

Data Appendix for
“Unhealthy Insurance Markets:
Search Frictions and the Cost and Quality of Health Insurance”

By

Randall D. Cebul, James B. Rebitzer, Lowell J. Taylor and Mark E. Votruba*

The following Appendix provides details pertaining to our use of the Employer Health Insurance Survey (EHIS) to measure price dispersion in the fully insured and self-insured segments of the group insurance market, as referred to in pp16-17 of our manuscript. Statistics reported in Table 9-11 of our manuscript reflect the estimated premiums residuals from the GLM models reported in Appendix Table A1 (below). The iterated simulation procedure (results reported in Table 12) also draws on these residuals.

1. Data Selection Criteria

The premium data are obtained from the 1997 Robert Wood Johnson Foundation Employer Health Insurance Survey (EHIS). The dataset consists of a stratified sample of establishments. Throughout our analysis, the “establishment” is treated as the unit of analysis. Sampling “establishment weights” provided in the EHIS dataset are used throughout the analysis to provide nationally representative estimates.

The following describes the selection criteria applied to construct the analytic sample:

- Private establishments offering at least one general medical plan (N=13,716).
- Restricted to establishments with at least 3 permanent employees (N=12,840). This was done because the characteristics of workers in the establishment were constructed over permanent employees only, and the characteristics of permanent employees were found to be strongly predictive premiums.
- Restricted to establishments with a single “dominant” health plan, that covered both inpatient and outpatient services (N=10,391). We defined a health plan as dominant when at least 90% of health plan enrollees from an establishment were enrolled in the same general medical plan. In the premium prediction models, the plan characteristics used as covariates were those associated with the dominant plan for that establishment.
- Restricted to establishments where funding of the dominant plan was recorded as either “fully insured” or “self insured” (N=10,329).
- Restricted to establishments where the single monthly premium recorded for the dominant plan was *not* imputed (N=7,578).

- Excluded establishments if the dominant plan was an HMO plan (N=5,261). HMO plans were excluded for two reasons. First, fewer than 5% of all HMO plans were self insured, compared with 17% of PPO/POS plans and 25% of indemnity plans. Second, premiums for HMO plans are expected to vary for reasons not well-captured by the recorded plan characteristics.

The application of these selection criteria yielded a sample of 5261 establishments offering a non-HMO plan as their dominant plan option. The variable of interest in the subsequent analysis was the single monthly premium recorded for the establishment's dominant plan. An alternative premium measure – the family monthly premium – was not recorded as frequently, specifically not in cases where a family plan was not offered by the employer.

2. Covariates Used in the Premium Prediction Models

The same covariates were included in the two premium prediction models, with the exception of three covariates that were not applicable to self insured establishments.

- Plan characteristics:
 - plan type is PPO/POS (indicator)
 - must enroll with gatekeeper (indicator, applies to PPO/POS plans only)
 - deductible level (quadratic)
 - copayment for typical office visit (quadratic, in some cases inferred by EHIS on basis of reported coinsurance level)
 - any catastrophic cap (indicator)
 - catastrophic cap level
 - includes prescription drug coverage (indicator)
 - includes mental health coverage (indicator)
 - includes vision care coverage (indicator)
 - includes dental coverage (indicator)
 - includes coverage for preventive dental care and orthodontics (indicator)
 - family coverage option available (indicator)
 - no exclusions for health reasons (indicator)
 - enrollees must report medical history (indicator)
 - no waiting period for enrollment (indicator)
 - contract includes guaranteed renewal (indicator)
 - contract include minimum participation requirement (indicator)
- Establishment characteristics:
 - establishment size (5 categories)
 - percent workers, permanent

- percent workers, full-time
- percent workers, female
- percent workers, age 30-39
- percent workers, age 40-49
- percent workers, age 50+
- mean payroll per worker (quadratic)
- any union workers (indicator)
- firm size (5 categories)
- firm industry (7 categories, agriculture collapsed with construction due to small sample size)
- firm years in business (4 categories)
- establishment part of insurance purchasing coalition
- premium source (4 categories)¹
- state dummies

¹ The source for the single monthly premium was identified as one of the following: (a) reported premium, (b) COBRA, (c) reported same premium for single and family premium, or (d) derived from aggregate paid premiums.

