on the sample of stayers. Panel B conditions on the sample of new enrollees. All the models include consumer demographics and diagnoses, as well as market, service, and insurer fixed effects. Robust standard errors in parenthesis. ***p<0.01, **p<0.05, *p<0.1.

Appendix Table 4: Robustness on regression specification: Selection on baseline costs

	(1) $ihs(service\ cost^{2010})$	(2) Two-part model	
		$\mathbf{1}\{service\ cost > 0\}$	$\log(service\ cost^{2010})$
$H^{2010}_{j'mk} - H^{2011}_{j'mk}$	0.005*	0.34***	0.24***
	(0.003)	(0.02)	(0.02)
Switch	-0.09***	-0.36***	-0.14**
	(0.02)	(0.08)	(0.06)
Switch × $(H^{2010}_{j'mk} - H^{2011}_{j'mk})$	-0.25***	-1.74**	0.56
	(0.09)	(0.79)	(0.58)
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. Column (1) presents results of an OLS regression of the inverse hyperbolic sine of 2010 total costs per service on a switching indicator and the difference between network breadth in 2010 and network breadth in 2011 for the insurer chosen in 2011 (j'). Column (2) shows results of a two-part model of total costs per service. 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. 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 5: Selection on current enrollees' baseline costs

	$\log(total\ cost^{2010}_{ijmt}+1)$ (1)	$any\ claim^{2010}_{ijmt}$ (2)
$(H^{2010}_{jmk} - H^{2011}_{j'mk})$	0.004*	-0.0001
	0.002	0.0002
Switch	-0.09***	-0.007***
	0.02	0.001
Switch $\times$ $(H^{2010}_{jmk} - H^{2011}_{j'mk})$	-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 6: Selection on baseline costs conditional on switchers

	$\log(total\ cost^{2010}_{ijmt}+1)$ (1)	$any\ claim^{2010}_{ijmt}$ (2)
$(H^{2010}_{j'mk} - H^{2011}_{j'mk})$	-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.

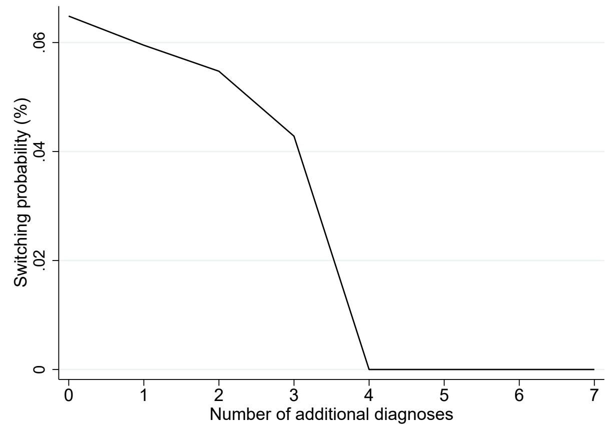
Appendix Table 7: New enrollees' risk scores and network breadth levels

	$\log(risk\ transfer^{2011}_{new})$
H^{2011}_{jk}	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 5 Evidence of inertia

(a) Switching by age



(b) Switching by diagnosis

Appendix Figure 4: Switching probability

Appendix Table 8: Insurer shares for current and new enrollees in 2011

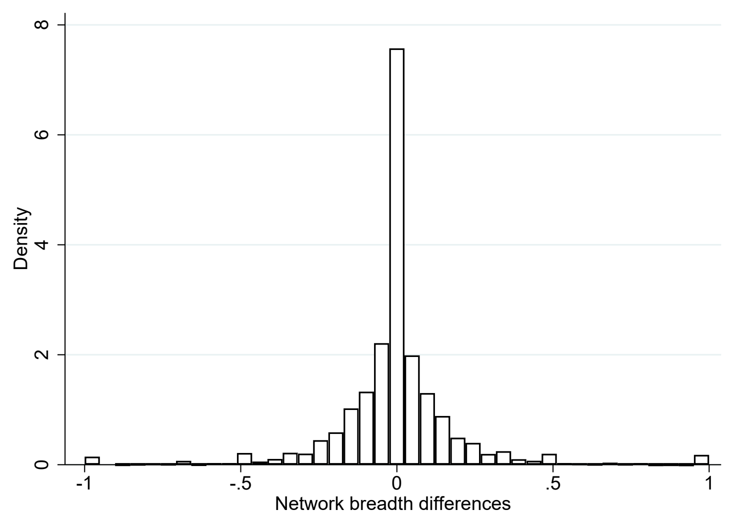
Insurer	Current	New
EPS001	0.06	0.14
EPS002	10.01	6.48
EPS005	21.48	18.30
EPS013	30.84	37.04
EPS016	0.14	0.09
EPS017	0.00	0.01
EPS018	0.01	0.00
EPS037	37.45	37.94

