The MIT Undergraduate Journal of Economics

Volume XXIV

2024-2025

Income Gaps and Political Divides: Local Inequality's Effect on Local Election Polarization
Fara Alade

The Impact of Autism Therapy Coverage Mandates on Outcomes for Autistic Children
Peter Berggren

Assessing the Impact of Project Green Light: A Google Sustainability Initiative to Reduce City Air Pollution
Kenneth Byrne

Motherhood and the Market: Paid Maternity Leave Policy Impacts on Women's Wages and Labor Force Participation
Mackenzie Bivin

The Effect of Quotas on Wages: The Homegrown Player Rule in the Premier League
Kate Ellison

The Effect of Changes in Fuel Prices on Adoption of Farming Mechanization: A Cross-Country Analysis
Lukas Hanson-Puffer

Shuffling the Deck: Did NYC's Lottery Admissions Reshape School Demographics?
Sonia Seliger

The MIT Undergraduate Journal of Economics Volume XXIV

2024-2025

Mailing Address:

The MIT Undergraduate Journal of Economics
Massachusetts Institute of Technology, Building E52-301
Cambridge, MA 02139

Foreword

"Money… must always be scarce with those who have neither wherewithal to buy it, nor credit to borrow it."

- Adam Smith

As MIT undergraduate economics students progress through their coursework, they are continuously introduced to new economic topics, constantly learning the ideas and models of established economists, and relentlessly being challenged to think differently about the observable phenomena around them. It is this enthusiasm for learning that led undergraduates at MIT to proceed in their own research—to experience the excitement of asking a question and striving to answer it. We hope that this year's papers highlight the vigor with which our undergraduate students pursue economic research and the rigor with which they present their ideas.

The publication of this Journal is made possible by the support of many people. We especially thank Professor Dave Donaldson for selecting the articles for this year's publications.

These relevant student papers demonstrate the enduring importance of rigorous economic research in the days ahead.

The MIT Undergraduate Journal of Economics Volume XXIV

2024-2025

Contents

Income Gaps and Political Divides: Local Inequality's Effect on Local Election Polarization
Fara Alade

The Impact of Autism Therapy Coverage Mandates on Outcomes for Autistic Children
Peter Berggren

Assessing the Impact of Project Green Light: A Google Sustainability Initiative to Reduce City Air Pollution
Kenneth Byrne

Motherhood and the Market: Paid Maternity Leave Policy Impacts on Women's Wages and Labor Force Participation
Mackenzie Bivin

The Effect of Quotas on Wages: The Homegrown Player Rule in the Premier League
Kate Ellison

The Effect of Changes in Fuel Prices on Adoption of Farming Mechanization: A Cross-Country Analysis
Lukas Hanson-Puffer

Shuffling the Deck: Did NYC's Lottery Admissions Reshape School Demographics?
Sonia Seliger

Income Gaps and Political Divides:

Local Inequality's Effect on Local Election Polarization

Abstract: A variety of literature has examined the rise of polarization in the U.S. and sought to identify the driving factors behind it. This paper explores the relationship between county-level income inequality and political polarization in senatorial and presidential elections from 2008 to 2020, using a lagged Gini coefficient as an instrument to address endogeneity. County-level polarization is assessed by the closeness of vote shares between Democratic and Republican candidates. The findings reveal a negative relationship between income inequality and polarization, indicating greater party dominance in regions with higher inequality. This relationship is statistically significant in senatorial elections but not in presidential elections. However, due to potential mechanisms such as media dynamics, voter mobilization, and historical policy shifts, the exclusion restriction does not hold, preventing us from identifying a causal relationship. The results offer evidence of an association that contributes to the literature on the link between income inequality and political polarization at a local level. Future research should investigate and account for alternative mechanisms to establish a causal relationship between county-level income inequality and political polarization.

Fara Alade

14.33

1. Introduction

As seen with the 2020 and 2024 presidential elections, elections with close margins of victory have become more common in US politics. This close partisanship could be seen as a positive development since the two parties represent close to half of the voting population. However, an increase in narrow elections can also lead to less of the population feeling represented, which encourages “the losing party to double down on existing policies rather than build a different and bigger coalition”. (Lindsay 2024)

These close races may worsen the recent false claims of ‘stolen’ elections and election interference. This closeness also often leads to calls for recounts and decreases public faith in election integrity. At the local level, closeness may also provide an interesting view of polarization. Navarrete et. al, find that an election is polarized if voters’ preferences are highly clustered geographically and there is a high turnout. Increased polarization in the US has been shaped by multiple factors, including financial crises, fragmented media coverage, and changes in party positions among others (Mian et. al. 2014, Kubin and von Sikorski 2021, Voorheis et. al 2015).

The literature has examined income inequality and economic perceptions as drivers of vote choice, establishing a link between rising inequality and the success of more extreme parties in elections (Proaño, Peña, and Saalfeld 2022). Proaño, Peña, and Saalfeld examine the relationship between changing income shares and the success of moderate and far-right parties. In this example, parliament remains relatively moderate while there is some increased success of far-right parties. Crises and other negative economic situations can further entrench people’s views into a particular party's message or economic framework. Greater income inequality also concentrates economic interests in smaller groups with the ability to contribute significantly to campaigns and shape the voting behavior of an area (Duca et al. 2014). Other factors like rising inequality due to globalization may also shape how people vote and further antagonize the other party. These factors affect how local areas vote, as income inequality may affect who people decide to vote for by either pushing them to the less popular party or strengthening a party's position. When people experience income inequality in their community, they may be frustrated by the more popular party or continue to vote for the party they believe will strengthen their community economically. Additionally, a county-level view helps provide a future understanding of how people consider inequality in their surrounding areas when voting in national-level elections. This paper contributes to the literature by providing a localized view of the national-level established relationship between income inequality and polarization (Duca et al. 2014).

This study uses an instrumental variable approach to determine the causal relationship between county-level inequality, measured by a Gini coefficient using American Community Survey (ACS) randomized income data and county-level polarization in Presidential and Senatorial elections from 2008 to 2020. This period was chosen as we hope to provide an empirical analysis of the most current relationship between the two values. The financial crisis and the period after not only affected inequality but also public awareness of inequality, which may affect voters' choices. Here, polarization is taken as a measure of election closeness. The instrument utilized in the analysis is the Gini coefficient from the 5 years before the election. This deals with the endogeneity that arises within an election year and assumes that people vote by considering past and present inequality.

This analysis provides insight into local election closeness, which provides a view into regional sorting and the impact of local income inequality on partisan divides. However, the instrument used in this analysis does not satisfy the exclusion restriction as there are other channels through which past inequality could affect political polarization, such as shifts in media, voter mobilization efforts, or policies shaped by prior inequality. The instrumental variable method helps address confounding factors, such as inequality after the election and shorter-term shocks. We recognize a negative relationship between local income inequality and local political polarization, meaning a party is more dominant within a county as inequality decreases. This relationship is only statistically significant for senatorial elections and not for presidential elections. This relationship may highlight a consolidation toward a particular party with a decrease in inequality. Additionally, the discrepancy between the presidential and senatorial elections may point to candidates and the public being more aware of counties' needs regarding inequality.

Section 2 provides a brief survey of relevant literature linking income inequality to political polarization and divides and reviews local-based vote choice theory and the polarization measure used in this paper. Section 3 describes the data sources of the analysis as well as the dataset construction and characteristics of the data. In Section 4, we present the empirical identification strategy used to identify the causal relationship between local income inequality and partisanship. We show the results of the analysis in Section 5. In Section 6, we discuss the results and possible mechanisms for the results. We conclude in Section 6 with the possible limitations of the analysis and potential future research and policy implications.

2. Literature Review

The literature has explored the relationship between economic inequality and polarization or other political views. Across advanced economies in Europe, over the last 20 years, net average income inequality has been linked to the increased success of far-right parties in parliamentary elections (Proaño, Peña, and Saalfeld 2022). They define polarization as increased "electoral success of far-right and far-left parties" as a portion of the population remains moderate. This paper finds that the increased success of far-right parties is largely based on the income share of the bottom 10% of income earners and increased income inequality. These changes are viewed as a reaction to economic and social grievances. However, income inequality has been found not just to affect vote choice but also the political views of prominent parties. In the United States, Duca and Saving 2014 found a bi-directional relationship between income inequality and polarization in the House and the Senate, where polarization is measured by increased average separation of Democratic and Republican policies over time. On a sub-national scale, Voorheis et. al 2015 found that at the state level, economic inequality makes the Democratic party in the state more liberal, moves state legislatures to the right, and replaces moderate Democrat legislators with Republicans.

However, our paper focuses on the impact of income inequality on the divide of voters' choices at a local level rather than entire parties. People often consider the characteristics of their immediate communities in their vote. In Rogers 2014, the results found that people consider the economic conditions of their local areas in their views of government officials at all levels to varying degrees. This paper tries to extend this line of inquiry to see if inequality in people's local areas affects political behavior. This gap has also been recognized as a topic that needs to be addressed. In a Review of Political Science and Political Economy literature regarding research on local government Anzia 2020 writes that scholars should examine when connections between local changes and economic situations affect local-level voting. This paper aims to contribute to the understanding of elections at the local level and examine if the effect of inequality and partisanship mirrors the trend seen on a large scale in voters' choices in elections.

The measurement for polarization in this paper is based on the polarization measurement in Chen et. al. 2022, Political Polarization = 1 - |DemVoteRatio − RepVoteRatio|. In this paper, a higher value meant a higher degree of political polarization, meaning that there were close elections, which the writers viewed would cause a lower flexibility for governments. This formula provides a useful view of local splits and possible clustering of voting patterns, which can be powerful indicators for broader political polarization (Navarrete et al. 2023). When counties tend to vote towards a particular party, it can point towards increased partisan sorting due to inequality, an important metric for nationwide polarization. Using this metric for polarization at a local level can provide an understanding of how votes in a county change with broader polarization and contribute to insight into national-level trends, particularly arguments about increases in locational partisan sorting (Navarrete et al. 2023).

We provide an addition to existing literature on the subject by providing a more local and election-focused lens to previous research on the relationship between income inequality and polarization.

3. Data

The income inequality data in this paper is constructed using the five-year and one-year estimates from the ACS data, compiled and extracted from IMPUS USA. For the counties in the sample using this estimate, we calculated the Gini coefficient using the fastgini function in Stata for each year from 2008 to 2020 for the 1-year samples. For the 5-year samples, we calculated the Gini coefficient for the five years before the election. For example, for the 2012 election, we took the five-year estimates for 2007 to 2011. The only exception to this was the 2008 election, which did not have a five-year estimate before the election year. The sample for this election is from 2004 to 2008. The demographic data, including race, gender, age, education, and veteran status, is also from the 1-year estimates from the ACS.

For the election and voting behavior data, we will be using the National Neighborhood Data Archive (NaNDA): Voter Registration, Turnout, and Partisanship by County, United States, 2004-2022. This data was created to explore the relationship between voter engagement, partisan political leanings, community health, and public policy. This dataset contains the following variables that are used in the analysis: votes for Democratic presidential candidates, votes for Republican presidential candidates, votes for Democratic Senate candidates, and votes for Republican Senate candidates.

Additionally, the dataset contains the following ratios for each election:

$$(1) \qquad \text{Presidential Democratic Ratio} = \frac{\text{Presidential Dem Votes}}{(\text{Presidential Dem Votes } + \text{ Presidential Rep Votes})}$$

$$(2) \qquad \text{Presidential Republican Ratio} = \frac{\text{Presidential Rep Votes}}{(\text{Presidential Dem Votes } + \text{ Presidential Rep Votes})}$$

$$(3) \qquad \text{Senatorial Democratic Ratio} = \frac{\text{Senatorial Dem Votes}}{(\text{Senatorial Dem Votes } + \text{ Senatorial Rep Votes})}$$

$$(4) \qquad \text{Senatorial Republican Ratio} = \frac{\text{Senatorial Rep Votes}}{(\text{Senatorial Dem Votes } + \text{ Senatorial Rep Votes})}.$$

The polarization metric is also based on these ratios. Here, we take polarization as,

$$(5) \qquad \text{Polarization} = 1 - |\text{Dem Ratio} - \text{Rep Ratio}|$$

This measure is from Chen et. al, whereas the polarization metric increases, there is more local-level polarization as the county is more evenly split. Finally, for each county, the data set includes the number of registered voters, ballots cast in general elections for all races, the citizen voting age population, percentage of eligible voters registered, percentage of eligible voters registered, percentage of eligible voters casting ballots, and percentage of eligible voters casting ballots. The partisanship index is taken as the percentage of votes cast for a party in the past 6 years, providing an understanding of a county's longer-standing political trends.

Once we had the election data, income data, and demographic data, we merged the datasets based on the FICPS code and year. For the person samples from the ACS, we only included people who were of voting age and did not have negative income according to the data set. Not every county in the ACS data was present in the election data and vice versa; therefore, after the matching process, 474 counties in total had complete data for all three datasets. There are 46 states represented in the dataset, with the highest concentrations in California, Texas, and Florida, as displayed in Table 1. The dataset is largely representative of the population distribution in the United States. From these counties, 2747 different election polarization measurements are being taken. As displayed in Table 2, on average, the counties are about evenly split by gender with an average of 48.5% male. The counties are, on average, white and middle-aged, with the average being 47 and 77% white. The sample has a slightly higher occurrence of veterans at 8.5% compared to the national average of 6%. The average for bachelor’s degree completion in the dataset is 28%, whereas the national average is 37% (DeSilver 2024). The unemployment rate is about average at 4.4%. Considering that the dataset is mostly white and without a college degree, one would expect the votes to also lean Republican in line with national trends (Nadeem and Nadeem 2024). However, in elections overall, the sample is about even with a Democratic partisan index of 49.8% and a Republican Partisan Index of 50.2%. When divided by type of election, the sample leans very slightly democratic for Presidential elections, with an average ratio of 50.3%. For senatorial elections, the sample leans slightly Republican, with an average Republican ratio of 52.6%.

4. Empirical Method

The identification strategy in this paper uses county-level inequality, county-level demographics, and county-level election polarization. The model is similar to the model presented by Voorheis et. al 2015 with an instrument of past inequality rather than national trends.

The basic model is

$$\text{Polarization}_{i,t} = \alpha + \beta ELECGINI_{i,t} + \gamma X_{i,t} + \epsilon_{i,t} \quad (6)$$

here $X_{i,t}$ is a matrix of the county-specific demographics controls for the election year. The error term is then described by

$$\epsilon_{i,t} = \alpha_i + e_{i,t}. \quad (7)$$

However, this model may lead to significant amounts of endogeneity. When estimating the effect of inequality in the same year, many variables may affect voters' choices, such as the media, policy, and other effects. Therefore, to counteract some of this, we use an instrument of the five-year past inequality in the years before the election. This instrument is the constructed Gini coefficient based on the five-year income data from the ACS. With this instrument, our first stage is as follows,

$$\text{First Stage: } ELECGINI_{i,t} = \alpha_0 + \alpha_1 LAGGEDGINI_{i,t} + X'_{it}\alpha + \epsilon_{i,t} \quad (8)$$

The X' value is a vector of the demographic controls from the ACS data. These controls are the average age, gender, population density, race (percentage white), veteran status, population under 25 of voting age, unemployment rate, population over 55, and educational attainment (population with a college degree) and are from the 1-year ACS estimate for the relevant year for each county. Here, the coefficient of interest is α_1, which, when statistically significant, would indicate that the instrument is relevant, meaning that the lagged Gini coefficient is correlated with the election-year coefficient. In Table 3, we see that this is true as the coefficient is close to one and statistically significant at the 5% level. This coefficient also means that the past lagged coefficient is very similar to the election year coefficient, allowing for strong instrument relevance.

