Using Difference in Differences to Understand the Effect of the Affordable Care Act (ACA) on the demographics of Federally Qualified Health Centers (FQHCs)

Aparajita Kaphle

Abstract

Recent studies have argued that the 340B program and subsequent expansions that increase eligibility has actually lead to health centers (FQHCs) opening up in wealthy areas rather than the low-income areas they were designed to help. Using the Affordable Care Act as a treatment variable, I utilize a difference in differences test to understand if adoption of the ACA and its expansions of FQHC eligbility has lead to more FQHCs opening in wealthier areas. I find that in states where the ACA was implemented, it has lowered the income of where FQHCs are located compared to states where ACA was not implemented.This indicates that perhaps FQHC eligibility expansion will allow the 340B program to work as intended.

1 Introduction

(See github link: https://github.com/akaphle01/AQRDFINALPROJ)

In this investigation I plan to employ a difference in differences model to analyze the effects of the Affordable Care Act (ACA) on the demographics of where FQHCs are opening up. I will be comparing the outcomes between states that adopted the ACA vs States that did not adopt the ACA using 2011 and 2019 as my before and after treatments. The ACA which was passed in 2010 but implemented in 2014 was adopted by some states and not by others giving me the experimental setting to examine if FQHCs opened in 2019 in ACA adoption states were located in wealthier areas than their non-ACA adoption states. To do so I will be treating ACA implementation as the treatment, the treatment group to be the States that implemented ACA and the control group as the non-ACA group. It should be acknowledged that the states that did not implement the ACA were largely Southern states and to try and control for this regional variation I have also included controls that take into account the rural/urban demography of an FQHC location as well as other variables such as unemployment. I also implemented a fixed effects model but lacking a lengthy enough panel data, I ultimately decided upon relying on One-way Linear Regression that later I carried out validation tests.

2 Literature Review

Enacted in 1992 to support rural hospitals and health centers that serve disadvantaged communities, the 340B program has come under fire in recent years as many hospitals and health centers have abused the program for their profit.[1] The passage of the Affordable Care Act allowed for an expansion of what health centers qualify as an FQHC and as a result, the number of FQHCs has rocketed. With this increase, some journalistic and empirical research indicates that FQHCs registered in the 340B program have steadily risen in wealthy and affluent neighborhoods (Conti, 2014).

The 340B program requires pharmaceutical companies to provide discounts for healthcare facilities offering care to low-income and disadvantaged communities. These discounts are most useful when treating uninsured residents as many patients are not able to pay out of pocket for some medication. However, if healthcare facilities within the 340B program treat someone with insurance, the hospital is able to charge insurance for the full cost of the drug and reap the profits between the discounted drugs and the insurance payment. This profit was originally intended to assist healthcare facilities in their day-to-day upkeep.

However, over the years, lack of reporting on the investments of healthcare facilities and expansions of eligibility into the 340B program has allowed healthcare facilities to open clinics in wealthy areas, serving primarily affluent communities and pocketing the profits generated from a heavily-insured population. Since its implementation, the program has only expanded, allowing Family Planning Centers to be eligible in 1998, children's hospitals in 2006. By 2017 there were over 12,000 covered entities and as of 2015, 40% of all US hospitals were enrolled as 340B entities (Mulcahy, 2014).

Under the Affordable Care Act, Federally Qualified Health Center (FQHCs) eligibility for 340B program was expanded. FQHCs differ from hospitals as they are explicitly defined as: "primary care clinics that receive federal funds to provide healthcare services to underserved communities. They operate in both rural and urban areas

[1]"A Little-Known Windfall for Some Hospitals, Now Facing Big Cuts" 2019. *nytimes.com*. https://www.nytimes.com/2018/08/29/upshot/a-little-known-windfall-for-some-hospitals-now-facing-big-cuts.html.

designated as shortage areas."[2] While this policy was originally intended to serve vulnerable populations, recent work of journalists have shown that 340B hospital-affiliated clinics were likely to be in more affluent areas with higher rates of insurance coverage. Previous investigations conducted by journalists at the New York Times have uncovered many nonprofit hospitals benefiting from the 340B program such as Bon Secours nonprofit hospital and beneficiary from the 340b program reaping over $100 million dollars in 2021.[3]

As this is a part of a greater independent project, the next section will focus briefly on explaining the dataframe and the data in more depth as it is a rather messy dataset that had gone through various stages of cleaning.

