The Effect of Competition among Health Insurance Providers on Health Premiums.

Alfred Zhu, The University of Texas at Austin

Abstract

The cost of healthcare and health insurance in the United States is higher than other comparable nations while the overall market concentration in health insurance companies has increased over time. This paper evaluates the effect of concentration in the private insurance market on the average price of annual health insurance premiums using a linear regression model with several control variables using data in the 2021 timeframe. The results show that while the effect of market concentration on insurance premiums is statistically significant, the effect's magnitude is small. Additionally, the empirical model demonstrates that increasing market concentration can result in lower health premiums when market concentrations are below a certain level. This suggests that antitrust agencies should be more lenient in evaluating mergers between private health insurance firms under certain conditions.

Introduction

The debate over health care and health insurance continues to be a major topic of issue in United States politics. A February 2024 survey published by the Statista Research Department found that 14% of voters find healthcare the most important issue in the United States. The concern over health care in the U.S. is not without merit. Sen (2015) found that the U.S. suffers from expensive health care due to asymmetric information between patients and health insurance providers as insurance providers do not fully know a patient's condition. This results in a lack of a competitive market equilibrium and inefficiencies in private healthcare insurance as insurance firms would not know the true risk they have taken and would like to exclude "high-risk" patients given the opportunity if it weren't for regulations Sen (2015). This source of market inefficiency can be seen when comparing healthcare spending between the

U.S. and other comparable nations. Wager et al. (2024) concluded that per capita health expenditures in the U.S. were $4,000 higher than in any other high-income nation.

Asymmetric information is not the only cause of market inefficiencies within the U.S. healthcare system. There are significant barriers to entry into the health insurance market due to high start-up costs and government regulations, reducing competition between health insurance providers and increasing market concentration (Schansberg 2014). Furthermore, concentration within the commercial health insurance market has increased over time with the estimated market share of the four largest insurers growing to 83% by 2014 (Dafny, 2015). Additionally, market concentration in the health insurance market has not decreased since 2017, according to data shared by the KFF. The increase in market concentration could potentially lower competition among firms and further increase the cost of health insurance premiums for people.

With the issue of potential lower competition due to increasing market concentration, consumers could become worse off when paying for health insurance premiums. If a higher market concentration of insurance providers does increase prices, more stringent antitrust law regulations should be enacted to prevent people from experiencing higher prices due to less competition. Therefore analyzing the relationship between health insurance premiums and health insurance provider competition will be vital in determining future antitrust policies among lawmakers and potentially reduce a source of upward pricing pressure for health insurance premiums. This paper will analyze the effects of market competition between health insurance firms on the cost of insurance premiums using an empirical regression model that models this relationship while controlling for other variables.

Literature Review

Previous literature on the relationship between competition between insurance providers and the costs of health insurance premiums has some conflicting results. Studies that quantify

competition by the number of choices offered by an employer are mixed on how it affects prices. Frank & Lamiraud (2009) found that in Switzerland, more choices in insurance inhibit people from switching plans, reducing the effectiveness of consumer choice-making, and suggests that reducing the number of choices may result in more price competition, lowering prices.

For the U.S., Bradford & Mobley (2004) have a similar conclusion that more choices do not have a competitive effect. Additionally, they found that having no other choice for an insurance plan leads to a lower premium paid, but higher margins over medical expenses (Bradford & Mobley 2004). This would conflict with a later study by Ho & Lee (2017) where they found that removing a health insurance issuer will typically increase the predicted premium, and consumer welfare would fall by as much as $200 per capita.

Apart from the studies that look at competition through the number of choices an employee can select from an employer, Dafny et al. (2012) found that increasing market concentration does increase health insurance premiums, making their findings more consistent with the results from Ho & Lee (2017). Overall the effects of market competition between health insurance firms on prices are ambiguous due to differing results.

This study will contribute to the previous literature on the effect of competition on prices by examining the effects of competition through market concentration on the cost of health insurance premiums as previous literature mostly used the number of choices in employer health insurance plans as competition rather than overall market concentration. This is because government agencies and departments use measures of market concentration when evaluating mergers in antitrust cases as seen in the U.S. Merger Guidelines. The study will use data from 2021 where overall market concentration is greater than previous time frames. I will use the Herfindahl-Hirschman Index to measure market concentration and analyze its effect on the cost of health insurance premiums. The unit of analysis will be at the state level.