The second stage is as follows,

$$\text{(9)} \qquad \text{Second Stage: Polarization}_{i,t} = \beta_0 + \beta_1 ELEC\hat{G}INI_{i,t} + X'_{it}\beta + \eta_{i,t}$$

Here β_1 captures the causal effect of income inequality on polarization. So our final reduced form is

$$\text{(10)} \qquad \text{Final Reduced Form: Polarization}_{i,t} = \gamma_0 + \gamma_1 LAGGEDGINI_{i,t} + X'_{it}\gamma + \nu_{i,t}$$

The final coefficient we will be examining is β_1. This identification strategy requires the assumption that people vote based on past inequality through its effect or comparison to the inequality in their county today. Rogers 2014 provides evidence that people often account for the economic conditions of their communities when voting. Additionally, in The Timeline of Presidential Elections, the authors find evidence that voters often filter their past economic experiences, such as their perceptions of their income, through those closer to the election (Erikson and Wlezien 2012, 122). This supports the assumption needed for the exclusion restriction that, ideally, the past Gini coefficient should only affect vote choice through the current Gini coefficient.

The exclusion restriction assumes that past inequality affects voting only through current inequality, which is unlikely to hold. Other mechanisms would violate the restriction and are beneficial avenues for future research to consider to provide stronger insights into the relationship between local income inequality and local political polarization. The historical Gini coefficient could affect migration flows, which impacts polarization without affecting polarization through the current coefficient. This is a valid concern, however, Liu et. al., find that the pattern largely affects counties that are already politically extreme rather than more moderate counties. Additionally, the paper finds that high rates of migration are also explained by similar industrial and institutional compositions (Liu, Andris, and Desmarais 2019). However, future research could account for migration by controlling for or analyzing where the inbound migrants are coming from and their effect on the political makeup of a country.

Additionally, there are other less direct mechanisms through which past inequality could affect current voting. Inequality before the election could affect the type of media (traditional or online) people are exposed to. Additionally, past inequality may affect the shape of the electorate as a whole, affecting election outcomes. For example, laws and other policies may have been put in place or not implemented properly, making it more difficult for some people to vote. Long-standing inequality or long-standing changes in inequality in the past may also affect how people vote, which the identification strategy does not account for. These mechanisms would be compelling for future studies to consider their effect on local-level polarization.

Therefore, we believe the exclusion restriction would likely not hold but will present the results and possible mechanisms if there were a causal effect.

5. Results

We first examine the relationship between county-level income inequality and senatorial election polarization. Our polarization measurement reflects local election competitiveness, measured as a smaller gap between parties in a county. Since the exclusion restriction is not satisfied, results only indicate an associative relationship rather than a causal one. Table 4 highlights the main result from the second stage regression for the relationship between county-level income inequality and polarization for senatorial elections. Table 4 shows a significant negative relationship between income inequality and polarization, with a coefficient of -1.325. This implies less competitive elections as inequality rises. While a one-unit increase in the Gini coefficient is unrealistic, a 0.01 increase corresponds to a -0.01325 decrease in polarization. This means that the share of dominance of the party in the county would be -0.01325%. The standard error (.397) and t-value (-3.33) confirm the relationship's statistical significance.

Figure 1 shows a decline in county-level polarization, reaching a low in 2016 before rising in presidential and senatorial elections. Figure 2 reveals fluctuations in the Gini coefficient, which increased until 2015 and then decreased, possibly reflecting the financial crisis and the recovery afterward.

Table 5 highlights the effect of income inequality on presidential election polarization. Similar to senatorial elections, the coefficient is negative (-0.48) but not statistically significant (standard error: .287; z-value: -1.67). Although not significant, the result points in the direction of a negative relationship with local income inequality. These results appear related to the earlier proposed idea that as nationwide political polarization increases, it has been coupled with greater locational sorting by political parties. Additionally, this result points to an interesting possible mechanism for decreased in-county polarization and a shift in the electorate.

6. Discussion

Despite the exclusion restriction not being satisfied, we find a positive association between county-level income inequality and party dominance in a county. This relationship was statistically significant for senatorial elections and was similarly negative for presidential elections but not statistically significant. These results align with Navarrete et al.'s findings, which link increased political polarization to locational sorting.

Beyond geographic sorting, other mechanisms may explain why inequality relates to less evenly split counties and greater party dominance. One possible explanation is that greater income inequality also concentrates economic interests in smaller groups (Duca et al. 2014). These groups may then be able to use their wealth and influence to strengthen their party's position by supporting their efforts and targeting the county's particular economic, social, and political concerns when campaigning.

Increased inequality can mobilize previously disadvantaged voters to become politically active. Recent 'get out the vote' efforts have proven effective at increasing turnout among marginalized groups (Schein et al. 2020, Middleton and Green 2008). Increasing turnout for these people may result in a stronger position of a party within a county. Across the world, the highest income inequality has been shown to reduce income bias in turnout (Matsubayashi and Sakaiya 2020). Outside of direct get-out-the-votes efforts, the experience of increased inequality and comparing it to their economic situation in the election year may also spur voters to vote. Although some people may attribute themselves to a specific party, experiencing and viewing increased inequality in their county may encourage them to make a different choice towards making a party more prevalent in their area. Future studies could analyze the relationship between local income inequality and turnout and the number and effectiveness of 'get-out-the-vote' efforts with changes in inequality. Anecdotally, this was a reason many people expressed when deciding whom to vote for in the 2024 presidential election. Some felt they were worse off in the election year and at the end of Biden's presidency than they were under Trump's presidency and decided to vote for Trump and other Republicans even if they had not voted for him previously. A similar effect may also have contributed to the decrease in local polarization in 2016, in which a party dominated a county more. The economy and people's perceptions of it were important to voters in the 2024 elections to a similar level as that during the 2008 election during the Great Recession (Brenan 2024). Economic grievance voting has been common in elections, particularly for incumbent candidates (Lewis-Beck and Martini 2020). These grievances with economic conditions could also affect how people perceived the inequality and changes in inequality in their community, leading to greater party dominance.

Additionally, in recent American history, there has been a shift in party support (Marble 2024 and Zacher 2023). Traditionally, the Republican party has been the party for the wealthy, while the Democratic party was known to be more representative of the working class. However, a shift has begun to occur, particularly but not exclusively for white voters (Marble 2024). This effect has been explained through cultural reasons but also has a straightforward income component and the importance that a group of voters places on economic conditions. If the relationship was causal, this shift in the electorate may mean that as income inequality increases, residents who feel they have been negatively impacted may lean toward the Republican party.

We also see the difference in the statistical significance of the effect of income inequality on polarization for senatorial elections and presidential elections in counties. The relationship may be significant for Senate elections as candidates are more able to discuss and address their plans to improve economic conditions at a more local level. This may result in shifts that are not possible at the Presidential level as in these campaigns it is difficult for candidates to address specific counties and communities and the factors in the income inequality in the county. Additionally, those who vote in Senatorial elections in off-presidential election years may be more aware of the political situations in their counties and states and more aware of which candidate or party they believe will improve the county. The turnout for these elections is significantly lower than for the Presidential election years, and voters tend to be older and whiter (Nadeem and Nadeem 2024). These voters are often retired and, therefore, more aware of their communities and have a stronger set of political affiliations, leading to less polarized elections.

7. Conclusion

Future research is needed to establish a causal relationship between local income inequality and county-level polarization at a larger and more nationally representative scale, which will provide a deeper understanding of local-level vote choice. These results in the paper point towards a possible positive relationship between the two values. It suggests that with local increases in income inequality, a less even distribution of political support within counties rises. If this relationship is causal, it may be due to shifting voter preferences or increased engagement from previously underrepresented groups. Additionally, the calculation of the Gini coefficient is based on the ACS randomized samples rather than tax data, which would include almost every person in the county. Further research using tax data may provide a more comprehensive view of income inequality within a county. The paper also does not address the impact of migration between counties on the change in elections.

As discussed previously, some factors violate the exclusion restriction necessary for the lagged Gini coefficient instrument presented in the paper to hold. Further research addressing these factors would contribute significantly to establishing a causal relationship between vote choice and polarization literature. These factors may be influenced by inequality and include the effect of media, voter suppression and other policies, and migration. Income inequality and other factors may affect the traditional and new media that people consume, which may then affect their vote choice. Areas that have high voter suppression may have been affected by inequality, which also would affect vote choice. Research examining the effect of these factors on inequality and polarization would provide useful implications for understanding the mechanisms behind the relationship between local inequality and vote choice.

Although there was no causal relationship established in this paper, if the relationship held causally, there would be useful applications for campaigns and redistributive policies. In states with high county-level income inequality, candidates who address inequality and economic concerns may be able to strengthen their strongholds in counties. Even though this possible effect is weaker at the presidential level, addressing inequality considerations may also help close gaps.

The results also point to the increase in national-level polarization coupled with increased locational sorting described in Navarrete et al. The increase in locational sorting may continue to be a concern, as it is easier to negatively characterize supporters of another party if they are not in your community. Reducing inequality could help create more balanced elections that are less impacted by locational sorting.

This paper's results provide non-causal preliminary findings about the relationship between inequality and local party dominance. The paper's localized focus contributes to the understanding of the rise of national-level political polarization, which may be coupled with increases in local partisan sorting related to income inequality. As the findings are not causal, future research to establish a causal relationship must understand the drivers of local polarization and how to address them in the United States.

Appendix

The counties in this data set have a population greater than 65,000 due to the ACS requirements. This data is compiled using the random sample weighted sample method used in the ACS. We will use 3-year estimates, which take a 3 in 100 random sample of the population, 1-year estimates, which take a 1 in 100 sample of the population, and the 5-year estimate, which takes a 5 in 100 sample of the population.

Table 1. States in Dataset

State	State FICPS	Number of Counties in Dataset
AL	1	8
AZ	4	5
AR	5	4
CA	6	34
CO	8	3
CT	9	8
DE	10	3
FL	12	32
GA	13	23
HI	15	2
ID	16	3
IL	17	18
IN	18	16
IA	19	6
KS	20	4
KY	21	5
LA	22	11
ME	23	3
MD	24	12
MA	25	3
MI	26	17
MN	27	9
MS	28	4

MO	29	7
MT	30	1
NE	31	3
NV	32	2
NH	33	1
NJ	34	20
NM	35	5
NY	36	22
NC	37	25
ND	38	1
OH	39	21
OK	40	3
OR	41	8
PA	42	23
RI	44	3
SC	45	9
TN	47	10
TX	48	38
UT	49	5
VA	51	11
WA	53	9
WI	55	13

Table 2. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Male	2857	.485	.013	.44	.599
White	2857	.773	.149	.118	.979
Mean Age	2857	46.941	2.787	35.967	59.618
Population Under 25	2857	.133	.044	.047	.445
Population Over 55	2857	.335	.062	.169	.641
Veteran	2857	.085	.031	.016	.234
College Educated	2857	.282	.097	.086	.752
Unemployed	2857	.044	.018	.001	.142
Population Density	2857	2748.351	6777.063	99.5	95901.138
Gini	2857	.481	.034	.371	.63
Lagged Gini	2554	.481	.031	.383	.609
Pres Dem Ratio	1662	.503	.152	.099	.957
Pres Rep Ratio	1662	.497	.152	.043	.901
Sen Dem Ratio	1294	.474	.181	0	1
Sen Rep Ratio	1294	.526	.181	0	1
Pres EC	1662	.756	.181	.086	1
Sen EC	1294	.711	.225	0	1
Partisan Index Dem	2857	.498	.146	.136	.925
Partisan Index Rep	2857	.502	.146	.075	.864

Note: The Partisan Index is the percentage of votes cast for each party in the past 6 years. Demographic Data is from 1-year ACS from 2008 to 2020. Election data is for Presidential and Senatorial elections from 2008 to 2020.

Table 3. First Stage F-Test

Gini	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
lag_gini	.803	.017	46.19	0	.769	.838	***
Male	.002	.03	0.08	.936	-.056	.06	
White	-.003	.003	-0.99	.324	-.009	.003	
Mean Age	.004	.001	4.59	0	.002	.006	***
Population Under 25	.153	.023	6.79	0	.109	.198	***
Population Over 55	-.107	.031	-3.39	.001	-.168	-.045	***
Veteran	.021	.014	1.44	.151	-.008	.049	
College Educated	.049	.005	9.46	0	.038	.059	***
Unemployed	.255	.023	10.84	0	.208	.301	***
Population Density	0	0	1.07	.283	0	0	
Constant	-.105	.037	-2.82	.005	-.177	-.032	***

Mean dependent var	0.481	SD dependent var	0.034
R-squared	0.723	Number of obs	2421
F-test	627.633	Prob > F	0.000
Akaike crit. (AIC)	-12612.320	Bayesian crit. (BIC)	-12548.609

*** *p<.01,* ** *p<.05,* * *p<.1*

Table 4. IV Senatorial Election Polarization - Second Stage

SEN_POLAR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Gini_hat	-1.325	.397	-3.33	.001	-2.105	-.545	***
Male	.043	.016	2.68	.008	.012	.075	***
White	.038	.544	0.07	.945	-1.03	1.105	
Mean Age	.029	.055	0.53	.596	-.079	.137	
Population Under 25	-.23	.264	-0.87	.383	-.747	.287	
Population Over 55	.451	.101	4.46	0	.253	.649	***
Veteran	1.615	.457	3.53	0	.717	2.512	***

College Educated	0	0	-6.08	0	0	0	***
Unemployed	1.688	.434	3.89	0	.837	2.539	***
Population Density	-.869	.57	-1.52	.128	-1.988	.25	
Constant	-.819	.677	-1.21	.227	-2.148	.51	

Mean dependent var	0.710	SD dependent var	0.227
R-squared	0.075	Number of obs	1141
F-test	9.142	Prob > F	0.000
Akaike crit. (AIC)	-212.194	Bayesian crit. (BIC)	-156.758

**** p<.01, ** p<.05, * p<.1*

Table 5. IV Presidential Election Polarization - Second Stage

PRES_POLAR	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
gini_hat	-.48	.287	-1.67	.095	-1.043	.084	*
Male	.051	.011	4.60	0	.029	.073	***
White	.896	.366	2.45	.014	.179	1.614	**
Mean Age	.104	.038	2.76	.006	.03	.178	***
Population Under 25	.071	.191	0.37	.711	-.305	.447	
Population Over 55	.142	.075	1.90	.057	-.004	.289	*
Veteran	1.473	.36	4.09	0	.767	2.179	***
College Educated	0	0	-8.68	0	0	0	***
Unemployed	1.595	.317	5.02	0	.972	2.218	***
Population Density	-1.307	.392	-3.33	.001	-2.076	-.538	***
Constant	-1.805	.461	-3.92	0	-2.709	-.901	***