3 Preliminary Data Discussion

The dataframe is composed of the names of FQHCs opened in 2011 and 2019, with corresponding census data referring to the county in which the FQHC opened up. To begin analysis, I selected for only a small subset of variables and identifiers that were necessary for the analysis at hand. The FQHC data was manually downloaded and cleaned from the Office of Health Resources and the Census data was retrieved via an API. This is an ongoing project that aims to integrate more years into the research and thus generate a more panel-level analysis akin to Scheve and Schavage. As it currently stands, this paper is largely inspired by the minimum wage analysis conducted by David Card and Alan Krueger in their paper: "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania" in addition to the OSHA case that looked into various prediction methods.

For the purpose of the OLS/fixed effects, we created new variables such as "state_status" which indicates if a state is belonging to a category of states that adopted the ACA vs states that did not adopt it.

4 Analysis

The intuition behind my research proposal can be seen in the set of visualizations in Figure 1 as well as the maps in Figure 2 and Figure 3 that depict the openings of FQHCs between 2011 and 2019.

[2]"Community Health Center Overview" *portal.ct.gov/*. https://portal.ct.gov/DPH/Family-Health/Community-Health-Centers/Community-Health-Center-Overview#:~:text=Federally%20Qualified%20Health%20Centers%20(FQHC)%20are%20health%20centers%20that%20receive,designation%20from%20the%20U.S.%20Dept..

[3]"How a Hospital Chain Used a Poor Neighborhood to Turn Huge Profits" *https://www.nytimes.com*. https://www.nytimes.com/2022/09/24/health/bon-secours-mercy-health-profit-poor-neighborhood.html.

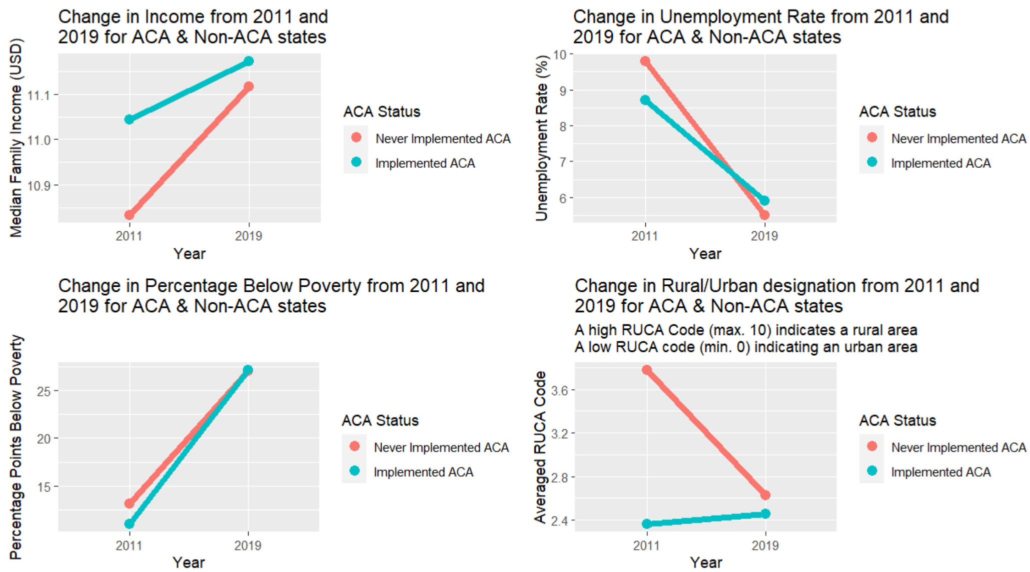

Figure 1: FQHC County Median Income compared between 2011 and 2019

From the visualizations, one can notice can notice that the median family income for the FQHCs in 2019 has grown for both states that implemented vs have not yet implemented the ACA. However, many events may have triggered this change including the potential that the states themselves got richer and therefore it appears that the FQHCs are in wealthier areas comparatively rather than relatively. What I expected to see if my hypothesis was correct was not only the FQHCs demographic get wealthier for ACA implemented states in 2019 to non-ACA implemented states but that compared to 2014, the difference between the two state categories is signifcantly higher.

