Vol. 2
Spring 2024

The Emerging Economist

The Undergraduate Journal of Economics at St. Olaf College

ST. OLAF COLLEGE

Omicron Delta Epsilon

International Honor Society in Economics
Beta Chapter of Minnesota at St. Olaf College

About Omicron Delta Epsilon

With nearly 700 chapters, Omicron Delta Epsilon is one of the largest honor societies in the world. The purpose of the society is to encourage and recognize high scholastic attainment in economics and to establish closer ties between students and professors both within and across institutions of higher learning.

St. Olaf's chapter, the Beta Chapter of Minnesota, is dedicated to recognizing and celebrating outstanding work in the Department of Economics. This includes publishing an independent in-house journal that showcases papers produced by students in St. Olaf economics classes.

Executive Board 2023-2024

The Emerging Economist

The Undergraduate Journal of Economics at St. Olaf College

Spring 2024

Acknowledgements

St. Olaf College's chapter of Omicron Delta Epsilon, the Beta of Minnesota, is delighted to present the second edition of our in-house journal, *The Emerging Economist*. We are excited to give students the opportunity to be recognized for their outstanding work during the course of the year, and we hope to set an example for years to come on the quality of student work that Oles can produce.

We would like to thank Professor of Economics Ashley Hodgson for encouraging the work done on this journal, and Michelle Potter-Bacon for her administrative support in helping to make this journal publication a reality. We would also like to acknowledge the tireless efforts of our executive team and the group of student reviewers who carefully read and deliberated on each submission. We hope you enjoy this second edition of *The Emerging Economist*.

Thank you!

Hannah Peschel, *Co-Editor*
Benjamin Reinhard, *Co-Editor*

Contents

Nominal Gas Price Shocks and Inflation Expectations: An Analysis of the United States and the European Union

Jenna Peschel

I. Introduction

The relationship between gas prices and inflation expectations has been of interest in economic study for many years, and their relationship is generally well known. In the post-COVID-19 era, we have seen an extreme volatility in gas prices across the entire United States, and more recently, we have seen a stark increase in gasoline prices. The European Union has experienced similar situations in their post-COVID-19 economy, but they have also faced extreme volatility in gasoline prices since the Russian invasion of Ukraine. Both of these factors indicate that the relationship between gas prices and inflation expectations must continue to be studied. In light of this volatility, the most up-to-date information available must be used to correctly specify this relationship and its policy implications. Furthermore, the salience of gas prices to consumers across the globe makes this study relevant. Gas prices are highly salient to consumers, not only because they purchase and use this resource frequently, but also due to the large signs advertising their prices. Because of how exposed consumers are to the price of gasoline, gas prices are often used as an indicator of economic expectations. In this paper, I will analyze the effect of nominal gas price shocks on short-run and long-run inflation expectations in both the United States and the European Union. I will compare the results of these analyses to see what conclusions can be drawn about similarities or differences between consumer behavior and expectations in these two economies.

I find that gas price shocks significantly impact both 1 and 5 year inflation expectations in the U.S. The effect is larger in magnitude in the short-run but is more stable in the long-run. I also find that after 1 year, gas price shocks explain more variation in 1 year inflation expectations than in 5 year expectations. For the EU, I find that shocks to gas prices significantly affect inflation expectations at the 1 year but not at the 5 year mark, and more variation in 1 year inflation expectations is explained by gas prices shocks 1 year after the shock than 5 year inflation expectations.

The paper is structured as follows. Section 1 gives an overview of the existing literature and research done on this relationship and indicates how my research will fill in gaps in the current work. Section 2 introduces the variables with which I will be working and offers exploratory data analysis and descriptive statistics on variables of interest. Section 3 contains information about the empirical methods; I will outline the structural VAR models and their assumptions used to analyze the relationships I have described above. Section 4 presents results and a discussion of these results, with Section 5 concluding.

II. Literature Review

Existing literature has analyzed the relationship between gas prices and inflation expectations at the U.S. State and National levels (Binder 2018; Kilian and Zhou 2023), as well as internationally in European and Eastern Asian countries (Kilian and Zhou 2023). These analyses historically have included both core and headline inflation in their models, typically finding that gas prices explain less of the variability in core inflation than for headline inflation. This makes sense, as gas is not included in core inflation, but is included in headline. Kilian and Zhou (2023) further explain that rising gas prices are a symptom, rather than a cause, of high inflation in the U.S. Using this language, they analyze how gas prices, apparently being a symptom of high inflation in the U.S., alter consumers' expectations of inflation. Their main finding is that there is no evidence in the U.S. that energy price shocks (which include shocks to gas prices) have materially changed long-run inflation expectations. Binder (2018) also analyzes the effect gas prices have on consumer inflation expectations, but she differentiates between inflation expectations and inflation perceptions, a difference that other analyses don't distinguish as obviously.

Both Binder (2018) and Zhou and Kilian (2023) use the Michigan Consumer Survey Data on inflation expectations, which is reported on a monthly frequency. Binder and Makridis (2020) use Gallup Survey Data to investigate the effect of gas price changes on what they call "consumer sentiment." This data has a daily frequency but is also microdata collected on the state level. Binder and Makridis discuss that daily data is more preferred for gas prices because of its volatility; they argue that since gas prices change so frequently, current and frequent data is needed to truly analyze their association with other economic measures.

The most recent use of the Michigan Consumer Survey data to investigate this relationship is by Kilian and Zhou (2023), which uses data up until 2023. My analysis will use data through August 2023. While this may not seem like a significant difference, the volatility of gas prices– especially as the weather changes, as described by Mu (2007)– makes these additional few months valuable in finding the true association.

This volatility of gas prices has greater significance to this analysis than simply demonstrating the need for current data. Binder (2018) explains how volatility of gas prices may lead to consumers placing a disproportionately large weight on gas prices in their belief on inflation expectations. She ultimately finds that consumers do not disproportionately take gas prices into account in determining long-run inflation expectations. In fact, she finds that the impact of gas prices on inflation expectations fade quickly in forecasts. Furthermore, Binder (2018) and Georganas et. al. (2014)

discuss another important property of gas prices in determining their effect on inflation expectations: salience. Binder argues that gas is one of the most salient products to a consumer because of how frequently it is purchased, and thus they are much more aware of changes to the price. This is termed the "frequency hypothesis." When consumers purchase things more frequently, they are more aware of their price changes and are more likely to use them when they form expectations about the future (Bruine de Bruin, 2011). This hypothesis, even before it was formally termed as such, has been the catalyst for many of the recent studies into the association of gas prices and inflation expectations, both in the U.S. and the rest of the world.

The recent analysis done by Boeck and Zörner (2023) on this relationship in the European Union is helpful to understanding my analysis. To account for the lack of explicit inflation expectation data in the EU, Boeck and Zörner use the harmonized index of consumer prices (HICP) measure, which comes from the Survey of Professional Forecasters Forecasts. The HICP measures how prices across the entire European Union change. They use the forecasts for the HICP as a measure of inflation expectations. See Section 3 for more specifics on this. Boeck and Zörner use this data along with a structural VAR model to estimate how natural gas price shocks impact both short and long run inflation expectations. They find that shocks to natural gas prices in the EU cause inflation expectations to increase, but these increases are much more pronounced in the short-term than the long term, which is consistent with what other research has shown.

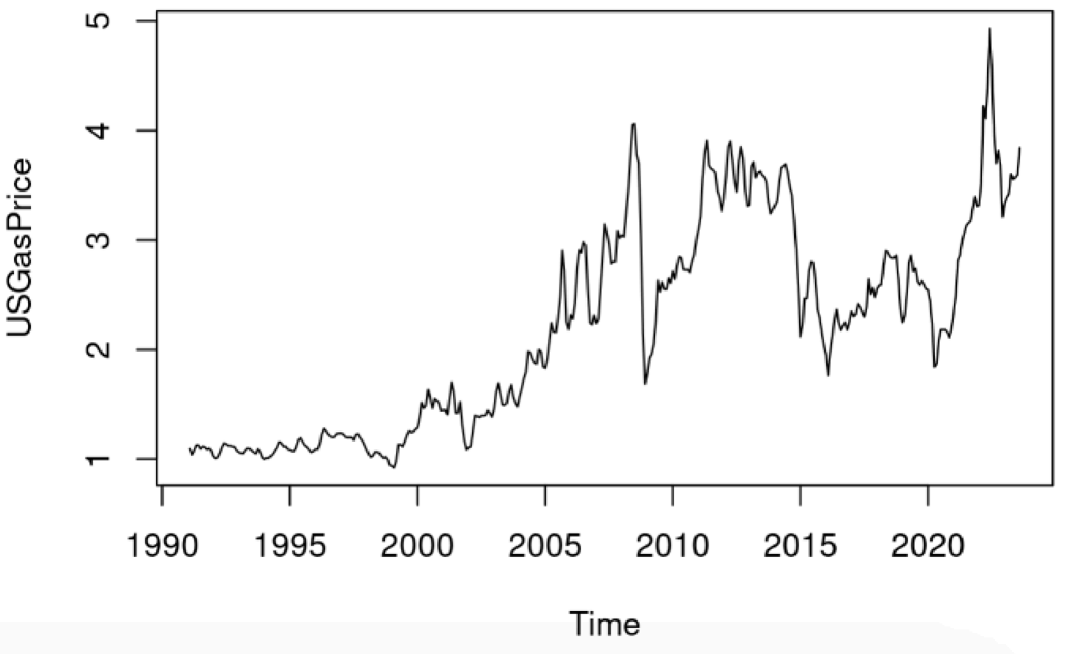
In this preliminary research, the consensus is that while gas is salient to consumers and the price of gas is volatile compared to other resources, gas prices do not significantly impact consumers inflation expectations above and beyond other non-gas goods in the long run. However, there is some evidence that inflation expectations are very sensitive to crude oil prices. The differentiation between gas prices and crude oil prices is periodically taken into account in existing research. In my analysis, I will only test gas prices and their effect on the short-run and long-run inflation expectations, using up-to-date data provided by the Michigan Consumer Survey. This will allow me to investigate if, and by how much, evidence has changed since the previous analyses were run using the data. I hypothesize that the gas prices will have more of an impact on short-run inflation expectations, but this impact will not continue into the long run forecasts of inflation expectations.

III. Data Description

A. U.S. Analysis

For my U.S. analysis, I will use gas price data from the U.S. Energy Information Administration (EIA). This data is collected weekly from a representative sample of 1000 gasoline outlets and reported in dollars per gallon (EIA 2018). I will use the data on a monthly frequency, starting in February 1991 and going through August 2023. According to the EIA website, "[M]onthly and annual averages are simple averages of the weekly data contained therein. For months and years with incomplete weekly data series, the monthly and/or annual averages are not available" (EIA 2018). I remove observations where the price averages are not available. Figure 1 illustrates how gas prices have changed in the time frame that I will be working with. We see sharp exponential growth between 2000 and 2008, with a dramatic drop happening in 2008. This was largely due to the U.S. Economic Recession that happened in 2008. In July of 2008, the reported gas price was $4.06, whereas by March of 2009, the price had dropped to $1.92. We see another stark decrease around 2014, and an increase around 2021. This increase is associated with the economic growth the U.S. experienced following the COVID-19 Pandemic. The maximum in gas prices occurs in June of 2022, with a reported average price of $4.93. Table 1 below presents descriptive statistics. Overall, the mean gas price in this data is $2.19, with a standard deviation of $0.94.

Figure 1. U.S. Gas Price (dollars/gallon), 1990-2023



Minimum	**Q1**	**Median**	**Q3**	**Maximum**	**Mean**	**SD**	**n**
0.921	1.25	2.19	2.87	4.93	2.19	0.94	391

Table 1. Summary Statistics for U.S. Gas Prices

I will also use the Michigan Consumer Survey for data on inflation expectations. I use monthly data in the same window as the gas price data, February 1991 to August 2023. The survey asks, "[d]uring the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?" and, if the respondent said 'up,' "[b]y what percent do you expect prices to go up, on the average, during the next 12 months?" (University of Michigan Survey Research Center, 2023). The survey also collects data on 5 year price expectations, asking the following questions: "What about the outlook for prices over the next 5 to 10 years?" and "[b]y about what percent per year do you expect prices to go up or down, on the average, during the next 5 to 10 years?" (University of Michigan Survey

Research Center, 2023). As my measure for inflation expectations, I will use the mean value that survey respondents believed prices would go up in the next year or in the next 5 to 10 years, depending on which analysis I am running. Figures 2a and 2b illustrate how inflation expectations have changed over time, both for 1 year expectations and 5 year expectations.

Figure 2a. U.S. 1- Year Inflation Expectations

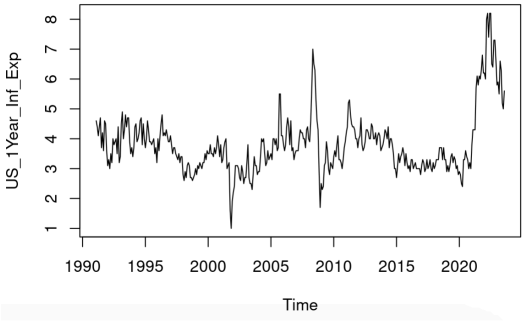
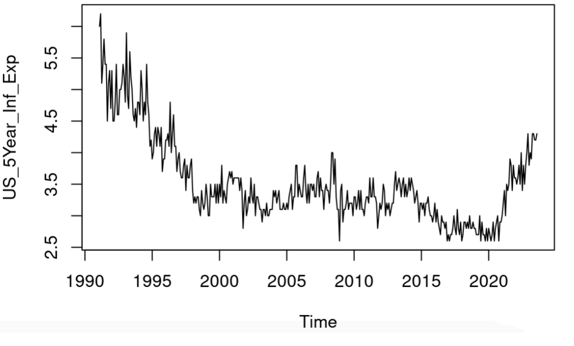


Figure 2b. U.S. 5-Year Inflation Expectations



In the 1 year expectations, we see abnormal activity around 2001 and 2008. The 2001 activity corresponds to the recession that came with the introduction of the internet, or the "dotcom bubble" as it has come to be known in the U.S. The 2008 trends again correspond with the recession that hit the U.S. in 2008. We also see a large increase after 2020, which reflects the U.S. economy coming out of the global pandemic and the consumer belief that an increase in economic activity will lead to an increase in prices. With the 5 year expectations, we see a general decrease in expectations from the start of our data in 1991 until 2020; each year, the mean amount that people believe prices will increase in the next 5 years decreases. Again, we see abnormal activity around 2008 and 2020, for similar reasons that I have outlined above.

B. EU Analysis

For the analysis on the European Union, I use petrol price data from the UK government data site. The data in this set come from the European

Commission Oil Bulletin for non-UK countries, and the Department for Energy Security and Net Zero Fuel Surveys for the UK data. The data set contains monthly data on the average price of petrol in EU countries starting in January 1990. Because the inflation expectation data for the EU only has a quarterly frequency (explained below), I modify the gas price data to be on a quarterly frequency as well. I take the average across all countries to get my variable of interest, which is the average price of petrol across countries in the European Union. Note that I will use the price of petrol with taxes included, and the prices are reported in pence per liter. I end up with a window of analysis of 2001Q1 to 2023Q2. Figure 3 reflects the progression of EU Petrol prices over time. We see that generally, prices have increased over time. The increase was relatively linear and stable between 2001 and 2008, but at around 2008 there appears to be a sharp decrease in price before it rises again until around 2014, where we see another drop. This was when the U.S. started producing oil, which increased the supply of oil worldwide, and thus the price fell in many markets across the globe. This decrease is sustained until the middle of 2016, where it increases again until 2020. In 2020, similar to the U.S., we see a decrease in gas prices due to the COVID-19 Global Pandemic; but the price steadily increases in the post-pandemic era as expected.

Figure 3. EU Petrol Prices (pence/liter), 1990-2023

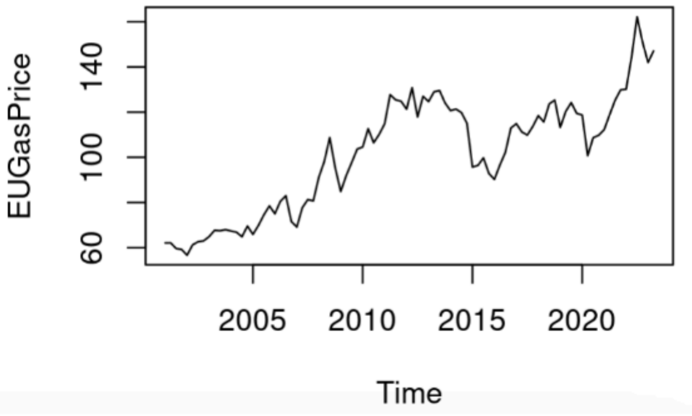


Table 2. Summary Statistics for EU Petrol Prices

Minimum	Q1	Median	Q3	Maximum	Mean	SD	n
56.66	77.90	107.57	120.598	162.226	101.21	25.63	90

Table 2 presents descriptive statistics for the EU Petrol Price data. We see that the minimum price was 56.66, which was in Q1 of 2001, while the maximum of 150.98 took place in Q4 of 2022, which makes sense as this was a period of great economic growth following the COVID-19 global pandemic. The overall mean price is 101.21, with a standard deviation of 25.63. Note that we have 90 observations, significantly less than the U.S. data, but this makes sense because the two differ in frequency.

As I mentioned briefly in Section 1, inflation expectations data is not as available for the European Union as for the U.S. Thus, following the work of Boeck and Zörner (2023), I will use the Survey of Professional Forecasters Forecasts for the Harmonised Index of Consumer Prices as a proxy for inflation expectations. The HICP is an index used to measure consumer price inflation, and it is harmonized because it measures the inflation across the EU as a whole, rather than individualized countries. This data starts in 1999 and is reported quarterly through 2023 Q4, but as with the EU Price Data, I utilize 2001Q1 to 2023Q2 as my window of analysis. The data offers forecasts for several time periods in the future; in my analysis I will use the forecast for "one year ahead" as the 1 year inflation expectations, and the forecasts for "longer term" as the 5 year inflation expectations. Figures 4a and 4b represent how both the 1 year and 5 year trends of the HICP have changed over time. Generally, we see pretty stable numbers for both the 1 year and 5 year data until 2020. In both cases, we see a spike in expectations after 2020; the 1 year expectations reach a high of 5.8 in Q4 of 2022, while the 5 year expectations reach their maximum of 2.2 in Q3 of the same year.

Figure 4a. 1-Year EU Inflation Expectations

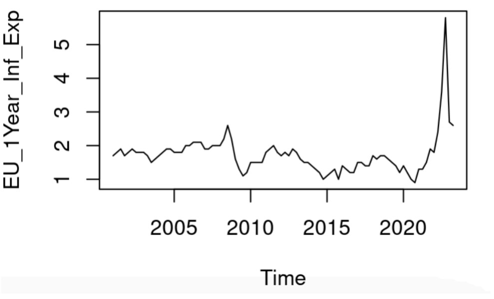
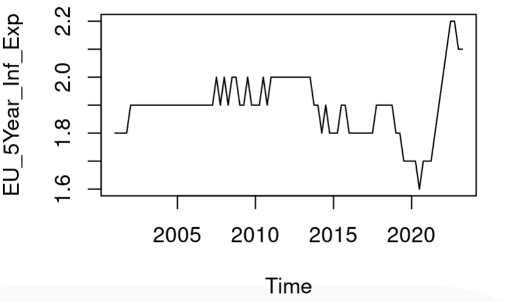


Figure 4b. 5-Year EU Inflation Expectations



C. Controls

Furthermore, in both analyses, I will use interest rates and industrial production as control variables at a monthly and quarterly frequency for the U.S. and EU analyses respectively. I use the Federal Funds Rate data for interest rates in the United States, and the Growth Rate of Industrial Production for my measure of industrial production, both of which come from the Organization for Economic Cooperation and Development (OECD). For interest rates in the EU, I use short term interest rates as defined and provided by the OECD. I use Total Industry Production excluding construction for the Euro Area for the measure of industrial production, which comes from the European Central Bank. See the appendix for more descriptive statistics regarding the controls.