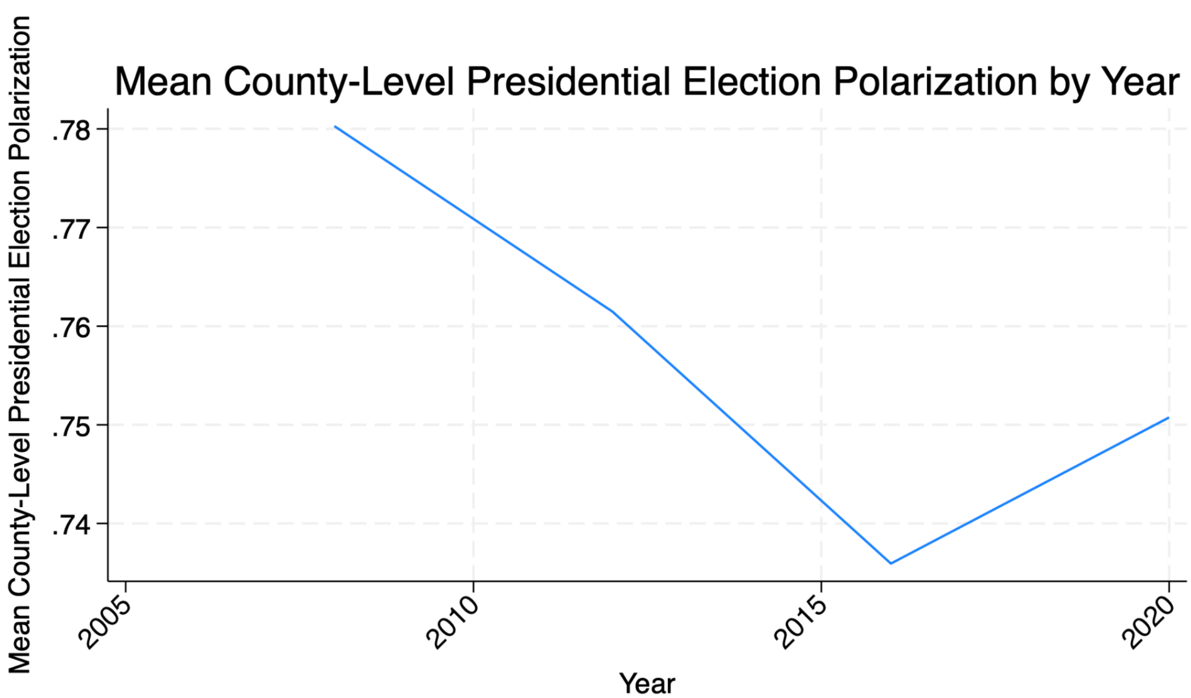
Mean dependent var	0.750	SD dependent var	0.184
R-squared	0.175	Number of obs	1188

F-test	25.007	Prob > F	0.000
Akaike crit. (AIC)	-864.405	Bayesian crit. (BIC)	-808.525

**** p<.01, ** p<.05, * p<.1*

Figure 1. Mean County-Level Polarization Over Time

A. Presidential Elections



B. Senatorial Elections

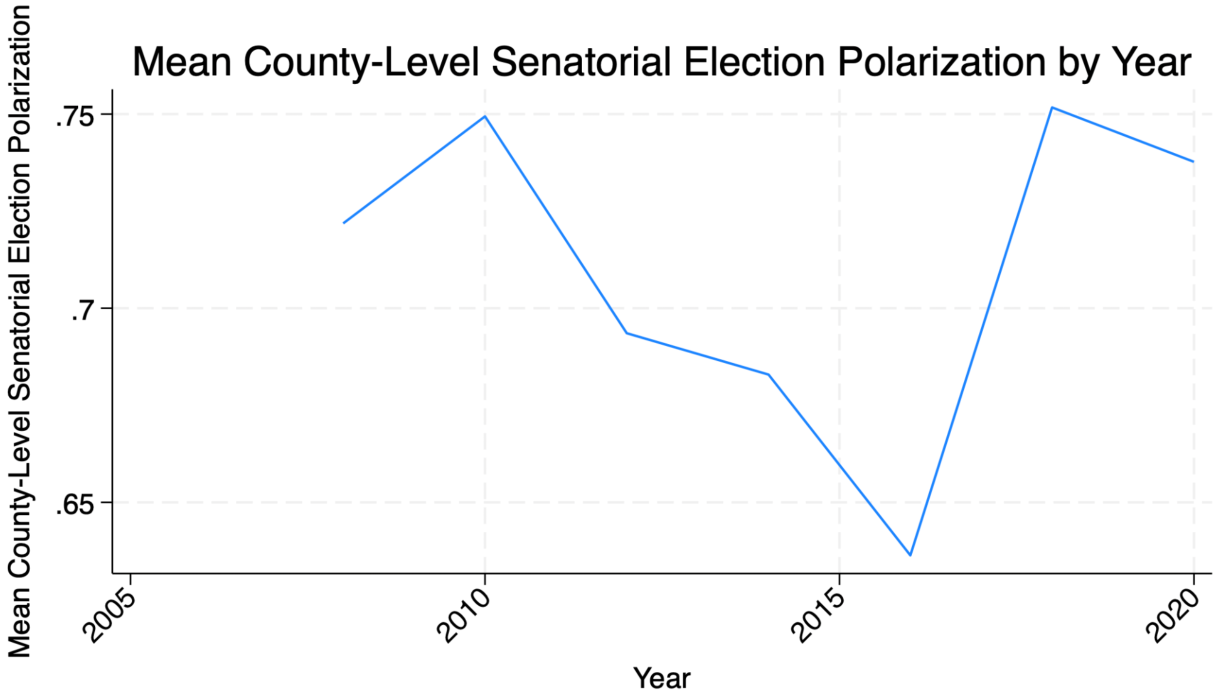
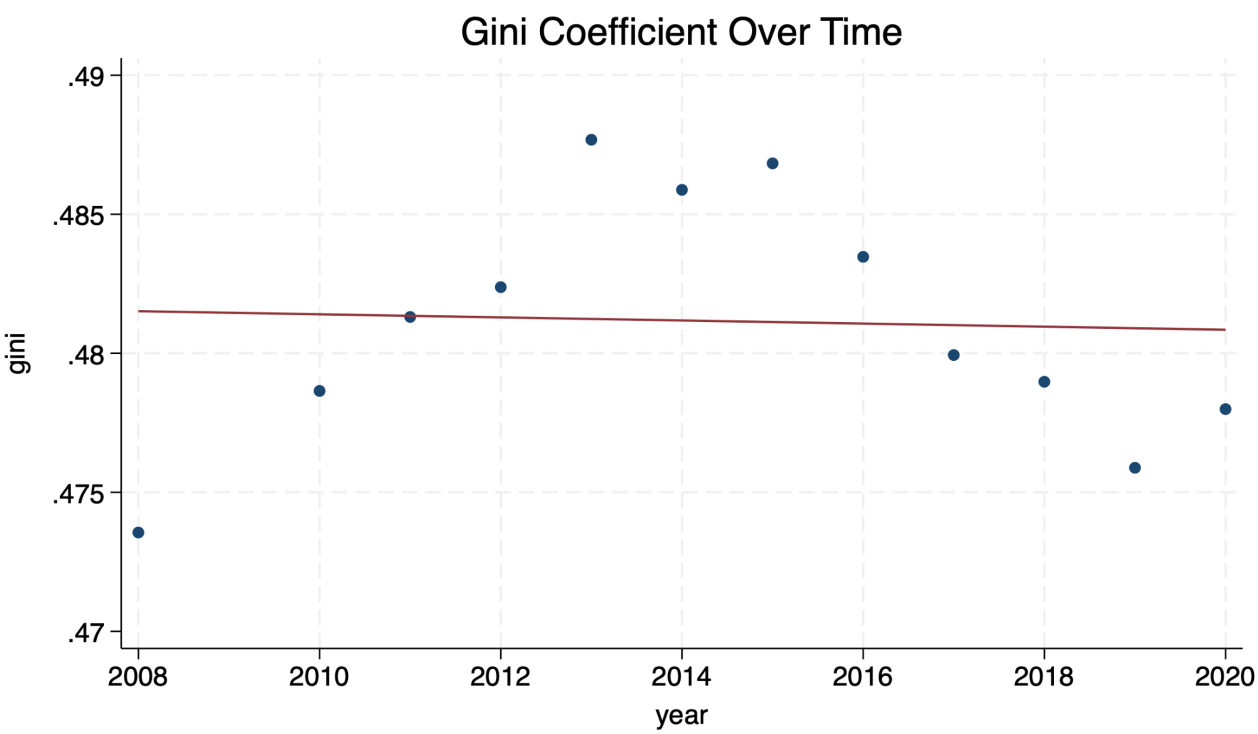


Figure 2. Average County 1-year Gini Coefficient Over Time



References

Anzia, Sarah F. 2020. “Party and Ideology in American Local Government: An Appraisal.” *Annual Review of Political Science* 24 (1): 133–50. https://doi.org/10.1146/annurev-polisci-041719-102131.

Brenan, By Megan. 2024. “Economy Most Important Issue to 2024 Presidential Vote.” *Gallup.com*, November 15, 2024. https://news.gallup.com/poll/651719/economy-important-issue-2024-presidential-vote.aspx.

Chen, Zhiwei, Zhaoyuan Li, and Sibo Liu. 2022. “The Price of Political Polarization: Evidence From Municipal Issuers During the Coronavirus Pandemic.” *Finance Research Letters* 47 (March): 102781. https://doi.org/10.1016/j.frl.2022.102781.

DeSilver, Drew. 2024. “10 Facts About Today’s College Graduates.” *Pew Research Center*, April 14, 2024. https://www.pewresearch.org/short-reads/2022/04/12/10-facts-about-todays-college-graduates/.

Duca, John V., Jason L. Saving, Federal Reserve Bank of Dallas, and Southern Methodist University. 2014. “Income Inequality and Political Polarization: Time Series Evidence Over Nine Decades.” Working Paper. *Working Paper 1408*. https://www.dallasfed.org/~/media/documents/research/papers/2014/wp1408.pdf.

Erikson, Robert S., and Christopher Wlezien. 2012. *The Timeline of Presidential Elections: How Campaigns Do (and Do Not) Matter*. http://ci.nii.ac.jp/ncid/BB10818189.

Inter-university Consortium for Political and Social Research. 2024. "National Neighborhood Data Archive (NaNDA): Voter Registration, Turnout, and Partisanship by County, United States, 2004-2022." *Www.Icpsr.Umich.Edu*, October. https://doi.org/10.3886/ICPSR38506.v2.

Kubin, Emily, and Christian Von Sikorski. 2021. "The Role of (Social) Media in Political Polarization: A Systematic Review." *Annals of the International Communication Association* 45 (3): 188–206. https://doi.org/10.1080/23808985.2021.1976070.

Lewis-Beck, Colin, and Nicholas F. Martini. 2020. "Economic Perceptions and Voting Behavior in US Presidential Elections." *Research & Politics* 7 (4): 205316802097281. https://doi.org/10.1177/2053168020972811.

Lindsay, James M. 2024. "Election 2024: Close Presidential Elections Have Become the Norm." *Council on Foreign Relations*, February 23, 2024. https://www.cfr.org/blog/election-2024-close-presidential-elections-have-become-norm.

Liu, Xi, Clio Andris, and Bruce A. Desmarais. 2019. "Migration and Political Polarization in the U.S.: An Analysis of the County-level Migration Network." *PLoS ONE* 14 (11): e0225405. https://doi.org/10.1371/journal.pone.0225405.

Marble, William. 2024. "What Explains Educational Realignment Among White Americans?" *SocArXiv Papers*, June. https://doi.org/10.31235/osf.io/2e3jp.

Matsubayashi, Tetsuya, and Shiro Sakaiya. 2020. "Income Inequality and Income Bias in Voter Turnout." *European Journal of Political Economy* 66 (October): 101966. https://doi.org/10.1016/j.ejpoleco.2020.101966.

Mian, Atif, Amir Sufi, and Francesco Trebbi. 2014. "Resolving Debt Overhang: Political Constraints in the Aftermath of Financial Crises." *American Economic Journal Macroeconomics* 6 (2): 1–28. https://doi.org/10.1257/mac.6.2.1.

Middleton, Joel A., and Donald P. Green. 2008. "Do Community-Based Voter Mobilization Campaigns Work Even in Battleground States? Evaluating the Effectiveness of MoveOn's 2004 Outreach Campaign." December 16, 2008. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1316927.

Nadeem, Reem, and Reem Nadeem. 2024. "1. Voter Turnout, 2018-2022." Pew Research Center. June 21, 2024. https://www.pewresearch.org/politics/2023/07/12/voter-turnout-2018-2022/.

Navarrete, Carlos, Mariana Macedo, Viktor Stojkoski, Marcela Parada-Contzen, and Christopher A Martínez. 2023. "Mapping Election Polarization and Competitiveness Using Election Results." arXiv.Org. August 16, 2023. https://arxiv.org/abs/2308.10862.

Proaño, Christian R., Juan Carlos Peña, and Thomas Saalfeld. 2022. "Inequality, Macroeconomic Performance, and Political Polarization: A Panel Analysis of 20 Advanced Democracies." *Review of Social Economy* 82 (3): 396–429. https://doi.org/10.1080/00346764.2022.2047768.

Rogers, Jonathan. 2014. "A Communotropic Theory of Economic Voting." *Electoral Studies* 36 (September): 107–16. https://doi.org/10.1016/j.electstud.2014.08.004.

Schein, Aaron, Keyon Vafa, Dhanya Sridhar, Victor Veitch, Jeffrey Quinn, James Moffet, David M. Blei, and Donald P. Green. 2020. "A Digital Field Experiment Reveals Large Effects of Friend-to-Friend Texting on Voter Turnout." *SSRN Electronic Journal*, January. https://doi.org/10.2139/ssrn.3696179.

Solt, Frederick. 2010. "Does Economic Inequality Depress Electoral Participation? Testing the Schattschneider Hypothesis." *Political Behavior* 32 (2): 285–301. https://doi.org/10.1007/s11109-010-9106-0.

Voorheis, John, Nolan McCarty, and Boris Shor. 2015. "Unequal Incomes, Ideology and Gridlock: How Rising Inequality Increases Political Polarization." *SSRN Electronic Journal*, January. https://doi.org/10.2139/ssrn.2649215.

Zacher, Sam. 2023. “Polarization of the Rich: The New Democratic Allegiance of Affluent Americans and the Politics of Redistribution.” *Perspectives on Politics* 22 (2): 338–56. https://doi.org/10.1017/s1537592722003310.

Peter Berggren

12/6/2024

14.33: Research and Communication in Economics

Instructor: Tobias Salz

The Impact of Autism Therapy Coverage Mandates on Outcomes for Autistic Children[1]

Abstract

Despite significant efforts towards enacting legal mandates for health insurers to cover autism therapies, there has been limited research into the effect of these mandates on outcomes such as diagnosis rates, uptake of behavioral therapy, healthcare affordability, and childhood flourishing metrics. Using a difference-in-difference analysis on Alabama's passage of such a mandate in 2017, this paper finds that this mandate led to a significant increase in the rate of autistic children receiving behavioral therapy, but no significant effect on diagnosis rates, healthcare affordability, or childhood flourishing metrics. The policy implications, as well as avenues for future research, are then discussed.

1. Introduction

Autism spectrum disorder (also known as autism or ASD) is a neurological condition which has a broad range of associated symptoms. These include challenges with social interaction, including difficulties making eye contact, engaging in conversation, and taking other people's perspectives. These also include intense and often overly focused interests on very specific topics, as well as repetitive movements and speaking patterns (National Institute of Mental Health 2024). Because autism leads to mental and behavioral challenges, it is categorized as a developmental disability (CDC 2024b).

[1]See Appendix C for details on the use of language surrounding autism in this paper.

There are many therapies available for autism, among the most notable being applied behavior analysis (ABA) (Autism Speaks, n.d.-b). ABA is a behavioral therapy based on recognizing the purpose of inappropriate behaviors (such as a child hitting people as a way to get their attention) and using positive reinforcement to teach more appropriate alternative behaviors (such as saying “Excuse me” or the person’s name) (Kahng and Stichter, n.d.).

Because therapies such as ABA, like other therapies, are often expensive, there have been many laws passed over the past few decades mandating that health insurers cover treatments for autism including ABA (National Conference of State Legislatures 2021). For purposes of brevity, these laws will be called “autism coverage mandates,” “coverage mandates,” or simply “mandates” throughout the paper. However, despite significant advocacy for the enactment and expansion of these laws (Autism Speaks 2019a), a literature review of studies on the impact of such mandates, despite finding a wide range of literature addressing related questions, found no studies making use of identification methods such as differences-in-differences to identify the effect of these mandates on therapy utilization and other indicators at a state level.