Empirical Model

This study will analyze the effects of market concentration on health insurance firms and health insurance premiums using a linear regression model with multiple independent regressors. The empirical regression model is:

$$Insurance\ Premium\ = \beta_0 + \beta_1 Concentration + \beta_2 Concentration^2 + \beta_3 Patient + \beta_4 Poverty + \beta_5 Unemployed + \beta_6 Price\ Level + \beta_7 Income + \beta_8 Race + \beta_9 Region + u$$

The insurance premium variable represents the average annual premium that is paid between the employer and employee by the state for employer-based health insurance measured in dollars. The concentration variable is the market concentration of the large group health insurance market captured using the Herfindahl-Hirshman Index (HHI). According to the 2023 U.S. Merger Guidelines, markets with an HHI between 1,000-1,800 are considered moderately concentrated, and markets with an HHI above 1,800 to be highly concentrated. The HHI is calculated by summing the squares of the market share of each competing firm in a market. A quadratic is used as its partial effects may differ at different concentration levels.

The patient variable is the total number of hospital outpatient visits per 1,000 people. This variable is used to control for potential differences in demand for medical services. The poverty variable measures the proportion of people that fall below the poverty line measured in percentage. The unemployment variable measures the percentage of people unemployed within the civilian non-institutional population. The price level variable is the regional price parity which measures the difference in price level of all consumption goods and services expressed as a percentage of the overall national price level. The race variable is a group of variables that measure the percentage distribution of race/ethnicity that includes the population percentages for White, Black, Hispanic, Asian, and multiple races. White is excluded from the regression model. The region variable is a group of dummy variables that determine if a state is in the

South, West, Midwest, or Northeast. South is excluded from the regression model. The income variable is the median annual household income measured in inflation-adjusted dollars.

Data

The insurance premium data is obtained from the KFF which was sourced from the Agency for Healthcare Research and Quality, Center for Financing, Access and Cost Trends, and can be found on this webpage. The data is at the state level in the 2021 time frame.

The HHI data for health insurance is obtained from the KFF which was sourced from KFF analysis from Health Coverage Portal, a market database maintained by Mark Farrah Associates, which includes information from the National Association of Insurance Commissioners. The data can be found on this webpage and is at the state level in the 2021 time frame.

The number of outpatients per 1,000 data is obtained from the KFF which is sourced from the AHA annual survey. The data can be found on this webpage and is at the state level in the 2021 time frame.

The poverty rate is obtained from the U.S. Economic Research Service from the USDA. The data is sourced by the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program which provides annual estimates of poverty statistics. The data can be found on this webpage and is at the state level in the 2021 time frame.

The unemployment rate is obtained from the KFF and is sourced from the Bureau of Labor Statistics. The data can be found on this webpage and is at the state level in the December 2021 time frame. The percentages are in decimal form.

The regional price parity data is obtained from the Bureau of Economic Analysis. The data can be found on this webpage and is at the state level in the 2021 time frame.

The percentage distribution of race is obtained from the KFF and is sourced from the 1-year estimates from the Census Bureau’s American Community Survey (ACS). The data can

be found on this webpage and is at the state level in the 2021 time frame. American Indian/Alaska Native, Native Hawaiian/Other Pacific Islander, and multiple races are summed into the other category. The percentages are in decimal form.

The median annual household income is obtained from the KFF and is sourced from the US Census Bureau's 2021 American Community Survey. The data can be found on this webpage and is at the state level in the 2021 time frame.

Table 1. Summary Statistics

Variable	**Mean**	**Std. dev.**	**Min**	**Max**
Insurance Premium	7388.31	535.05	6340	9037
Concentration	4713.75	1944.05	1194	9276
Patient	2779.67	1116.832	910	6218
Poverty	12.57	2.65	7.4	19.5
Unemployed	0.041	0.01	0.023	0.061
Price Level	97.18	7.08	86.6	113.23
Income	69243.76	11314.42	48716	90203
White	0.66	0.16	0.20	0.91
Black	0.1	0.1	0.004	0.43
Hispanic	0.13	0.11	0.017	0.503
Asian	0.04	0.06	0.007	0.38
Other	0.07	0.05	0.038	0.306
West	0.25	0.44	0	1
South	0.33	0.47	0	1
Midwest	0.24	0.43	0	1
Northeast	0.18	0.39	0	1

Table 1 shows that the mean total insurance premium (that is paid between the employee and employer) is $7,388.31 per person with a standard deviation of $535.05. Arkansas has the lowest total premium of $6,340 per person while Alaska has the highest of $9,037. The mean HHI concentration is 4,713.75 with a standard deviation of 1,944.05. New York has the lowest HHI of 1,194 which is moderately concentrated while Alaska has the highest HHI of 9,276 which is close to the maximum of 10,000 points. Aside from New York and Wisconsin, all of the states' HHI are considered highly concentrated.