D. Stationarity Testing

Lastly, I will test for stationarity in our variables of interest; I do this to avoid the issue of spurious regressions. To test for stationarity, I conduct the Phillips-Perron test on each variable. Table 3 specifies the results of each test for my variables of interest. See the appendix for the results of the Phillips-Perron test for the control variables.

Table 3. Phillips-Perron Test Results

Variable	Phillips-Perron Test Statistic	Critical Value	Result
U.S. Gas Prices	-1.7388	-2.87	Non-stationary
EU Gas Prices	-0.9667	-2.89	Non-stationary
U.S. 1-Year Inflation Expectations	-4.0204	-2.87	Stationary
EU 1-Year Inflation Expectations	-3.7316	-2.89	Stationary
U.S. 5-Year Inflation Expectations	-4.1858	-2.87	Stationary
EU 5-Year Inflation Expectations	-2.3914	-2.89	Non-stationary

Since some of the variables are determined to be non-stationary, I will transform each of the variables in the system to ensure stationarity throughout the system, and to assist in ease of interpretation in the results section. For those variables that are reported as a growth rate, I take the first difference; for those reported as an index or a true number, I take the log difference. After running the Phillips Perron Test again, I find that all transformed variables follow a stationary process. See the appendix for specific results from this second stationarity test. Confirming stationarity allows me to utilize time series methods to test variable associations; it is to these methods I now turn.

IV. Methods

For my analysis of nominal gas price shocks on inflation expectations, both in the U.S. and in the EU, I will use a VAR model. Existing literature has utilized block recursive VAR models (Clark and Terry 2010; Kilian and

Lewis, 2011) to estimate relationships between energy prices and how expectations about the economy respond to them. However, Killian and Zhou (2023) argue that the use of a block recursive VAR model with oil prices is significantly different than using gasoline prices, and the distinction between which is important in the conclusion of any causal relationship. More literature has also used structural VAR models with both gas prices and crude oil prices (Killian and Zhou 2023; Boeck and Zorner 2023; Casoli et. al 2022). For my purposes, I will use gasoline prices in a structural VAR, similar to the work of Killian and Zhou (2023), and Boeck and Zorner (2023). For the U.S. analysis, I will estimate the structural VAR with interest rates and industrial production as control variables. I order gas prices first, implying that shocks to gas prices contemporaneously affect the interest rate, industrial production, and inflation expectations, but the only shock that contemporaneously affects gas prices is its own. I order inflation expectations last, which assumes that shocks to industrial production, interest rates, and gas prices all contemporaneously affect inflation expectations. The residual matrix illustrating the identifying assumptions is shown below. I use the Akaike Information Criterion to select the number of lags to include in my VAR model estimations. I will include 12 lags for the 1 year estimations and 5 lags for the 5 year estimates for the U.S. analysis.

Matrix 1. Residual Matrix indicating ordering of variables for each VAR model

$$e_t = \begin{bmatrix} e_t^{gas} \\ e_t^{FFR} \\ e_t^{ind\ pro} \\ e_t^{inf\ exp} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \beta_{21} & 1 & 0 & 0 \\ \beta_{31} & \beta_{32} & 1 & 0 \\ \beta_{41} & \beta_{42} & \beta_{43} & 0 \end{bmatrix} \begin{bmatrix} \epsilon_t^{gas} \\ \epsilon_t^{FFR} \\ \epsilon_t^{ind\ pro} \\ \epsilon_t^{inf\ exp} \end{bmatrix}$$

For the EU analysis, I will estimate similar structural VARs as in the U.S. analysis. I use industrial production and interest rates as controls in these systems as well, and order the variables in the same way such that the residual matrix depicted above represents residuals for both the U.S. and the EU VAR models. Using the Akaike Information Criterion, I determine that both the 1 and 5 year models will have 12 lags included.

Before presenting results from the model selection and estimation, I test for Granger causality between gas prices and inflation expectations on both geographical areas of interest and for both 1 year and 5 year expectations. Table 4 presents the results from these tests.

Table 4. Granger Causality Test Results

Response	**Explanatory**	**F-Statistic**	**p-value**
U.S. 1-Year Expectations	U.S. Gas Price	2.3844	0.000009756
U.S. 5-Year Expectations	U.S. Gas Prices	1.8749	0.001409
EU 1-Year Expectations	EU Petrol Prices	2.0224	0.1106
EU 5-Year Expectations	EU Petrol Prices	1.8222	0.009207

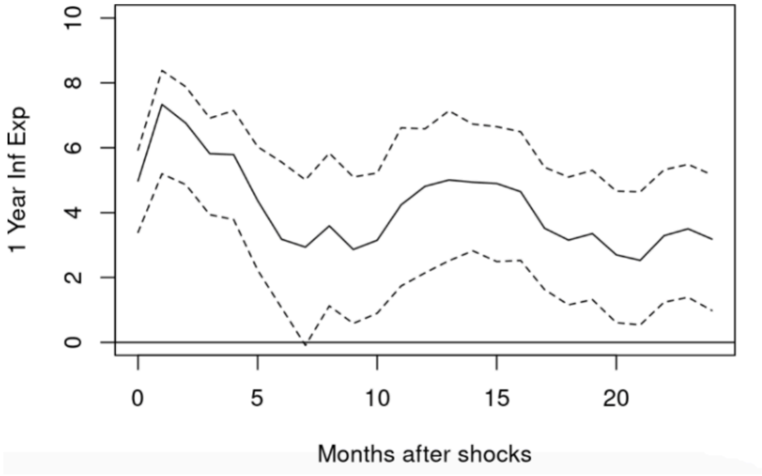
I find that U.S. gas prices do Granger cause both 1 year and 5 year inflation expectations, after controlling for industrial production and interest rates. This indicates that in analyzing the impulse response functions from our VAR system, we expect to see that a shock to gas prices will cause a significant change to both 1 year and 5 year inflation expectations. Note that we have much stronger evidence for the short-run than in the long-run, which is consistent with the literature and previous results in this space. Interestingly, we do not find evidence that gas prices in the EU Granger cause 1 year inflation expectations, after controlling for industrial production and interest rates, but we do find evidence that they Granger cause 5 year inflation expectations at the 90% confidence level. I expect to see that the short-run impacts of gas prices on inflation expectations are more significant than long-run impacts; however, that is not what we observe in this initial test. I hypothesize that we may still find significance when we analyze the impulse response functions from the VAR models due to the increased robustness of the VAR models as compared to the Granger causality test. Since I now have an indication of what I expect to see in the VAR system, I now turn to discussing the results of these models.

V. Results

A. U.S. Analysis: 1-Year Expectations

I begin by analyzing the effect of a gas price shock in the U.S. on consumers' 1 year inflation expectations after controlling for interest rates and industrial production. Plot 5 illustrates this effect and how it changes over time. We see that a shock that increases gas prices by 5.04 percentage points contemporaneously increases inflation expectations by 4.97 percentage points, after accounting for interest rates and industrial production. This effect dies out slightly as time progresses. After 1 year, the 5 percentage point shock to gas prices causes inflation expectations to increase by 4.81 percentage points, and after 2 years, by 3.18 percentage points. These findings are significant at the 95% confidence level. Furthermore, using forecast error variance decomposition, we find that after 1 year, 22.5% of the variation in inflation expectations is driven by gas price shocks. We will see how this effect is different for 5 year expectations, and from 1 year expectations in the EU.

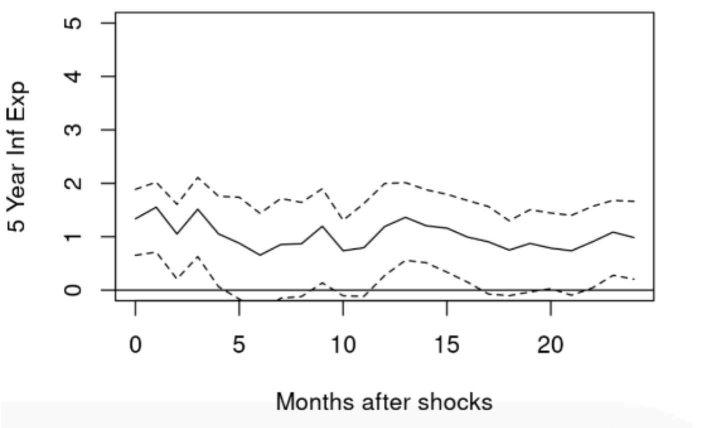
Figure 5. Effect of shock to Gas Prices on U.S. 1-Year Inflation Expectations



B. U.S. Analysis: 5-Year Expectations

Looking at the impulse response functions for the 5 Year U.S. Inflation Expectations, we see in Figure 6 that a shock to gas prices has a significant but minimal effect on these consumer inflation expectations. Similar to the 1 year analysis, the shock initially causes a slight increase in inflation expectations, but this increase is minimal, and after 4 months, inflation expectations have become much more stable than we saw in the 1 year analysis. Specifically, we find that a shock to gas prices that increase them by 4.99 percentage points contemporaneously causes inflation expectations to increase by just 1.36 percentage points, which is much less than what we saw in the 1 year analysis. We also see that after 1 year, the shock that increases gas prices by almost 5 percentage points causes inflation expectations to increase by 1.2 percentage points, and after 2 years, the same shock is associated with just a 0.98 percentage point increase, after accounting for interest rates and industrial production. This is consistent with what we would expect to see based on the current literature and economic intuition. As I have explained above, gas prices are very salient to consumers, and their changes are observed at a high frequency, so we would not really expect a shock to gas prices today to substantially impact consumers' expectations about what prices will look like 5 years from now. Further illustrating this point, we use forecast error variance decomposition and determine that after 1 year, just 4.85% of the variation in 5 year inflation expectations can be explained by shocks to the gas price.

Figure 6. Effect of shock to Gas Prices on U.S. 5-Year Inflation Expectations

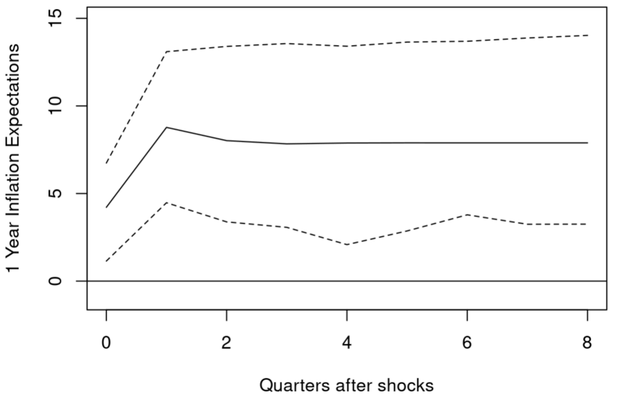


C. EU Analysis: 1-Year Expectations

We now turn to looking at this relationship in the EU. Based on the impulse response function depicted in Figure 7, we see that a shock of 6.52 percentage points to gas prices contemporaneously causes 1 year inflation expectations to increase by 4.2 percentage points. Interestingly enough, as we continue out into the forecast horizon, the effect of the shock to gas prices on 1 year inflation expectations increases, before finally stabilizing around 3 quarters after the shock. We see that the shock to gas prices of 6.52 percentage points causes 1 year inflation expectations to increase by 7.8 percentage points one year later, and this effect is sustained two years into the future. These numbers are significant at the 95% confidence level after controlling for industrial production and interest rates, which contradicts our findings from the Granger Causality tests. However, given the marginal nature of the conclusion from the Granger Causality test (p-val = 0.1; significant at the 90%

confidence level) and the increased robustness of the VAR compared to the Granger Causality tests, this contradiction is not entirely unsurprising, and the results from the VAR are in line with current literature. As with the U.S. analysis, we use forecast error variance decomposition to find that after 1 year, 14.5% of the variation in 1 year inflation expectations can be explained by shocks to gas prices. This variation is the same after 2 years, indicating sustained stability.

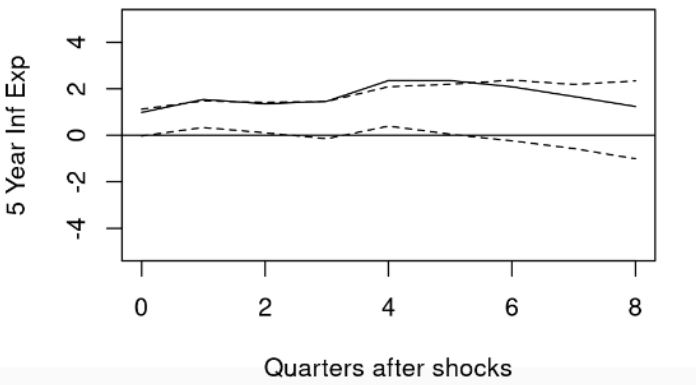
Figure 7. Effect of shock to Petrol Prices on EU 1-Year Inflation Expectations



D. EU Analysis: 5-Year Expectations

Lastly, we turn our attention to the impulse response functions for the EU 5 year expectations. Here, we find rather volatile results. We see that a shock to gas prices of 6.966 percentage points contemporaneously causes 5 year inflation expectations to increase by just 0.98 percentage points. After 5 quarters, we see that the shock to gas prices has caused the inflation expectations to increase by 2.35 percentage points. But then after 2 years, the effect decreases again, with the same shock causing 5 year inflation expectations to increase by just 1.24 percentage points. Not only are these results rather volatile, they are also insignificant. As we can see from the confidence bands in Figure 8 below, the effect of a gas price shock to 5 year inflation expectations in the EU is not significant after controlling for interest rates and industrial production. While this is inconsistent with our results in the Granger Causality tests, it is not all that surprising given the existing literature and results. As I mentioned in the U.S. 5 year analysis, there is no real economic reasoning that a shock to gas prices today would impact how consumers expect inflation to change 5 years from now, so the volatility of 5 year inflation expectations and the insignificance of their response to a gas price shock makes intuitive sense. Using forecast error variance decomposition, we find that after 1 year, 11.5% of the variation in 5 year inflation expectations can be explained by gas price shocks.

Figure 8: Effect of shock to Petrol Prices on 5-Year EU Inflation Expectations



E. Comparisons

I have observed that shocks to gas prices more substantially impact one year inflation expectations in the EU than in the U.S. This heterogeneous response is understandable when thinking about the different characteristics of the EU and U.S. economies. The EU is an importer of oil; other than Norway, Finland, and Sweden, who are not included in this analysis data, countries in the EU are not big producers of oil. They rely more heavily on imports to bolster their oil supply. If the production of oil changes, EU countries are going to feel that impact more heavily on their gas prices, and because that impact is stronger, their perception of prices overall will also be stronger. In the U.S., because the U.S. produces oil, there are some sectors that benefit from the increase in gas prices, so the overall effect felt from the increase is weaker

The relative stability of the 5 year inflation expectations in both the U.S. and the EU is in line with economic thinking. As I have outlined above, due to the volatility and the salience of gas prices to consumers on a daily basis, consumers understand that a shock to gas prices currently is probably not going to impact prices 5 years into the future, and we see this in the analysis for both economies. There is not a justification for gas prices now to impact prices 5 years in the future, but there is more reasoning as to gas prices now impact prices one year into the future, so the variance that we see across the one and five year analysis for both economies is justified.

VI. Conclusion

I have analyzed the relationship between nominal gas price shocks and inflation expectations to determine if variation in inflation expectations can be explained by shocks to gas prices. Using structural VAR models with zero restrictions, I found that in the U.S., gas price shocks do have an effect on inflation expectations at both the 1 year and the 5 year levels, but the effect is greater and more long-lasting at 1 year than at 5 years. In the EU, gas price shocks significantly impact one year inflation expectations, but do not significantly affect 5 year inflation expectations. The lesser impact and lack of significance I found in both the 5 year analysis is consistent with current literature and economic intuition. My study used the HICP measure as a proxy for inflation expectations in the EU, and while this is justifiable, real data on inflation expectations for the EU would increase the robustness of the results. Furthermore, the use of the structural VAR makes restricting assumptions that could be relaxed with more advanced methodologies. Further research could use a rolling regression or a threshold model to analyze the dynamic relationship between gas prices and inflation expectations over time, and if any substantial economic events caused the relationship to change drastically

References

Binder, Carola Conces. 2018. "Inflation expectations and the price at the pump." *Journal of Macroeconomics* 58:1-18. https://doi.org/10.1016/j.jmacro.2018.08.006

Binder, Carola and Christos Makridis. 2022. "Stuck in the Seventies: Gas Prices and Consumer Sentiment." *The Review of Economics and Statistics* 104(2): 293-305. https://doi-org.ezproxy.stolaf.edu/10.1162/rest_a_00944

Board of Governors of the Federal Reserve System (US), Federal Funds Effective Rate [FEDFUNDS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/FEDFUNDS

Board of Governors of the Federal Reserve System (US), Industrial Production: Total Index [INDPRO], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/INDPRO

Boeck, Maximilian and Thomas O. Zörner. 2023. "Natural Gas Prices and Unnatural Propagation xEffects: The Role of Inflation Expectations in the Euro Area" *Social Science Research Network.* http://dx.doi.org/10.2139/ssrn.4376796

Casoli, Chiara, Matteo Manera, and Daniele Valenti. "Energy Shocks in the Euro Area: Disentangling the Pass through From Oil and Gas Prices to Inflation." FEEM Working Paper no. 45, July 2023. http://dx.doi.org/10.2139/ssrn.4307682

De Bruin, Wandi Bruine, Wilbert van der Klaauw, and Giorgio Topa. 2011. "Expectations of Inflation: The Biasing Effect of Thoughts about Specific Prices" *Journal of Economic Psychology* 32(5): 834-845. https://doi.org/10.1016/j.joep.2011.07.002

Department for Energy Security and Net Zero. “International Road Fuel Prices.” GOV.UK, Last Updated November 30, 2023. https://www.gov.uk/government/statistical-data-sets/comparisons-of-industrial-and-domestic-energy-prices-monthly-figures

EIA. 2018 “Methodology for EIA Weekly Retail Gasoline Estimates.” US Energy Information Administration. https://www.eia.gov/petroleum/gasdiesel/gas_proc-methods.php

European Central Bank. “Inflation and Consumer Prices.” European Central Bank, June 23, 2023. https://www.ecb.europa.eu/stats/macroeconomic_and_sectoral/hicp/html/index.en.html.

Georganas, Sotiris, Paul J. Healy, and Nan Li. 2014. “Frequency bias in consumers’ perceptions of inflation: An experimental study.” *European Economic Review* 67: 144-158. https://doi.org/10.1016/j.euroecorev.2014.01.014

Kilian, Lutz and Xiaoqing Zhou. “Oil Price Shocks and Inflation.” Federal Reserve Bank of Dallas Research Department Working Papers no. 2312, Federal Reserve Bank of Dallas, August 2023. https://doi.org/10.24149/wp2312

Mu, Xiaoyi. 2007. “Weather, storage, and natural gas price dynamics: Fundamentals and volatility.” *Energy Economics* 29(1): 46-63. https://doi.org/10.1016/j.eneco.2006.04.003

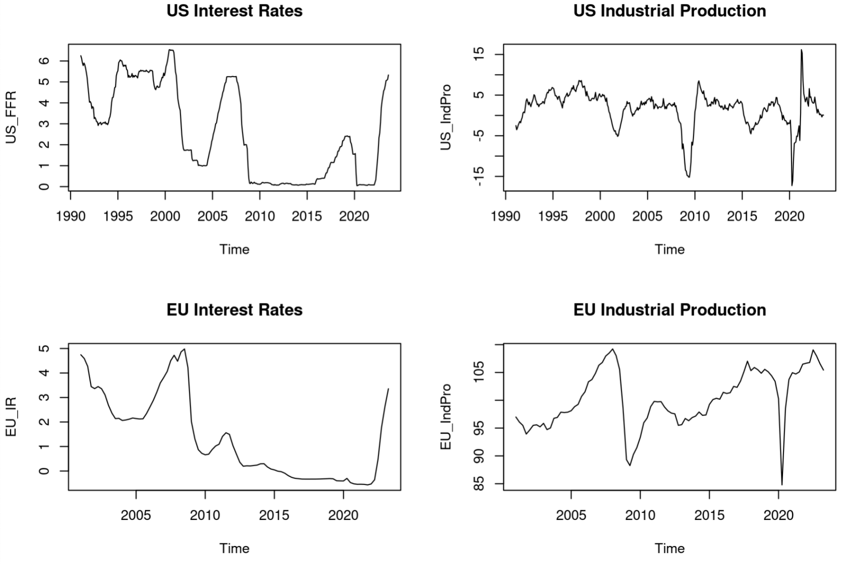
Organization for Economic Co-operation and Development, 2023. Production: Industry: Total Industry Excluding Construction for Euro Area (19 Countries) [PRINTO01EZQ661S], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PRINTO01EZQ661S.