Despite the lack of studies directly investigating autism coverage mandates, there are existing studies using “natural experiment” designs to investigate the effect of ABA. Such literature is common in the field because, since ABA is an established therapeutic methodology, it would likely be considered unethical to deny therapy to a control group (Anderson and Carr 2021). A study by Dixon et al. (2021), which used people on a waitlist to receive treatment as the control group, is typical of this body of research, and found noticeable effects on indicators such as cognitive test scores, with effects that are more pronounced when a comprehensive plan is used that draws insights from more modern research, rather than solely from the original research underpinning ABA. However, for practical reasons, many of these studies have small sample sizes (in the case of Dixon et al. (2021), there were 17 treated participants divided into two groups and 11 waitlisted control participants) and operate out of a single clinic. Although results may be highly statistically significant despite the small sample size, the small sample size still makes studies vulnerable to publication bias and potentially non-generalizable outside of the context of the particular clinic where they took place.

Additionally, there is some literature on autism coverage mandates, such as Acton, Imberman, and Lovenheim (2021) and Choi et al. (2022), but it focuses on topics such as comparing legislation across states (in the case of Choi et al. (2022)) and examining the possibility of spillovers on school-based services (in the case of Acton, Imberman, and Lovenheim (2021)), rather than directly evaluating the effects of autism coverage mandates on therapy utilization, healthcare costs, and other indicators (e.g. childhood flourishing). In addition, although the Acton study on spillovers used a differences-in-differences approach similar to the one used in this paper, it used students without autism as a control group, which makes it more difficult to establish parallel trends than it would be when using another control group (such as autistic children in a neighboring state in the case of this paper's analysis).

Finally, there are a wide range of studies on the effect of health insurance on health outcomes. The RAND Health Insurance Experiment, conducted between 1971 and 1986, was a randomized controlled trial which examined the effects of different types of health insurance on healthcare utilization, finding, among other things, that plans with cost sharing reduced the amount of health care used, but that this had little effect on health outcomes outside of a few areas (Brook et al. 2006). More recently, when Oregon expanded Medicaid, eligibility for new programs was initially provided by lottery, allowing the situation to be analyzed as a natural experiment (Finkelstein et al. 2012).

This paper aims to determine the impact of insurance laws, particularly those mandating that insurers cover therapies such as applied behavior analysis (ABA), on autistic children. This impact is measured through indicators such as the proportion of children diagnosed with autism, the proportion of autistic children who are currently receiving behavioral therapies (of which ABA is the most notable example; note that behavioral therapies as measured by the utilized datasets will be collectively referred to as "ABA" throughout the paper), indicators of healthcare utilization, and indicators of childhood flourishing. This allows determination of whether these laws increase uptake of ABA, make ABA more affordable for families already receiving it, both, or neither, as well as the effects of ABA therapy received as a result of these laws.

The childhood flourishing metrics (which are parents' answers to three questions about whether the child shows interest and curiosity in learning new things, whether the child works to finish the tasks they start, and whether the child stays calm and in control when faced with a challenge) (Child and Adolescent Health Measurement Initiative (CAHMI), n.d.-a) are of key importance to this study, as they indicate, in combination with the relevant healthcare utilization metrics, an effect of autism diagnosis and treatment on quality of life for autistic children. Though Dixon et al. (2021) investigated the effect of ABA on cognitive test scores, they did not investigate the effect on metrics more directly related to quality of life, and were limited to a particular clinic. As such, an assessment of childhood flourishing in this context is essential in determining results from a mandate, and serves to investigate the effects of autism diagnosis and therapy in a way which more closely relates to quality of life than cognitive test scores do.

There are significant policy implications to potential outcomes of this study. A finding that autism coverage mandates have no effect would imply that it is a waste of effort to continue advocating for them in environments where they may not be in effect yet (such as in other countries). A finding that the mandates increased autism diagnosis rates would show that the unaffordability of diagnosis was keeping children from being diagnosed with autism, and the implications if the mandates increased behavioral therapy rates would be similar, but for access to ABA. A finding that the mandates led to reductions in out-of-pocket healthcare spending (or increased perceptions of coverage adequacy among families of autistic children) would show an effect of making healthcare more affordable to people who already were getting newly covered therapies, but were likely paying for them out of pocket. Additionally, any effect of the mandates on childhood flourishing could be used to show an effect of either increased diagnosis or increased treatment on childhood flourishing, depending on the effect sizes of the mandates on diagnosis and treatment.

The dataset used here comes from the National Survey of Children's Health, a large survey conducted of families across the country (CAHMI, n.d.-b). It has been conducted on an annual basis since 2016, and at the time that this project started, the most recent data release was in 2022. As a result, data from 2016 to 2022 were used. In addition, data on which state laws were in effect at which times came from Autism Speaks and the National Conference of State Legislatures (National Conference of State Legislatures (NCSL) 2021) (Autism Speaks, n.d.-a) (Autism Speaks 2019b).

The identification strategy used here is differences-in-differences (DID), based on a regression performed on data from Alabama and Mississippi. This was performed by linear regressions of outcome variables of interest on year dummies, a state dummy, and a dummy for whether or not an autism coverage mandate was in place, as detailed in the Methods section of this paper. These two states were chosen because, while Mississippi had an autism coverage mandate in effect as of the start of 2016 (Autism Speaks 2019b), Alabama's mandate only went into effect in October 2017 (Autism Speaks, n.d.-b). These states also had many superficially similar characteristics, supporting an assumption of parallel trends despite the lack of observable pre-trends. The focus on Alabama and Mississippi in particular also contrasts this study with Acton, Imberman, and Lovenheim (2021), which studied a mandate in Michigan (in a different region of the country from Alabama and Mississippi) and did not make inter-state comparisons.

This analysis found no significant effects on diagnosis rate, out-of-pocket healthcare spending, perception of adequacy of health insurance coverage, enrollment in special education, or childhood flourishing, but did find a significant effect on utilization of behavioral therapy among autistic children equal to 38.4 percentage points. The lack of an effect on diagnosis rate or enrollment in special education suggests that the primary effect of the mandate was to improve access to therapy among children already diagnosed with autism, rather than to increase the number of children diagnosed. Additionally, the lack of an effect on out-of-pocket healthcare spending or perception of adequacy of health insurance coverage suggests that few people, before the mandate, were paying for autism therapies out-of-pocket. Finally, the lack of an effect on childhood flourishing suggests either that effects of ABA received as a result of coverage mandates do not manifest on these flourishing metrics, that the population covered by these metrics does not match the population most impacted by ABA, that the effects on flourishing do not happen close enough to the start of ABA therapy for the difference-in-difference strategy to capture them, or that the effects on flourishing were camouflaged by a change in the question phrasing that happened in 2018 (as discussed in greater detail in the Methods section).

These findings suggest that autism coverage mandates are an effective way to increase utilization of autism therapies, but do not significantly affect diagnosis rates or healthcare affordability among people already receiving these therapies. Additionally, these findings suggest that any effects of behavior therapies such as ABA (at least on the population that is not already getting ABA, which would likely tend to be people with lower incomes, less access to services through the school system, and less outwardly severe autism that may not be recognized by the school system) apply to metrics other than the childhood flourishing reports in the NSCH, or at least take more than a year to apply to flourishing. This suggests that, while advocacy for autism coverage mandates (likely in other countries, as virtually all US states already have such mandates in place both for private insurers and for Medicaid) (NCSL 2021) would likely be an effective way to improve access to autism therapies in populations that are not already getting them, more research is needed as to what effect these therapies have on the population served by these mandates.

2. Data

For information on state laws, Autism Speaks' database of state laws on insurance mandates was used to identify Alabama and Mississippi as two states with similar insurance laws that were enacted two years apart, and within the range of years covered by the National Survey of Children's Health (Autism Speaks, n.d.-a) (Autism Speaks 2019b). In addition, a database run by the National Conference of State Legislatures was used to corroborate the names and dates of important autism-related insurance laws across states (NCSL 2021). These datasets provide valuable perspective on the timelines of insurance mandates for autism therapies across US states, as well as being useful in corroborating each other (the NCSL dataset provides citations of specific laws, while the Autism Speaks dataset summarizes key information about the laws).

The dataset used in this paper for health outcomes came from the National Survey of Children's Health (CAHMI, n.d.-b). The National Survey of Children's Health dataset is a large dataset assembled by the Health Resources and Service Administration Maternal and Child

Health Bureau, which is part of the Department of Health and Human Services. This dataset includes autism diagnosis rates, as well as behavioral treatment rates for autism and a selection of indicators related to insurance affordability and mental health across US states. This dataset also includes a set of general indicators related to child behavior and cognition (including questions about behavioral problems, intellectual disabilities, and how often a child's medical conditions prevent them from doing things other children do), which can be used in conjunction with data on whether surveyed children are diagnosed with autism to see how these outcomes are affected among autistic children in particular.

In the years ranging from 2016 to 2022, the number of observations in the dataset (i.e. people surveyed) ranged from a low of 21,599 in 2017 to a high of 54,103 in 2022. Each observation in the dataset represents a child who was surveyed as part of the NSCH, and so the variables in question are child-specific.

After merging all the data files into a single dataframe and restricting the scope of analysis to people in Alabama and Mississippi, the dataset includes a total of 9,605 observations across all years, ranging from a low of 858 in 2017 to a high of 1,776 in 2021. Restricting the scope to children who currently have autism reduces the number of observations further to 273 across all years, ranging from a low of 27 in 2016 to a high of 56 in 2021. A brief summary of the data here is included in Appendix A as Table 1.

3. Methods

The empirical test for this paper involved a difference-in-difference regression performed on autism-related outcomes in Alabama and Mississippi. These two states were chosen because they, according to Autism Speaks, have similar insurance laws surrounding autism, but Mississippi's mandate for insurance coverage of autism therapies went into effect in July 2015

(Autism Speaks 2019b), while Alabama's went into effect in October 2017 (Autism Speaks, n.d.-a). For this reason, deviations in 2016 and 2017 from parallel trends in outcomes between the states can be attributed to the differences in timing of the legislation.

The regression here used year dummies to accommodate non-linear parallel trends, as part of a difference-in-difference analysis. This regression can be represented by the equation $Y = \beta_0 + \Sigma(\beta_t Year_t) + \beta_s(State) + \beta_3(Mandate) + \varepsilon$, where Y is the dependent variable, β_0 is the constant term, β_t are the coefficients on year dummies for years t, β_s is the coefficient on the state dummy, β_m is the coefficient on the mandate dummy, and ε is the error term. This specification was chosen because it captures the difference-in-difference concept of analysis in a regression which can easily be performed by software, while not having so many coefficients that the specification would be vulnerable to selective reporting of significant results (i.e. "p-hacking"), which a specification which allows time leads and lags on each year separately could facilitate.

A brief summary of the process behind difference-in-difference analysis is as follows. Suppose that two states, all else equal, have parallel trends in certain indicators (e.g. the proportion of autistic children who are receiving behavioral therapy). These parallel trends can be captured by a regression on year dummies and a state dummy, regardless of the year-by-year pattern. Suppose also that some policy change (such as a new law in one of the states) leads to a change to that state, but allows parallel trends to otherwise continue. Then a regression specification that also includes a dummy variable for whether that law is in effect will capture the true effect of the law.

The key assumption for the difference-in-difference test is the assumption of parallel trends between the states. A key shortcoming of this analysis is that the National Survey of

Children's Health data only extends back to 2016, making it impossible to observe parallel trends pre-treatment (as Mississippi's law went into effect in 2015). The qualitative observation of parallel trends post-treatment (i.e. when both states had autism coverage mandates in place) in the paper's figures, including qualitatively similar changes to flourishing coinciding with both a change in specification in 2018 (as discussed later in this section) and the COVID-19 pandemic, serves to mitigate this challenge, as does the close geographic similarity of Alabama and Mississippi. Nevertheless, the lack of pre-trends suggests an avenue for further research, if datasets can be found that cover both states before any autism coverage mandate was passed in either state.

Another challenge to parallel trends is, if there were other events happening in Alabama but not Mississippi in late 2017 and early 2018, we cannot assume that the autism insurance mandates in Alabama were the sole contributor to any deviation from parallel trends. Third, one key exception to the similarity between Alabama and Mississippi laws here is that Alabama laws passed in 2017 specifically require Alabama Medicaid and CHIP to cover autism therapies (NCSL 2021), while no similar Mississippi law seems to exist (unless the 2015 law on insurance coverage includes Medicaid implicitly). However, Mississippi Medicaid does currently cover ABA, which is among the most notable autism therapies (Full Spectrum ABA 2024). For purposes of this analysis, we assume that no significant changes were made to Mississippi Medicaid coverage of autism therapies from 2016 to 2022. This seems not unreasonable, given that in many cases (such as Alabama's), laws ensuring Medicaid coverage of autism therapies were passed at the same time as analogous laws for private insurance, and given that Medicaid will, broadly speaking, cover treatments that are medically necessary (Autism Speaks, n.d.-b).

Finally, given that both Alabama and Mississippi are located in the southern United States, it is unclear whether any findings that apply to these two states can be extrapolated to the rest of the country. Nevertheless, the southern states are the ones for which this research has the most policy applicability, as they have historically been the ones most at risk from Medicaid defunding and other reductions of government interventions that were designed to ensure affordable provision of health care (Hawkins 2024).

All the variables listed here, except for the flourishing metric, were binary variables derived from simple manipulations of National Survey on Children's Health variables performed in accordance with the NSCH codebooks (CAHMI 2020). The flourishing metric for children age 6 to 17, however, was defined based on how many questions out of three had "positive" answers as defined in the NSCH codebooks. These three questions were on whether the child shows interest and curiosity in learning new things, whether the child works to finish the tasks they start, and whether the child stays calm and in control when faced with a challenge. Note that in 2016 and 2017, parents were asked "how true" these statements are and asked to rank these statements on a scale of "definitely true," "sometimes true," or "not true," while for 2018 onwards, they were asked "how often" these statements were true and asked to rank them on a scale of "always," "usually," "sometimes," or "never" (CAHMI, n.d.-a). For 2016 and 2017, the number of "definitely true" answers out of a possible three was counted, while for 2018 onwards, it was the number of "always" or "usually" answers. A value of 1 was assigned if 0 or 1 questions had those answers, a value of 2 was assigned if 2 questions had those answers, and a value of 3 was assigned if all 3 questions had those answers.

If we assume that the change in specification had an equal effect on both states, we can still assume parallel trends, preserving the validity of the estimator. Nevertheless, if it had a differential effect on the two states for any reason unrelated to the difference in autism coverage mandates, the assumption of parallel trends would not hold for the flourishing metric.

4. Results

Table 2 (found in Appendix A) displays the results of the difference-in-difference regression specification. The year dummy coefficients and p-values are not reported, as many of them are significant at different levels, and the objective of this study was to investigate the effects of mandates and not to analyze year-over-year trends. The only significant effects found in regression are on the behavioral treatment rate. Being in Alabama increases the behavioral treatment rate by 13.1 percentage points (significant at the $p < 0.05$ level) while having a mandate in place increases the behavioral treatment rate by 38.4 percentage points (significant at the $p < 0.01$ level).[2]

Figure 2 (found, with the other figures, in Appendix B) gives a visualization of the behavioral treatment rate among autistic children, as well as comparing and contrasting actual values with the estimates provided by the difference-in-difference approach (estimates which yield the main significant finding of this project). Figure 1 gives a comparable visualization for the rate of children who have ever been diagnosed with autism, and shows a surprising spike in this rate in Alabama in 2017, a fact explained in the Discussion section. Figure 3 shows the DID method's strength in analyzing nonlinear flourishing trends. While there were no significant results involving flourishing, these figures provide a valuable example of how the DID regression specification can accommodate changes in flourishing that apply to both states. These figures also show parallel trends in the relevant characteristics between states (with the exception

[2] See Appendix D for details on the subgroup that was chosen to be examined (children who have ever been diagnosed with autism) for all results other than diagnostic rate.

of the spike in diagnosis rates in Alabama 2017), both before and after the implementation of the Alabama insurance mandate, supporting the use of a difference-in-difference analysis.