To get a better understanding of the demographics, we also created two summary tables that allowed us to see both regional variation between states that implemented the ACA vs states that did not implement the ACA. Nevertheless, the tabular representation of regional variation in Figure 4 and Figure 5 helps us understand further the discrepancies in demographics such as median family income, unemployment, rural/urban community.

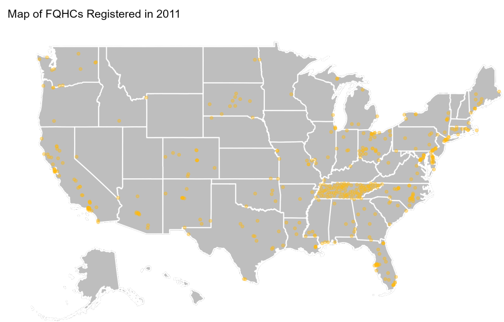

Figure 2: Map of FQHCs opened in 2011

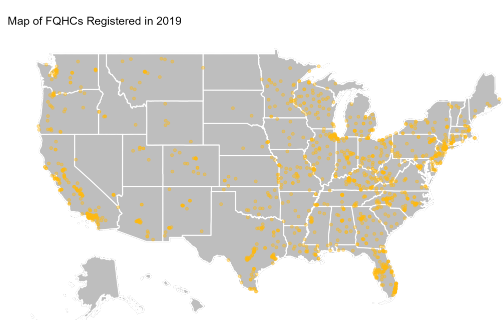

Figure 3: Map of FQHCs opened in 2019

FQHC and Census Data by Region and Implementation of ACA

	Median White Population Percentage (%)	Median Black Population Percentage (%)	Median Percentage Below Poverty (%)	Median Estimated Family Income (USD)	Median Percentage On Cash Assistance (%)	Median Unemployment Rate (%)	Average Population	Number of Observations
Never Implemented ACA								
North Central	91.15	1.45	26.05	62,238.50	2.00	4.6	147097.2	130
Northeast	93.20	0.90	10.00	67,897.00	4.20	5.5	135730.2	15
South	73.50	17.40	25.20	61,112.00	1.90	6.2	620668.6	688
West	89.65	0.75	21.65	80,248.00	1.80	4.2	265883.0	18
mean	86.88	5.12	20.73	67,873.88	2.48	5.12	292,344.76	—
sd	9.03	8.19	7.40	8,767.36	1.15	0.90	226,658.16	—
Implemented ACA								
North Central	85.60	7.25	27.70	67,977.00	2.50	6.2	858553.3	286
Northeast	73.00	12.90	25.10	76,930.00	3.20	6.2	1025380.7	309
South	79.60	11.60	31.20	59,205.00	2.00	7.1	148717.2	187
West	70.70	4.00	22.60	74,092.00	3.40	6.3	2442834.1	475
mean	77.22	8.94	26.65	69,551.00	2.77	6.45	1,118,871.34	—
sd	6.74	4.08	3.68	7,844.05	0.64	0.44	961,003.39	—