Appendix 6 Steady state assumptions

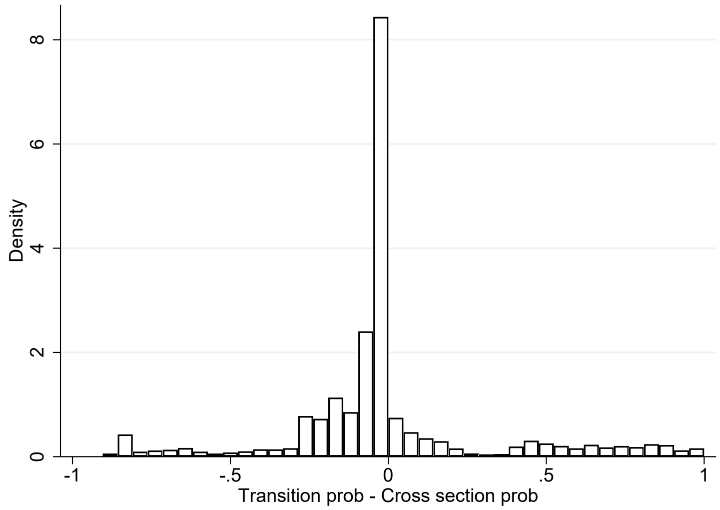
Steady state conditions require that choice variables are not changing over time. In my case, this means that service-level network breadth in 2010 should be equal to its value in 2011. In figure 5 I observe significant dispersion over time in my measure of coverage, which is inconsistent with a steady state argument.

Another way to test for steady state is to compare the transition probability from 2010 to 2011 of each consumer type θ into each of the possible diagnoses l, to the cross-sectional probability in 2011 that the consumer type θ has diagnosis l. In steady state, these two probabilities should be equal. I define a consumer type θ as a sex and age combination, and focus on a list of 7 exhaustive and mutually exclusive list of diagnoses in l: cancer only, cardiovascular disease only, diabetes only, renal disease only, other disease only, several comorbidities, and no diseases. More explicitly, in steady state, the transition probability of a 40 year old healthy man into a 41 year old man with cardiovascular disease should equal the cross-sectional probability that a 40 year old man has cardiovascular disease.

Figure 6 shows the distribution of the difference between these two probabilities. Although there is a high density at zero, the distribution exhibits long tails. This means that even if I focus on a full commitment equilibrium in networks and on new enrollees' myopic decisions, to appropriately capture insurers' incentives to offer broad networks, I need to model their future profits as a function of networks and consumer type-specific transition probabilities to different health states.