The increased uptake of behavioral therapy in Alabama suggests a significant improvement in access to behavioral therapy as a result of the autism coverage mandate, suggesting that it fulfills its intended purpose in this case. However, the lack of an effect on diagnosis rates suggests either that the previous system was able to get children diagnosed (such as through evaluations performed through the school system) or that even the current system may not be sufficient to diagnose everyone (e.g. if many people have difficulties in finding a doctor who will take their insurance). Furthermore, the lack of a significant effect on flourishing suggests either that the effects of ABA on flourishing are too small to be detected by this test, or that they occur over a time period longer than a year.

5. Conclusion

As a result of the analysis comparing Alabama and Mississippi after Alabama passed its mandate for insurance coverage of therapies for autism, we find that, while there is a significant effect on whether autistic children have received behavioral therapies in the last twelve months, there is no significant effect on any other variable of interest, including autism diagnosis rates, perceptions of coverage adequacy, proportion spending over $1,000 out of pocket per year, or flourishing.

The fact that the mandate did not affect perceptions of affordability or out-of-pocket expenses for families of autistic children implies that very few people who had behavioral treatments such as ABA were paying for it out of pocket; instead, they were likely getting it through insurers that covered it, through their children’s schools, or through other methods. However, the fact that the mandate did significantly increase behavioral treatment uptake implies that many autistic children who were good candidates to receive ABA were not getting it prior to the mandate because their insurance did not cover it. This suggests that these mandates fulfill their purpose of expanding access to ABA, while not expanding access to autism diagnostic evaluations (as shown by the lack of a mandate effect on diagnosis rates) and not increasing the affordability of ABA for families already receiving it.

This finding suggests that advocacy for these mandates would likely result in an increase in behavioral therapy. However, given how many US states have these mandates in place, much of the debate over these mandates would apply to other countries, which this study's findings do not necessarily extend to. A key area for further research would be to investigate other countries (or potentially, if earlier data can be obtained, other US states with similar characteristics to other countries of interest). Another key area would be to investigate potential long-term effects of ABA with other study methodologies, such as by examining longitudinal (rather than cross-sectional) data, examining larger datasets that can accommodate analyses that account for leads and lags on the effect, or conducting larger studies across multiple clinics that make use of waitlisted controls in the style of Choi et al. (2022).

6. Discussion

The results of this research, while interesting in their own right, also leave quite a few avenues for future research. The most obvious is that future research could make use of more autism-specific datasets with more observations, which could lead to reduced standard errors in regression and greater statistical power to detect small changes in the outcomes. Larger datasets would also allow for more sophisticated analyses, such as introducing time leads and lags to possible effects, to return results that would likely have lower p-values (due to reduced variance), mitigating the risk of inadvertent selective reporting of results. A larger sample could also include more children under 6 years old, from which flourishing data could potentially be collected. Additionally, future research could corroborate self-reports in surveys with clinical data, to reduce the effect of salience-related biases on self-reports. In particular, the bump in autism diagnosis rates in 2017 in Alabama, right when the insurance coverage mandate went into effect, was likely due to such a salience effect (e.g. pediatricians were thinking about it and so possibly talking with parents about it more), as this bump does not continue into subsequent years.

Additionally, future research could investigate the effects of autism coverage mandates on healthcare provision mechanisms other than private health insurance, including Medicaid, other types of public health insurance, or universal healthcare systems of the type found in many other countries. In particular, if a country had universal healthcare and began rolling out autism therapies such as ABA as part of a pilot program based either on a lottery system or to particular regions of the country before others, this could present a valuable opportunity for a natural experiment on outcomes, as well as possibly allowing a larger, more autism-specific survey to take place in applicable populations.

Note also that any findings on the effects of insurance coverage mandates for autism therapies apply to the marginal therapy recipient. That is, any effect that these therapies have on people who would be getting them whether there is an insurance mandate or not would not affect the result of this analysis. This means that this finding would not be applicable in, for instance, deciding whether a school should use a particular autism therapy methodology or not. However, it would be applicable in determining whether an autism insurance coverage mandate should come into effect, at least in environments sufficiently similar to Alabama or Mississippi.

Appendix A: Tables

Table 1: NSCH Summary Statistics on Autism Indicators in Alabama and Mississippi		
Note: all statistics other than the rate of children who have ever been diagnosed with autism and the rate of children who currently have autism are expressed for the subgroup of children who have ever been diagnosed with autism.		
Variable	**Alabama**	**Mississippi**
Panel A: Primary Outcomes		
Children who have ever been diagnosed with autism (per 1,000 children)	31.0	30.0
	(2.5)	(2.5)
Children currently diagnosed with autism (per 1,000 children)	29.3	27.8
	(2.4)	(2.4)
Children receiving behavioral treatment in last 12 months (%)	56.4	52.1
	(4.1)	(4.2)
Panel B: Insurance Coverage Outcomes		
Reporting over $1,000 in out-of-pocket expenses (%)	23.0	17.3
	(3.5)	(3.2)
Reporting mental/behavioral health needs “always” or “usually” covered (%)	74.2	76.6
	(3.9)	(4.1)
Panel C: Educational Services		
Currently in SPED or EI plan (%)	59.2	63.6
	(4.1)	(4.1)
Panel D: Flourishing		
Average flourishing metric (ages 6-17)	1.44	1.54
	(0.07)	(0.07)

Table 2: Difference-in-difference regression coefficients						
Note: all statistics other than the autism diagnostic rate are expressed for the subgroup of children who have ever been diagnosed with autism. Year coefficients are not reported, as there are a large number of them.						
*** = significant at p = 0.1; ** = significant at p = 0.05; *** = significant at p = 0.01**						
	Ever diagnosed with autism (per 1,000)	**Received behavioral treatment in the last 12 months (%)**	**Paid over $1,000 out of pocket last year (%)**	**Reporting that mental/behavioral health care is "always" or "usually" covered (%)**	**Currently in SPED (%)**	**Average flourishing metric among 6-17-year-olds**
Alabama dummy (Alabama = 1)	-1.442 (4.043)	13.1** (6.67)	1.59 (5.39)	1.86 (6.49)	-3.66 (6.57)	-0.175 (0.112)
Mandate dummy (True = 1)	-11.32 (8.25)	38.4*** (14.5)	-15.3 (11.7)	21.0 (14.2)	9.33 (14.5)	-0.400 (0.243)
Constant	26.38*** (6.813)	33.4*** (12.4)	35.2*** (9.99)	50.9*** (13.0)	68.5*** (12.2)	1.44*** (0.209)

Appendix B: Figures

In these analyses, "Actual" results are based on empirical data for the year and state in question (i.e. the "Alabama - Actual" result on the 2019 line in Figure 2 represents the proportion of children in Alabama in the 2019 NSCH who have ever been diagnosed with autism that are currently receiving behavioral treatment). "Predicted" results are based on the difference-in-difference estimate of the outcome variable for the year and state in question, as described in the Methods section of the paper.

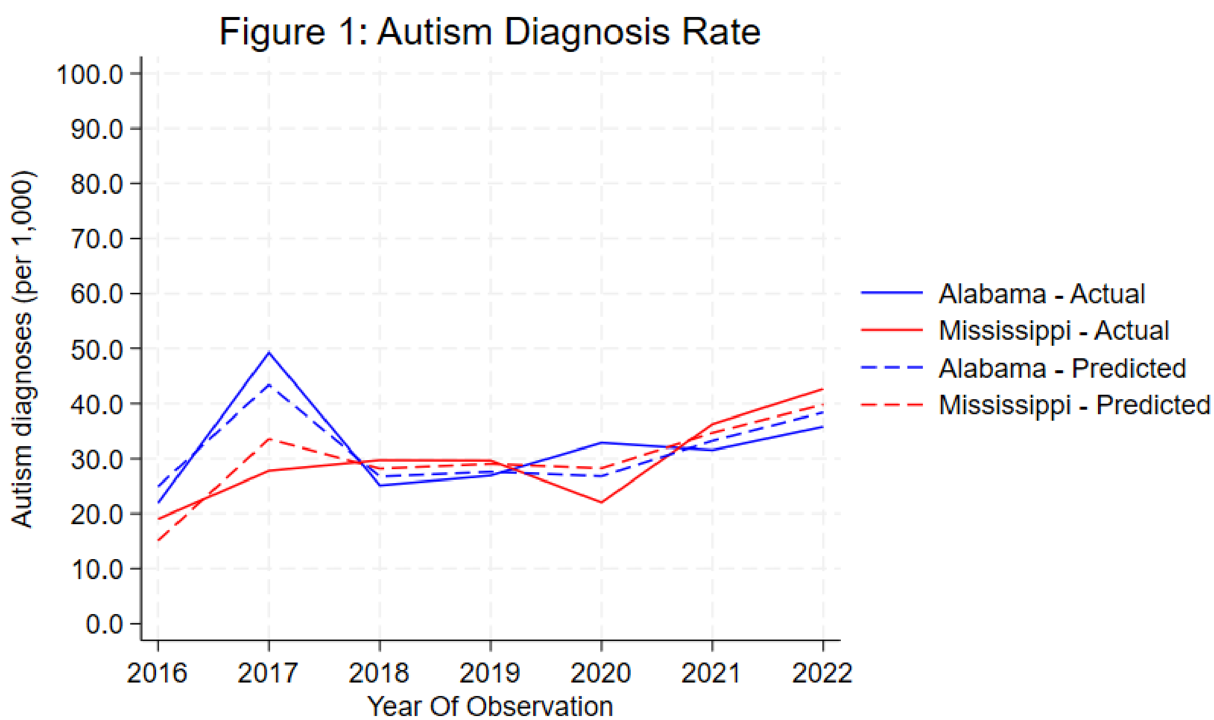


Figure 2: Behavioral Treatment Rate of Autistic Children

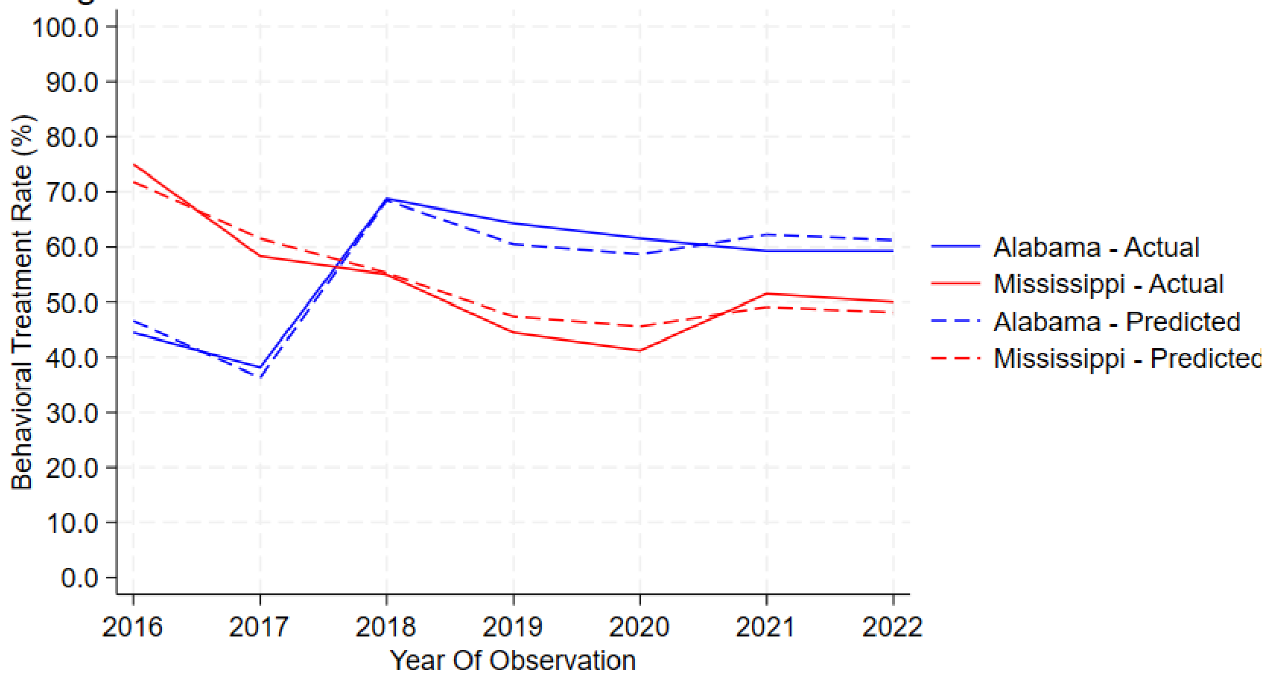
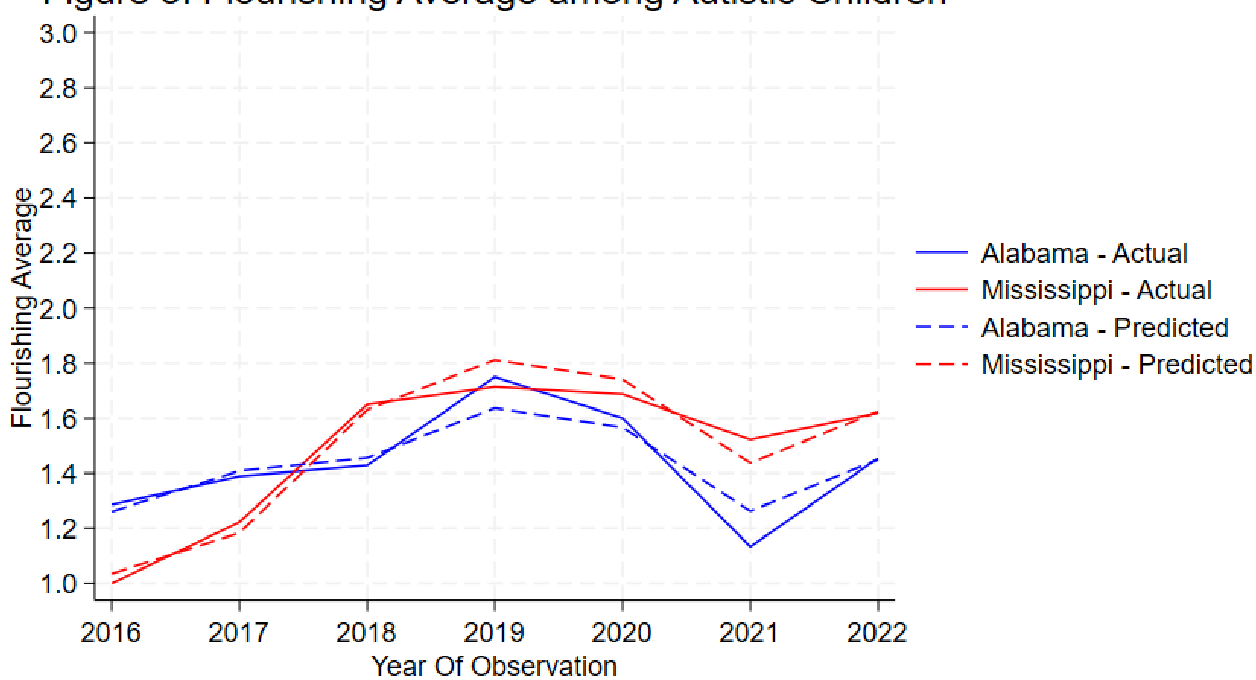


Figure 3: Flourishing Average among Autistic Children



Appendix C: A Note on Language Surrounding Autism

Within discussions of autism, there is significant debate over whether to use person-first language (e.g. “person with autism) or identity-first language (e.g. “autistic person”). Though person-first language is considered the standard within most disability communities (Wooldridge 2023), many members of the autistic community have expressed a preference for the use of identity-first language, as they see autism as an identity and not just as a medical condition (Brown 2012). However, many other figures who have relevant lived experiences relating to autism, including many parents of children with autism, prefer person-first language. As this paper discusses a survey of autistic children which was also conducted with the help of parents, there was some uncertainty as to which language to use. Ultimately, in the interest of readability and fair representation of all stakeholders, this paper uses a combination of person-first and identity-first language, in keeping with the current internal style at Autism Speaks (Autism Speaks, n.d.-d), and with the choice of which to use in a particular instance determined first and foremost by readability.