Organization for Economic Co-operation and Development, 2023. Short-term interest rates: https://doi.org/10.1787/2cc37d77-en

University of Michigan Survey Research Center, 2023. "Time Series Data Tables." University of Michigan Survey of Consumers, 2023. https://data.sca.isr.umich.edu/data-archive/mine.php.

Appendix

Appendix Figure 1. Plots for Control Variables



Appendix Table 1: Phillips Perron Results for Control Variables

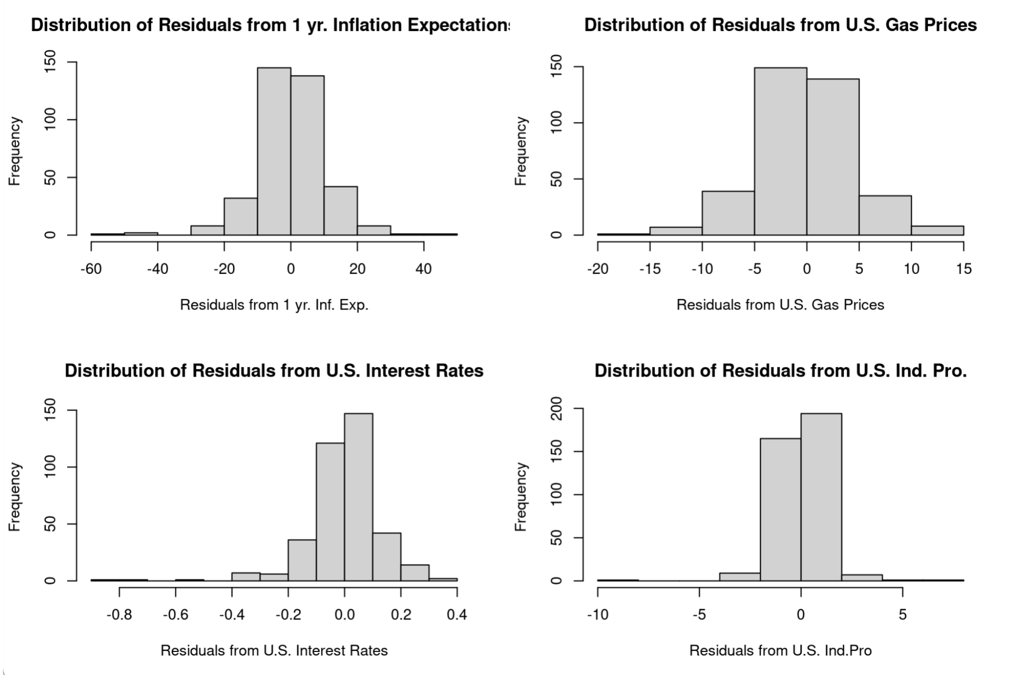
Variable	Phillips-Perron Test Statistic	Critical Value	Result
U.S. Interest Rates	-1.7456	-2.87	Non-stationary
E.U. Interest Rates	-2.0308	-2.89	Non-stationary
U.S. Industrial Production	-3.9301	-2.87	Stationary

E.U. Industrial Production	-2.7929	-2.89	Stationary

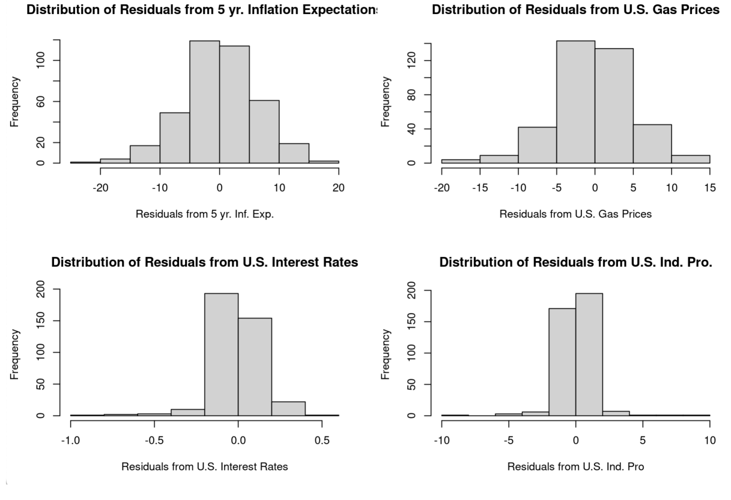
Appendix Table 2: Phillips Perron Results for All Transformed Variables

Variable	Phillips-Perron Test Statistic	Critical Value	Result
U.S. Gas Prices	-12.2057	-2.87	Stationary
EU Gas Prices	-9.4238	-2.89	Stationary
U.S. 1 Year Inflation Expectations	-21.8523	-2.87	Stationary
EU 1 Year Inflation Expectations	-9.3325	-2.89	Stationary
U.S. 5 Year Inflation Expectations	-39.3833	-2.87	Stationary
EU 5 Year Inflation Expectations	-11.3662	-2.89	Stationary
U.S. Interest Rates	-9.411	-2.87	Stationary
E.U. Interest Rates	-15.2786	-2.87	Stationary
U.S. Industrial Production	-3.9231	-2.87	Stationary
E.U. Industrial Production	-9.1542	-2.89	Stationary

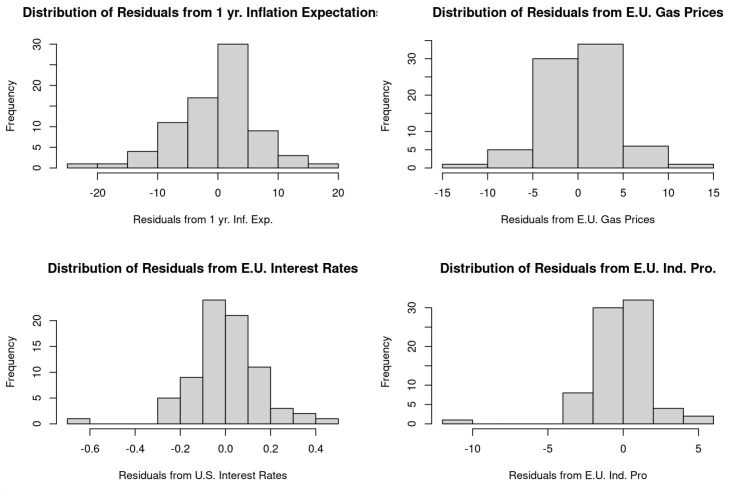
Appendix Figure 2: U.S. 1 Year Expectations Model Residual Diagnostic Plots



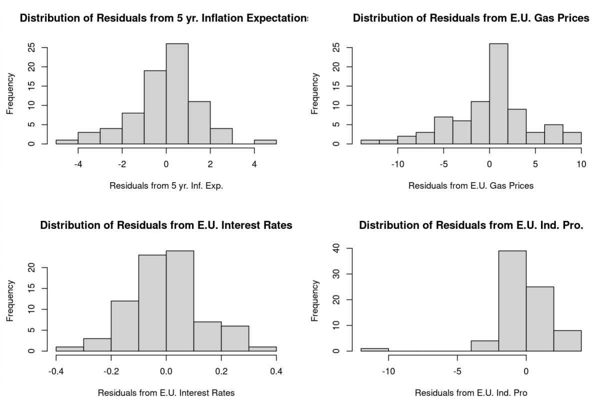
Appendix Figure 3: U.S. 5 Year Expectations Model Residual Diagnostic Plots



Appendix Figure 4: EU 1 Year Expectations Model Residual Diagnostic Plots



Appendix Figure 5: EU 5 Year Expectations Model Residual Diagnostic Plots



The Impact of Disparate Compensation Structures in the NFL and the EPL on Salary Plateaus

Lukas Haugen and Thomas Hillman

I. Introduction

Compensation structures have a profound impact on both workers and firms. In 2021, the National Football League (NFL) generated total revenues of $17.19 Billion (Gough 2022). Over the same period, the English Premier League (EPL) generated $6.20 Billion in revenue (Poindexter 2022). Specifically, NFL franchises (1st) and EPL clubs (2nd) are the most profitable sports ventures in the world, earning $537,187,500 and $310,000,000 per team in the past year. Both leagues rely on disparate payroll regulations, making it quite difficult to compare "football" to "football." As a result, an attempt to evaluate the financial implications of this reality is unprecedented within the literature. In this paper we ask if disparate compensation structures within the NFL and EPL lead to a common result; the existence of a salary plateau for the average veteran player.

"The NFL provides an ideal laboratory to analyze decisions about the equality of pay distributions" (Mondello and Maxcy 2009). The salary cap plays an invaluable role as the foundational difference in compensation structure between the NFL and the EPL. In 1993, the NFL and the National Football Players Association (NFLPA) approved a new Collective Bargaining Agreement (CBA) in which the players earned the right to free agency and the owners implemented a "hard" salary cap, meaning that no team is allowed to exceed the cap for any reason (Mondello and Maxcy 2009). This had a resounding positive effect on league competition and success. From 1994 to 2022, the NFL salary cap ballooned from $34.6 million to $208.2, outpacing inflation by over 400% over the same time period (Spotrac). As a result, the average value of an NFL franchise has skyrocketed from $165 million to $4.47 billion (Koons 2022) NFL players currently receive 48% of revenue sharing profits, making the NFL a classic “allocation of scarce resources” decision that provides a unique opportunity to study organizational decision making under restrained resources (DeCort 2022; Mullholland and Jensen 2019; Borghesi 2018).

When examining the impact of salary dispersion on NFL team success, Mondello and Maxcy (2009) explore the tradeoffs between two fundamental compensation structures. The first is a “hierarchical pay structure,” where a larger portion of pay is concentrated on fewer individual employees. The second is a “compressed pay allocation,” where there is minimal dispersion in compensation levels across the firm. “Success” is defined in two ways: on-the-field wins and off-the-field revenues. Specifically, Mondello and Maxcy conclude that a hierarchical pay structure is more efficient for generating off-the-field revenue; whereas a compressed pay allocation is more effective for improving on-the-field performance. This situation creates a conflict of objectives where a uniform salary structure with incentive bonuses for performance is the most productive solution.

Borghesi (2008) agrees with Mondello and Maxcy when he presents evidence in favor of a compressed pay allocation. Specifically, Borghesi contends that teams who compensate players the most inequitably are likely to underperform, as “franchises taking the superstar-approach to personnel decisions perform worse on average.” This evaluation critiques teams who sacrifice a balanced roster by spending major percentages of their limiting available resources on individual positions such as the quarterback. Borghesi’s work is defended by recent results, as the average cap percentage for every quarterback to make it to the conference championship or further since 2010 has only been 7.5% (Faber 2022). Most importantly, Borghesi concludes that the salary cap had led to an exponential increase in reliance on incentive-type bonus payments (Figure 1) and a worsening income inequality within the NFL (Figure 2).

Figure 1. Base and Bonus Compensation in the NFL across Season

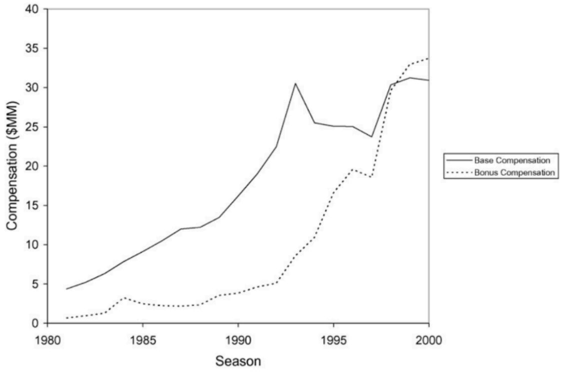
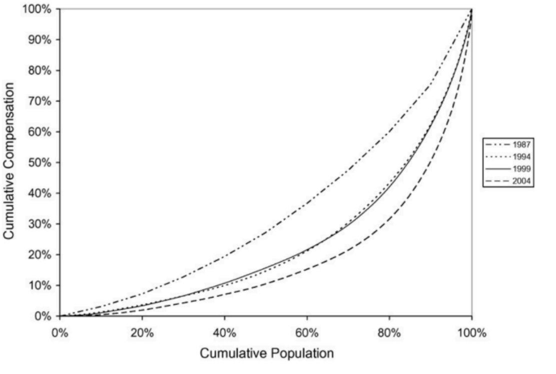


Figure 2. Population vs. % of Compensation in the NFL



Mulholland and Jensen (2018) offer a radically divergent conclusion. Accounting for win contributions as a measure of individual player performance, they recommend an optimal allocation strategy that is highly non-uniform. This is an unprecedented opinion within the literature. They substantiate their argument by emphasizing the growing importance of the quarterback position. Most notably, Mullholland and Jensen introduce the revolutionary assumption that the NFL is not an efficient labor market. As a result, they expose the NFL "Rookie Wage Scale" as the optimal way to capitalize on labor market inefficiencies.

The link between salary dispersion and team success is just as evident in the English Premier League literature. The absence of a salary cap within the EPL has created a vast, unique payroll disparity between clubs. Specifically, the top four clubs have an average total payroll of $187,587,250; whereas the bottom four clubs have an average total payroll of $22,786,200. (Spotrac.com 2022). Due to this lack of compensation regulation, the EPL is plagued by two pressing problems: increasing debt levels and a persisting lack of competition.

The combined debt of all EPL franchises was ￡3.1 Billion during the 2009-2010 season, with Manchester United the most indebted club at a total of ￡700 million in debt (Davies 2010). Freestone and Manoli (2017) suggest that the "Financial Fair Play" (FFP) regulations passed in 2011 serve as an alternative model to a salary cap. The FFP rules were marketed as a direct response to the profitability problem faced by England's top clubs in an attempt to "shift the focus of sporting competition away from financial strength more towards natural means of competition such as efficiency, innovation, and good management."

Davies (2010) offers a compelling explanation and justification for the apparent inability of EPL clubs to remain profitable in comparison to their NFL counterparts. Specifically, Davies argues that NFL franchises are more financially incentivized than EPL clubs due to the nature of their unique competition models. On one hand, the NFL is a "closed competition market" in which firms cannot freely enter or exit the market. This allows NFL franchises to be "profit-maximizing firms" as their participation in the league is guaranteed in the future.On the other hand, the EPL is an "open competition market" where firms can be promoted or relegated from year-to-year depending upon on-field performance. This forces EPL clubs to be "utility-maximizing firms" as winning matters above all else.

The impact of FFP has been minimal to date. As of 2021, the combined net debt of the EPL has grown throughout the FFP era to over ￡4 billion (Statista Research Department 2022, "EPL.."). Fortunately, FFP restrictions are due to expand significantly in 2023. As Dan Sheldon explains in his article, the new FFP rules will not resemble a true salary cap where all teams are held to the same restricted maximum payroll value. However, "clubs will be limited to spending a set percentage of their revenue in a calendar year on transfers, agents' fees and player wages." These new rules, called the "Financial Sustainability and Club Licensing Regulations"

(FSCLR), represent a major change for the EPL. The ultimate goal is that they allow the UEFA to continually monitor clubs who are in financial peril. “The limit in 2023 will be 90 per cent before dropping to 80 per cent in 2024 and 70 per cent from 2025 onwards” (Sheldon 2022). Perhaps these updates are the solution to the EPL’s financial plight. However, as these rules do not go into effect until next year, we consider the current FFP regulations within our study.

Carmichael, McHale, and Thomas (2010) introduce the idea of a “causal link between revenue earned and competitive imbalance via investments in players.” Furthermore, they find that investment in players’ skills and abilities buys on-field success, and that as a result rich teams get richer by building on past success and acquiring players from poorer clubs. This is directly proven by the lack of competitive balance within the EPL evidenced by the disparity of club championships from 1989 to 2022 compared to the revenue for each club, as seen in Figures 3 and 4 below. This criticism serves as an explanation for the top-heavy success of the EPL that is driven by major disparities in salary payrolls amongst teams.

Figure 3. Premier League Championships by Club, 2023

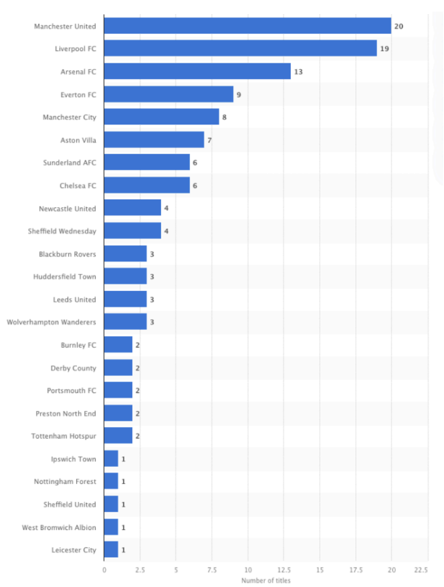
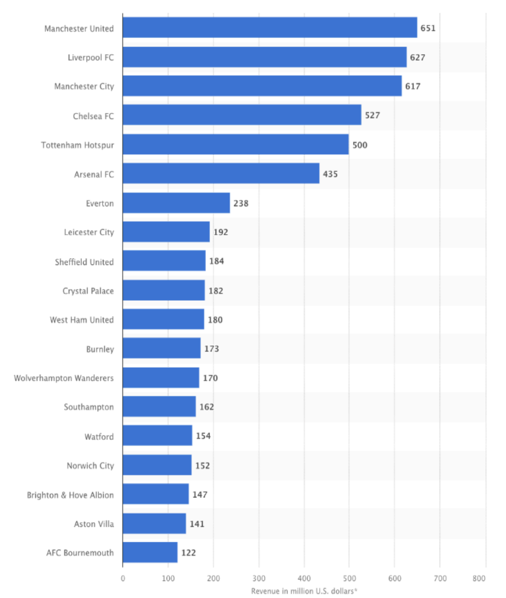


Figure 4. Premier League Revenue by Club, 2023



As explained above, the literature is rich regarding salary dispersion in the NFL and the EPL. Both leagues support the "Efficiency Wage Theory" that higher payrolls are positively correlated with improved performance and winning percentage (Mondello and Maxcy 2009). However, there is no literature regarding the individuals who are most impacted by this salary dispersion. Mondello and Maxcy acknowledge that there was inadequate attention paid to the effects of the salary cap on individual team decisions and performance. They recommend that "future research could also involve a mixed methods approach to help gain an additional understanding." By using both the NFL and the EPL, we fill this gap in the literature by relying on a mixed methods approach to provide three major contributions. First of all, we emphasize the importance of player position as a primary factor affecting average annual wages. Second of all, we discover differing modes of player acquisition to be responsible for major payroll disparities. Third and finally, we provide evidence of an unquestionable salary plateau for veteran players in both the NFL and EPL.

We compiled cross-sectional panel data on both NFL and EPL players from *Spotrac*, the largest sports team and player contract resource in the world. Using *Spotrac,* we collected data such as AAV, or Average Annual Value, which serves as the dependent variable in our model. This statistic, consistent within both the NFL and the EPL, allows us to accurately gauge differences in salary dispersion between leagues. We also used *Spotrac* to provide data for key independent variables such as Position, NumContract, AvgLContract, Experience, Experience Squared, Age, First Round Pick, NumTransfers, Rookie/ Veteran, and Team. The data includes 1502 active NFL contracts and 499 active EPL contracts. As a result, the data is cross-sectional given the fact that it only considers active contracts for the 2022-2023 NFL and EPL league seasons. See Table 1 below for the variables and their descriptions.

Using these variables, we construct two OLS models to discover the primary factors responsible for average annual contracts in both leagues. Due to their differing compensation structures, we examine the NFL and EPL in separate models, with nearly-identical independent variables. For example, we use First Round Pick in the NFL and Number of Transfer in the EPL as independent variables which account for player acquisition. As a result, we determine the independent variables with the largest impact on AAV, and conclude that they are responsible for a compressing effect on the salary of mid-tier players.