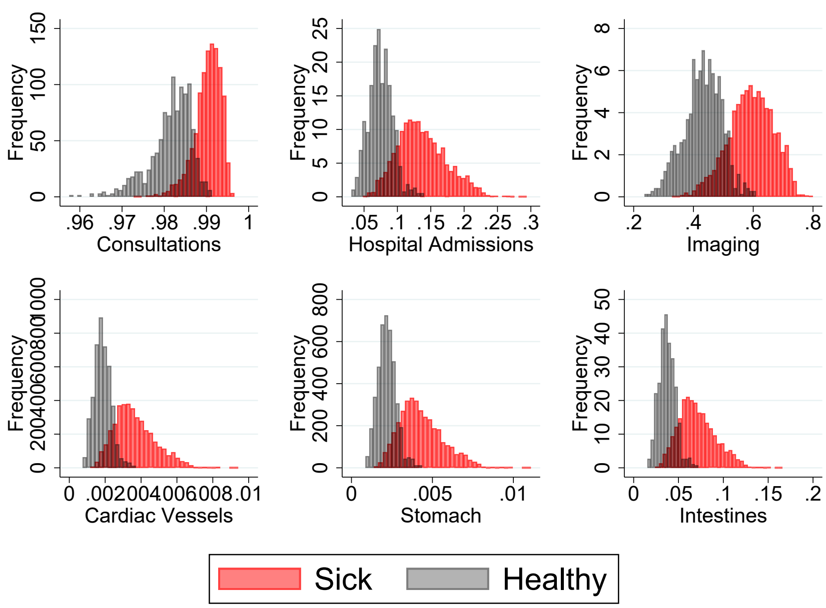


Appendix Figure 5: Differences in network breadth over time



Appendix Figure 6: Steady state diagnosis probability

Appendix 7 Robustness checks on demand



Appendix Figure 7: Distribution of service claims probability

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

$$u_{ijt} = \beta_i^D \sum_{l \in L} \gamma_{\theta(i)l(i)} \sum_m \gamma_{\theta(i)l(i)mt} H_{jmt} - \alpha_i \sum_{l \in L} \gamma_{\theta(i)l(i)} c_{\theta(i)l(i)y(i)jt} + \delta_j + \varepsilon_{ijt}$$

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

Appendix Table 9: Insurer demand with expectation over diagnoses

Variable		Coefficient	Std. Error
Network		2.33***	0.01
OOP spending		-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)
N		5,800,610	
N 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.

Appendix Table 10: Insurer demand in small markets

Variable		Coefficient	Std. Error
Network		1.14***	0.02
OOP spending		-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 11: Insurer demand with additional quality measures

Variable		Coefficient	Std. Error
Network		2.32***	0.01
OOP spending		-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)
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 12: Insurer demand with alternative network measures

Variable		(1) Only large hospitals Coef.	SE	(2) All providers Coef.	SE
Network		2.37***	0.01	2.54***	0.01
OOP spending		-6.64***	0.21	-4.90***	0.16
Interactions					
Network	Demographics				
	Male	0.30***	0.01	0.14***	0.01
	Age	-0.01***	0.00	0.00***	0.00
	Diagnoses				
	Cancer	-0.34***	0.02	-0.71***	0.02
	Cardiovascular	-0.33***	0.01	-0.50***	0.02
	Diabetes	-0.44***	0.04	-0.65***	0.06
	Renal	-0.61***	0.08	-0.64***	0.10
	Other	-0.53***	0.02	-0.69***	0.03
	>=2 diseases	-0.64***	0.02	-0.93***	0.02
	Healthy	(ref)	(ref)	(ref)	(ref)
	Location				
	Normal	0.05***	0.01	-0.25***	0.01
	Special	0.73***	0.04	-0.59***	0.02
	Urban	(ref)	(ref)	(ref)	(ref)
OOP spending	Demographics				
	Male	0.05	0.09	-0.02	0.06
	Age	-0.01***	0.00	-0.01***	0.00
	Diagnoses				
	Cancer	5.36***	0.22	4.41	0.15
	Cardiovascular	5.89***	0.18	4.57***	0.15
	Diabetes	5.77***	0.31	4.81***	0.25
	Renal	6.27***	0.22	4.82***	0.17
	Other	5.56***	0.20	4.32***	0.15
	>=2 diseases	5.89***	0.18	4.61***	0.15
	Healthy	(ref)	(ref)	(ref)	(ref)
	Location				
	Normal	1.15***	0.11	0.74***	0.08
	Special	0.67	0.43	0.92***	0.16
	Urban	(ref)	(ref)	(ref)	(ref)
N		5,800,610		5,800,610	
N enrollees		500,000		500,000	
Pseudo-R^2		0.17		0.12	

Note: This table presents results of the insurer choice model under alternative specifications of the network breadth variable. Column (1) reports coefficients and standard errors of a model where network breadth is constructed based on 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 the market. Column (2) presents coefficients and standard errors of a model where network breadth is constructed using all institutional provider and stand-alone doctors. Includes insurer fixed effects. Robust standard errors reported. ***p<0.01, **p<0.05, *p<0.1.