Appendix D: A Note on the Subgroup Chosen for the Analysis

For all values of interest other than the autism diagnosis rate, the sample was constrained to the subgroup of children who have ever been diagnosed with autism. Additionally, the autism diagnosis rate was expressed in terms of the proportion of children who have ever been diagnosed with autism. This was done, rather than constraining to the subgroup of children who currently have autism, because if ABA caused children who received it to no longer meet the diagnostic criteria for autism (such children will be referred to subsequently as “formerly autistic children”), then that would cause a negative effect on the proportion of children who currently have autism. However, the extent to which ABA can do this is outside the scope of this paper, and so to limit our analysis to the mechanism of autism coverage mandates potentially increasing diagnosis rates by improving access to diagnosis, the diagnosis rate analysis was performed in terms of the proportion of children who have ever been diagnosed with autism. Additionally, this population was used as the subject of the analyses for other variables, both because it was seen as desirable to capture any effects on indicators (notably flourishing) among formerly autistic children and to ensure consistency between these results and the results on diagnosis rates. Note also that the population in question is referred to throughout the paper as “autistic children.” This is not to imply that they all currently meet the diagnostic criteria for autism; only that the vast majority of them do (as the proportion currently diagnosed with autism and the proportion ever diagnosed with autism are within one standard deviation of each other in both Alabama and Mississippi), and that any formerly autistic children likely did meet these criteria when receiving any potential effects from this treatment (as a child who no longer met the diagnostic criteria for autism would likely not be receiving therapies for autism anymore).

References

Acton, Riley, Scott Imberman, and Michael Lovenheim. 2021. "Do Health Insurance Mandates Spillover to Education? Evidence from Michigan's Autism Insurance Mandate." *Journal of Health Economics* 80 (December):102489. https://doi.org/10.1016/j.jhealeco.2021.102489.

Anderson, Angelika, and Monica Carr. 2021. "Applied Behaviour Analysis for Autism: Evidence, Issues, and Implementation Barriers." *Current Developmental Disorders Reports* 8 (4): 191–200. https://doi.org/10.1007/s40474-021-00237-x.

Autism Speaks. 2019a. "Health Insurance Coverage for Autism." September 2019. https://www.autismspeaks.org/health-insurance.

———. 2019b. "Mississippi State-Regulated Insurance Coverage." December 2019. https://www.autismspeaks.org/mississippi-state-regulated-insurance-coverage.

———. 2024. "Autism Diagnosis on the Rise, According to Trends Study." November 1, 2024. https://www.autismspeaks.org/science-news/why-autism-increasing.

———. n.d.-a. "Alabama State-Regulated Insurance Coverage | Autism Speaks." Accessed October 2, 2024. https://www.autismspeaks.org/alabama-state-regulated-insurance-coverage.

———. n.d.-b. "Applied Behavior Analysis (ABA)." Accessed November 26, 2024. https://www.autismspeaks.org/applied-behavior-analysis.

———. n.d.-c. "Autism Spectrum Disorder (ASD)." Accessed November 26, 2024. https://www.autismspeaks.org/what-autism.

———. n.d.-d. "Describing Autism and People on the Spectrum." Accessed November 27, 2024. https://www.autismspeaks.org/describing-autism-and-people-spectrum.

Brook, Robert H., Emmett B. Keeler, Kathleen N. Lohr, Joseph P. Newhouse, John E. Ware, William H. Rogers, Allyson Ross Davies, et al. 2006. “The Health Insurance Experiment: A Classic RAND Study Speaks to the Current Health Care Reform Debate.” RAND Corporation. https://www.rand.org/pubs/research_briefs/RB9174.html.

Brown, Lydia. 2012. “Identity-First Language - Autistic Self Advocacy Network.” *Https://Autisticadvocacy.Org/* (blog). March 2, 2012. https://autisticadvocacy.org/about-asan/identity-first-language/.

CDC. 2024a. “Autism Spectrum Disorder (ASD).” Autism Spectrum Disorder (ASD). February 22, 2024. https://www.cdc.gov/autism/index.html.

———. 2024b. “Developmental Disability Basics.” Child Development. July 30, 2024. https://www.cdc.gov/child-development/about/developmental-disability-basics.html.

Child and Adolescent Health Measurement Initiative (CAHMI). 2020. “NSCH Codebooks.” 2020. https://www.childhealthdata.org/learn-about-the-nsch/nsch-codebooks.

———. n.d.-a. “Full-Length NSCH Survey Instruments.” Accessed November 27, 2024. https://www.childhealthdata.org/learn-about-the-nsch/survey-instruments.

———. n.d.-b. “National Survey of Children’s Health.” Accessed October 2, 2024. https://www.childhealthdata.org/learn-about-the-nsch/NSCH.

Choi, Kristen R., Bhumi Bhakta, Elizabeth A. Knight, Tracy A. Becerra-Culqui, Teri L. Gahre, Bonnie Zima, and Karen J. Coleman. 2022a. “Patient Outcomes After Applied Behavior Analysis for Autism Spectrum Disorder.” *Journal of Developmental and Behavioral Pediatrics : JDBP* 43 (1): 9–16. https://doi.org/10.1097/DBP.0000000000000995.

———. 2022b. “Patient Outcomes After Applied Behavior Analysis for Autism Spectrum Disorder.” *Journal of Developmental & Behavioral Pediatrics* 43 (1): 9–16. https://doi.org/10.1097/DBP.0000000000000995.

Choi, Kristen R., Elizabeth A. Knight, Bradley D. Stein, and Karen J. Coleman. 2020a. “Autism Insurance Mandates in the US: Comparison of Mandated Commercial Insurance Benefits Across States.” *Maternal and Child Health Journal* 24 (7): 894–900. https://doi.org/10.1007/s10995-020-02950-2.

———. 2020b. “Autism Insurance Mandates in the US: Comparison of Mandated Commercial Insurance Benefits Across States.” *Maternal and Child Health Journal* 24 (7): 894–900. https://doi.org/10.1007/s10995-020-02950-2.

Dixon, Mark R., Dana Paliliunas, Becky F. Barron, Ayla M. Schmick, and Caleb R. Stanley. 2021. “Randomized Controlled Trial Evaluation of ABA Content on IQ Gains in Children with Autism.” *Journal of Behavioral Education* 30 (3): 455–77. https://doi.org/10.1007/s10864-019-09344-7.

Douglas, Megan D., Teal W. Benevides, and Henry Carretta. 2017. “Analyzing State Autism Private Insurance Mandates for Allied Health Services: A Pilot Study.” *OTJR : Occupation, Participation and Health* 37 (4): 218–26. https://doi.org/10.1177/1539449217730355.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. 2012. “The Oregon Health Insurance Experiment: Evidence from the First Year.” *The Quarterly Journal of Economics* 127 (3): 1057. https://doi.org/10.1093/qje/qjs020.

Full Spectrum ABA. 2024. “ABA Therapy in Mississippi.” Full Spectrum ABA. 2024. https://www.fullspectrumaba.com/mississippi.

Hawkins, Drew. 2024. “The Politics Holding Back Medicaid Expansion in Some Southern States.” *KFF Health News* (blog). August 8, 2024. https://kffhealthnews.org/news/article/politics-medicaid-expansion-mississippi-south/.

Kahng, Sungwoo, and Janine Stichter. n.d. “Can ABA Help a Teen on the Mildly Affected End of the Autism Spectrum?” Accessed November 26, 2024. https://www.autismspeaks.org/expert-opinion/aba-teens.

National Conference of State Legislatures (NCSL). 2021. “National Conference of State Legislatures (NCSL) Autism and Insurance Coverage Database .” August 24, 2021. https://www.ncsl.org/health/autism-and-insurance-coverage-state-laws.

National Institute of Mental Health. 2024. “Autism Spectrum Disorder.” February 2024. https://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-asd.

Optum, Inc. 2022. “Mississippi CAN/CHIP Autism Program Provider Training.” https://public.providerexpress.com/content/dam/ope-provexpr/us/pdfs/clinResourcesMain/autismABA/msaba/MScanCHIPprovTrainrev032019.pdf.

“States With Specific Autism Mandates.” n.d. American Speech-Language-Hearing Association. Accessed October 2, 2024. https://www.asha.org/advocacy/state/states-specific-autism-mandates/.

US Census Bureau. n.d. “National Survey of Children’s Health (NSCH).” Census.Gov. Accessed October 2, 2024. https://www.census.gov/nsch.

Wang, Li, Junyi Ma, Ruchita Dholakia, Callie Howells, Yun Lu, Chen Chen, Runze Li, Michael Murray, and Douglas Leslie. 2019. "Changes in Healthcare Expenditures After the Autism Insurance Mandate." *Research in Autism Spectrum Disorders* 57 (January):97–104. https://doi.org/10.1016/j.rasd.2018.10.004.

Wooldridge, Shannon. 2023. "Writing Respectfully: Person-First and Identity-First Language." National Institutes of Health (NIH). April 12, 2023. https://www.nih.gov/about-nih/what-we-do/science-health-public-trust/perspectives/writing-respectfully-person-first-identity-first-language.

Assessing the Impact of Project Green Light: A Google Sustainability Initiative to Reduce City Air Pollution

Kenneth Byrne

14.33: Research and Communication in Economics

December 6, 2024

Abstract

In early 2020, Google Research began a sustainability initiative titled Project Green Light to create AI-powered recommendations for existing traffic lights to reduce emissions and traffic congestion in partnering cities. This paper utilizes a stacked difference-in-differences approach to separately investigate the effects of Project Green Light traffic light recommendations on eight different polluting gases at the intersection level in three cities across the world - Boston, Budapest, and Seattle. Ultimately, the results suggest that Project Green Light has an insignificant impact on pollutant concentrations at participating intersections.

Introduction

Cities around the world are struggling with combating high traffic congestion and air pollution, and due to the inextricable link of these two issues, improving traffic congestion goes hand-in-hand with improving air pollution. The noteworthy financial consequences of poor traffic flow in cities is quantifiable and serves as the driving force to motivate cities to improve traffic congestion. For instance, both Project Green Light participating cities in the US (Boston and Seattle) are included in the sample, and drivers in these cities lose on average 88 and 58 hours per year due to traffic congestion, respectively. According to the traffic data insights company INRIX, such delays result in economic damages of $4.5 billion per year.

The research question addressed in this paper regards assessing the treatment effect of implementing Google's traffic light recommendations at intersections on eight different air pollutant concentrations. It is hypothesized that Green Light treatment will decrease polluting gas concentrations at treated intersections in the short term.

Transportation and Air Pollution

Transportation is a major contributor to greenhouse gas emissions and air pollution. Road transport accounts for approximately 75 percent of all transport emissions globally (Ritchie, 2020). Although Google has asserted that the Project Green Light initiative has the “potential for up to 30% reduction in stops and 10% reduction in greenhouse gas emissions,” empirical research into the traffic relieving or environmental benefits of this particular project is non existent. Considering that pollution at city intersections can be up to 29 times greater compared to open roads, where approximately half of these increased emissions are a result of stop-and-go traffic, Project Green Light attempts to tackle an issue with significant room for improvement (Marini, 2024).

According to the US EPA, there are six criteria air polluting gasses[1] (ozone, particulate matter, carbon monoxide, lead, sulfur dioxide, and nitrogen dioxide), and vehicles primarily emit three of these - carbon monoxide (CO), nitrogen oxides (NOx), and particulate matter (PM). Rossi et al. (2020) studied the effect of road traffic on city air pollution in Italy, and found that vehicle traffic significantly impacts the pollutant concentrations of nitrogen oxides (NOx), yet their research found no clear relationship between traffic flows and particulate matter pollutant concentrations.

Google Project Green Light

Google Research began Project Green Light in 2020 as a sustainability initiative to reduce traffic congestion and air pollution in cities. Since the first pilot in 2021, Project Green Light has expanded to over 70 intersections across 13 total cities today, with Boston recently added in April 2024. Participation in the initiative is completely free for cities, and once enrolled, cities receive access to a customized dashboard which provides AI-powered traffic light recommendations for lights throughout their city. Google generates these recommendations for local traffic engineers to approve of and implement by utilizing extensive amounts of historical driving data from Google Maps. According to Laura Wojcicki, an engineer at Seattle’s Department of Transportation, recommendations can be implemented in as little as five minutes (Tracy, 2024).

[1] https://www.epa.gov/criteria-air-pollutants

The majority of traffic signals in the US are ”fixed time” signals which operate according to a pre-set pattern. Traffic lights running on fixed timings tend to have one pattern for peak times such as rush hour windows, and one other pattern for off-peak times (Mims, 2024). According to (Aleko and Djhael, 2020), they allege that fixed time signals will be suboptimal for combating the increasing congestion levels in cities as they lack the ability to respond to ever-changing traffic flows.

For cities intending to improve the traffic flow at lights in their city, there are a couple of main options. First, cities can transform their fixed time lights into adaptive lights. Existing forms of adaptive traffic control technologies require cities to purchase hardware in the form of cameras and sensors or invest in new software. To make the transition, it can cost upwards of $250,000 to modernize an intersection with the required cameras, sensors, and infrastructure to implement such adaptive technologies. A second and more common alternative is for cities to run traffic studies which involve counting the number of vehicles that pass through an intersection at various times of the day. Such studies cost approximately $5,000, and for this reason, cities tend on average to reassess the timing of traffic lights once every five years.

Given the hefty costs to improve traffic flows at intersections via modernization of infrastructure or re-optimizing light timing, the zero cost aspect of participating in Project Green Light is a major benefit, especially for cash-strapped municipalities that lack sufficient resources. While Green Light is not a dynamic traffic control system as it is meant to re-optimize existing fixed lights, the ability of Google to monitor intersections in real-time and provide frequent assessments of traffic flows effectively serves the role of modernizing participating intersections.

Data

Air Pollution Data[2]

Granular air pollution data has been obtained from the OpenWeather Historical Air Pollution API. With this API, air pollution data for eight major polluting gasses can be retrieved for specified latitude and longitude coordinate locations down to one-hour granularity. These polluting gases include carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), particulate matter less than 2.5 micrometers ($PM_{2.5}$), particulate matter less than 10.0 micrometers (PM_{10}), ammonia (NH_3), and ozone (O_3). The units for each polluting gas are $\mu g/m^3$. It is important to note that these eight polluting gases are not greenhouse gases. Greenhouse gases occur naturally and refer to any gases that can trap heat in the atmosphere such as carbon dioxide (Gulec, 2023). Polluting gases include any gas that can negatively impact the environment.