Appendix Table A1. Premium Prediction Models

	GLM		OLS	
	FI Plans (1)	SI Plans (2)	FI Plans (3)	SI Plans (4)
A. Establishment characteristics				
Employee characteristics				
fraction permanent	0.11 (0.08)	0.01 (0.14)	18.93 (14.70)	25.10 (22.81)
fraction full-time	0.05 (0.04)	-0.02 (0.07)	4.08 (7.76)	0.47 (12.70)
fraction female	0.07 (0.04)	0.03 (0.08)	10.35 (7.70)	0.18 (13.71)
fraction age 30-39	0.04 (0.05)	0.06 (0.07)	5.52 (9.84)	4.10 (12.10)
fraction age 40-49	0.25 (0.07)**	0.15 (0.06)*	45.27 (12.96)**	23.13 (11.96)+
fraction age 50+	0.36 (0.07)**	0.06 (0.09)	61.43 (12.56)**	7.56 (16.85)
mean annual pay (/10 ⁵)	0.24 (0.09)*	-0.17 (0.10)+	39.53 (19.90)+	-26.18 (20.31)
mean annual pay sqrd (/10 ¹⁰)	-0.12 (0.05)**	0.05 (0.05)	-19.16 (9.69)+	4.04 (10.42)
any union employees	0.18 (0.04)**	0.05 (0.08)	36.26 (8.11)**	7.82 (14.48)
Establishment size ^a				
8-20	0.00 (0.02)	0.06 (0.04)	-0.25 (3.82)	13.55 (8.24)
20-50	0.01 (0.03)	-0.13 (0.05)*	-0.08 (4.70)	-26.25 (10.92)*
50-300	-0.03 (0.04)	-0.08 (0.07)	-3.55 (8.50)	-12.58 (11.94)
300+	-0.04 (0.05)	0.17 (0.09)+	-8.69 (9.27)	36.01 (20.62)+
Firm size ^b				
10-35	-0.03 (0.03)	0.22 (0.14)	-5.91 (4.94)	31.74 (22.63)
35-250	-0.03 (0.03)	0.15 (0.11)	-6.07 (5.34)	19.18 (20.10)
250-5000	-0.12 (0.04)**	0.23 (0.10)*	-24.10 (7.86)**	35.74 (15.91)*
5000+	-0.03 (0.10)	0.19 (0.11)+	-2.45 (17.13)	28.76 (17.44)
Industry category ^c				
mining and manufacturing	0.00 (0.04)	-0.05 (0.08)	-0.78 (7.16)	-19.88 (16.35)
transportation/communications	-0.01 (0.05)	0.01 (0.08)	-3.77 (8.68)	-5.17 (17.73)
wholesale trade	0.07	-0.15	15.87	-35.15

	(0.07)	(0.09)+	(15.23)	(21.09)+
retail trade	0.02	-0.12	4.46	-26.06
	(0.03)	(0.12)	(6.24)	(22.96)
finance/insurance/real estate	0.02	-0.09	2.56	-24.94
	(0.03)	(0.08)	(4.91)	(15.76)
professional services	0.05	0.01	8.09	-8.65
	(0.04)	(0.11)	(6.09)	(23.15)
other services	-0.02	0.05	-4.65	5.07
	(0.04)	(0.09)	(8.08)	(17.18)
Years in business ^d				
5-10 years	-0.03	-0.04	-4.03	-9.32
	(0.06)	(0.08)	(12.03)	(14.49)
10-20 years	-0.00	0.07	-0.91	8.57
	(0.04)	(0.07)	(6.97)	(12.20)
20-50 years	0.01	0.10	1.22	16.49
	(0.04)	(0.07)	(7.49)	(10.31)
50+ years	0.06	0.07	11.60	7.39
	(0.03)+	(0.08)	(6.89)+	(11.36)
Establishment involved in a purchasing arrangement to buy HI	0.04	0.08	6.30	17.70
	(0.02)*	(0.06)	(3.37)+	(12.64)
B. Plan characteristics				
Plan type				
PPO plan	-0.02	0.07	-3.82	13.24
	(0.03)	(0.05)	(5.37)	(11.08)
PPO plan w/ gatekeeper	-0.04	-0.08	-6.70	-12.66
	(0.02)*	(0.07)	(3.85)+	(12.63)
Plan offers family coverage	0.06	-0.49	10.00	-121.18
	(0.05)	(0.11)**	(9.04)	(25.01)**
Cost sharing				
deductible (/10 ²)	0.00	-0.04	1.16	-8.45
	(0.01)	(0.03)	(2.00)	(5.01)+
deductible sqrd (/10 ⁴)	-0.00	0.00	-0.10	0.42
	(0.00)	(0.00)	(0.21)	(0.39)
copay level (/10)	-0.02	-0.02	-3.43	-4.77
	(0.04)	(0.07)	(7.40)	(13.61)
copay level sqrd (/10 ²)	-0.02	-0.00	-3.14	-0.11
	(0.01)	(0.01)	(2.50)	(2.74)
plan has catastrophic cap	0.10	0.14	17.84	25.99
	(0.03)**	(0.06)*	(5.85)**	(10.88)*
catastrophic cap amount	-0.11	-0.01	-20.04	3.67
	(0.09)	(0.22)	(18.27)	(37.81)
Coverage included				
outpatient prescription drugs	0.08	-0.04	12.87	-6.81
	(0.03)*	(0.12)	(6.01)*	(20.80)
mental health services	-0.05	-0.18	-11.87	-24.27
	(0.07)	(0.12)	(15.63)	(18.77)