Appendix Table 13: Insurer demand on adults aged 19 or older

Variable			Coef.	SE
Network			2.28***	0.01
OOP spending			-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. Includes insurer fixed effects. Robust standard errors reported. ***p<0.01, **p<0.05, *p<0.1.

Appendix 8 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 γ is the probability that consumer i of type (θ, 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:

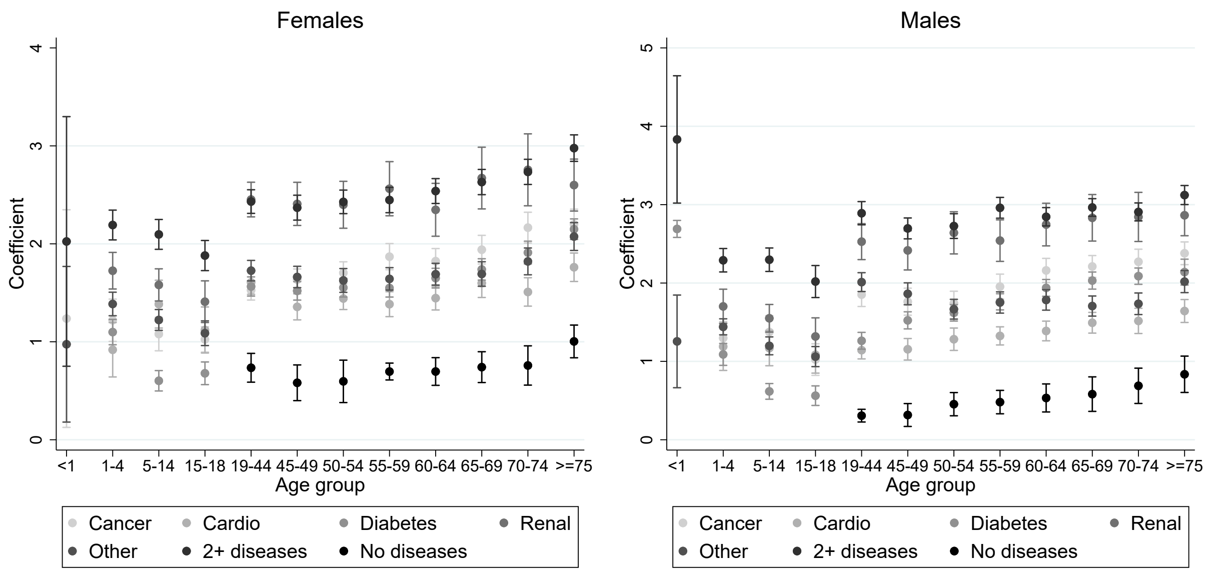
$$AC_j(H_j) = \sum_i \left(\sum_m \gamma_{\theta(i)l(i)m} \underbrace{\sum_{h \in H_{jm}} p_{jhm} \frac{exp(\delta_{ihm})}{1 + exp(\delta_{ihm})} \left(\frac{1 + exp(\eta(H_j))}{exp(\eta(H_j))} \right)}_{A_m} \right) \frac{exp(\eta_i(H_j))}{1 + exp(\eta_i(H_j))}$$

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

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

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

Appendix 9 Additional average cost results



Appendix Figure 8: Consumer type fixed effects

Appendix Table 14: Predicted change in females' average cost by type of carrier and diagnosis

	F, 19-44, Healthy		F, 19-44, Cancer	
Service	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 15: Predicted change in males' average cost by type of carrier and diagnosis

	M, 19-44, Healthy		M, 19-44, Diabetes	
Service	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 10 Dropout and transition probabilities