Hourly-level data observations for a four hour sample period each day (6:00AM-9:00AM) over a five month pre-treatment window and up to a five month post-treatment window have been obtained from the API for each polluting gas. I have chosen such a window as this is presumably when the effect of Project Green Light would be most pronounced. Slightly less than half of the treated intersections in Boston were implemented less than five months ago, so I obtained the maximum length post-treatment sample window for those intersections. The minimum length post-treatment window is six weeks, which applies to four intersections in Boston.

For each data observation, I determined the number of weeks it was relative to Green Light treatment (e.g. five months before Green Light treatment is 20 weeks prior). For each week relative to Green Light treatment in the five month pre-treatment window up to a five month post-treatment window, I calculated a weekly average pollutant concentration, where I excluded weekends and holidays.

[2] https://openweathermap.org/api/air-pollution

Treated Intersections in Participating Cities

Attempts to contact each of the transportation departments in the 13 active Green Light participating cities throughout the world have yielded varying levels of success. Three cities - Budapest, Boston, and Seattle - have responded with detailed lists of intersections in which Green Light traffic recommendations have been implemented, the initial dates of implementation, and or other background information such as specific intersections in which Green Light recommendations were quickly reverted due to worsened traffic flow.

Budapest implemented Green Light recommendations at five total intersections beginning October 04, 2022. Seattle implemented recommendations at four intersections at varying times, beginning July 21, 2022. Boston has implemented recommendations at 26 intersections since April 2024. Please find an attached table in the Appendix with a complete list of all actively treated intersections across these three cities, the geographic coordinates of the intersections, and the implementation dates and end dates, if applicable.

Data Exclusions

I have excluded three treated intersections from the sample. One treated intersection was located in Seattle at 5th Ave N and Denny Way where a Green Light implementation was made on 7/21/2022 and then timing change reverted back on 8/23/2022 based on the findings. The remaining two treated intersections excluded were in Boston at Columbia RD and Dorchester Avenue and Columbia Road and Pond Street. These were the first two recommendations which Boston implemented prior to April School Vacation and the recommendations were reverted back to the original timings the following week due to poor performance.

Methods

The main empirical test in this paper is a stacked difference-in-differences (DID) strategy to account for the varying treatment dates of intersections within cities. The unit of observation is at the intersection level and the frequency of observations is weekly.

The stacked DID allows for differences in baseline outcome levels between treated and control units within each constructed group. While the implementation of Green Light recommendations is fully at the discretion of transportation engineers within the participating cities as Google's role is limited to simply providing recommendations via a dashboard, I argue that it is unlikely that local city transportation engineers selected Green Light treatment on trends in the outcome variable (i.e. pollutant concentration), or even on trends in variables highly correlated with the outcome variables such as traffic congestion. Such treatment selection is unlikely as transportation engineers are concerned with implementing recommendations to maximize traffic flow without compromising safety.

Control and Treatment Group Construction

The sample contains 32 total treated units and 66 control units (Table 3). Groups of treated and control intersections within each city have been constructed based on the initial implementation date of a Green Light recommendation in order to restrict each group's observations to the same calendar time to mitigate confounding effects. The control intersections are given a fake treatment date that is the same as that of the corresponding treated intersections in their group. Using OpenStreetMap[3], the roadway classification of each treated unit was determined to ensure that the corresponding control units compose a representative sample of roadway classifications to the treated units in their group. Across the three cities in our sample, there are 17 total groups. Each group contains twice as many control units as treated units and groups are stacked according to

[3] https://www.openstreetmap.org

their week relative to treatment.

Due to the relatively low number of treated units at each city, control units have been selected from the same city in which their corresponding group of treated unit(s) belong to for two primary reasons. First, the low number of traffic lights receiving treatment at each city suggests that it is unlikely that Project Green Light has significant effects on overall city air pollution, but rather more localized effects at the intersection level as hypothesized. Therefore, the bias on our treatment effect estimate is presumably minimal as long as the distance between control units and the corresponding treated units in their group is sufficiently large. A radius of approximately 1 mile between treated and control units in each group has been imposed. Second, although selecting control units in different cities of the treated units in their group would eliminate the potential spillover effects from treated to control units inherent in a roadway network, the current design mitigates the potential bias inflicted from variation in the seasonality of traffic flows across cities which would exist if we sampled control intersections from different cities.

I argue that choosing control intersections from different cities would inflict a greater bias on our treatment effect estimate compared to the potential spillover effect bias as confounding factors such as construction projects, new transportation developments, significant weather events, or large-scale events such as sporting events or parades could vary greatly across cities. However, I'll assume for this analysis that the roughly 1 mile radius between control and treated units in each group is sufficient to mitigate spillover effects, but that it is not large enough to the point that factors as mentioned above have significantly different effects on intersections in the same city.

Stacked Difference-in-differences Model

The DID model can be expressed as follows:

$$Y_{igw} = \beta_0 + \beta_1 \cdot T_{ig} \times \text{Post}_{gw} + \lambda_i + \lambda_{gw} + \varepsilon_{igw} \tag{1}$$

Y_{igw} is the outcome variable (pollutant gas concentration) for intersection i in group g at week w. Group g refers to treated intersections in a city that first received treatment on the same date and their associated control intersections. T_{ig} is a binary indicator that equals 1 if intersection i is one of the treated intersections in group g and 0 otherwise. Post_{gw} is a binary indicator that equals 1 for the post-treatment period and 0 for the pre-treatment period for group g in week w. β_0 is the intercept. β_1 is the DID estimate on the $T_{ig} \times \text{Post}_{gw}$ interaction term that captures the treatment effect of Project Green Light at the intersection level. Last, the model captures the intersection level fixed effects in λ_i as well as time varying group effects as λ_{gw} captures the fixed effect of group g in the w^{th} week relative to Green Light implementation. Standard errors are clustered at the group level.

The saturated stacked DID model is robust to many concerns of identification as the intersection level fixed effects account for baseline differences in pollutant concentrations, and the time fixed effects for groups in each week control for time varying differences.

Event Study Model

An event study model has also been applied to estimate the dynamic effects in each weekly period up to 20 weeks prior to Green Light treatment and up to 20 weeks post-treatment.

$$\text{Y}_{igw} = \sum_{w=\underline{w}, w\neq 0}^{\bar{w}} \beta_w \cdot T_{ig} \cdot I_{gw} + \lambda_i + \lambda_{gw} + \varepsilon_{igw} \tag{2}$$

The event study model contains similar fixed effects terms to the stacked DID. The post-treatment indicator is replaced with I_{gw} which is a binary variable that equals 1 when observations in group

g are w weeks away from treatment. [$\underline{w}$, w] is the range of time periods at which the dynamic treatment effects are assessed. Here, [$\underline{w}$, $\bar{w}$] = [-20, 20]. Week 0, or the week when treatment begins, is imposed to be the reference week in our event study model, as we assume there are no anticipatory effects as drivers must obey traffic laws and react to the traffic light cycle timing changes in real-time.

A key satisfying assumption for identification is the parallel trends assumption, and we have tested whether each $\beta_w = 0$ for $-20 \leq w < 0$ for each of the eight polluting gases, which can be found in the Appendix. For each of the eight polluting gases, we have found significant coefficients for approximately three-fourths of the pre-treatment periods, which provides evidence against the parallel trends assumption.

Furthermore, a joint significance test was conducted on the pre-treatment window event study coefficients for each gas. For each gas, the coefficients in the pre-treatment window are all jointly significant (Table 10). The violation of parallelism is some time periods serves as a bias to our results.

Results

The estimates of the stacked DID find that intersections receiving Green Light treatment largely have insignificant changes in air pollutant concentrations. In Table 5, although the interaction term coefficients on NO_2 and O_3 are both statistically significant at the 1 percent level, the size of their coefficients is still relatively small as both are less than one-sixth of one standard deviation of their respective gas. The sign of the interaction term coefficient on NO_2 is positive, while the sign on the O_3 interaction term coefficient is negative.

While the finding of a positive sign on the treatment effect for NO_2 is unexpected, a surface level examination of the negative sign on the treatment effect for O_3 appears contradictory. However, the relationship between the differing signs of these two coefficients is corroborated by empirical scientific research due to a naturally occurring process in the atmosphere known as NO_x titration. First, it is necessary to clarify that O_3 is not directly emitted by vehicles, but instead forms as a result of ultraviolet radiation and mixture with other pollutants. Jhun et. el. (2015) researched the impact of NO_x concentrations on changing O_3 concentrations in the US, and explains that increasing NO_x concentrations is causally linked to decreasing O_3 concentrations as the amount of O_3 quenched via NO_x titration is greater.

Potential Sources of Bias

The OpenWeather Historical Air Pollution API utilizes historical sensor data to estimate air pollutant concentrations at specified coordinates. The power of the econometric analysis is limited due to the fact that data at a particular intersection may be derived from sensors far away from the actual treated intersection. This inflicts measurement bias in a manner that supports our findings of an insignificant treatment effect. The direction of measurement bias on our estimates is unknown, but it likely serves to nullify some of the true treatment effect of Project Green Light.

The data collected in this analysis spans a morning rush hour window fro 6:00AM-9:00AM. It is possible that Green Light implementations are optimizing traffic flows for other windows of the day outside of this period, such as evening rush hour. If this were true, it would corroborate our findings of an insignificant treatment effect during the morning rush hour window.

Last, since controls are selected in the same cities as their treated units, there will be a spillover effect bias to a certain degree, but as discussed, this is plausibly much smaller than the potential bias introduced from picking controls from different cities where seasonality of traffic flows can vary greatly among different cities. In this paper, we assume that the one mile buffer radius between controls and treated units inflicts a negligible bias on our estimates.

Robustness Checks

Varying Sample Window Lengths

All treated intersections in both Budapest and Seattle have post-treatment window lengths of at least 20 weeks. However, Boston is the most recent city in the world to adopt Project Green Light, and of Boston's 24 treated intersections, eight intersections began treatment less than 20 weeks ago. The newest treated intersections in Boston have post-treatment window lengths of six weeks. Consequently, I have re-estimated the stacked DID model for the same 20 week pre-treatment period up to a six week post-treatment period.

In Table 6, we can generally see similar findings to the ones from the stacked DID for a post-treatment window length of 20 weeks. However, the coefficient on NO is statistically significant at the 1 percent level, but the coefficient is still statistically small in size at approximately one-fifth the value of its standard deviation. The coefficients on O_3 and NO_2 are also both statistically insignificant now. This suggests it takes a longer time period than six weeks post-treatment to capture the treatment effect for O_3 and NO_2.

City by City Analysis

As discussed, the decision to implement a Green Light traffic light recommendation is at the discretion of local transportation engineers in the participating city. It is plausible that the ability of transportation departments to foresee potential issues with implementing a Green Light recommendation varies across cities. Furthermore, many of Boston's implementations occur as pairs of intersections in close proximity which is unlike Budapest or Seattle. Perhaps Boston is better equipped to implement more complex recommendations that the other two cities are unwilling to attempt. Re-estimating the stacked DID for each city allows us to see how the treatment effect varies in each city. Surprisingly, none of the three cities have significant interaction term coefficients for CO, although that is one of the primary polluting gases that vehicles emit. As we would expect, the coefficients for SO_2 and NH_3 are statistically insignificant for all three cities as vehicles do not primarily emit these gases.

Boston

Boston has implemented the greatest number of Green Light recommendations across a wide variety of road classifications, so they have the been the greatest adopter of the three cities in the sample. Boston also appears unique in the sense that many of their implementations have been pairs of closely related intersections. According to the National Association of City Transportation Officials, coordinated signal timing is typically applied to intersections located within a one-quarter mile proximity on corridors, which is likely why Boston has implemented many of these pairings.

The only interaction term with a statistically significant coefficient for Boston is O_3. Once again, although it is statistically significant, it's size is still quite small when compared to its standard deviation. The treated units in Boston account for three-fourths of all treated units, so it makes sense that the signs on the air pollutants in the Boston sample are similar to the signs on the air pollutants in the overall sample. For example, we can see that the same inverse relationship exists between NO_2 and O_3 in the Boston sample.

Budapest

Budapest was one of the earliest adopters of Project Green Light as they began implementing recommendations in October 2022.

The interaction terms with statistically significant coefficients for Budapest are NO, O_3, both particulate matters, and NH_3. Once again, although they are statistically significant, they are all still quite small when compared to its standard deviation. In Budapest, the same inverse relationship exists between NO_2 and O_3 as in the overall sample.

Seattle

While Seattle was also one of the earliest adopters of Project Green Light, they have the lowest number of treated intersections of any city in the sample at just three.

None of the interaction term coefficients for Seattle are significant. It is of note that all three of Seattle's treated intersections are in residential neighborhoods where the effect of Project Green Light is likely least pronounced due to presumably lower traffic volumes on average.

Discussion

The results are consistent with empirical research by Rossi et al. (2020) in that a significant interaction term coefficient was measured for NO_2, but treatment effects for particulate matter concentrations were both insignificant at the 5 percent level. Despite all interaction term coefficients being small in a real-world sense, we must consider the reasons why a city is using Green Light as a means to provide an explanation for such results. First and foremost, cities are utilizing Green Light to optimize traffic flow as this is most easily quantifiable for transportation departments. It is most likely that these cities are not measuring the effect of Green Light implementations on air pollutant concentrations at all. Therefore, the significant interaction term coefficient on NO_2 could be due to improved traffic flow at intersections which leads to more vehicles driving on those treated intersections, resulting in air pollutant concentrations increasing.

While Boston has exhibited active engagement with using Project Green Light for updating timing at its traffic lights, the stagnant and sparse implementation of Green Light recommendations by Budapest and Seattle brings into question the motivation behind participating in such a program. As two of the earliest cities in the program, it is plausible that these cities enrolled partially for publicity purposes due to the sustainability aspect of Project Green Light, as news of their enrollment was widespread across major news outlets. Enrollment in such an innovative service also demonstrates ambition and progressive behavior on behalf of city officials in attempting to tackle a major problem with a technologically advanced and cost-conscientious solution. For this reason, cities such as Budapest and Seattle may have been inclined to enroll, but are not wholeheartedly committed to utilizing Green Light. While participation in the program may be attributed to the superficial reasons mentioned prior, the other explanation for minimal engagement may be that cities lack sufficient guidance from Google to successfully implement recommendations, or they lack the time and resources to continually implement, monitor, and reassess implementations.

Last, the following observation provides evidence against rejecting the merits of Project Green Light. Aside from the three treated intersections that were excluded from the analysis as the traffic engineers at Boston and Seattle explained that the implemented recommendations were all reverted within one month due to poor performance, it is notable that all other treated units have remained treated until the present day. Some of these recommendations have remained in effect for over two years, which alone provides support for the efficacy of the Green Light initiative. At the very least, it suggests that Green Light recommendations have not resulted in poor enough outcomes to end participation altogether.

Conclusion

While the stack DID model in this paper is robust to many concerns of identification, the degree to which measurement error is biasing our treatment estimates is largely unknown. Due to the high granularity of the intersection level observations, it is quite plausible that the measurement error is nullifying the true treatment effect of Project Green Light. Furthermore, such measurement error provides some explanation for the statistically small and generally insignificant treatment estimates we found in the stacked DID.

Google's Project Green Light is a high-potential initiative that poses the ability to greatly impact air pollution and traffic congestion in cities around the world. Providing the ability to easily implement recommendations under existing traffic systems at no charge, the initiative has the possibility to scale quickly and garner widespread adoption. The growth in the number of participating cities since the first pilot in 2021 has been steady, and the active engagement of Boston, the newest Green Light city, is encouraging for promoting adoption across other US cities and cities throughout the world. Furthermore, cities that can benefit from Project Green Light are not limited by their existing infrastructure or geographic region, so the wide accessibility of the program also supports scalability.