The remainder of this paper is organized as follows: in Section 2 we provide the methodology behind our comparative model. In Section 3, we present our empirical model and an explanation of the variables we chose. In Section 4, we discuss our results and the robustness of our findings. In Section 5, we conclude with our contribution to the literature and recommendations for future research.

Table 1. Shared (NFL and EPL) Variables and Descriptions

Variables	Definition	Source
Position	Game position held by a player (ex. Quarterback or Goalkeeper)	Spotrac
NumContract	Number of league contracts held by a player	Over the Cap, Spotrac
AvgLContract	The average length of a player's league contract	Over the Cap, Spotrac
Experience	The number of years a player has played within their league	Spotrac
Experience^2	The squared value of the experience variable, representing the diminishing returns on	Spotrac

	player experience	
Age	The age of the player when signing their active contract	Spotrac
FirstRoundPick	For NFL players, an indicator of whether the player was a first round pick during their rookie draft	Pro-Football-Reference
NumTransfers	For EPL players, the number of transfers between EPL teams	Spotrac
Rookie/Veteran	An indicator of whether a players is considered a Rookie (first year in the league) or a Veteran (4+ years of experience)	Spotrac
Team	Team of a player's active contract	Spotrac

Position and Team are represented through a series of dummy variables. Experience was calculated by taking the difference of the first contract start year and current year, and only accounts for years spent in the NFL and EPL leagues specifically.

II. Methodology

Mondello and Maxcy (2009) identify two key assumptions when discussing the importance of a salary cap on competitive balance and salary dispersion. We adopt these assumptions to be the center of our methodology. The first assumption states that the implementation of a salary cap "theoretically creates a relatively equal distribution of talent across league members." The second assumption states that "hypothetically, this mechanism attempts to restrict teams with the most financial resources from dominating their respective leagues by consistently accumulating better talent" (Mondello and Maxcy 2009).

As a result, we assume that the existence of a salary cap within the National Football League is primarily responsible for the evidence of a salary plateau for middle-class players. Using this same framework, we predict that the absence of a uniform salary structure within the English Premier league is to credit for the absence of a salary plateau and less equitable compensation among players.

In an attempt to conduct a proper comparative analysis of the NFL and EPL, we created two nearly-identical OLS models. We selected general factors such as number of contracts, average contract length, age, and experience variables as relevant measures of a player's talent and value within their respective league. By relying upon comparable factors and not individual player statistics, we provide an objective determination of the effect of a salary cap on payroll distribution for individual players.

Despite accounting for the impact of individual positions, we also included individual team payrolls to highlight the vast differences in compensation structures between the NFL and EPL. According to the tables included below, team variables are an appropriate representation of overall compensation structure within the league

III. Empirical Strategy and Model

The data for this study considers currently active contracts for the 2022-2023 NFL and EPL league seasons. Therefore, this data is cross-sectional as it contains a number of observations over a singular time frame. Additionally, as we examine two different league salary structures, we use two separate regression models to determine if a plateau exists in player salary based on average experience and talent due to the NFL's salary cap.

A successful comparative analysis is dependent upon key differences and commonalities. Beyond their differing compensation structures explained above, the NFL and EPL share a number of fundamental variables responsible for variations in average annual contracts for individual players.

A. *Shared Dependent Variables*

Average Annual Value (AAV) is defined as the average annual value of a player's active contract. As a result, AAV is calculated by taking the total value of a contract and dividing by its length. We chose *Log (AAV)* as our dependent variable in both leagues for two reasons. First of all, *AAV* is a commonly comparable measure of payroll disparity in both leagues. This is due to the fact that the EPL does not adhere to the same salary cap system as the NFL. NFL wages are often calculated as a percentage of the overall cap value, which would be irrelevant to the EPL in which teams spend dramatically different amounts on players. Second of all, *AAV* allows us to view the effects of Team spending on the player's wages as a part of the whole league, showing how the cap system affects wages. Second of all, by taking the *Log (AAV)*, we are able to interpret our results in percentages, allowing us to more easily compare salary dispersion across leagues. By dealing with percentages rather than monetary values, we can ignore exchange rates and simplify our analysis.

We denote the NFL dependent variable as *Log(AAVNFL)* and expect to find evidence of a salary plateau due to the salary cap system. We denote the EPL dependent variable as Log(AAVEPL) and do not expect to discover a salary plateau for EPL players. This is due to the lack of a salary cap in the English Premier League and our prediction that less payroll regulation will allow AAV to be more consistently scaled with talent and experience.

B. Shared Independent Variables

The independent variables explained below are identical within both models. This symmetry grants credibility, verifiability to our comparative analysis.

Position is defined as the specific role each player is compensated for within their respective teams. Both the NFL and EPL have unique position dummy variables that are all accounted for within our refined model. Our research leads us to expect a highly significant correlation between *Position* and *Log (AAV)* within both leagues. Highly-valued skill positions such as quarterback (AAV NFL) and forward (AAV EPL) will likely share a positive relationship with annual compensation due to the importance of offense and the star power associated with the position. Less valuable special teams positions such as punter (AAV NFL) and goaltender (AAV EPL) will likely correlate negatively with average salaries. It is important to note that we did not include individual player statistics in our data. This is due to the differing nature of the sports, as there is no realistic conversion between "touchdowns" in the NFL and "goals" in the EPL. By omitting individual player statistics, we successfully avoid subjective valuations of on-field performance.

NumContract is defined as the sum of a player's contracts in their respective league. This independent variable accounts for important factors such as staying power and reputation by accounting for total contracts signed. The literature predicts a positive correlation between *NumContract* and *Log (AAV)* in both the NFL and EPL. With each additional contract signed, players can expect a pay raise as their past performance has justified an extension. We adopt the "Efficiency Wage Theory," and predict that only players whose performance matches or exceeds their compensation will remain in the league over time. However, *NumContract* could exhibit a negative relationship with *Log (AAV)* if veteran players sign multiple short-term contracts due to their inability to procure a long-term agreement. Due to its inability to differentiate between short-term and long-term contracts, *NumContract* predicts that a player who signs multiple one-year contracts will be better off than the player who signs a single long-term contract. This is not necessarily true, as signing a long-term contract provides wage stability that might be in the best interest of both the team and the player.

AvgLContract is defined as the average longevity of a player's contracts throughout their career. It is included to account for the longevity shortfall associated with the *NumContract* variable. The literature is divided on the predicted impact of *AvgLContract* on *Log (AAV)*. This is due to the fascinating tradeoff between contract value and contract length. On one hand, a player who values long term wage stability will be motivated to sign a long-term contract with a slightly lower AAV. Patrick Mahomes is a perfect example. Mahomes and the Kansas City Chiefs recently broke the NFL contract mold when they agreed to a 10-year contract worth a total of $450 million with $141 million fully guaranteed. This is the longest contract in NFL history, and an example of the premium teams are willing to pay for talented, promising young players in skill positions like quarterback. On the other hand, a player who is willing to "bet on themselves" would sign a short-term contract with a higher AAV hoping to cash in now. Kirk Cousins is an example of capitalizing on wage uncertainty. After playing back-to-back seasons on the single year "Franchise Tag," Cousins agreed to a 3-year contract worth a total of $84 million fully guaranteed with the Minnesota Vikings. Over the past seven years, Cousins has earned $198.9 million, far more than he would have earned had he tried to negotiate a seven-year contract back in 2016. With that being said, we predict *AvgLContract* to have a significant, positive relationship with both *Log (AAV)* and $Log(AAV_{EPL})$as only the most talented players will be offered attractive long-term contracts by their respective franchises and clubs.

Experience is defined as the number of seasons a player has competed in their respective league. We expect a significant positive coefficient for *Experience* in the NFL model. This is due to the "Rookie Wage Scale," which was implemented with the 2011 CBA. This policy set a binding salary cap on the total value of a player's contracts during their first four seasons. As a result, teams placed a premium on rookie contracts and draft picks. We do not anticipate as large of an effect in the EPL model, as there is no limit on rookie compensation. Additionally, the EPL's unique system of player acquisition allows veteran players from other professional leagues to be considered EPL rookies within our model. While we assume a positive correlation between *Experience* and $Log(APY_{EPL})$, this relationship may be weakened by "inexperienced" players who transferred into the EPL after long professional careers in foreign leagues.

Experience Squared is defined as a player's number of years in the league times itself. This variable is included to account for the diminishing returns that occur as a player's body deteriorates due to increased competition over time. We expect *Experience Squared* to have a negative coefficient within both the NFL and EPL models. This is due to the fact that highly experienced players often receive lower salaries as they near the end of their playing careers. By including this variable, we are able to evaluate the marginal returns generated by each additional year of league experience.

Age is defined as the player's age when they signed their active contract. We predict a contradiction within the *Age* coefficient in relation to *Log (AAV)* in both models. For players younger than 27, we presume a significant positive coefficient. This is driven by the assumption that salaries will increase as player's prove their ability to consistently perform at a high level. However, as players age into their early 30's we expect this coefficient to become significant and negative as players face the reality of diminishing marginal returns on performance. In other words, we anticipate that younger veteran players will sign more lucrative contracts whereas older veteran players will sign less favorable contracts. This is due to the added element of risk associated with older players who possess an increased likelihood of injury, retirement, and declining performance. Playing in the NFL and the EPL is extremely difficult, and often older players lose the edge they once had. This concern is evident in our research, as 75% of NFL players are out of the league before 27 years old. The exceptions to this trend are players whom the league finds valuable despite their older age. The perfect example, and ultimate anomaly, is Tom Brady, who is 45 years old yet still made $30 million in 2022.

Rookie/ Veteran is defined as a dummy variable that is used to identify a player's contract status. Within our model, a *Rookie* is any player who has four or less years of total experience in their respective league, whereas a *Veteran* has more than four years of league experience. NFL literature strongly suggests a significant relationship between *Rookie/ Veteran* and annual income. Specifically, we confidently assume that *Rookie* will boast a negative coefficient whereas *Veteran* will display a positive coefficient. This is directly tied to the "Rookie Wage Scale" which creates a uniform salary case system for draft picks. NFL players are free to renegotiate their contract via an extension with their current team or free agency at the end of their rookie contract. In either scenario, NFL players can reasonably expect much more fair compensation once they achieve veteran status. EPL literature does not provide evidence of a significant relationship between *Rookie/ Veteran* and $Log(APY_{EPL})$. This is most likely due to the fact that there is no limit on rookie compensation within the EPL.

Team is defined as the respective franchise or club that employs a player. In our refined OLS model adjusted for heteroskedasticity, *Team* is split into dummy variables to account for differences in spending by each respective franchise in the NFL and club in the EPL. This is done to analyze the importance of *Team* on a player's annual contract. Specifically, we anticipate *Team* to be insignificant within the NFL due to the existence of a hard salary cap and floor which enforces nearly-equal payroll expenses between franchises. We predict that *Team* will have a significant coefficient within the EPL model due to the vast payroll disparities between clubs.

C. Different Independent Variables

First Round Pick is a dummy variable within our NFL model which accounts for a player's draft status upon entering the league. First round picks are extremely valuable in the NFL, as teams have the opportunity to select one of the 32 best players available every season. We reasonably expect *First Round Pick* to have a significant positive coefficient within our NFL model. This is due to the reality that first round picks are often used on players who have a higher degree of skill or higher likelihood of success than the other rookies being drafted. Additionally, first round picks often play premium skill positions and are given greater opportunities than other rookies as teams don't want to embarrass themselves. Historically, first round picks earn higher salaries. As of 2022, first round picks make, on average, three times the amount of other players during their career ($9 million vs. $3 million).

NumTransfers is defined as a player's total number of transfers to an EPL club. This independent variable is included as a parallel to the NFL's *First Round Pick* dummy variable. Generally, we assume that teams will only pay the transfer fee for a highly talented, sought-after player. As a result, we expect *NumTransfers* to have a significant positive relationship with *Log (AAV)* within our EPL model.

D. Model Specification

$$Log(AAVNFL) = \beta_0 + \beta_1 Position + \beta_2 NumContract + \beta_3 AvgLContract + \beta_4 Experience + \beta_5 Experience^2 + \beta_6 Age + \beta_7 FirstRoundPick + \beta_8 Rookie/Veteran + \beta_9 Team + \epsilon$$

$$Log(AAVEPL) = \beta_0 + \beta_1 Position + \beta_2 NumContract + \beta_3 AvgLContract + \beta_4 Experience + \beta_5 Experience^2 + \beta_6 Age + \beta_7 NumTransfers + \beta_8 Rookie/Veteran + \beta_9 Team + \epsilon$$

Our Specific OLS Comparative Models are nearly identical to our General Model. However, the specific models include dummy variables for every unique position and team in both the NFL and EPL.We also refine our Specific OLS model by removing insignificant independent variables. All three models predict a salary plateau in the NFL due to allocative sacrifices driven by the salary cap, whereas the EPL should not exhibit the same trend.

IV. Results

A. Summary Statistics

Table 2. Summary Statistics for NFL Players

NFL Variables	Observations	Mean	Std. Dev	Min	Max
Log(AAV)	1502	14.673	1.0500	12.240	17.733
NumContract	1502	2.3868	1.9910	1.000	15.000
AvgLContract	1502	3.3485	0.89328	1.000	7.000
Experience	1502	3.6924	3.2273	0.000	22.000
$Experience^2$	1502	24.043	37.762	0.000	484.00
Age	1502	24.722	3.5063	20.000	44.000
FirstRoundPick	1502	0.1738	0.37904	0.000	1.000
Rookie/Veteran	1502	0.45806	0.45806	0.000	1.000

Table 3. Summary Statistics for EPL Players

NFL Variables	Observations	Mean	Std. Dev	Min	Max
Log(AAV)	499	14.459	1.1974	9.903	17.104
NumContract	499	1.7515	0.94549	1.000	6.000
AvgLContract	499	4.1935	1.0649	1.000	7.000
Experience	499	3.7074	2.6272	0.000	17.000
Experience^2	499	20.633	26.544	0.000	289.00
Age	499	24.080	4.1101	17.000	39.000
NumTransfers	499	0.49098	0.55376	0.000	3.000
Rookie/Veteran	499	0.48497	0.50028	0.000	1.000

The most notable takeaway from our summary statistics is the fact that the EPL possesses a greater standard deviation in Log (AAV). This is significant because it tells us that the overall level of salary dispersion is worse within the EPL. This result is not surprising as we expected the salary cap to improve competitive balance and payroll disparity. Next, we ran three distinct OLS regression models in an attempt to isolate the most significant factors affecting average annual contracts for both NFL and EPL players.

A. General OLS Comparative Model Results

Table 4. General OLS Model Results

General OLS Model (HC1)		
Dependent Variable = Log (AAV)		
Variables	NFL	EPL
Constant	11.0184***	11.1135***
	(.4216)	(0.5411)
Position	-0.0070*	.1130***
	(.0037)	(.0398)
NumContract	-.0434	.0671
	(.0287)	(.0832)
AvgLContract	.4871***	.2911***
	(.0468)	(.0535)
Experience	.0768**	-.0670
	(.0311)	(.0457)
Experience^2	-.0063**	.0052
	(.0031)	(.0038)
Age	.0503***	.1115***
	(.0144)	(.0127)
FirstRoundPick	.7989***	NA
	(.0554)	NA
NumTransfers	NA	.2221***
	NA	(.0696)
Veteran	1.3162***	.1192
	(.1071)	(.1472)
Team	.0039	-.0902***
	(.0020)	(.0075)
Observations	1502	499
Adjusted R^2	0.5728	0.4816

Our first model is a General OLS regression model corrected for heteroskedasticity. Our results were strong, as we achieved an Adjusted R^2 of 0.5728 in the NFL and 0.4816 in the EPL. Most importantly, this general model tells us that position is significant in both models, but that team is only significant in the EPL model. We attempt to explain this interesting result including all unique position dummy variables and unique team dummy variables in the Specific OLS Comparative Model below.

A. *Specific OLS Comparative Model Results*

Our second model is a Specific OLS Comparative Model that includes dummy variables for both position and team in both leagues. Unsurprisingly, we find a number of independent variables to be insignificant within both the NFL and EPL models. In an attempt to isolate only the most significant independent variables, we eliminate all insignificant variables in our Specific Refined OLS Comparative Model below.

Table 5. NFL OLS Model Results

Variables	NFL	Variables	NFL	Variables	NFL	Variables	NFL
Constant	11.1977***	Right Guard	-.0585	Jets	-.0874	49ers	-.1469
	(.4091)		(.0827)		(.1629)		(.1770)
NumContract	-.0511*	Right Tackle	-.0383	Lions	-.0743	Bears	-.1988
	(.0264)		(.0827)		(.1499)		(.1439)
AvgLContract	.4728***	Running Back	-.1745***	Packers	-.1146	Bengals	-.0713
	(.0430)		(.0626)		(.1484)		(.1465)
Experience	.0770**	Tight End	.0303	Panthers	-.1211	Bills	-.0682
	(.0299)		(.0744)		(.1485)		(.1629)
Experience^2	-.0067**	Safety	-.0613	Patriots	-.1080	Broncos	-.1005
	(.1571)		(.0675)		(.1571)		(.1543)
Age	.0497***	Wide Receiver	.2491***	Rams	-.1934	Browns	-.2492
	(.0136)		(.0759)		(.1496)		(.1561)
FirstRoundPick	.7525***	Cornerback	.0423	Ravens	-.3060*	Buccaneers	-.0785
	(.0545)		(.0575)		(.1608)		(.1626)
Rookie/Veteran	1.3387***	Quarterback	.4372***	Saints	-.0171	Cardinals	-.0892
	(.1020)		(.1196)		(.1553)		(.1614)
Center	.0420	Punter	-.3679***	Texans	-.1454	Chargers	-.0926
	(.1061)		(.1335)		(.1470)		(.1499)
Edge Rusher	.0398	Colts	-.1716	Vikings	-.0964	Chiefs	-.2357
	(.0640)		(.1454)		(.1552)		(.1499)
Fullback	.0800	Falcons	-.2601	Seahawks	.0208	Steelers	-.0862
	(.2631)		(.1597)		(.1588)		(.1649)
Long Snapper	-.8970***	Commanders	-.1370	Jaguars	.0743	Giants	-.1530
	(.1174)		(.1526)		(.1547)		(.1448)
Kicker	1.1721***	Cowboys	-.3437**	Titans	-.0274		
	(.1814)		(.1574)		(.1619)		
Left Guard	.0297	Dolphins	.0495				
	(.0855)		(.1497)				
Left Tackle	.1026	Eagles	.0178				
	(.0900)		(.1530)				
						Observations	1502
						Adjusted R^2	0.6151

Table 6. EPL OLS Model Results

Variables	EPL	Variables	EPL	Variables	EPL
Constant	8.9145***	Midfielder	.7893***	Leeds United FC	-.6481**
	(.5288)		(.1384)		(.3012)
NumContract	.0761	AFC Bournemouth	-.4981	Leicester City	.2706
	(.0756)		(.3282)		(.2629)
AvgLContract	.2880***	Arsenal FC	.9287***	Liverpool FC	.8773***
	(.0518)		(.2588)		(.2809)
Experience	-.0395	Aston Villa FC	.3329	Manchester City FC	1.1770***
	(.0461)		(.2721)		(.2725)
Experience^2	.0016	Brentford FC	-.4484*	Manchester United FC	1.3160***
	(.0037)		(.2664)		(.2763)
Age	.1299***	BrightonHoveAlbion	.1792	Newcastle United FC	-.0067
	(.0124)		(.2622)		(.2677)
Rookie/Veteran	.0706	Chelsea FC	1.3076***	Southampton FC	.1932
	(.1424)		(.2572)		(.2820)
NumTransfers	.1547**	Crystal Palace	.3884	TottenhamHotspur FC	.7290***
	(.0670)		(.2759)		(.2770)
Defender	.7156***	Everton FC	.3293	Nottingham Forest FC	-.4004
	(.1402)		(.2781)		(.3324)
Forward	.9874***	Fulham FC	-.1577	WestHam United FC	.4347
	(.1474)		(.2908)		(.2764)
				Observations	499
				Adjusted R^2	0.537384

B. Specific Refined OLS Comparative Model Results

Our third model is a Specific Refined OLS Comparative Model which includes dummy variables for every unique position and team within both the NFL and EPL. This final model enables us to identify the most significant independent variables that affect Log (AAV). As a result, all of the independent variables included above are significant at at least the 10% level. Somewhat surprisingly, eliminating insignificant variables improved the overall health of our model. For example, our Adjusted R-squared rose from .5728 to .6191 in the NFL Model, and from .4816 to .5355 in the EPL Model. As a result, we improve our model through simplification without losing overall effectiveness.