vision care	0.00 (0.03)	0.04 (0.05)	1.65 (5.09)	7.12 (9.08)
any dental	0.05 (0.03)+	-0.01 (0.04)	7.04 (5.14)	-0.96 (7.80)
premium dental ^c	0.11 (0.04)**	0.09 (0.07)	22.39 (7.53)**	15.25 (12.75)
Coverage restrictions				
any employees excluded because of health conditions	0.01 (0.02)	0.04 (0.04)	2.59 (3.94)	6.05 (7.98)
employees required to provide medical history	-0.00 (0.02)	--	-0.43 (5.14)	--
no waiting period for coverage	0.01 (0.02)	-0.01 (0.03)	1.86 (4.51)	-2.10 (6.21)
contract for plan is guaranteed renewal	-0.01 (0.03)	--	-1.12 (5.54)	--
contract includes minimum participation requirement	-0.00 (0.03)	--	-1.11 (5.16)	--
Basis for reported premium				
COBRA	-0.09 (0.07)	0.03 (0.03)	-15.12 (13.88)	6.37 (6.59)
reported that single premium did not differ from family premium ^f	0.09 (0.03)**	0.16 (0.23)	16.89 (5.67)**	27.87 (43.14)
derived from aggregate paid premiums	-0.11 (0.05)*	-0.01 (0.12)	-15.97 (9.75)	-10.53 (19.62)
State Fixed Effects	YES	YES	YES	YES
R^2	--	--	.2015	.5304
Sample size	4287	974	4287	974
Premium mean	176.2	177.7	176.2	177.7
Premium variance	6418.2	3741.4	6418.2	3741.4
Residual premium variance	5156.1	1736.0	5124.8	1757.1

Notes: Data is drawn from the 1997 Robert Wood Johnson Foundation Employer Health Insurance Survey (EHIS). Sample consists of 5261 private establishments meeting the following criteria: at least 3 permanent employees at establishment; firm offers at least one general medical plan to employees at establishment; at least 90 percent of participating employees were enrolled in the same non-HMO plan (the establishment's "dominant" health plan); and single plan premium was not imputed. Dependent variable is single monthly premium recorded for establishment's dominant plan. Plan characteristics refer to establishment's dominant plan. Prediction model in columns 1 and 2 are estimated using generalized linear model (GLM) with log link and gamma distributional family. Coefficients are interpreted as changes in the log of predicted premium. The distributional family was determined by way of the modified Park Test. Prediction model in columns 3 and 4 are estimated using linear OLS. Establishment-level weights applied to produce nationally representative estimates. Robust standard errors in parentheses, adjusted for clustering by sampling strata (+ significant at 10%; * significant at 5%; ** significant at 1%).

^a omitted category: <8 employees

^b omitted category: <10 employees

^c omitted category: construction, plus agriculture/fishing/forestry (one establishment)

^d omitted category: <5 years

^e dental coverage that includes both preventive and orthodontic services

^f omitted category: reported premium

^g recorded single premium same as reported family premium in these cases