To estimate the marginal cost of network formation in the third step of my model, I first need to compute the probability that consumer type (θ, l) drops out of the contributory system and the probability that consumer type (θ, 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 (θ, l), I calculate the probability that she drops out of the system non-parametrically as the number of individuals of type (θ, l) observed only in 2010 but not 2011, divided by the total number of type (θ, l) individuals in 2010. Table 16 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 θ to θ' is deterministic, I only need to compute transition probabilities across diagnoses. The probability that type (θ, l) transitions into $(theta', l')$ equals the number of type (θ, l) in 2010 that end up with diagnosis l' in 2011, divided by the number of type (θ, l) individuals in 2010. Table 17 presents the mean and standard deviation in parenthesis of transition probabilities from having cancer, cardiovascular disease, diabetes, renal disease, other diseases, 2 or more diseases, and no diseases in period t to having each of these 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 (≥ 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 16: 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 17: Summary statistics of transition probabilities across diagnoses

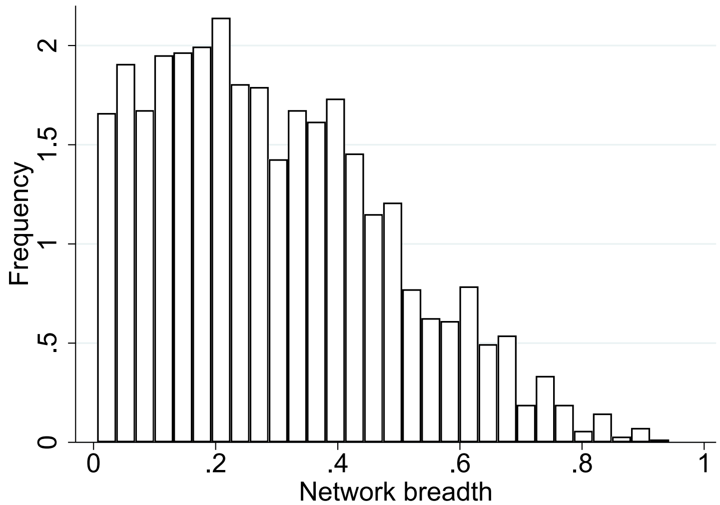
Diagnosis t/t+1	Cancer	Cardio	Diabetes	Renal	Other	≥2 disea.	No disea.
Cancer	0.789	0.000	0.000	0.000	0.000	0.211	0.000
	(0.136)	(0.000)	(0.000)	(0.000)	(0.000)	(0.136)	(0.000)
Cardio	0.000	0.774	0.000	0.000	0.000	0.226	0.000
	(0.000)	(0.148)	(0.000)	(0.000)	(0.000)	(0.148)	(0.000)
Diabetes	0.000	0.000	0.654	0.000	0.000	0.346	0.000
	(0.000)	(0.000)	(0.209)	(0.000)	(0.000)	(0.209)	(0.000)
Renal	0.000	0.000	0.000	0.643	0.000	0.357	0.000
	(0.000)	(0.000)	(0.000)	(0.167)	(0.000)	(0.167)	(0.000)
Other	0.000	0.000	0.000	0.000	0.711	0.289	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.171)	(0.171)	(0.000)
>=2 diseases	0.000	0.000	0.000	0.000	0.000	1.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No diseases	0.034	0.087	0.007	0.004	0.044	0.054	0.770
	(0.038)	(0.086)	(0.015)	(0.016)	(0.028)	(0.084)	(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 18: 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 marginal variable profits in the left-hand side of equation (5). Measured in millions of Colombian pesos per service per market.



Appendix Figure 9: Distribution of network breadth for FOC

Appendix Table 19: First stage regression of network breadth

H_{jmk}	Coefficient	Std. Error
H_{jmk}^{t-1}	0.76***	0.01
$\overline{\gamma}_{female,m,k}$	33.94***	8.45
$\overline{\gamma}_{healthy,m,k}$	14.31***	4.20
$\overline{\gamma}_{age\ 19-44,m,k}$	-55.33***	13.63
$H_{jmk}^{t-1} \times \overline{\gamma}_{age\ 19-44,m,k}$	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 (6). H_{jmk}^{t-1} is the network breadth in 2010. $\overline{\gamma}_{i,m,k}$ 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 20: 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