Table 7. NFL Refined OLS Model Results

Variables	EPL	Variables	EPL
Constant	9.10453***	Brentford FC	-0.663992***
	(0.448753)		(0.150236)
AvgLContract	0.278183***	Chelsea FC	1.13410***
	(0.0519503)		(0.129992)
Age	0.133532**	Leeds United FC	-0.829779***
	(0.0111181)		(0.209467)
NumTransfers	0.167674**	Liverpool FC	0.707050***
	(0.0652274)		(0.168494)
Defenders	0.706060***	Manchester City FC	1.01564***
	(0.141105)		(0.150601)
Forward	0.988564***	Manchester United FC	1.13882***
	(0.146481)		(0.164004)
Midfielder	0.781385***	TottenhamHotspur FC	0.562081***
	(0.140471)		(0.155910)
AFC Bournemouth	-0.680328***	Nottingham Forest FC	-0.594169**
	(0.242104)		(0.250738)
Arsenal FC	0.744772***		
	(0.127510)		
		Observations	499
		Adjusted R^2	0.535487

NumContract is statistically significant at the 5% level (p<0.05) within the NFL model. While this is expected, we were shocked to see that the coefficient was negative as this tells us that NFL players can anticipate a 5.48% decrease in salary with every contract they sign. This seemingly goes against the NFL Rookie Wage Scale that predicts a massive raise for veteran players. However, this is most likely explained by the fact that NFL players receive lower wages as they age as teams are less willing to award older players with massive contracts due to their greater risk of injury, retirement, and decreasing production. *NumContract* has been dropped from the EPL for a lack of significance. This result was not surprising, as the labor acquisition model of the EPL allows clubs to buy experienced professionals from other leagues. As a result, a rookie could be a 20-year old prospect, or a 32-year old established star. This vast disparity leads to the fact that EPL players experience a massive disparity in first contract compensation.

Table 8. EPL Refined OLS Model Results

Specific	Refined	NFL OLS	Model
Variables	NFL	Variables	NFL
Constant	11.0732***	Long Snapper	-.8992***
	(0.389007)		(.1164)
NumContract	-.0548**	Punter	-.3908***
	(.0261)		(.1297)
AvgLContract	.4698***	Quarterback	.4167***
	(.0425)		(.1134)
Experience	.0792***	Running Back	-.1857***
	(.0296)		(.0550)
Experience^2	-.0067**	Wide Receiver	.2353***
	(.0029)		(.0691)
Age	.0514***	Cowboys	-.2358**
	(.0134)		(.0996)
Rookie/Veteran	1.3270***	Jaguars	.1867**
	(.1013)		(.0946)
FirstRoundPick	.7632***	Ravens	-.1942*
	(.0537)		(.1016)
Kicker	1.1661***		
	(.1746)		
		Observations	1502
		Adjusted R^2	0.619089

NumContract is statistically significant at the 5% level ($p<0.05$) within the NFL model. While this is expected, we were shocked to see that the coefficient was negative as this tells us that NFL players can anticipate a 5.48% decrease in salary with every contract they sign. This seemingly goes against the NFL Rookie Wage Scale that predicts a massive raise for veteran players. However, this is most likely explained by the fact that NFL players receive lower wages as they age as teams are less willing to award older players with massive contracts due to their greater risk of injury, retirement, and decreasing production. *NumContract* has been dropped from the EPL for a lack of significance. This result was not surprising, as the labor acquisition model of the EPL allows clubs to buy experienced professionals from other leagues. As a result, a rookie could be a 20-year old prospect, or a 32-year old established star. This vast disparity leads to the fact that EPL players experience a massive disparity in first contract compensation.

Average Contract Length is statistically significant ($p<.01$) within both models. Specifically, NFL players can expect a 46.98% increase in salary for every additional year on their contract. This is representative of the Rookie Wage Scale, and the idea that NFL teams are willing to pay a premium for long-term wage stability, rather than renewing contracts more often. This is likely due to the danger of free agency, which makes player retention extremely difficult. Additionally, EPL players can expect a 27.82% raise for each additional contract year. This represents the same idea, that young talented prospects are incentivized to sign long contracts with their employers. These results were expected, but the discrepancy between them is likely explained by NFL teams being more willing to reward consistent talent with longer, more lucrative contracts. This is almost certainly due to the threat of free agency that is nonexistent in the EPL's transfer system. Additionally, short-term contracts like one-year deals often go to veterans earning the league minimum in the NFL. By comparison, the EPL teams find both reliability and renewal rates important factors to consider when choosing how much players should be paid.

Experience, Experience Squared, and *Veteran* Status are all only statistically significant with the NFL model. Contrary to our null hypothesis, it is clear that the NFL significantly values player experience. For every additional year of experience, NFL players can expect to receive a 7.92% salary increase. On the other hand, the diminishing marginal return on experience leads to a 0.67% decrease in salary annually. Most notably, NFL players receive a 132.70% raise upon achieving veteran status after four years in the league. The absence of these factors in the EPL model is likely due to the EPL's transfer system which places a large emphasis on foreign investment in labor. Additionally, the EPL's promotion and relegation system allows teams to set their spending limits which leads to vast payroll disparities. As a result, factors such as *Experience* are insignificant relative to these compensation structure factors.

Both *First Round Pick* and *Transfers Count* are statistically significant in their respective models at the 1% significance level ($p<.01$). NFL first round draft picks should expect to hold salaries 76.32% higher than the average over the course of their careers. In comparison, EPL players should find a 16.77% increase upon transferring between teams. Clearly, the reputation of a first round pick carries massive weight in the NFL, as teams are willing to give these highly regarded prospects a number of chances to prove their ability to perform. As these factors are supposed to be reflective of how the leagues introduce new talent, the differences in values were to be expected. However, the significance of the EPL transfer count shows that if a player wishes to increase their pay, the advised approach for a player should be to transfer to a new team.

Position appears as the most significant shared independent variable between leagues. This particular significance of *Position* reflects the respective value of particular positions between leagues. Within the NFL model, most of the significant positions are offensive occupations, with the rest being less highly-valued or niche positions. *Long Snapper*, *Kicker, and Punter* are all specialist positions, meaning that special teams positions are valued in a very different way than traditional positions. Specifically, kickers can expect to make 116.61% more than the average player, whereas long snappers and punters make 89.92% and 39.08% less respectively. As the league moves to a pass-oriented offense, *Quarterback*, and *Wide Receiver* are the key offensive positions, and they are valued accordingly. As a result, quarterbacks make 41.67% more than average, whereas wide receivers experience a positive 23.53% benefit. EPL data finds that relative to the suppressed Goalkeeper position, all other positions have some significance. Specifically, *Defenders* (+70.61%), *Midfielders* (+78.14%), and *Forwards* (+98.86%), make significantly more than *Goalkeepers*. This difference is likely related to how the NFL has prioritized players in skill positions. For example, a star quarterback throwing a touchdown pass or a world-class forward scoring a goal is far more exciting and interesting to watch than a right guard blocking the opposition or a defender breaking up an odd-man rush. The EPL's generalized four positions lead more to a positionally balanced game for spectators, as the defenders have to pass the ball to the midfielder and then to the forwards. This natural progression of the game leads to an active involvement by all positions, which contributes to the overall viewer experience. Contractually, as every position is paid more than the Goalkeeper, the EPL devalues purely defensive positions for viewer experience.

Team is the final independent variable included in our model and appears far more significantly in the EPL model. Specifically, ten of the twenty teams in the EPL registered as statistically significant. *AFC Bournemouth (-68.03%), Brentford FC (-66.40%), Leeds United FC (-89.98%), and Nottingham Forest (-59.42%)* all hold negative coefficients, meaning that players on these clubs make significantly less money than the average EPL player. On the other hand, *Arsenal (+74.48%), Chelsea FC (+113.41%), Liverpool FC (+70.71%), Manchester City (+101.56%), Manchester United (+113.88%), and Tottemham Hotspur (+56.21%)* have the most significant, positive coefficients. These results reflect the vast disparity in salary payrolls in the EPL, with a difference of £211,305,000 between the richest and poorest teams. This vast wage expense disparity within the EPL is responsible for the overshadowing of variables like *Experience* and relative *Veteran* status.

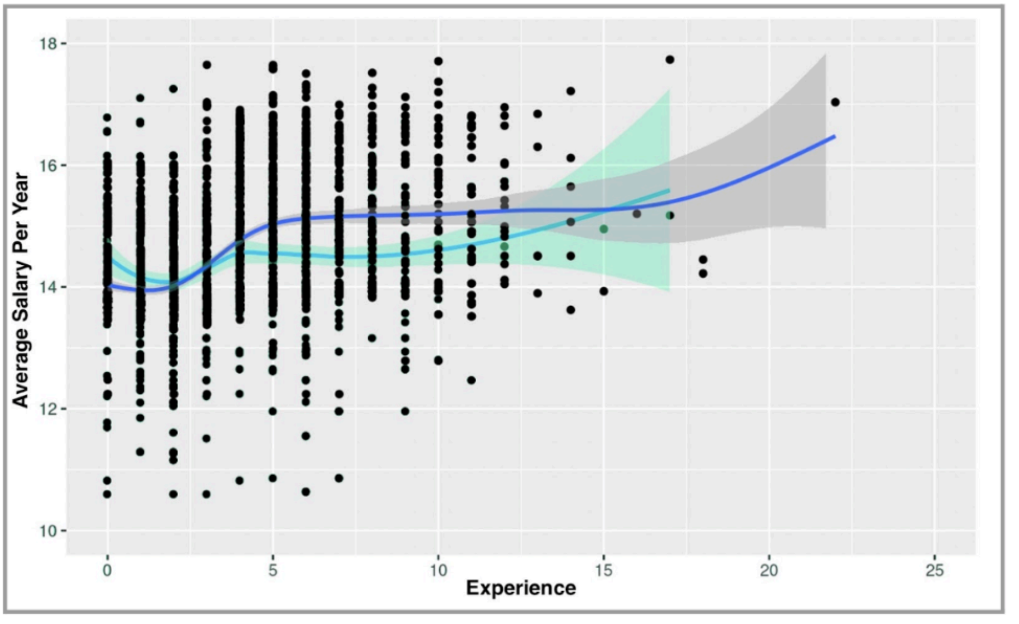
Furthermore, despite having twelve more teams than the EPL, only three of the NFL's teams showed any significance; the Cowboys ($p < .05$), Jaguars ($p > .05$), and Ravens ($p > .10$). The fact that 29 out of 32 NFL teams appear insignificantly within the model clearly shows that the salary cap system's intent of creating wage parity and fair competition between teams is successful.

V. Conclusion

In conclusion, the existence of a salary cap in the NFL (blue curve) does in fact lead to a plateau in the salary of mid-tier players. This result was somewhat predictable, as the NFL "Rookie Wage Scale" leads to a preference for rookies due to their affordable contracts and young age. With that being said, position does have a significant impact on average salary, perhaps more so than the salary cap itself.

The EPL (green curve) results reflect a surprising parallel. Specifically, we find that mid-tier players in the EPL experience an even more dramatic salary plateau than the NFL. This is possibly due to the nature of the transfer system and the absence of a rookie wage scale. We conclude with three key contributions. First of all, we advance the literature by identifying three modern factors that most significantly contribute to salary dispersion; veteran status, individual position, and team. Second of all, we highlight the importance of player acquisition as a means of capitalizing on inefficiencies in international labor markets benefits. Third of all, we provide groundbreaking, shocking, unquestionable evidence of a salary plateau within both the NFL and EPL.

Figure 5. Average Salary per Year by Experience



We recommend that future research builds on these ideas by examining the potential effects of implementing a true salary cap in the EPL. This would most likely be impossible without an amateur, or foreign player, draft. While there are certainly a number of shortcomings with this proposal, most notably the challenge posed by the EPL relegation system, it would almost certainly lead to increased parity in the leagues.

References

Borghesi, Richard. 2008. “Allocation of Scarce Resources: Insight from the NFL Salary Cap.” *Journal of Economics and Business* 60(6): 536-550. https://doi.org/10.1016/j.jeconbus.2007.08.002

Carmichael, Fiona, Ian McHale, and Dennis Thomas. 2010. “Maintaining Market Position: Team Performance, Revenue and Wage Expenditure in the English Premier League.” *Bulletin of Economic Research* 63(4): 464–497. https://doi.org/10.1111/j.1467-8586.2009.00340.

Davies, Chris. 2010. “The Financial Crisis in the English Premier League: Is a Salary Cap the Answer?” *European Competition Law Review* 31(11): 442-448.

DeCort, Callie. 2022. “NFL Salary Cap Explained.” *The 33rd Team*, May 8. https://www.the33rdteam.com/explaining-nfl-salary-cap/

Faber, Tyler. 2022. “Realizing the True Value of an NFL Quarterback in Today's Game.” *FanSided*, Apr. 27, https://fansided.com/2022/04/27/realizing-true-value-nfl-quarterback/.

Freestone, John Christopher and Argyro Elisavet Manoli. 2017. “Financial Fair Play and Competitive Balance in the English Premier League.” *Sport, Business, and Management: An International Journal* 7(2):175-196., https://doi.org/10.1108/SBM-10-2016-0058.

Gough, Christina. 2022. “NFL Revenue 2021.” *Statista*, Sep 7, https://www.statista.com/statistics/193457/total-league-revenue-of-the-nfl-since-2005/#:~:text=In%202021%2C%20the%2032%20teams,of%2017.19%20billion%20U.S.%20dollars.

Koons, Zach. 2022. “Cowboys Top NFL Franchise Valuation List at $8 Billion, Per Report.” *Sports Illustrated*, Aug 22, https://www.si.com/nfl/2022/08/22/cowboys-lead-nfl-franchise-valuation-list-8-billion-forbes#:~:text=NFL%20franchises%20are%20now%20worth,in%202021%20(%243.48%20billion).

Miragaia, Dina, João Ferreira, Alexandre Carvalho, and Vanessa Ratten. 2019. “Interactions between Financial Efficiency and Sports Performance.” *Journal of Entrepreneurship and Public Policy*, 8(1): 84–102. https://doi.org/10.1108/jepp-d-18-00060.

Mondello, Mike, and Joel Maxcy. 2009. “The Impact of Salary Dispersion and Performance Bonuses in NFL Organizations.” *Management Decision* 47(1): 110–123. https://doi.org/10.1108/00251740910929731.

Mulholland, Jason, and Shane Jensen. 2019. “Optimizing the Allocation of Funds of an NFL Team under the Salary Cap.” *International Journal of Forecasting* 35(2): 767-775. https://doi.org/10.1016/j.ijforecast.2018.09.004.

Poindexter, Owen. 2022. “Premier League Dominates Soccer in Revenue.” *Front Office Sports*, Aug. 8, https://frontofficesports.com/premier-league-dominates-soccer-in-revenue/#:~:text=Premier%20League%20teams%20earned%20a,in%20the%202021%2D22%20season.

Statista Research Department. 2023. “Clubs that have won the most Premier League titles as of 2023.” *Statista*, 26 Oct. 2023, https://www.statista.com/statistics/383696/premier-league-wins-by-team/.

Statista Research Department. 2022. “EPL and Championship Club Debt England 2021.” *Statista*, 29 Sept. 2022, https://www.statista.com/statistics/1336333/net-debt-premier-league-championship-clubs/#:~:text=The%20combined%20net%20debt%20of,been%20just%20four%20years%20earlier.

Statista Research Department. 2022. “Premier League Clubs by Revenue 2020.” *Statista*, Dec. 8, https://www.statista.com/statistics/566666/premier-league-clubs-by-revenue/.

Sheldon, Dan. 2022. “Explained: How UEFA's FFP Rules Work.” *The Athletic*, Aug 25, https://theathletic.com/3532690/2022/08/25/arsenal-ffp-rules-uefa/.

Spotrac.com. *EPL Active Player Contracts.* Retrieved December 13, 2022, from https://www.spotrac.com/epl/contracts/.

Spotrac.com. *NFL Active Player Contracts*. Retrieved December 13, 2022, from https://www.spotrac.com/nfl/contracts//.

Right to Work Laws in Politics: Assessing Current Literature

Jonah Nelson

I. Introduction

Right to Work (RTW) laws are laws that are passed at a state level, through a bill, a constitutional amendment or even a referendum. These laws state that no workplace can legally require newly hired employees to join a union, even if every employee up to that point is a union member, however, the union still will have to provide the protections afforded to members to employees who opt not to join the union. (Jacobs & Dixon 2006)

In 1935, the Federal Government passed the Wagner Act or the National Labor Relations Act, which officially legalized collective bargaining and action by employees, and required that employers negotiate with the unions created by the employees. However, in 1947, this Act was amended by the Taft-Hartley Act, a law that attempted to limit union power in a few ways. First, unions were banned from certain striking practices and union membership policies, for example, unions were prohibited from striking to support another union and from charging “excessive” union dues. Additionally, the law definitively outlined that states were allowed to pass RTW laws. By the time the Taft-Hartley Act passed, nine states had already passed RTW laws, however, within fifteen years thereafter, another nine states also passed similar legislation. After this flurry of RTW laws, the passage of such legislation slowed dramatically. Between 1963 and 2011, only three states adopted RTW laws. However, in recent years, there has been a new surge in RTW laws being passed. Since 2012, five states have adopted RTW policies.

Although RTW laws have been passed in 28 states and currently are in place in 27, their actual impact is still not entirely clear. Traditional economic theory indicates that the laws should cause a decrease in union density. RTW laws disincentivize joining a union by providing employees with an ability to freeride on collective action taken by the unions. Additionally, there is a strong correlative relationship between the presence of RTW laws and lower union density. Collins (2014) finds that states which currently have RTW laws have only a third of the union density that states without the laws have. RTW states sit at an average of 5% union density, while states which do not have RTW laws are at roughly 15%.

Despite the clear correlation between RTW laws and lower union density, studies assessing whether there is a relationship between the laws and union density have been largely inconclusive. Studies like those performed by Lunn (2023) and Fortin, Lemieux and Lloyd (2022) find that RTW laws do, in fact, cause union density to decline. However, this is challenged by a study performed by Newman and Moore (1987) which finds that the laws are mostly a symbolic gesture, and they don't have a significant impact on union density.

RTW laws seem to have significant impact outside of union density, though. Feigenbaum, Hertel-Fernandez and Williamson (2018) find that the laws do have a significant effect on partisanship. They find that RTW laws cause the percentage of the vote received by Democrats to decrease by 3.5%. The study also concludes that voter participation decreases by between 2-3% after passage of said legislation. This study is not perfect, though; it is highly limited by the data utilized in the study. The study only looks at years dating back to 1980, while over two thirds of RTW laws were passed before then. Radcliff and Davis (2000) also address this question, albeit indirectly, by looking at union density in relation to voter participation. They find that there is a correlation between lower union density and lower voter turnout, indicating that for every 1% decrease in union density, voter turnout goes down by .25%.

These studies taken together paint a bit of a contradictory picture. There is not a clear answer to whether RTW laws actually impact union density, however. Research has not yet provided a clear finding about the relationship between union density and RTW laws. Some studies claim that RTW laws cause union density to decrease, while others contradict this statement. This lack of clarity also can be seen in the research on how the laws impact politics. While there is not a great deal of literature on this topic, the literature that does exist is highly limited by the years that it assesses.