Figure 4: Summary Statistics 1

FQHC and Census Data by Region

	Median White Population Percentage (%)	Median Black Population Percentage (%)	Median Percentage Below Poverty (%)	Median Estimated Family Income (USD)	Median Percentage On Cash Assistance (%)	Median Unemployment Rate (%)	Average Population	Number of Observations
2011								
North Central	90.20	3.60	10.00	57,909.00	2.80	8.6	413890.3	73
Northeast	80.30	8.80	8.80	66,846.00	3.20	8.4	615854.3	53
South	75.40	16.20	13.20	50,611.00	2.20	9.8	321700.9	236
West	75.30	3.05	12.40	62,595.00	3.40	9.2	2564000.9	112
mean	80.30	7.91	11.10	59,490.25	2.90	9.00	978,861.60	—
sd	7.00	6.10	2.05	6,954.33	0.53	0.63	1,063,875.58	—
2019								
North Central	86.20	4.70	28.90	68,824.00	2.30	5.0	683542.0	343
Northeast	71.60	13.50	25.10	78,786.00	3.30	5.9	1056230.0	271
South	75.10	17.40	27.50	65,393.00	1.80	5.8	592971.4	639
West	72.50	3.50	24.10	76,673.00	3.40	6.1	2304367.5	381
mean	76.35	9.78	26.40	72,419.00	2.70	5.70	1,159,277.70	—
sd	6.73	6.76	2.19	6,348.86	0.78	0.48	789,277.46	—

Figure 5: Summary Statistics 1

From the summary table, we can notice some patterns that is worth investigating further, for example, with regards to Figure 4, the states that never implemented the ACA compared to the states that eventually did, the median income of the FQHCs is high in all regions with the exception of the South where it decreased from \$61,112 to \$59,205.It appears however that states that implemented ACA have a higher median income (\$69,551) than states that did not (\$67,873.88). Interestingly however is that unemployment and percentage on cash assistance is also higher for the counties served by the FQHCs in areas where the ACA was implemented. Although this is interesting, it is not enough to deduce a causal effect from. Furthermore, looking at Figure 5 we notice that compared to 2011 (when ACA had not been passed federally), 2019 had a notable increase in median income across all regions. Prior to the enacting of the ACA, the average median income is around 59,490 dollars, a figure that grew to 72,419 after the implementation of ACA. Looking more regionally, we can notice that the South is notably ranked the lowest in median income compared to other regions and while the FQHCs in the South grew in income in 2019 (Post-ACA), it is not nearly as high as the other regions. This is also interesting to note as the many southern states did not implement the ACA and yet still have a increase in FQHC income (\$50,611 median in 2011 vs \$65,363 median in 2019). Other statistics of note is that median percentage below poverty grew Post-ACA across all regions from 11.10% to 26.40% while median unemployment rate decreased from 9% to 5.70% (although this may be due to the aftermaths of the 2008 housing crisis).

In order to address the variation present in income, we employed a difference in differences model. While we did also run a fixed effects model and doing so is feasible (as soon in the equations below), we opted to focus primarily on the OLS models for ease of interpretability.

The equations we employed for OLS were:

$$I_{it} = \alpha + \beta_1(ExpandedStates)_i + \beta_2(PostACA)_t + \beta_3(ExpandedStates_i x PostACA_t) + \gamma X + \epsilon,$$

where I_{it} is the log median family income, β_1 represents the coefficient for when a state has expanded and implemented the ACA, in that case, i = 1 when referring to expanded states and i = 0 when referring to non expansion states, β_2 represents the time period and in reference to the Federal Passage of the ACA (not to confused with the state level adoption). In this case, t = 0 to Pre-ACA and t = 1 to Post-ACA. Finally β_3 is the interaction between the states that have implemented the ACA in the post-ACA time period. γ represents a vector of coefficients for controls such as rural/urban and unemployment rate.

A fixed effects equation that would capture the same thing would be:

$$I_{it} = \alpha_i + \theta_t + \beta_{it}(PostACA) + \gamma_{it}X + \epsilon_{it}$$

Where i in this case indicates the state, the t is referring to the time. We used α to represent state fixed effects and θ to represent time fixed effects. Ultimately however due to the 2 by 2 nature of the data, we opted primarily to use the OLS form in running our equations.