Empirical Results

Table 2. Linear Regression Model

Source	SS	df	MS	Number of obs	=	51
				F(14, 36)	=	8.58
Model	11013542.6	14	786681.614	Prob > F	=	0.0000
Residual	3300220.39	36	91672.7885	R-squared	=	0.7694
				Adj R-squared	=	0.6798
Total	14313763	50	286275.26	Root MSE	=	302.78

premium	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
hhi	-.2710025	.1173978	-2.31	0.027	-.5090963	-.0329087
hhi_sq	.0000359	.0000112	3.22	0.003	.0000133	.0000585
patient	.1130164	.0697656	1.62	0.114	-.0284748	.2545075
poverty	37.55462	49.08135	0.77	0.449	-61.98697	137.0962
unemployed	11218.6	6698.894	1.67	0.103	-2367.39	24804.58
pricelevel	65.70975	19.95896	3.29	0.002	25.2311	106.1884
income	.0081708	.0143212	0.57	0.572	-.020874	.0372156
black	-1976.16	1016.62	-1.94	0.060	-4037.96	85.64089
hispanic	453.9356	788.9147	0.58	0.569	-1146.058	2053.929
asian	-6351.484	1479.262	-4.29	0.000	-9351.566	-3351.401
other	2654.03	1467.398	1.81	0.079	-321.992	5630.052
west	-482.6857	201.8549	-2.39	0.022	-892.0664	-73.30501
midwest	46.61272	175.0875	0.27	0.792	-308.4812	401.7066
northeast	-53.82515	213.2612	-0.25	0.802	-486.339	378.6887
_cons	-100.1165	1582.65	-0.06	0.950	-3309.879	3109.646

The linear regression model has an R^2 value of 0.7694 which explains around 77% of the variance in the average health insurance premium paid.

The variables HHI concentration and HHI concentration squared are both statistically significant at the 5% level. The quadratic component of HHI concentration is statistically significant at the 1% level which means that the partial effect of market concentration is likely to differ at different concentration levels. The effect of a point increase in HHI, holding all else constant, would be $-\ 0.27\ +\ 0.000072(HHI)$. This means when the HHI is at 5000, a point increase in HHI, holding all else constant, results in a $0.09 increase in insurance paid. The coefficient for concentration is negative while the coefficient for concentration^2 is positive. This means that the partial effect of concentration on the predicted insurance premium is negative at first but the effect will eventually turn positive at a point. The point when the partial effect becomes zero and turns positive would be when HHI is at 3,774. This makes economic sense as a firm growing in market share could take advantage of economies of scale and enjoy lower marginal costs which could be passed down to the consumer. At higher concentrations, however, a firm could use its monopoly power to raise prices.

The economic significance of the effects of market concentrations should also be considered. According to the 2023 Merger Guidelines from the DOJ, a change in HHI of more than 100 points is a significant increase. Even when the market concentration is very high with an HHI of 9000, a 100-point increase in HHI results in a predicted increase of insurance premiums of approximately $38. This is less than the standard deviation of $535 for insurance premiums in the data.

The price level variable is also significant at the 1% level. Holding all other variables fixed, a one-point increase in Regional Price Parity increases the health insurance premium by $65.71.

The Black population percentage is significant at the 10% level. Holding all other variables fixed, a one percentage point (0.01) increase in the Black population relative to the White population decreases the health insurance premium by $19.76.

The Asian population percentage is significant at the 1% level. Holding all other variables fixed, a one percentage point (0.01) increase in the Asian population relative to the White population decreases the health insurance premium by $63.51.

The other race population percentage is significant at the 10% level. Holding all other variables fixed, a one percentage point (0.01) increase in the other population relative to the White population increases the health insurance premium by $26.54.

The West variable is significant at the 5% level. Holding all other variables fixed, a state in the West has its insurance premium lowered by $482.69 relative to the South.

Diagnostic Tests

Table 3. Heteroskedasticity Test - White Test

```
. regress ehatsq yhat yhatsq
```

Source	SS	df	MS	Number of obs	=	51
				F(2, 48)	=	0.71
Model	1.3878e+10	2	6.9391e+09	Prob > F	=	0.4950
Residual	4.6678e+11	48	9.7245e+09	R-squared	=	0.0289
				Adj R-squared	=	-0.0116
Total	4.8065e+11	50	9.6131e+09	Root MSE	=	98613

ehatsq	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
yhat	198.6575	623.3865	0.32	0.751	-1054.745	1452.06
yhatsq	-.015344	.0411155	-0.37	0.711	-.0980122	.0673242
_cons	-562134	2357567	-0.24	0.813	-5302340	4178072

The White Test F-statistic shows that we fail to reject the null hypothesis of homoscedasticity at the 5% level. Therefore there is no evidence of heteroskedasticity.