This paper attempts to investigate the findings of the previous research to determine the actual impact of RTW laws by looking at a broader dataset than previous studies relied upon, and by using a recently developed method of a staggered difference-in-difference (DID). In a series of recent studies by Callaway and Sant'Anna (2021), Sun and Abraham (2021) and Chaisemartin and D'Haultfoeuille (2022), it is shown that previous methods of staggered DID produced results that don't actually have any real meaning. Each of these papers produced their own method for circumventing the issues presented. In this paper, I adopt the method proposed by Callaway and Sant'Anna (2021). To address the issue faced by previous researchers of limited data, this paper utilizes data that has recently been synthesized by Farber et al. (2021) which provides estimates on union density dating back to 1937. Because the most significant research on the impacts of RTW laws on politics, Feigenbaum, Hertel-Fernandez and Williamson (2018), only looks at data after 1980, this analysis attempts to verify whether their findings on voter participation and partisanship are consistent with RTW laws outside of their dataset.

The actual tests performed are a series of staggered DIDs that utilize the method proposed by Callaway and Sant'Anna (2021). The tests are each performed twice, once for the years 1937 to 1979 and second time for 1980 to 2020. These tests indicate that RTW laws broadly have a significant and negative relationship to the number of Democrats in state legislatures.

Additionally, RTW laws passed between 1937 and 2001 broadly don't have a significant effect on union density, while the laws cause voter participation to increase. The RTW laws passed after 2001, though, all cause both union density and overall voter participation to decrease. What this means broadly is that the apparent contradictions in previous studies are largely explained by a shift in the actual impact of RTW laws after 2001.

II. Theory

In his book, *The Logic of Collective Action*, Mancur Olson (1965) challenged what, at the time, was considered to be the standard view on democracy and collective action within that democracy. It had been believed that if all members of a group had a common goal, they would unite to achieve that goal. Olson challenged this assertion. He said that when the common goal is a public good, the incentive to freeride is too great and thus, the goal is not likely to be achieved. Olson proposed that in order to achieve a goal through collective action, the rewards of the collective action must be exclusive to participants in the group. Elinor Ostrom built on this by proposing a set of eight guiding principles that would allow for collective action to take place in a successful manner. These principles included instructions, for example, boundaries for who is and who isn't part of the group should be very clearly defined. Another instruction proposed that these organizations should have a right to self-determination clearly expressed by authorities, in particular, governmental authority. At the most basic level, unions are simply an organization that attempts to perform collective action. Unions are essentially a group of workers all organizing in order to achieve a common goal for all of the workers. Because unions are the organization of individuals to perform collective action, they are susceptible to the pitfalls suggested by Olson. To counteract this, unions have historically listened to Ostrom, structuring themselves in a way that, at the very least, makes a concerted effort to comply with her eight guiding principles.

RTW laws get in the way of this, though. They make unions susceptible to the pitfalls suggested by Olson and undermine the ability of unions to properly adhere to multiple of Ostrom's principles. The laws remove the ability of unions to define boundaries that prevent freeriding. In states without RTW laws, unions can require new hires into a workforce they represent to join a union. This leads to a system where the benefits secured by unions only apply to union workers, removing the concern of freeriding. However, when a RTW law is passed, it gets in the way. The collective action toward a common goal becomes non-exclusive, which, according to Olson's theory, should impose the problem of free riding on the group. There is little incentive for any new employee to join a union because, while they will receive the same benefits offered by the union regardless of membership, they don't have to pay the union dues if they refuse to join. Following this logic, RTW laws should cause union density to decline.

As found by Radcliff and Davis (2000), this decline in union membership should, in theory, cause a decline in voter participation. They discuss that unions actively engage in voter mobilization measures for both their members and in the broader community. Additionally, Kim (2016) suggests that union membership serves to increase social consciousness in a similar manner to education, which contributes to the increase in voter participation when unionization rates are higher. What this means in a broader sense is that so long as union density does decrease with the passage of RTW laws, the RTW laws should cause voter participation to decrease. Following this logic, if the findings of Feigenbaum, Hertel-Fernandez and Williamson (2018) are accurate, we should see that RTW laws cause a decrease in union density, at least after 1980.

III. Data, Testing, and Results

A. Data

This paper utilizes panel data that looks at metrics sorted by state and by year. It utilizes data on union membership collected and synthesized in the paper, *Unions and Inequality over the Twentieth Century: New Evidence from Survey Data* written by Farber et al. (2021). The data gathered by the paper provides the best estimate of union membership in each state before 1973, the year the Census Bureau started collecting that data. To do so, the authors synthesized Gallup poll data on union membership by household and combined it with the census data to create a complete record of union membership in each state. Additionally, data is used from the Michigan State University Correlates of State Policy Project, a project that works to compile and disseminate data that is relevant to assess and analyze policy in all 50 states. This dataset provides several key variables for my research. Also utilized in this study is voter turnout data compiled by the American Presidency Project at the University of California Santa Barbara as well as voter turnout data compiled by the MIT Election and Data Science Lab.

For a variety of reasons, a total of ten states had to be excluded from the analysis performed in this study. The period being investigated in this study starts in 1937, however, neither Alaska nor Hawaii gained statehood until 1959. New Mexico and Ohio were eliminated because both have unique laws that implement RTW ideas in a limited capacity, which means that they cannot be clearly classified as RTW or as non-RTW states. Lastly, Louisiana and New Hampshire had to be dropped because both of them have passed and later repealed RTW laws. With the DID method proposed by Callaway and Sant'Anna (2021) that is being utilized, it is not possible to untreat a unit once it has already been treated. As a result, it is not possible to test these states in this study.

B. Method

For this research, the staggered DID proposed by Callaway and Sant'Anna (2021) was the best form of analysis. The staggered DID model is used to quantify the relative impact of identical treatment being applied to different individuals at different times in order to determine if there is causality. The method is useful, in part, because it is able to account for exogenous variables in a manner that many other methods cannot. This study is one that utilizes staggered DIDs well because not only is there a clear-cut treated and non-treated group, as well as clear treatment times, but the treatment remains the same across cases. Additionally, the impact of legislation is very likely to have a wide variety of exogenous variables, some of which are obvious, but many of which are not. This problem is largely solved by using a DID.

Staggered DIDs are not perfect, though; they are very conditional, relying on a number of conditions being met. Among other things, they don't work if there are anticipation effects. Additionally, if there is any inconsistency in treatment or any uncertainty of whether a group is treated, the method simply will not work. Even within the realm of staggered DIDs, certain methods are better than others depending on the circumstances, for example, the method proposed by Callaway and Sant'Anna (2021) does not grant free movement in and out of treatment for variables while Abraham and Sun's (2021) method does.

Before the tests were performed, the data was split into two groups of years, 1937 to 1980, and 1981 onward. This was done because there is a clear split between RTW laws passed in the mid 20th century and those codified afterward. Between the years 1965 and 2000, there were only two RTW laws passed, while before 1965, 20 RTW laws were passed and since 2000, another six have been passed. This indicates that there may be a potential difference in motivation and outcome of the passage of the law. Additionally, one key study that is being investigated is the study on the impacts that RTW have on politics by Feigenbaum, Hertel-Fernandez and Williamson (2018). This study only looks at data from 1980 onward to verify the accuracy of this study, it makes sense to only compare results from the same time period.

In this testing, I specifically look at the effects of RTW laws on unity density, voter participation and the proportion of Democrats and Republicans in state legislatures because these are domains addressed by previous studies which can be tested. There are other variables that would likely be valuable to the study, however, because this study is using data going back to 1937, there are restrictions on available possible metrics. The amount of data being collected at the state level in 1937 was very limited. This fact was relevant for previous studies as well, because even metrics like current union members in a state were not officially tracked until 1973.

C. *Testing*

The first test performed looks at union density from 1937 to 1979 (Graph 1). The totals of all the results from every treated group in the sample indicate that RTW laws don't significantly impact union density. That result generally holds when looking at the results by year RTW was implemented, otherwise known as cohorts. This is not applicable to every cohort, however. There is no clear pattern, and most cohorts that are significant only consist of a single state. These results indicate that, at least in this window, RTW laws did not have a clear causal effect on union density.

Union density after 1979 had a more significant effect (Graph 2). Seven states passed RTW laws in this window, however, even among those seven, there is a clear divide between the laws passed in 1985 and 2001 as compared to those passed from 2012 to 2017. For the first two states to pass RTW laws in this window, the results indicate that the laws caused a significant and positive effect on union density. This signals that the earlier trend of RTW laws not having a detrimental effect on union density continued until as late as 2001. However, this is the last that is seen of this trend. The results show that in the five states treated between 2012 and 2017, RTW laws significantly caused a decrease in union density. While it certainly is not definitive, this shift seems to indicate that the trend of RTW laws not hurting union density came to an end.

Figure 1. Average RTW Effect on Union Density 1937-1979

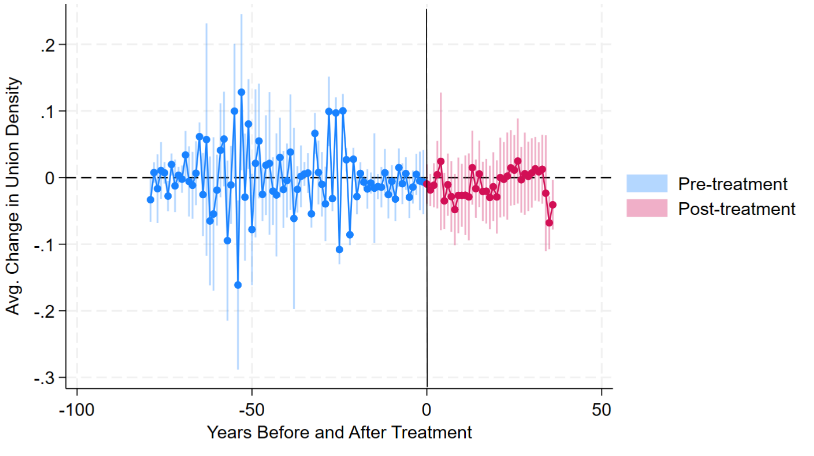
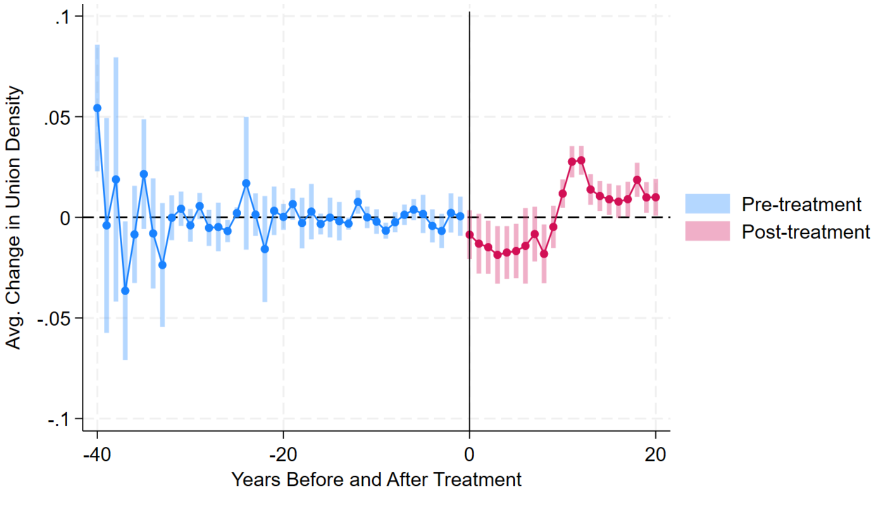


Figure 2. Average RTW Effect on Union Density 1980-2020



Based on current literature, because union density is not significantly impacted by RTW laws between 1937 and 1979 (Graph 3), voter participation should not be significantly impacted by RTW laws. The results don't reflect this, though. Instead, with the exception of the 1946 cohort, every cohort of RTW laws until 1955 had a positive effect on voter participation. From 1955 onward though, the impact became insignificant.

The impact of RTW laws shifting voter participation in the negative direction continued after 1980 (Graph 4). Of the seven states that passed RTW laws after 1980, five of them experienced a significant negative impact on voter participation as a result. At first glance, this seems to be consistent with the predictions, however, a closer look at which states experienced the decrease in participation suggests that there may be more going on. Both states that gained union density because of RTW laws simultaneously experienced a decrease in voter participation, and two of the five states that saw a decrease in union density did not experience a decrease in voter participation.

Figure 3. Average RTW Effect on Voter Participation 1937-1939

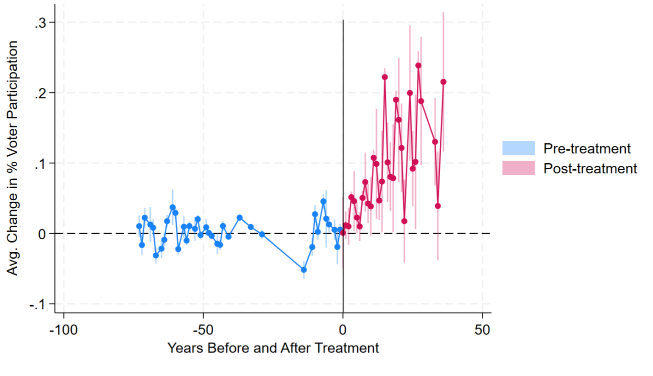
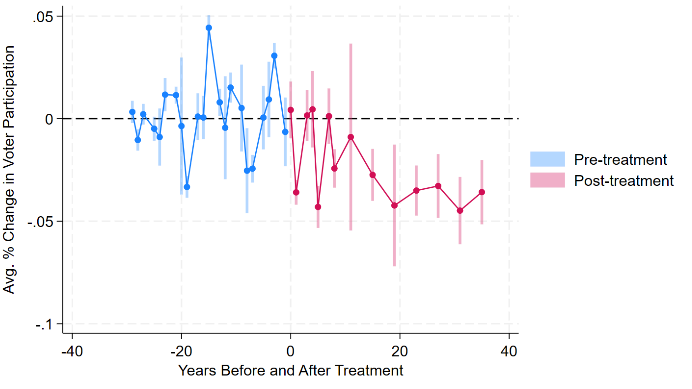


Figure 4. Average RTW Effect on Voter Participation 1980-2020



The last variable analyzed in relation to the passage of RTW laws is the percentage change in state house and senate seats held by Democrats (Figures 5 and 6 below). This measure is consistent across all years. With the exception of a few cohorts across the entire dataset which are insignificant, nearly every cohort has the same outcome. RTW laws cause Democrats to lose seats in the state legislature. In other words, in spite of the impact that RTW laws have on union density and even voter participation, the one factor that remains consistent is that Republicans will gain ground as a result of the law.

Figure 5. Percentage of House Seats Held by Democrats, All Years

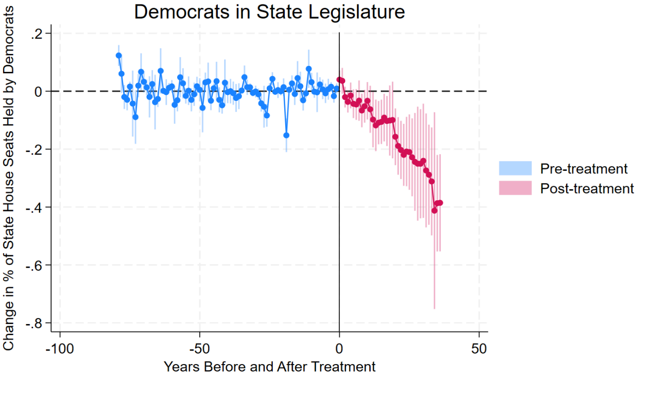
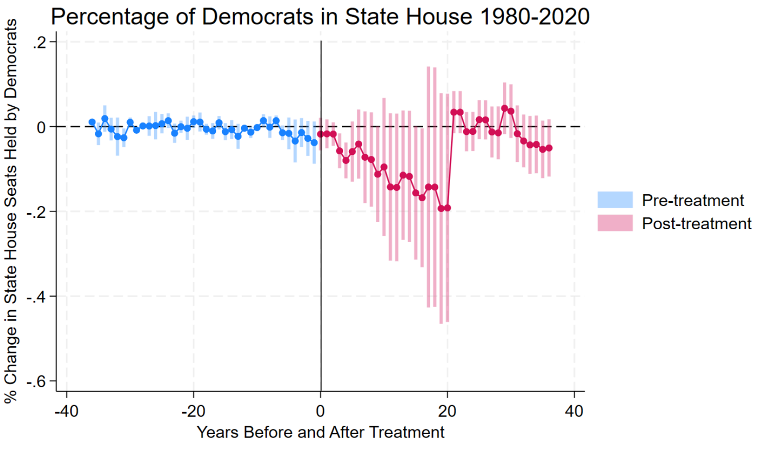


Figure 6. Percentage of House Seats Held by Democrats, 1980-2020



D. Implications

These findings have broad implications for the previous research. RTW laws appear to now have a negative effect on union density, however this has not always been the case. This seems to bridge the existing gap in literature on RTW laws. Previous research that appears to be contradictory may not in fact be so. Recent studies like Lunn (2023) and Fortin, Lemieux and Lloyd (2022) that find that RTW laws cause a decrease in union density may actually be correct as they only look at recent RTW laws while the study done by Newman and Moore (1987) would have been correct in saying that the laws are relatively insignificant. Interestingly though, this finding does not appear to be consistent with implications of the theory devised by Olson (1965) and Ostrom. This is not to say that these theories are wrong, there are certainly other potential explanations for why the theory does not match with outcomes, however, it is to say that there is likely more that is going on.

The implications of RTW laws for politics are also noteworthy. The clear decrease in voter participation as a result of RTW laws passed after the year 2001 does seem to largely confirm the findings of Feigenbaum, Hertel-Fernandez and Williamson (2018) which indicate that after the passage of RTW laws, voter participation should decrease. Interestingly, the average totals over the period are significantly positive, however, the separation of cohorts provides a clearer picture of what is actually happening. While the study certainly provides results that appear accurate, the study fails to tell the whole story. This is not true with partisanship, though. The research finds that RTW laws cause a significant increase in votes for Republican candidates, which tracks with the results found in this study.

IV. Conclusion

While many of the studies appear to have been correct, none seem to be able to explain the full story in regard to any of the indicators investigated in this study. Each of the previous studies has had an explanation for why their result found is consistent, however, this study indicates that the consistency assumed in previous studies is not entirely accurate. There is something more going on. This could be due to any number of factors. It's possible that WWII had an impact on the results or that the Great Society had an impact. It's also possible that the rise of the internet has had a real impact on why the impact of RTW laws has changed. No matter the reason,though, it is clear that RTW laws have changed in some capacity. To actually figure out what is causing this shift in impact, further research on the impact of RTW laws is necessary. This research is also particularly important because of the apparent impact that the laws have on politics. In order to understand why this is happening, it is essential to understand what is happening to unions and to voter participation. Especially in a time when RTW laws have renewed significance, this type of research is necessary.

References

Acemoglu, Daron and James A. Robinson. 2013. "Economics versus politics: Pitfalls of policy advice." *Journal of Economic Perspectives* 27(2): 173–92.