To run a difference in differences model we have to make some significant assumptions of our findings which unfortunately based on the available data set are difficult to confirm. Firstly, we have to make the parallel trends assumption whereby we suppose that absent treatment (ACA), the states that would've implemented the ACA would have trended the same way (in parallel) to the states that did not implement ACA. As our data is not yet part of a larger panel series, we do not have a method to confirm this assumption (although it should be noted that it is impossible to fully confirm the parallel trends assumption). Potential methods to do so however as the project possesses, would require utilizing a lags and leads, looking to see if leads, specifically indicate certain trends that would be happening absent of treatment.

5 Results

Below we can notice the results of the regression.

Model Results for Log Median Family Income				
	OLS Short Regression Model	OLS Regression Model with Unemployment Controls	OLS Regression Model with Rural/Urban Controls	OLS Long Regression Model
(Intercept)	10.849***	11.470***	11.017***	11.547***
	(0.016)	(0.024)	(0.016)	(0.021)
States where ACA was Adopted	0.195***	0.162***	0.132***	0.113***
	(0.022)	(0.018)	(0.020)	(0.016)
Post-ACA Federally Passed	0.244***	-0.001	0.192***	-0.019
	(0.018)	(0.017)	(0.017)	(0.015)
Effect of Post-ACA Federally Passage on States where ACA was Implemented	-0.085***	-0.035+	-0.029	0.006
	(0.025)	(0.021)	(0.022)	(0.019)
Unemployment Rate (%)		-0.065***		-0.058***
		(0.002)		(0.002)
RUCA Designation			-0.045***	-0.037***
			(0.002)	(0.002)
R2	0.178	0.440	0.342	0.550
RMSE	0.24	0.20	0.21	0.18
+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001				

Figure 6: Table2: OLS Model Summary

The intercept of 10.849 indicates that on average States that would not have implemented the ACA prior to 2014, would have on average a median log family income of 10.849. The difference across time between states that have not implemented the ACA is captured by the coefficient reflecting the 2014 ACA federal adoption which was .243. This indicates that from 2011 and 2019, States that were not going to implement the ACA saw a 24.3% increase in income in FQHC counties. The difference in differences was captured by the Effect of Post-ACA Federal Passage on States where ACA was Implemented which was -.0848 which indicates that the post the passage of the ACA, the states that implemented ACA actually had a decrease in FQHC county income of around 8.48%.

This does not support our hypothesis as we can notice that ACA implementation actually leads to a decrease in log median family income in FQHC counties. When pursued further controls, this pattern is still prevelant as this coefficient remains negative until the long regression where it is a small positive (however this is not signficant). In the long regression, we controlled for unemployment rate and Rural/Urban Designation and found that regardless of ACA impacts, the counties where the FQHCs are located are most impacted by unemployment or being rural/urban as a percentage point increase in unemployment rate, ceteris paribus, is associated with a -5.8% decrease in income and a rural county has an median income that is -3.7% less than their urban counterparts. Adjusting for these controls allowed us to recognize that the causal relationship we are supposing between ACA implementation on county income is not as strong under this experimental design as once accounting for variables such as unemployment rate and RUCA designation, the significance diminishes. In employing such controls we are able to understand some relationships that may be of note.

To test our model's accuracy we also split our data into training and test data, beginning first by separating out the years of the data and recombining as there are more FQHCs in 2019 than in 2011 and we didn't want to over represent one group than the other. The results of the split dataset can be seen in Figure 7 with the RMSE being the lowest

Testing and Training of OLS Models		
	RMSE	
OLS Models	Training	Testing
Short Reg	0.235	0.247
Short Reg with Unemployment Controls	0.193	0.205
Short Reg with Rural/Urban Controls	0.211	0.217
Long Reg	0.173	0.183

Figure 7: RMSE Table

for the long regression model that utilized all controls. Both the training and test RMSE were the lowest out of the dataset. This makes sense as we only have a small set of variables therefore we avoid the issue of overfitting while still making sure our prediction is as accurate as possible in predicting log income. If we want to choose a model that has a more interpretable coefficient β_3 we can opt for the model with unemployment controls as that had the second lowest training and testing errors. In this version the coefficient is still negative but is -.035 which translates to the ACA implementation leading to a -3.5% change in the log income.