Table 4. Multicollinearity Test - Variance Inflator Factor Test

```
. vif
```

Variable	VIF	1/VIF
hhi_sq	28.80	0.034724
hhi	28.41	0.035199
income	14.32	0.069831
pricelevel	10.90	0.091773
poverty	9.19	0.108763
black	5.76	0.173541
west	4.31	0.232276
hispanic	3.80	0.263201
northeast	3.68	0.271954
asian	3.66	0.273254
patient	3.31	0.302004
midwest	3.07	0.325878
other	2.89	0.345678
unemployed	2.47	0.405322
Mean VIF	8.90	

The mean VIF is under 10 therefore it can be concluded that the model passes the multicollinearity test. The model does not have near-perfect linear combinations of two variables.

Table 5. Normality Test - Skewness-Kurtosis All Normality Test

```
. sktest premium

Skewness and kurtosis tests for normality
```

				Joint test	
Variable	Obs	Pr(skewness)	Pr(kurtosis)	Adj chi2(2)	Prob>chi2
premium	51	0.0497	0.2158	5.23	0.0733

The insurance premium variable failed to pass the normality test as the skewness test is significant at the 5% level. The joint test chi^2 statistic is also significant at the 10% level. This can be remedied by taking the natural logarithm of the variable.

Table 6. Ramsey Test

```
Ramsey RESET test for omitted variables
Omitted: Powers of fitted values of premium

H0: Model has no omitted variables

F(3, 33) =    0.38
Prob > F = 0.7651
```

The F-statistic is not significant at the 5% level therefore we fail to reject the null hypothesis and the model has no omitted variables.

Alternate Specifications

Because the normality test failed, the natural log of the insurance price premium is used in the model. The new model now passes the Skewness-Kurtosis All Normality Test.

Table 7. Alternate Model Normality Test

```
. sktest lpremium

Skewness and kurtosis tests for normality
                                                    ——— Joint test ———
    Variable |        Obs    Pr(skewness)    Pr(kurtosis)   Adj chi2(2)   Prob>chi2
-------------+-------------------------------------------------------------------
    lpremium |         51          0.1841          0.4485          2.47      0.2915
```

The test for skewness and kurtosis are now both not statistically significant at the 5% level.

Table 8. Alternate Linear Regression Model

Source	SS	df	MS	Number of obs	=	51
				F(14, 36)	=	8.09
Model	.192632687	14	.013759478	Prob > F	=	0.0000
Residual	.06122009	36	.001700558	R-squared	=	0.7588
				Adj R-squared	=	0.6651
Total	.253852777	50	.005077056	Root MSE	=	.04124

lpremium	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
hhi	-.0000348	.000016	-2.18	0.036	-.0000673	-2.42e-06
hhi_sq	4.65e-09	1.52e-09	3.06	0.004	1.57e-09	7.73e-09
patient	.0000157	9.50e-06	1.65	0.107	-3.57e-06	.000035
poverty	.0044283	.0066849	0.66	0.512	-.0091292	.0179858
unemployed	1.531401	.912386	1.68	0.102	-.3190032	3.381806
pricelevel	.0087207	.0027184	3.21	0.003	.0032075	.0142339
income	1.01e-06	1.95e-06	0.52	0.606	-2.94e-06	4.97e-06
black	-.2632893	.1384631	-1.90	0.065	-.5441055	.0175269
hispanic	.0683578	.1074498	0.64	0.529	-.1495605	.2862761
asian	-.8338429	.2014748	-4.14	0.000	-1.242453	-.4252331
other	.3182194	.1998589	1.59	0.120	-.0871133	.7235521
west	-.0661124	.0274925	-2.40	0.021	-.1218699	-.010355
midwest	.0059961	.0238468	0.25	0.803	-.0423675	.0543597
northeast	-.0090256	.0290461	-0.31	0.758	-.0679338	.0498826
_cons	7.919154	.2155561	36.74	0.000	7.481986	8.356322

All the variables that were statistically significant in the initial model are still statistically significant in the alternate model, however, the intercept in the alternate model is now also statistically significant at the 1% level.