Bono-Lunn, Dillan. 2023. "The impacts of U.S. right-to-work laws on free riding, unionization, and compensation." *Southern Economic Journal* 90: 769-791. https://doi.org/10.1002/soej.12665

Caughey, Devin, and Christopher Warshaw. 2015. "The Dynamics of State Policy Liberalism, 1936-2014." *American Journal of Political Science* 60(4): 899-914. doi: 10.1111/ajps.12219

Collins, Benjamin. 2014. Right to Work Laws: Legislative Background and Empirical Research. *Congressional Research Service*. January 6. https://sgp.fas.org/crs/misc/R42575.pdf

Ellwood, David T. and Glenn A. Fine. 1983. "The Impact of Right-to-Work Laws on Union Organizing." NBER Working Paper Series No. 1116, National Bureau of Economic Research. https://doi.org/10.3386/w1116

Farber, Henry S. 1984. "Right-to-Work Laws and the Extent of Unionization." *Journal of Labor Economics*, 2(3): 319–352. http://www.jstor.org/stable/2534945

Feigenbaum, James, Alexander Hertel-Fernandez, and Vanessa Williamson. 2018. "From the Bargaining Table to the Ballot Box: Political Effects of Right to Work Laws." NBER Working Paper Series No. 24259, National Bureau of Economic Research. https://doi.org/10.3386/w24259

Flavin, Patrick and Benjamin Radcliff. 2011. "Labor Union membership and voting across nations." *Electoral Studies* 30(4): 633–641. https://doi.org/10.1016/j.electstud.2011.06.001

Fortin, Nicole, Thomas Lemieux, and Neil Lloyd. 2022. “Right-to-Work Laws, Unionization, and Wage Setting.” NBER Working Paper Series No. 30098, National Bureau of Economic Research. https://doi.org/10.3386/w30098

Henry S. Farber, Daniel Herbst, Ilyana Kuziemko, and Suresh Naidu. 2021. “Unions and Inequality over the Twentieth Century: New Evidence from Survey Data.” *The Quarterly Journal of Economics* 136(3):1325-1385. https://doi.org/10.1093/qje/qjab012

Ichniowski, Casey, and Jeffrey S. Zax. 1991. “Right-to-Work Laws, Free Riders, and Unionization in the Local Public Sector.” *Journal of Labor Economics* 9(3): 255–275. http://www.jstor.org/stable/2535144

Jacobs, David and Dixon, Marc. 2006. “The Politics of Labor-Management Relations: Detecting the Conditions that Affect Changes in Right-to-Work Laws.” *Social Problems 53*(1): 118–137. https://doi.org/10.1525/sp.2006.53.1.118

Kim, Dukhong. 2016. “Labor Unions and Minority Group Members’ Voter Turnout”. *Social Science Quarterly 97*(5): 1208–1226. https://www.jstor.org/stable/26612382

Moore, William. J., and Robert J. Newman. 1985. “The Effects of Right-to-Work Laws: A Review of the Literature”. *Industrial and Labor Relations Review 38*(4): 571–585. https://doi.org/10.2307/2523992

Olson, Mancur. 1977. *The Logic of Collective Action*. Harvard University Press: Cambridge

Radcliff, Benjamin and Patricia Davis. 2000. “Labor Organization and Electoral Participation in Industrial Democracies.” *American Journal of Political Science* 44(1): 132–141. https://doi.org/10.2307/2669299

Zullo, Roland. 2008. "Union Membership and Political Inclusion." *Industrial and Labor Relations Review 62*(1): 22–38. http://www.jstor.org/stable/25249183

Appendix

Appendix Table 1. Union Density, 1937-1979

Variable	Estimate	t-stat
GAverage	-0.00634	-0.45
G1944	-0.0460***	-3.32
G1946	-0.0301	-0.75
G1947	-0.0144	-0.70
G1953	0.0706***	5.67
G1954	0.0824*	2.01
G1955	-0.0224	-1.03
G1958	-0.0375*	-2.05

* p<0.05, ** p<0.01, *** p<0.001

Appendix Table 2. Union Density, 1980-2020

Variable	Estimate	t-stat
GAverage	-0.00940***	-5.93
G1985	0.0340***	7.64
G2001	0.0117***	4.64
G2012	-0.0103***	-5.44
G2013	-0.0164***	-6.79
G2015	-0.0316***	-10.78
G2016	-0.0210***	-8.34
G2017	-0.0315***	-12.88

* p<0.05, ** p<0.01, *** p<0.001

Appendix Table 3. Voter Participation 1937-1979

Variable	Estimate	t-stat
GAverage	0.0702***	5.99
G1944	0.120***	3.32
G1946	0.0194	0.61
G1947	0.0741***	3.74
G1953	0.143***	25.35
G1954	0.115***	8.53
G1955	-0.00357	-0.63
G1958	0.00747*	1.61

* p<0.05, ** p<0.01, *** p<0.001

Appendix Table 4. Voter Participation, 1980-2020

Variable	Estimate	t-stat
GAverage	-0.0133***	-5.83
G1985	-0.0156**	-3.05
G2001	-0.0281***	-6.83
G2012	-0.0125***	-4.10
G2013	-0.00190	-0.52
G2015	-0.0395***	-10.77
G2016	0.0155***	4.23
G2017	-0.0108*	-2.43

* p<0.05, ** p<0.01, *** p<0.001

Appendix Table 5. Percentage of State House Seats Held by Democrats, 1937-1979

Variable	Estimate	t-stat
GAverage	0.175***	-4.25
G1944	0.222***	-4.04
G1946	0.193	-0.86
G1947	0.183***	-3.11
G1953	-0.167***	-8.74
G1954	-0.204***	-7.26
G1955	-0.0919***	-5.07
G1958	-0.0246	-0.86

* p<0.05, ** p<0.01, *** p<0.001

Appendix Table 6. Percentage of State House Seats held by Democrats, 1980-2020

Variable	Estimate	t-stat
GAverage	-0.0663***	-8.94
G1985	0.00595	-0.32
G2001	-0.207***	-12.24
G2012	-0.120***	-9.08
G2013	0.00143	0.19
G2015	-0.0263***	-3.91
G2016	-0.0515***	-6.43
G2017	-0.0668***	-9.54

*p<0.05, ** p<0.01, *** p<0.001

The Effect of the No Surprises Act on Medical Out-of-Pocket Expenditures

Hannah Peschel

I. Surprise Billing and Healthcare Spending

On December 27, 2020, President Donald Trump signed the Consolidated Appropriations Act into law, an omnibus spending bill that incorporates compromises and provisions from both sides of the political aisle. Tucked in this 2000-page piece of legislation is the No Surprises Act, which aims to protect patients from surprise medical bills (unexpected bills sent by out-of-network providers or facilities) by increasing price and cost-sharing transparency ahead of time, as well as limiting the total amount a patient is required to pay to an out-of-network provider. Surveys find that an average of 18% of emergency department visits in the United States result in at least one surprise bill, and 67% of adults are concerned about being able to afford unexpected medical bills (Pollitz 2020). Surprise billing often arises from emergency situations when patients and families either do not have a choice or lack the time to make a choice regarding their provider of emergency services, such as anesthesia and air evacuation transportation methods such as air ambulances. Thus, this out-of-network charge can be burdensome for families and individuals who, if they are in an emergency medical care situation, are already facing other daunting concerns. This issue is compounded by the fact that the process for disputing medical claims is challenging, and the costs of disputing, both monetary and otherwise, are prohibitive for many families and individuals. The No Surprises Act endeavors to safeguard Americans from these issues at the federal level.

Prior to the passage of the No Surprises Act, some states attempted to alleviate these concerns for Americans by passing state-level protections against surprise billing from out-of-network providers. California, Florida, and New York were among the first to pass protections, and at the time of the federal No Surprises Act, 18 states had some form of legislation against surprise billing. The New York law specifically targeted surprise billing and enforced a new arbitration style to settle billing disputes between physicians and insurers. Anecdotal evidence indicates that balance billing is no longer a concern for residents of New York (Corlette and Hoppe 2019), but one empirical study found that provider out-of-network non-emergency billing actually increased in New York after the passage, which appears to be mostly driven by surgeons and surgical assistants (Gordon et. al 2021). The empirical analyses on state-level protections provide mixed results, and the federal No Surprises Act has yet to be empirically tested due to its fairly recent effective date and the lack of robust data on these most recent years.

My work will provide an early insight into the economic impacts of the federal No Surprises Act and whether or not Americans are actually seeing a decrease in their medical out-of-pocket expenditures as a result of the No Surprises Act. I hypothesize that states without previous legislation to protect against these expenses would see larger decreases in their medical out-of-pocket expenditures than states with pre-existing legislation. Using longitudinal household-level data from 2021-2022 and a difference-in-differences approach, I estimate the effect of this policy on family medical out-of-pocket expenditures. I ultimately find that this policy had the opposite effect, and actually coincided with increases in medical out-of-pocket expenditures for households where no previous protective legislation existed.

The paper proceeds as follows. Section 2 presents a review of the literature on state-level analyses of protection against surprise billing and further indicates how my analysis fills a gap in this literature. Section 3 reviews the methods I use to conduct my own analysis, Section 4 describes the data and model estimation strategy, Section 5 presents the results, and Section 6 concludes.

II. State-Level Analyses of Protections Against Surprise Billing

Due to the recency of the legislative passing of the No Surprises Act, studies have yet to emerge on the efficacy of the policy as a whole on reducing healthcare spending for Americans and on out-of-network spending on emergency services. The US Department of Health and Human Services' July 2023 report states that it will be important to understand the interaction of the federal No Surprises Act and other state-level protections because the policy itself does defer to pre-existing state protections in some areas, so some state variation in policy implementation is expected (Office of the Assistant Secretary for Planning and Evaluation, 2023). Nonetheless, a survey of the state-level analyses on these types of protections is necessary to understand the motivations for the No Surprises Act, as well as what might be expected in further empirical studies of the federal policy.

Most empirical studies examine the effect of a state-level protection on a specific targeted specialty in a given state. Fuse Brown et. al. (2019) focus on air ambulance services, which often come from out-of-network providers and burden the patient and their family with exceedingly high out-of-pocket payments for this service. Since air ambulances are used in emergency settings, the patient and their family do not get a say in their provider for the air ambulance, which means that surprise billing is more common here. This analysis focuses on Maryland and Maine due to their low ambulance service fees. Maryland runs a public air ambulance service, while Maine supports a non-profit, hospital-based (as opposed to network-based) air ambulance service. After the implementation of these systems, both states saw a significant decrease in their fees and charges from air ambulance services (Fuse Brown et. al. 2019). This analysis is not quantitatively robust in nature, but does provide evidence that certain state-level actions may work to limit healthcare spending on this emergency service. Another area where patients do not get to choose their own providers is anesthesiology.

La Forgia et. al. (2021) examine anesthesiologists in California, New York, and Florida, three states with some of the oldest surprise-billing protections. They measure differences in amounts paid to out-of-network anesthesiologists at in-network facilities before and after the protections were passed in each of these states. Using a differences-in-difference approach, the authors found that prices paid to out-of-network anesthesiologists did significantly decrease immediately after the passage of the state's protection policy in California and Florida. In New York, a significant decrease did not appear until the fourth quarter after the law was passed. These studies provide insight into where we might expect the No Surprises Act to have significant effects in other states, but it is also possible that their results are idiosyncratic to the targeted state and specialty combination.

Thus, it would be fruitful to attempt a more generalizable analysis that can provide a more illustrative picture about the effect of surprise-billing protections. Few studies have tackled this problem but Adler et. al. (2019) do attempt to do this for the state of California. Their analysis examines the effect of the legislation on the *amount* of care provided, rather than the amount spent on care, which is still relevant. If out-of-network care decreases, it can be assumed that this in-network care is being substituted in its place. They find that out-of-network care did decrease from the pre-policy period to the post-policy period, which indicates a shift from out-of-network to in-network providers due to the passage of the policy.

Adler et. al. (2021) attempts another generalization, this time with emergency services in Connecticut, another one of the pioneering states in passing state-level balance billing protections. Instead of looking at affected specialties, this study analyzes the impact of Connecticut's legislation on spending on emergency services across the state. Interestingly, they found a significant *increase* in the amount paid to emergency service physicians after the policy was passed. The authors attribute this to setting the minimum payment to out-of-network emergency physician providers too high, and instead advocate for Connecticut's reversal of this policy.

Overall, the empirical evidence on these programs is mixed. The effect of state-level protections against balance and surprise billing is mixed across certain states and specialties, and there does not appear to be a generalizable effect for these state laws on any of the dependent variables considered by the authors. This is likely due in part to a recency effect–these laws have only been in place for a few years at most, so understanding their full impact will take more time. A more relevant metric for evaluating the effect of these programs, however, might be overall out-of-pocket spending on healthcare for Americans. While it is important to understand the primary outcomes like the allowed amounts to physicians in certain specialities, as done by the studies reviewed above, no state level analyses has examined downstream effects of this policy on things like out-of-pocket expenditures on healthcare, quality of care, or access to care (Office of the Assistant Secretary for Planning and Evaluation, 2023).

Thus, my work seeks to fill two gaps in the literature on surprise billing protections. First, by conducting a preliminary analysis on the effect of the federal No Surprises Act and how this might vary between states who had state-level protections in place compared to states who did not have state-level protections, and second, by providing an analysis on the downstream effects of the No Surprises Act rather than the direct outcome.

III. Methods

Using the data described above, I employ a differences-in-differences approach to estimate the effect of the No Surprises Act. This study design estimates the effect of a particular policy by examining its differential effect on a group that was exposed to the “treatment,” in this case a policy, compared to the group that was not exposed to the treatment. The quintessential differences-in-differences analysis is conducted by Card and Krueger (1995). They examined the effect of a minimum wage increase for fast food workers in New Jersey and compared the results to New York, where the minimum wage remained the same. The relationship between New Jersey and New York, two states similar in many other regards but different in the policy effect, presented a natural experiment with a treatment and a control group, and this is what I attempt to replicate, albeit imperfectly, here.

The differences-in-differences methodology is especially useful in healthcare policy analysis because of its ability to control for background changes in patient outcomes that occur with time (Dimick and Ryan, 2014). This means that the analysis will not erroneously come to the conclusion that a policy is causing a certain outcome when the outcome may have resulted anyways, and this is done through the use of a natural experiment situation that presents itself. Not all natural experiments will be perfect, but they allow for the ability to control for state and time differences to determine the effect of a policy change. By using the differences-in-differences approach, I am able to compare the effect of the policy directly by examining a group that was exposed to the policy change and one that was not.

As discussed above, 18 states had passed their own protections against surprise billing prior to the federal No Surprises Act.[1] I will use households from these states as the control group, and households from states where no previous legislation was in place as the treatment group. If a state already had protective legislation on surprise billing, households in that state likely would not be as affected by the No Surprises Act as households in states where no legislation previously existed. For model estimation, this variable is coded as

[1] California, Colorado, Connecticut, Florida, Georgia, Illinois, Maine, Maryland, Michigan, New Hampshire, New Jersey, New Mexico, New York, Ohio, Oregon, Texas, Virginia, Washington

a binary indicator, which takes on a value of “1” if the state had previous legislation against surprise billing and a value of 0 if the state did not have previous legislation against surprise billing. Households with a value of 1 are thus a part of the control group, and households with a value of 0 are part of the treatment group. Thus, the two independent variables of interest are the previous law indicator and the time indicator, which takes on a value of 1 for the pre-policy period and a value of 1 for the post-policy period. I will examine the interaction of these two variables to determine whether the effect of the policy on medical out of pocket expenditures is different in states with previous legislation compared to those without previous legislation.

Formally, the null hypothesis is that there is no difference in the change in mean medical out-of-pocket expenditures from the pre-policy period to the post-policy period between households in states with previous legislative protections and states without previous legislation. The alternative hypothesis is that there *is* a difference in the change in medical out-of-pocket expenditures from the pre-policy period to the post-policy period for households in states with previous protective legislation compared to households in states without previous protective legislation. Specifically, if the No Surprises Act had an effect on medical expenditures, I would expect that states without previous legislation would see a greater decrease in medical out-of-pocket expenditures after policy implementation than states with legislation already in place, but I will use a two-sided alternative hypothesis to examine the possibility of both directions. See equations (1), (1a) and (2) below.

(1) $H_0\colon \mu_{treatment,\,pre} - \mu_{treatment,\,post} = \mu_{control,\,pre} - \mu_{control,\,post}$ or,

(1a) $H_0\colon \mu_{diff\ in\ treatment} - \mu_{diff\ in\ control} = 0$

(2) $H_A\colon \mu_{diff\ in\ treatment} - \mu_{diff\ in\ control} \neq 0$

Some limitations do exist with this method. The two primary assumptions that must be met to conduct a differences-in-differences analysis are the parallel trends assumption and the common shocks assumption. The parallel trends assumption states that the two groups must have been following the same trend before the policy intervention (i.e. their regression lines would be 'parallel'). For the Card and Krueger (1995) analysis, this condition was satisfied by the fact that New York and New Jersey had the same minimum wage prior to the change implemented by New Jersey. In this analysis, I would need to examine the trends of medical out-of–pocket expenditures over a longer period of time in each of the two groups. This would be an extension of the traditional model, since the original study was only a two-time period model. My data does not inherently allow for checking this condition, but I proceed with the difference-in-difference methodology.

Additionally, the common shocks assumption states that any economic shocks experienced by one group must also have been experienced by the other group. In my analysis, this means that I assume any economic shocks in the pre-policy period were felt by both the treatment and the control group. While states may each have their own smaller economic shocks, I would argue that any large-scale economic shocks would be consistent across all states in the U.S. regardless of their exposure to a certain policy or not, so while this assumption cannot be directly tested, we can proceed with this methodology since there is no strong reason for rejecting this assumption.

IV. Data

A. Data Source and Variable Descriptions

To test these hypotheses, I use data from the Current Population Survey (CPS) compiled by IPUMS. This is a monthly household survey run by the U.S Census Bureau to obtain information on demographics, education level, labor force statistics, and other social and economic characteristics of the U.S population. While this survey contains a vast amount of information, the available variables related to healthcare and specifically healthcare expenditures are relatively limited, especially when coupled with this study's need for 2022 data to properly analyze the effect of the new policy. Nonetheless, I obtained my primary dependent variable of interest from this survey, which is total family medical out-of-pocket expenditures. Medical out-of-pocket expenditures measures the total out-of-pocket payments made by a family towards medical expenses. Given that my dependent variable is measured at the family level, each observational unit in my data is a unique household-year combination.

Other control variables are necessary to guard against omitted variable bias. These variables are also taken from the CPS and include number of children under the age of 5 in the household, number of people in the household overall, household income, and average age of the household, and whether this household is in the metro area or not. While these are not the primary variables of interest for the study, it is necessary to control for these variables in our model estimation to properly estimate the effect of the policy on medical out-of-pocket expenditures.

A few other cleaning steps were necessary in preparation for the analysis. I dropped 361 households with an average income of 0 dollars for the year and four households that had medical out-of-pocket expenditures greater than 125,000. While these outliers in medical out-of-pocket expenditures may present interesting cases, I exclude them here to get a more representative sample of American households. To remedy those with $0 in medical out-of-pocket expenditures, I also estimate a left-censored Tobit model in a second set of specifications as outlined below.

I also dropped observations where the total number of people in the household was less than the number of children reported, as well as those where the number of children is reported as 0 and the number of children less than 5 was reported as greater than 0. After dropping these households, I am left with 16,139 observations for 2022, and 16,152 observations for 2021. This gives us a slightly unbalanced panel, but given the small difference between the two (13 observations out of about 16,000 for 0.08%), I proceed with model estimation using the slightly unbalanced panel.

B. Model Estimation

The dependent variable of interest is medical out-of-pocket expenditures in each year. The independent variables of interest are the previous law and time indicators, and we also have a vector of control variables. Thus, the model I will estimate takes the following form:

$$(3) \quad y_{it} = \beta_1 + \beta_2 previous\ law_i + \beta_3 time_t + \beta_4 previous\ law_i * time_t + \beta_5 X_i + \epsilon_{it}$$

where y is the medical out-of-pocket expenditure for individual i in time period t and X_i is a vector of the control variables listed above, and β_4 is the difference-in-differences estimator.

V. Results

A. Descriptive Statistics

After the dropped observations outlined above, I am left with a dataset of 32,291 households, with 16,152 in 2021 and 16,139 in time period 2. There are 14,613 households from the 18 states that did have previous protective legislation, and 17,678 households from the 32 states that did not already have previous legislation. The average age for households in this sample is 50.2 years, and the average household income is $100,323.30

For households in states with previous legislation, the mean medical out-of-pocket expenditures in the pre-policy time period was $4377.47, and in the post-policy time period, it was $4552.75, which gives us a difference in the treatment group of -$175.28 (see row 1 in Table 1 below). Notably, this difference is negative, which indicates that medical out-of-pocket expenditures actually increase in 2022 for the treatment group.