6 Evaluation

The findings of this paper were rather surprising given our hypothesis as well as previous work in the field analyzing the 340B program. From our initial visualizations it was clear that the number of FQHCs had increased substantively from 2011 to 2019. This increase also appeared to correlate with an increase in the income of the areas where the FQHCs were opening in. When looking at the summary statistics, we were able to notice that between 2011 and 2019, the counties where FQHCs were opening were noticeably more wealthy. We wanted to learn if this increase in wealth was due to the ACA decreasing limitation of qualifying FQHCs and therefore FQHCs could open in wealthier areas. We also had to be cognizant that the affects we were associating with the implementation of the ACA was not due to the fact that maybe the counties were already on their way to becoming wealthier regardless of ACA implementation (more economic growth).

To do this, we implemented a difference in differences model with controls that took into account if an area is urban/rural and their unemployment rate. It should be noted that there was some colinearity between urban/rural and unemployment that was not assumed to be significant enough to skew the analysis. When looking at the average log median income we noticed that in 2019 for both states that implemented the ACA and did not implement the ACA the log median income went up. Our difference in difference simple regression model indicates that implementation of ACA leads actually to a decrease of 8.5% in median income of the FQHCs. While the additional of controls such as unemployment and rural/urban makes this decrease smaller and more positive in the case of the long regression, it can be deduced that the ACA instead of leading to higher income environments of FQHCs, there is a subtle decrease, indicating that the FQHCs may actually be helping more vulnerable and low-income areas than previously thought.

This being said, our experimental design has several drawbacks that must be addressed to must adequately prepare for future experiments and derive more robust conclusions. The primary issue with a simple before and after experiment that does not have panel level data is that it neglects effects of time. The composition of US counties have been getting wealthier over time and therefore between 2011 and 2019, the counties of interest grew in wealth mirroring a national pattern. Therefore, in future studies and in fact where this study is expanding towards would take into account panel data and run an interrupted time-series analysis. In doing so, we can account for general increases in county wealth and more precisely accredit any variation to the ACA. Similarly, with more panel level data, our fixed effects would also take into account state level variation and be better able to account if variation was due simply to states.

7 Conclusion

Our original hypothesis purposed that ACA expansion which increased FQHC eligibility and thus lead to more FQHCs in wealthier eras. However our difference in differences paints a different story indicating that the ACA may be helping the FQHCs locate in less wealthy areas compared to the states that did not implement the ACA.

This provides support for ACA-based regulation of FQHC placement as perhaps expanding eligibility will not lead to excessive profit seeking behaviors but rather more FQHCs opening in areas where it would otherwise not be served. Moreover, this data-set uses county-level data which may not be accurately reflecting the income demographics of the FQHC locations as many health centers may be in a "less wealthy" county but are located in wealthy zipcodes or in areas where they are able to still engage in profit-seeking behaviors. In fact according to the Department of Commerce, geographic income inequality has increased more than 40% between 1980 and 2021.[4]

[4]"Local incomes have become more unequal over time" *https://www.commerce.gov/*. https://www.commerce.gov/news/blog/2023/06/geographic-inequality-rise-us#:~:text=Geographic%20income%20inequality%20has%20risen,of%20metropolitan%20and%20micropolitan%20areas..

As the ACA and other pieces of legislation are passed and FQHCs continue to pop up in number, it is nevertheless vital to continually monitor where these FQHCs are located in so that they can continue to serve the population they were designed to help.

References

Conti, R. M. (2014). The 340b drug discount program: hospitals generate profits by expanding to reach more affluent communities. *Health Affairs (Project Hope)*, 33(10):1786–1792. PMID: 25288423 PMCID: PMC4591849.

Mulcahy, A. W. (2014). The 340b prescription drug discount program: Origins, implementation, and post-reform future. Technical report.