For the HHI concentration variable, the point when the partial effect becomes zero and turns positive is when HHI is at 3,742 which is very close to the initial model. A 100-point increase in HHI when HHI is at 9000, holding all other variables fixed, increases the predicted log(premium) by approximately 0.00489 which is still a small partial effect.

Conclusions

The empirical results show that the Herfindahl-Hirschman Index for the health insurance market is a statistically significant predictor of the total health insurance premium. This conclusion is consistent with the results from Ho & Lee (2017) and Dafny et al. (2012). The

economic significance of market concentration on health insurance premiums however is small. Even when a market is heavily concentrated, giving a firm monopoly power, a 100-point increase in HHI still has a low partial effect on health insurance premiums.

The model's estimate of the effect of market concentration on insurance premiums also demonstrates that increases in concentration can result in lower price premiums up to an HHI of approximately 3,700 points. This could be the result of economies of scale from insurance companies as they may have better bargaining power and negotiate better prices from hospitals. This suggests that antitrust departments and agencies such as the Department of Justice and the FTC should be more lenient with evaluating mergers between private health insurance firms when market concentrations are below that point as it could result in lower prices.

The results of this study have some limitations. While this study does try to control for time by using data from 2021, estimates from government agencies can still have different timings during the year. Additionally, the data used for the unemployment rate is from December of 2021 which may not be accurate with the timings from the other data. Furthermore, because the data used are aggregated to the state level, it may mask some variances in the lower level. Because the data for health insurance premiums come from a survey of employer-sponsored offerings, there could be reporting bias in the results. There could also be reverse causality where lower health insurance premiums could attract certain demographics which may be seen in the effect of the Asian population percentage which was statistically significant. There may also be an endogeneity issue between market concentration and insurance premiums as higher demand for insurance premiums in the form of prices may be correlated with concentration. The distribution of health insurance premiums is also not normally distributed, however it can be remedied with a log transformation.

Further research on the effect of market concentration on insurance premiums across time may give more insight into determining the effect of concentration on insurance premiums.

Other factors that may affect insurance premiums such as public health should also be looked upon.

References

Statista Research Department, & 23, F. (2024, February 23). Most important voter issues U.S. 2024. *Statista*. https://www.statista.com/statistics/1362236/most-important-voter-issues-us/

Sen, A. (2015). Universal health care: The affordable dream. *HPHR Journal*, *2015*(5). https://doi.org/10.54111/0001e/2

Wager, E., McGough, M., Rakshit, S., Amin, K., & Cox, C. (2024, January 23). *How does health spending in the U.S. compare to other countries?*. Peterson-KFF Health System Tracker. https://www.healthsystemtracker.org/chart-collection/health-spending-u-s-compare-countries/#GDP%20per%20capita%20and%20health%20consumption%20spending%20per%20capita,%202022%20(U.S.%20dollars,%20PPP%20adjusted)

Schansberg, D. E. (2014). The Economics of Health Care and Health Insurance. *The Independent Review*, *18*(3), 401–420. http://www.jstor.org/stable/24563243

Dafny, L. S. (2015). Evaluating the impact of health insurance industry consolidation: Learning from experience. *The Commonwealth Fund*. https://doi.org/10.15868/socialsector.25054

KFF. (2023, December 6). Large Group Insurance Market Competition. *KFF*. https://www.kff.org/other/state-indicator/large-group-insurance-market-competition/?activeTab=graph&currentTimeframe=0&startTimeframe=8&selectedDistributions=herfindahl-hirschman-index-hhi&sortModel=%7B%22colId%22%3A%22Location%22%2C%22sort%22%3A%22asc%22%7D

Bradford, W. D., & Mobley, L. R. (2004). Employment-based health insurance and the effectiveness of Intrafirm competition between insurance providers. *Southern Economic Journal*, *70*(4), 1012. https://doi.org/10.2307/4135285

Ho, K., & Lee, R. S. (2017). INSURER COMPETITION IN HEALTH CARE MARKETS. *Econometrica*, *85*(2), 379–417. http://www.jstor.org/stable/44955124

Dafny, L., Duggan, M., & Ramanarayanan, S. (2012). Paying a Premium on Your Premium? Consolidation in the US Health Insurance Industry. *The American Economic Review*, *102*(2), 1161–1185. http://www.jstor.org/stable/23245448

Frank, R. G., & Lamiraud, K. (2009). Choice, price competition and complexity in markets for health insurance. *Journal of Economic Behavior & Organization*, *71*(2), 550–562. https://doi.org/10.1016/j.jebo.2009.04.005

U.S. Department of Justice and Federal Trade Commission. (2023) Merger Guidelines. *U.S. Department of Justice*