For individuals in states without previous legislation, the mean medical out-of-pocket expenditures was $4483.01 in the pre-policy time period and $4409.98 in the post-policy time period, which gives a difference of $73.03. Here, we see that medical out-of-pocket expenditures *decreased* in 2022 for the control group. Additionally, the control group spent an average of $105.54 *more* than the treatment group in the pre-policy time period, but in the post-policy time period, the control group spent an average of $142.77 *less* than the treatment group, again indicating that the treatment group saw their medical out-of-pocket expenditures rise more than the control group from 2021 to 2022. The difference-in-differences for these groups, then, is -$248.31. This tells us that the treatment group saw their medical-out-of-pocket expenditures decrease by -$248.31 more than the control group. In other words, the medical out-of-pocket expenditures for households in states without previous legislation increased by $248.31 more dollars than for households in states with previous legislation.

See Table 1 below for the means and differences in means for each group and time period. Based on this, we can see that the impact of the policy is different on states with previous legislations compared to states without previous legislation, and this difference in impact is what the difference-in-differences analysis will attempt to estimate.

Table 1. Difference-in-Differences Table

	Pre-Policy (time = 0)	Post-Policy (time = 1)	Difference
Households in states without previous legislation (treatment group) (previous = 0)	$4377.47 (n = 8844)	$4552.75 (n = 8834)	-$175.28 (n = 17678)
Households in states with previous legislation (control group) (previous = 1)	$4483.01 (n = 7308)	$4409.98 (n = 7305)	$73.03 (n = 14613)
Difference	-$105.54 (n = 16152)	$142.77 (n = 16139)	-$248.31 (n = 32,291)

B. Difference-in-Differences Model Estimation

I first estimate a simple difference-in-differences model with no control variables. Medical out-of-pocket expenditures significantly increased at the 10% level from 2021 to 2022
(t = 1.93, p = 0.053). The difference-in-differences estimator is also significant at the 10% level (t = -1.84, p = 0.065); however, it enters with the opposite sign from my hypothesis. Since this is a simple model, the coefficients in the model results should be the same as the differences outlined in the table above, and this is reflected in column (1) of Table 2 below. The coefficient on the differences-in-differences estimator is -248.31, which tells us that the policy had the effect of an additional $248.31 increase in medical-out-of-pocket expenditures in the treatment group as compared to the control group. Thus, the treatment group saw a decrease in medical out-of-pocket expenditures that was $248.31 less than the control group; in other words, they saw an *increase* in medical out-of-pocket expenditures where the control group saw a *decrease*, as above.

The R-squared on this model is extremely low, indicating that our model is not a great fit. In line with the literature on modeling medical expenditures, I include control variables for type of healthcare coverage, age of the individual, gross income, and a general measure of health to guard against omitted variable bias and increase model performance. Including each of these variables as an additional predictor in a stepwise fashion along with the simple model yields similar results, as seen in columns (2) through (6) in Table 2 below. The final model given in column (6) below incorporates all of the significant control variables and yields an R-squared value of just over 11%. Considering the initial value was under 1%, this set of control variables does a decent job of predicting medical out of pocket expenditures for households as compared to just the time and previous legislation variables.

Importantly, the differences-in-differences estimator is robust to the inclusion of control variables. In all specifications, the differences-in-difference estimator is negative and significant, with a t-value hovering around -2.13 and a p-value around 0.031. The No Surprises Act does appear to have had an effect, although the opposite effect of what I initially hypothesized. The fact that the difference-in-differences estimator is negative comes from the way that I conducted the subtraction, doing (pre-policy - post-policy). If medical out-of-pocket expenditures lowered after the policy implementation, the difference here would be positive. The negative value tells us that the treatment group saw a decrease in medical out of pocket expenditures that was \$270.90 less than the control group. In other words, households in states without previous legislation actually saw an increase in their medical out of pocket expenditures after the passage of the No Surprises Act, and households in states with previous legislation are the ones who saw a decrease in medical out of pocket expenditures in the post-policy time period. Thus, the No Surprises Act may have actually contributed to an increase in medical out of pocket expenditures for households with no other legislation in place to protect them from surprise billing.

Table 2. Diff-in-Diff Estimates for Medical Out-of-Pocket Expenditures, 2021-2022

	OLS Models						Tobit Models		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Previous Law in Place	105.55	-34.59	-33.92	-41.71	-91.04	-89.56	49.07	-137.28	-156.09*
	(95.26)	(93.29)	(93.28)	(93.13)	(91.70)	(91.71)	(100.85)	(96.00)	(95.59)
Year	175.28*	112.70	120.42	121.77	105.65	106.44	168.06*	92.78	94.95
	(90.64)	(84.89)	(84.87)	(84.80)	(83.26)	(83.25)	(92.05)	(86.83)	(86.51)
(Previous Law) x (Year)	-248.31*	-278.89**	-278.94**	-279.92**	-272.99**	-270.90**	-231.85	-258.38*	-256.17*
	(134.74)	(129.58)	(129.54)	(129.41)	(127.52)	(127.54)	(141.46)	(133.50)	(133.03)
Household Income (thousands)		15.29***	15.08***	14.78***	12.84***	12.84***		14.39***	13.60***
		(0.645)	(0.652)	(0.65)	(0.617)	(0.617)		(0.65)	(0.64)
Age			-7.29***	65.50***	166.97***	174.06***		26.03***	184.46***
			(1.62)	(8.91)	(9.60)	(9.86)		(1.81)	(10.13)
Age²				-0.737***	-1.40***	-1.47***		—	-1.55***
				(0.09)	(0.09)	(0.09)			(0.097)
Number of People in Household					922.29***	911.57***	—	873.28***	963.48**
					(34.71)	(34.66)		(34.66)	(36.44)
Number of Children under 5 in Household (2)						847.87***			875.19***
						(349.49)			(357.15)
Constant	4377.47	2945.46	3327.41	-2.76	-3222.18	-3381.92	4217.20	-3222.18	-3998.19
N	32,291	32,291	32,291	32,291	32,291	32,291	32,291	32,291	32,291
R^2	0.0001	0.0836	0.0841	0.0842	0.1137	0.1149	0.0001	0.0061	0.0065

1: OLS estimates are reported with robust Standard Errors. Estimates are significant at *p<0.1; **p<0.05; ***p<0.01
2: For Tobit models, the pseudo-R^2 value is reported.
3: The number of people in a household is reported here as a continuous variable. See the Appendix for the marginal effects analysis, which treats this as a discrete variable and conducts the analysis accordingly.

Importantly, the differences-in-differences estimator is robust to the inclusion of control variables. In all specifications, the differences-in-difference estimator is negative and significant, with a t-value hovering around -2.13 and a p-value around 0.031. The No Surprises Act does appear to have had an effect, although the opposite effect of what I initially hypothesized. The fact that the difference-in-differences estimator is negative comes from the way that I conducted the subtraction, doing (pre-policy - post-policy). If medical out-of-pocket expenditures lowered after the policy implementation, the difference here would be positive. The negative value tells us that the treatment group saw a decrease in medical out of pocket expenditures that was $270.90 less than the control group. In other words, households in states without previous legislation actually saw an increase in their medical out of pocket expenditures after the passage of the No Surprises Act, and households in states with previous legislation are the ones who saw a decrease in medical out of pocket expenditures in the post-policy time period. Thus, the No Surprises Act may have actually contributed to an increase in medical out of pocket expenditures for households with no other legislation in place to protect them from surprise billing.

Some interesting secondary results also emerge in the interpretation of the control variables. For average household income income level, I find that every $1000 increase in average household income corresponds with an $12.84 increase in medical out-of-pocket expenditures ($t = 20.82$, $p < 0.0001$), after accounting for average household age, number of people in household, time period, and whether the household is in a state with a prior law in effect. Household income appears to be an important predictor here–adding it to just the simple model increases the adjusted R-squared value by over 8%, as seen in specification (2). I also examined metro status and the number of children under the age of five, but these did not show up significantly in specification (5), and Wald postestimation tests confirmed this result for all except having two children under the age of five, so I include this level of the variable in my final model. See the Appendix for the Wald test results.

As expected, age enters positively and significantly in all model specifications. In the final model specification, each additional increase in age from the mean corresponds with a \$174.06 increase in medical out-of-pocket expenditures ($t = 17.86$, $p < 0.0001$) after accounting for household income, number of people in household, average income of the household, time period, and whether the household is in a state with a prior law in effect. This is given in the regression output, and it is the same as the result obtained from the marginal effects analysis (see Appendix Table 1 for the marginal effects results). I also include an age squared term in specifications (4) through (6). It would be expected that households and families would spend more money on medical expenditures as age increases, since medical expenses likely become more common and involved over time. When included, the age squared term enters negatively and significantly. This indicates that at a point, age actually decreases the amount of medical expenses. Looking at the marginal effects, this change appears to happen between the ages of 55 and 65. See Appendix Table 2a for these results. One potential explanation for this is that Medicaid kicks in at age 65, which significantly lowers the out-of-pocket expenditures that older adults have to spend on medical services. The fact that this kicks in between ages 55 and 65 makes sense when considering that this measure of age is an average; at an average household age between 55 and 65, there is likely at least one person in the household 65 or older and receiving Medicare, which could explain why medical expenditures start to decrease here. Individual-level analysis would likely want to exclude individuals aged 65 and up, but since this analysis is conducted at the household level and the age variable is an average across the household, I elected not to limit the average age variable so as to not exclude any households who might be living with an older relative who lifts up the average age of the household.

Finally, the number of people in a household enters positively and significantly, as expected. With each additional person added to the household, the final model predicts that the household will increase their medical expenditures by an additional \$911.57 ($t = 26.30$, $p < 0.0001$). The marginal effects analysis is relevant here as well; we cannot have any half people in a household, so the number of people in a household is technically a discrete variable. Wth this analysis, we can determine the marginal effect of an additional person in the household on medical out-of-pocket expenditures. The marginal effects for the final model tell us that the marginal effect of going from a single-person household to a two-person household on medical out-of-pocket expenditures is $1994.67, and it is $1063.04 for going from a two-person to a three-person household. See Appendix Table 3 for the marginal effects of increasing the number of people in a household for other numbers of people.

C. Tobit Model Estimation

The response variable in this analysis is medical out-of-pocket expenditures. Since a person cannot have a negative value for medical out-of-pocket expenditures, this can be considered a left-censored value for the data with a lower limit of 0, and I can estimate a series of left-censored Tobit models to the data. In my sample, 1,613 of the 32,291 observations are censored, which is a censorship rate of just under 5%. Due to this low censorship rate, I do not expect the Tobit model estimates to give drastically different estimates from the original difference-in-differences analysis. See specifications (7) through (9) in Table 2 below for the results from the Tobit regression

For the Tobit analysis, I again start with the simple model with no additional covariates beyond time period and the indicator for treatment/control group. As expected, the estimates here are quite similar to the simple OLS difference-in-differences model estimation. The variable for metro status is still insignificant, and although the difference-in-differences estimator is not significant in the simple model, it becomes negative and significant with the inclusion of other control variables like age and age squared, household income, and number of people in the household.

Similarly, the magnitude of the coefficient estimates are not substantially different between the OLS and the Tobit models. For the differences-in-differences estimator in the final OLS model, the treatment group experienced a decrease in medical out-of-pocket expenditures that was $270.90 less than that experienced by the control group; in the final Tobit model, this decrease was $258.40. Similarly, in the final OLS model, each increase in average age of the household corresponded to a $166.97 increase in medical out-of-pocket expenditures, and in the final Tobit model, each increase in average age of the household corresponded to a $177.07 increase in medical out-of-pocket expenditures after controlling for the other covariates. The marginal effects were also similar between the Tobit and OLS models; see Appendix Tables 1b, 2b, and 3b for the results from the Tobit marginal effects analysis.

VI. Conclusion

Overall, my analysis shows that the No Surprises Act did *not* have an effect on lowering the medical out-of-pocket expenditures. Instead, medical out-of-pocket expenditures actually significantly increased where there was no protective legislation already in place, which was the opposite effect of what I had initially hypothesized. This result was robust to the inclusion of control variables, and the magnitude of the coefficient estimate on the difference-in-differences estimator remained relatively stable across model specifications.

One important limitation of this study is the somewhat broad dependent variable. The No Surprises Act aims to limit out-of-pocket spending in emergency situations, which plays a part in medical out-of-pocket expenditures but is inherently more specific. As more data in the post-policy period becomes available, employing data like that from the Medical Expenditure Panel Survey may provide more insight into the true effect of the No Surprises Act on the aspect that it was designed to directly target.

Moving forward, I would not argue that this policy should be repealed based on the results discussed here. Importantly, this analysis only considers a two-time period model due to the recency of the passage of this legislation. It may be that with more time and more data in the post-policy time period, the No Surprises Act could lead to decreases in medical out-of-pocket expenditures. As I have shown, it appears that households in states with previous legislation are seeing their medical out-of-pocket expenditures decrease years after their policy went into effect, so it could be the case that this policy has a temporal component to seeing and experiencing its benefits. While the legislation went into effect on January 1 of 2022, it is unlikely that implementation is perfect even now, so future research should continue to monitor the situation and conduct further analysis as the true policy impact unfolds.

References

Adler, Loren, Erin Duffy, Bich Ly, and Erin Trish. 2019. "California saw reduction in out-of-network care from affected specialties after 2017 surprise billing law." *Brookings Institute*, September 26. https://www.brookings.edu/articles/california-saw-reduction-in-out-of-network-care-from-affected-specialties-after-2017-surprise-billing-law

Anjali A. Dixit, D. Lee Heavner, Laurence C. Baker, Eric C. Sun. 2023. Association between "Balance Billing" Legislation and Anesthesia Payments in California: A Retrospective Analysis. *Anesthesiology 139*:580–590 doi: https://doi.org/10.1097/ALN.0000000000004675

Card, David and Krueger, Alan B. 1994. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania*." American Economic Review 84*(4): 772-793.

Corlette, Sabrina and Olivia Hoppe. 2019. "New York's Law to Protect People from Surprise Balance Bills is Working as Intended, but Gaps Remain." Center on Health Insurance Reforms, Georgetown University. May 13. https://chirblog.org/new-york-law-surprise-balance-billing/

Cooper, Zack, Fiona Scott Morton, and Nathan Shekita. 2020. "Surprise! Out-of-network billing for emergency care in the United States." *Journal of Political Economy 128* (9): 3626-3677.

Dimick, Justin B., and Andrew M. Ryan. 2014. "Methods for Evaluating Changes in Health Care Policy: The Differences-in-Differences Approach." *Journal of the American Medical Association 312*(22): 2401–2402. doi:10.1001/jama.2014.16153

Fuse Brown Erin C., Alex McDonald, and Ngan T. Nguyen. 2020. "What States Can Do to Address Out-of-Network Air Ambulance Bills." *The Journal of Law, Medicine & Ethics 48* (3): 462–73. doi: 10.1177/1073110520958869.

Gordon, Aliza S., Ying Liu, Benjamin L. Chartock, and Winnie C. Chi. 2021. "Provider Charges and State Surprise Billing Laws: Evidence from New York and California." *Health Affairs 41* (9). https://doi.org/10.1377/hlthaff.2021.01332

La Forgia Ambar, Amelia M. Bond, Robert T. Braun, Klaus Kjaer, Manyao Zhang, and Lawrence P. Casalino. 2021. "Association of Surprise-Billing Legislation with Prices Paid to In-Network and Out-of-Network Anesthesiologists in California, Florida, and New York: An Economic Analysis." *JAMA Internal Medicine 181*:1324–1331. doi: 10.1001/jamainternmed.2021.4564

Lieneck, Cristian, Mario Gallegos, Madison Ebner, Hannah Drake, Emma Mole, and Kaitlin Lucio. 2023. “Rapid Review of "No Surprise" Medical Billing in the United States: Stakeholder Perceptions and Challenges.” *Healthcare 11*(5): 761. doi:10.3390/healthcare11050761

Office of the Assistant Secretary for Planning and Evaluation, U.S Department of Health and Human Services. 2023. *Evaluation of the Impact of the No Surprises Act on Health Care Market Outcomes: Baseline Trends and Framework for Analysis– First Annual Report*. https://aspe.hhs.gov/sites/default/files/documents/48b874b63796dc6a68a783cf079ba42a/aspe-no-surprises-act-rtc.pdf

Pollitz, Karen, Lunna Lopes, Audrey Kearney. 2020. “US Statistics on Surprise Medical Billing.” *JAMA Internal Medicine 323(6):* 498. doi:10.1001/jama.2020.0065

Appendix

Appendix Table 1a. Marginal Effects of Continuous Variables from Final OLS Model At Means

Variable	***dy/dx***	***Std. err.***	***t-stat***	***p-values***
HH Income	12.84	0.62	20.82	0.000
Average Age	26.65	1.74	15.35	0.000
Number in HH (see below for individual marginal effects)	911.57	34.66	26.30	0.000

Appendix Table 1b. Marginal Effects of Continuous Variables from Final Tobit Model At Means

Variable	*dy/dx*	*Std. err.*	*t-stat*	*p-values*
HH Income	13.60	0.64	21.13	0.000
Average Age	28.15	1.84	15.30	0.000
Number in HH (see below for individual marginal effects)	963.48	36.39	26.47	0.000

Appendix Table 2a. Marginal Effects of Age at Different Values from Final OLS Model

Age at:	*dy/dx*	*Std. err.*	*t-stat*	*p-values*
20	115.33	6.19	18.63	0.000
40	56.60	2.78	20.38	0.000
50	27.24	1.74	15.63	0.000
55	12.56	1.83	6.86	0.000
65	-16.80	3.05	-5.51	0.000
80	-60.85	5.616	-10.84	0.000

Appendix Table 2b. Marginal Effects of Age at Different Values from Final Tobit Model

Age at:	*dy/dx*	*Std. err.*	*t-stat*	*p-values*
20	122.19	6.53	18.70	0.000
40	59.91	2.93	20.45	0.000
50	28.77	1.85	15.59	0.000

55	13.20	1.95	6.78	0.000
65	-17.93	3.24	-5.53	0.000
80	-64.64	5.95	-10.86	0.000

Appendix Table 3. Marginal Effects of Number of People in a Household from Final OLS Model

Number in Household	*Mean*	*Difference (Marginal Effect)*
1	2542.765	—
2	4537.434	1994.669
3	5515.178	977.744
4	6578.214	1063.036
5	6851.346	273.132
6	6550.014	-301.332
7	5876.844	-674.014
8	6768.307	891.462
9	7053.073	284.766
10	6783.827	-269.246
11	5705.459	-1078.368

Appendix Table 4. OLS Model (5) Results with Metro Indicator and Wald Test Results

Variable	Coefficient	Robust SE	Test Stat	p-value
Previous Law	-81.8	92.62	-0.88	0.377
Time	105.57	83.27	1.27	0.205
Previous Law *Time	-273.07	127.52	-2.14	0.032
Household Income (Thousands)	12.86	0.62	20.75	0.000
Average Age	166.95	9.61	17.38	0.000
Average Age^2	-1.404	0.092	-15.28	0.000
Number of People in Household	921.81	34.75	26.53	0.000
Metro	-54.52	77.57	-0.70	0.483

Wald Test
Null Hypothesis: 1.metro_ind2 = 0
Test Stat: F(1, 32282) = 0.49
P-value = 0.4830

Appendix Table 5. OLS Model (5) Results with Number Children Less than 5 Indicator and Wald Test Results

Variable	Coefficient	Robust SE	Test Stat	p-value
Previous Law	-90.29	91.72	-0.98	0.325
Time	106.6	83.26	1.28	0.2
Previous Law *Time	-270.93	127.55	-2.12	0.034
Household Income (Thousands)	12.84	0.62	20.81	0.000
Average Age	174.97	10.46	16.72	0.000
Average Age^2	-1.47	0.098	-14.93	0.000
Number of People in Household	912.43	25.62	26.53	0.000
Number of Children Less than 5	—	—	—	—
1	57.20	156.65	0.37	0.715
2	859.49	358.26	2.40	0.016
3	-245.73	794.26	-0.31	0.757
4	-1171.97	1864.66	-0.63	0.530
5	-3235.43	4935.69	-0.566	0.512

Wald Test

Null Hypothesis:

1.nchlt5 = 0
3.nchlt5 = 0
4.nchlt5 = 0
5.nchlt5 = 0

F-Statistic: F(4, 32278) = 0.28
p-value = 0.8933

Omicron Delta Epsilon
Department of Economics
St. Olaf College
Northfield, MN

ST. OLAF COLLEGE