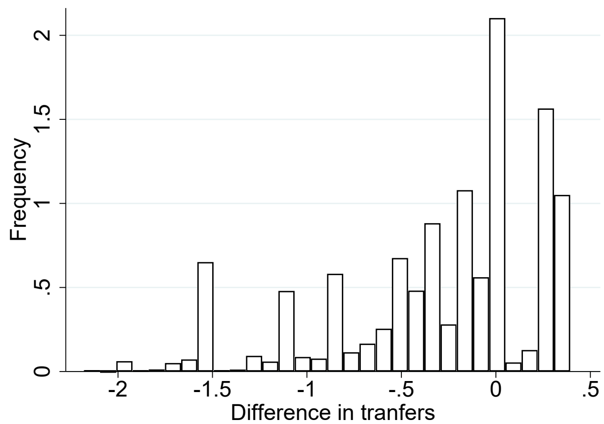
Note: Column (1) presents the predicted average total cost of network formation in millions of pesos across markets and column (2) presents this cost as a percentage of total variable profits.

Appendix Table 21: Decomposition of profit changes after network breadth increase by diagnosis

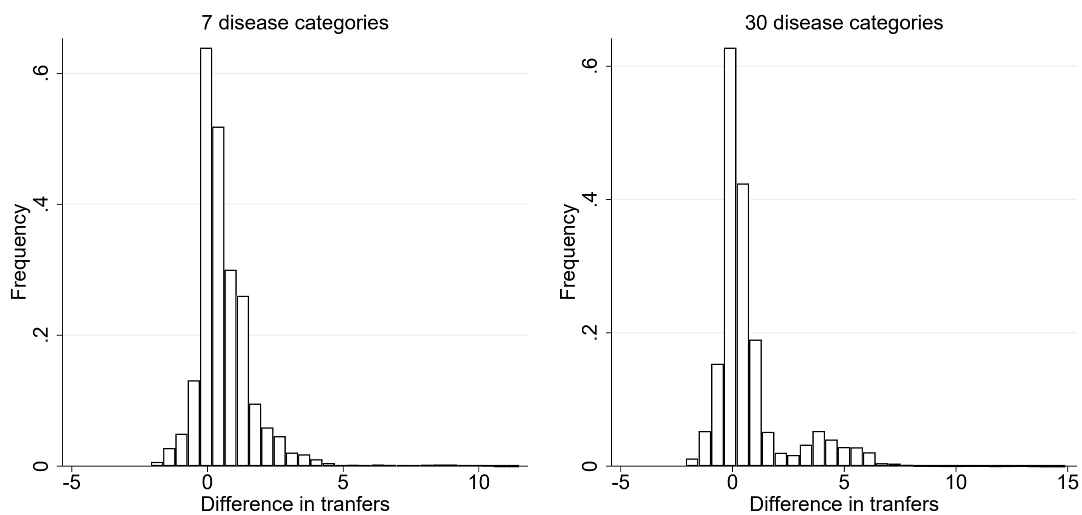
	Healthy		Renal		Other disease	
Service	$\%\Delta s_{ijk}$	$\%\Delta AC_{\theta ljk}$	$\%\Delta s_{ijk}$	$\%\Delta AC_{\theta ljk}$	$\%\Delta s_{ijk}$	$\%\Delta AC_{\theta ljk}$
Cardiac vessels	0.01	0.00	0.03	0.00	0.02	0.00
Stomach	0.01	0.00	0.03	0.00	0.03	0.00
Intestines	0.25	0.02	0.60	0.02	0.49	0.02
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 11 Additional counterfactual results



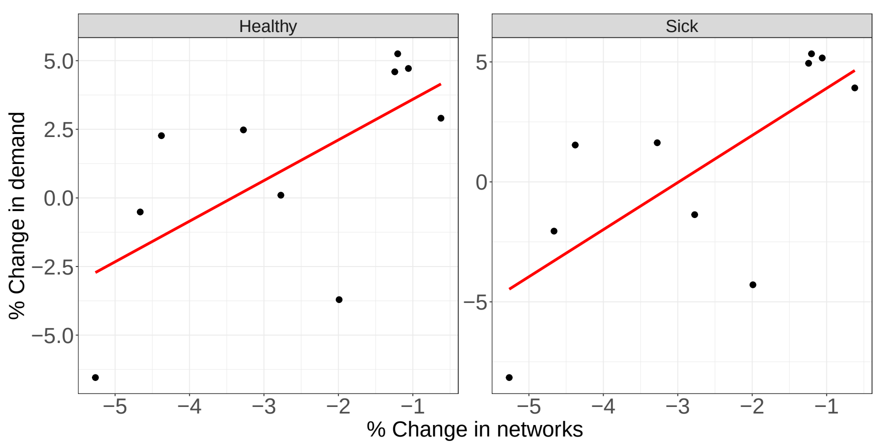
Appendix Figure 10: Distribution of counterfactual minus observed transfer under no risk adjustment



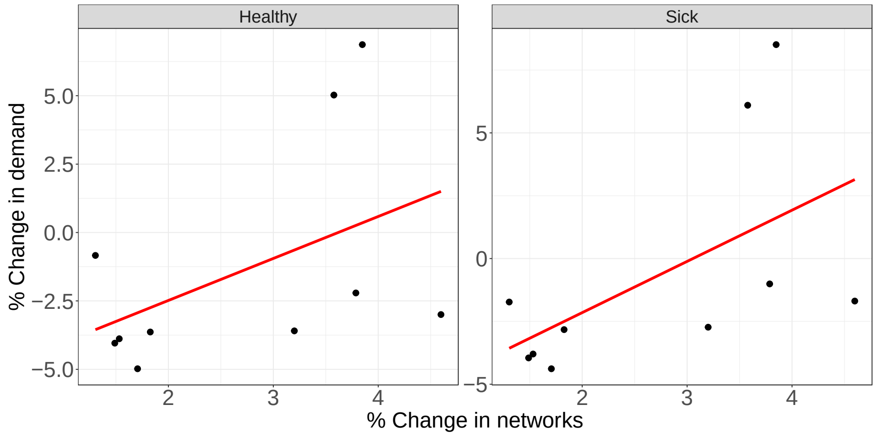
Appendix Figure 11: Distribution of counterfactual minus observed transfer under improved risk adjustment

Appendix Table 22: Disease categories

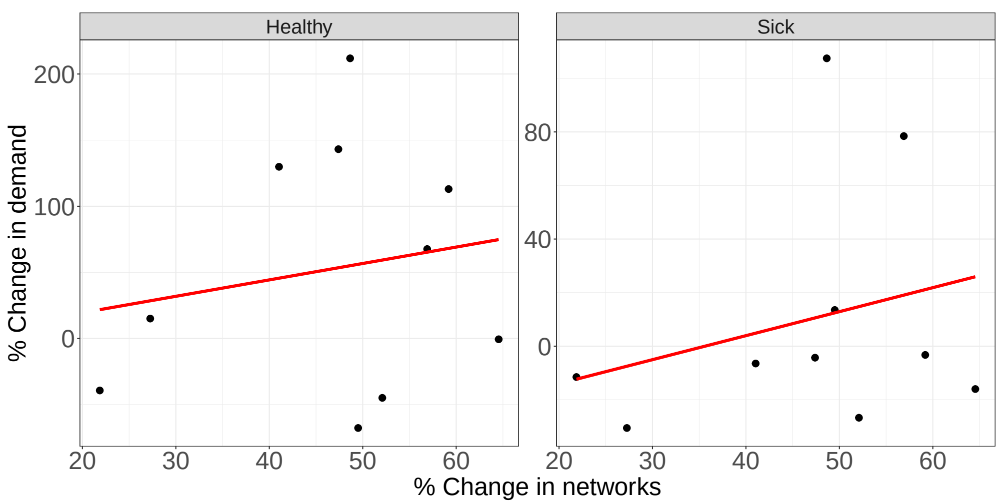
Arthritis
Arthrosis
Asthma
Autoimmune disease
Breast cancer
Cancer in digestive organs
Cancer in female genitalia
Cancer in male genitalia
Cancer therapy
Invasive cervical cancer
Local cervical cancer
Diabetes
Epilepsy
Genetic anomalies
HIV-AIDS
Hypertension
Cancer in respiratory organs
Lymphatic cancer
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



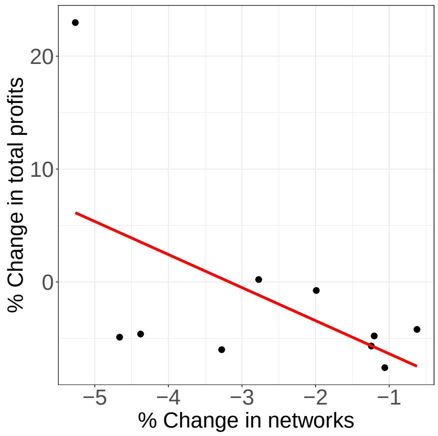
Appendix Figure 12: Correlation between network changes and changes in demand under no risk adjustment



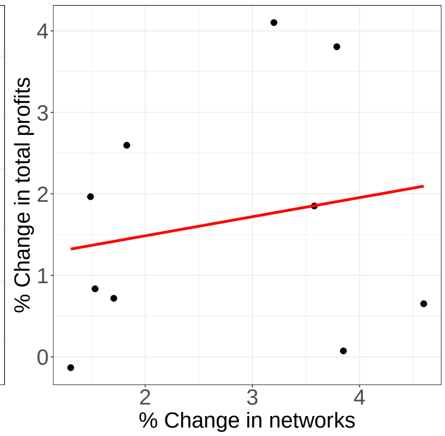
Appendix Figure 13: Correlation between network changes and changes in demand under improved adjustment



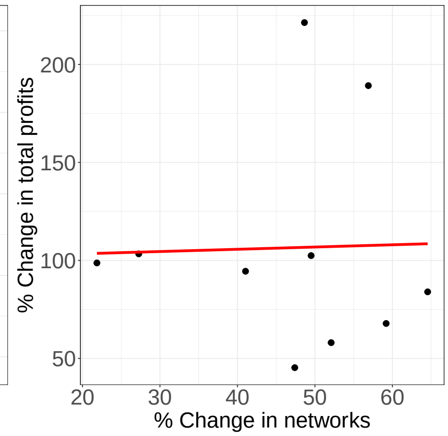
Appendix Figure 14: Correlation between network changes and changes in demand under premium deregulation



(a) No risk adjustment



(b) Improved risk adjustment



(c) Premium deregulation

Appendix Figure 15: Correlation between profit and network changes in counterfactual

Appendix Table 23: Welfare changes by consumer subgroups

	No risk adjustment (1)	Improved risk adjustment (2)	Premium deregulation (3)
Sex			
Female	-2.6	2.7	-62.0
Male	-2.6	2.8	-76.0
Age group			
5-14	-2.6	2.5	-75.1
15-18	-2.3	2.2	-69.5
19-44	-2.7	2.6	-5.2
45-49	-2.6	2.8	-73.3
50-54	-2.6	2.7	-15.9
55-59	-2.5	2.5	-67.5
60-64	-2.5	2.8	-88.0
65-69	-2.4	2.6	-53.9
70-74	-2.5	2.5	-51.1
>=75	-2.2	2.5	-55.7
Diagnoses			
Cancer	-2.6	2.4	9.6
Cardiovascular	-2.3	2.2	13.0
Diabetes	-2.2	2.1	10.8
Renal	-2.8	2.6	21.4
Other	-2.5	2.3	10.1
>=2 diseases	-2.6	2.3	14.6
No diseases	-2.6	2.8	-73.7

Note: This table shows the welfare change for subgroups of consumers under the counterfactual without risk adjustment in column (1), with the improved risk adjustment formula using the list of 30 disease categories in column (2), and under premium deregulation in column (3), relative to the model's predictions in the observed scenario.