Although the paper finds insignificant treatment effects for Project Green Light in the short-term, the Green Light program is still relatively new. Coupling Google's extensive data reserves with more time to refine their recommendation system strengthens the case that Green Light has the potential to pose significant long-term benefits.

References

[1] Aleko, Dex R., and Soufiene Djahel. "An Efficient Adaptive Traffic Light Control System for Urban Road Traffic Congestion Reduction in Smart Cities." *Information*, vol. 11, no. 2, 2020, p. 119. Available: https://www.researchgate.net/publication/339411789_An_Efficient_Adaptive_Traffic_Light_Control_System_for_Urban_Road_Traffic_Congestion_Reduction_in_Smart_Cities.

[2] Brinson, Linda, and Francisco Guzman. "How Much Air Pollution Comes from Cars?" *HowStuffWorks*, 7 July 2021. Available: https://auto.howstuffworks.com/air-pollution-from-cars.htm.

[3] Gulec, Ayse. "Greenhouse Gases and Air Pollution." *Airqoon - Cost Effective and Easy to Use Air Monitoring at Scale*, 30 Jan. 2023. Available: https://airqoon.com/resources/greenhouse-gases-and-air-pollution/.

[4] INRIX. "Global Traffic Scorecard." *Inrix*, 1 June 2024. Available: https://inrix.com/scorecard/#form-download-the-full-report.

[5] Jhun, Iny, Coull, Brent A., Zanobetti, Antonella, and Koutrakis, Petros. "The impact of nitrogen oxides concentration decreases on ozone trends in the USA." *Air Quality, Atmosphere and Health*, vol. 8, 2015, pp. 283–292. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC4988408/

[6] Marini, Ari. "How Google Uses AI to Reduce Stop-And-Go Traffic on Your Route — and Fight Fuel Emissions." *Google*, 29 July 2024. Available: https://blog.google/outreach-initiatives/sustainability/google-ai-project-greenlight/.

[7] Matias, Yossi. "Project Green Light's Work to Reduce Urban Emissions Using AI." *Google*, 10 Oct. 2023. Available: https://blog.google/outreach-initiatives/sustainability/google-ai-reduce-greenhouse-emissions-project-greenlight/.

[8] Mims, Christopher. "Google's Green Light System Aims to Optimize Traffic Lights." *The Wall Street Journal*, 27 Nov. 2024. Available: https://www.wsj.com/tech/personal-tech/google-green-light-traffic-light-optimization-992e4252.

[9] Ritchie, Hannah. "Cars, Planes, Trains: Where Do CO2 Emissions from Transport Come From?" *Our World in Data*, 6 Oct. 2020. Available: https://ourworldindata.org/co2-emissions-from-transport.

[10] Rossi, Riccardo, et al. "Effect of Road Traffic on Air Pollution. Experimental Evidence from COVID-19 Lockdown." *Sustainability*, vol. 12, no. 21, 29 Oct. 2020, p. 8984. Available: https://doi.org/10.3390/su12218984.

[11] Tracy, Ben. "How Google Is Using AI to Help One U.S. City Reduce Traffic and Emissions - CBS News." *CBS News*, 4 Jan. 2024. Available: https://www.cbsnews.com/news/google-project-green-light-seattle/.

Supporting Figures and Tables

Polluting Gas Data

Table 1: Polluting Gases Data Available with API

Polluting Gas	**Do Cars Directly Emit?**
Carbon Monoxide (CO)	Yes
Nitric Oxide (NO)	Yes
Nitrogen Dioxide (NO_2)	Yes
Sulfur Dioxide (SO_2)	Yes
Particulate Matter ($PM_{2.5}$)	Yes
Particulate Matter (PM_{10})	Yes
Ammonia (NH_3)	Yes
Ozone (O_3)	No

Google Project Green Light Participating Cities

Table 2: Green Light Cities

Number	**City**	**Country**
1	Seattle, WA	USA
2	Boston, MA	USA
3	Abu Dhabi	UAE
4	Bali	Indonesia
5	Bangaluru	India
6	Budapest	Hungary
7	Haifa	Israel
8	Hamburg	Germany
9	Hyderabad	India
10	Jakarta	Indonesia
11	Kolkata	India
12	Manchester	England
13	Rio de Janeiro	Brazil

Treated Intersections Table

Table 3: Treated Intersection Data

City	Intersection Name	Start Date	End Date	Coordinates
Seattle	15th Ave NW at NW Market St and NW 53rd St	12/21/2022	-	47.668811, -122.376311
Seattle	N 80th St and Greenwood Ave N	08/23/2022	-	47.687338, -122.355301
Seattle	2nd Ave Ext and Main St	05/21/2024	-	47.600219, -122.330483
Boston	Columbia Rd. & Washington St.	06/01/2024	-	42.305407, -71.080501
Boston	Columbia Rd. & Seaver St.	06/01/2024	-	42.304224, -71.082507
Boston	Carlton St. & Mountfort St.	07/15/2024	-	42.348924, -71.109024
Boston	Mountfort & St. Mary’s	07/15/2024	-	42.348560, -71.106860
Boston	St. Alphonsus St. & Tremont St.	06/09/2024	-	42.333015, -71.101318
Boston	Blakemore St. & Hyde Park Ave.	06/16/2024	-	42.285497, -71.118883
Boston	Cummins Hwy. & Hyde Park Ave	06/16/2024	-	42.279952, -71.118918
Boston	Rockland St. & Washington St.	05/08/2024	-	42.258878, -71.159927
Boston	Grove St. & Washington St.	05/08/2024	-	42.261235, -71.157063
Boston	Amory St. & Green St.	05/23/2024	-	42.310396, -71.107062
Boston	Cedar St. & Columbus Ave.	06/18/2024	-	42.328115, -71.096912
Boston	Centre St., Columbus Ave., & Ritchie St.	06/18/2024	-	42.322602, -71.098446
Boston	Prentiss St. & Tremont St.	06/23/2024	-	42.333034, -71.092520
Boston	Ruggles St., Tremont St., & Whittier St.	06/23/2024	-	42.334580, -71.089602
Boston	Frontage Rd. NB & South Boston Bypass Rd.	10/01/2024	-	42.333988, -71.063750
Boston	Frontage Rd. NB & Massachusetts Ave. Connector	10/01/2024	-	42.335413, -71.064759
Boston	Berkeley St. & Columbus Ave.	09/07/2024	-	42.348607, -71.071846
Boston	Berkeley St. & Stuart St.	09/07/2024	-	42.349547, -71.072227
Boston	Boylston St., Jersey St., & Yawkey Way	05/17/2024	-	42.344581, -71.097749
Boston	Boylston St. & Ipswich St	05/17/2024	-	42.345263, -71.095357
Boston	Faneuil St. & Market St.	04/25/2024	-	42.354831, -71.150074
Boston	Arlington St., Market St., & Sparhawk St.	04/25/2024	-	42.351644, -71.151920
Boston	Boylston St. & Hemenway St.	10/03/2024	-	42.346915, -71.089261
Boston	Boylston St. & Massachusetts Ave.	10/03/2024	-	42.347259, -71.087661
Budapest	Hősök tere (northern intersection on Dózsa György út)	10/04/2022	-	47.515544, 19.078667
Budapest	Kőbányai út / Szállás utca	10/04/2022	-	47.483426, 19.120752
Budapest	Fehérvári út / Dombóvári út	10/04/2022	-	47.470989, 19.045345
Budapest	M0 / 6	10/04/2022	-	47.399567, 19.009683
Budapest	Bajcsy Zsilinszky út / Alkotmány utca	10/04/2022	-	47.508031, 19.055014

Sample Summary Statistics by City

Table 4: Summary of Treated Intersections and Groups

Statistic	**Boston**	**Budapest**	**Seattle**	**Overall**
Number of Treated Intersections	24	5	3	32
Number of Control Intersections	50	10	6	32
Number of Treated Groups	13	1	3	17
Earliest Implementation Date	04/25/2024	10/04/2022	08/23/2022	08/23/2022
Latest Implementation Date	10/03/2024	10/04/2022	05/21/2024	10/03/2024

Pollutant Baseline Sample Summary Statistics

Treated Intersections - 5 month pre-treatment window

	co	**no**	**no2**	**so2**	**nh3**	**pm2.5**	**pm10**	**o3**
Mean	318.45	4.23	17.51	2.81	1.23	7.08	9.04	45.71
Median	308.34	2.45	15.69	2.68	0.84	5.99	7.89	44.97
Minimum	163.72	0.01	2.69	0.74	0.17	1.08	1.88	10.06
Maximum	684.05	27.97	53.77	8.00	6.91	25.80	32.76	78.71
Standard Deviation	88.99	4.69	9.07	1.18	1.10	4.48	5.30	14.67
Number of Observations	19,840	19,840	19,840	19,840	19,840	19,840	19,840	19,840

Treated Intersections - Up to 5-month post-treatment window

	co	**no**	**no2**	**so2**	**nh3**	**pm2.5**	**pm10**	**o3**
Mean	313.85	4.91	15.86	3.00	1.16	9.65	12.13	33.96
Median	300.99	2.39	15.50	2.71	0.61	6.44	9.19	34.52
Minimum	188.07	0.01	5.01	0.70	0.05	1.33	2.09	4.73
Maximum	727.32	30.93	43.01	14.73	27.86	64.06	73.47	70.02
Standard Deviation	76.65	5.22	6.48	1.95	2.72	9.02	10.21	13.82
Number of Observations	16,264	16,264	16,264	16,264	16,264	16,264	16,264	16,264

Stacked DID Regressions (all three cities)

Table 5: Stacked DID: 20 weeks pre-treatment to 20 weeks post-treatmemt

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	co	no	no2	o3	so2	pm2.5	pm10	nh3
TreatXPost	1.366	0.0807	0.964***	-1.767***	-0.0128	0.477*	-0.178	0.00334
	(2.205)	(0.194)	(0.313)	(0.458)	(0.0734)	(0.258)	(0.293)	(0.0287)
Constant	287.6***	3.045***	13.36***	47.56***	2.454***	7.617***	8.789***	0.866***
	(0.217)	(0.0198)	(0.0311)	(0.0437)	(0.00694)	(0.0245)	(0.0287)	(0.00274)
Observations	3698	3698	3698	3698	3698	3698	3698	3698

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Stacked DID: 20 weeks pre-treatment to 6 weeks post-treatmemt

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	co	no	no2	o3	so2	pm2.5	pm10	nh3
TreatXPost	4.255*	1.069***	0.230	0.647	0.272***	0.450*	0.125	-0.0299
	(2.376)	(0.276)	(0.309)	(0.529)	(0.0718)	(0.254)	(0.214)	(0.0403)
Constant	290.7***	3.073***	14.52***	46.50***	2.479***	7.839***	9.323***	0.925***
	(0.237)	(0.0257)	(0.0305)	(0.0520)	(0.00694)	(0.0238)	(0.0207)	(0.00382)
Observations	3698	3698	3698	3698	3698	3698	3698	3698

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Stacked DID Regressions (separated by city)

Table 7: Boston Stacked DID: 20 weeks pre-treatment to 20 weeks post-treatmemt

	(1) co	(2) no	(3) no2	(4) o3	(5) so2	(6) pm2.5	(7) pm10	(8) nh3
TreatXPost	-2.274	0.211	-0.347	0.748**	-0.0117	-0.0588	0.00102	-0.0266
	(1.544)	(0.185)	(0.217)	(0.334)	(0.0462)	(0.106)	(0.128)	(0.0169)
Constant	285.1***	2.949***	13.44***	45.03***	1.958***	6.113***	7.951***	0.585***
	(0.153)	(0.0182)	(0.0220)	(0.0335)	(0.00442)	(0.0105)	(0.0126)	(0.00165)
Observations	2738	2738	2738	2738	2738	2738	2738	2738

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Budapest Stacked DID: 20 weeks pre-treatment to 20 weeks post-treatmemt

	(1) co	(2) no	(3) no2	(4) o3	(5) so2	(6) pm2.5	(7) pm10	(8) nh3
TreatXPost	17.14	3.232***	1.103	-4.167***	0.713	2.517***	2.179***	0.198*
	(12.40)	(1.059)	(0.962)	(0.804)	(0.482)	(0.784)	(0.384)	(0.113)
Constant	313.7***	3.677***	12.09***	65.75***	4.572***	12.47***	9.406***	1.712***
	(0.963)	(0.0920)	(0.0871)	(0.0758)	(0.0321)	(0.0774)	(0.0404)	(0.0124)
Observations	600	600	600	600	600	600	600	600

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Seattle Stacked DID: 20 weeks pre-treatment to 20 weeks post-treatmemt

	(1) co	(2) no	(3) no2	(4) o3	(5) so2	(6) pm2.5	(7) pm10	(8) nh3
TreatXPost	0.382	-0.211	0.0947	1.542	0.256	-0.185	-0.0593	0.107
	(7.529)	(0.938)	(1.230)	(1.916)	(0.182)	(0.526)	(0.398)	(0.110)
Constant	287.0***	3.771***	19.43***	32.25***	2.540***	6.663***	8.716***	0.665***
	(0.721)	(0.0902)	(0.124)	(0.178)	(0.0186)	(0.0461)	(0.0395)	(0.00961)
Observations	360	360	360	360	360	360	360	360

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Event Study Results (Page 1 of 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	co	no	no2	so2	o3	pm2.5	pm10	nh3
F_1	-5.892***	-0.848***	-0.106	0.428	-0.00800	-0.166	-0.408	-0.0931***
	(2.084)	(0.212)	(0.176)	(0.913)	(0.0455)	(0.203)	(0.249)	(0.0308)
F_2	-4.372***	-0.347***	-0.453***	0.589	-0.0463	-0.675***	-0.773***	-0.0473**
	(1.566)	(0.125)	(0.123)	(0.858)	(0.0403)	(0.189)	(0.218)	(0.0201)
F_3	8.222***	-0.146	1.189***	-1.761**	0.00669	0.984***	1.261***	-0.0555**
	(1.724)	(0.0984)	(0.226)	(0.893)	(0.0544)	(0.199)	(0.251)	(0.0254)
F_4	12.40***	0.622***	1.102***	-2.807***	0.130***	2.192***	3.058***	-0.0140
	(1.692)	(0.0953)	(0.146)	(0.690)	(0.0396)	(0.310)	(0.448)	(0.0214)
F_5	11.23***	0.567***	0.802***	-3.084***	0.0907**	1.518***	1.790***	-0.0685**
	(1.775)	(0.114)	(0.114)	(0.656)	(0.0393)	(0.232)	(0.273)	(0.0285)
F_6	15.60***	0.785***	1.122***	-4.086***	0.172***	1.553***	1.819***	-0.0797***
	(2.317)	(0.161)	(0.158)	(0.765)	(0.0375)	(0.224)	(0.248)	(0.0203)
F_7	11.38***	0.990***	0.712***	-2.990***	-0.00575	0.870***	1.090***	-0.251***
	(2.332)	(0.194)	(0.167)	(0.902)	(0.0520)	(0.200)	(0.264)	(0.0783)
F_8	6.063***	0.279***	0.718***	-4.403***	-0.0172	-0.0148	0.182	-0.193***
	(1.834)	(0.0740)	(0.127)	(0.762)	(0.0413)	(0.288)	(0.316)	(0.0475)
F_9	-0.522	-0.373***	-0.322***	-2.222***	-0.0302	-0.228***	-0.232***	-0.248***
	(1.172)	(0.0737)	(0.102)	(0.464)	(0.0393)	(0.0759)	(0.0871)	(0.0541)
F_10	-2.015**	-0.561***	0.170**	-2.661***	0.0779*	-0.621***	-1.077***	-0.210***
	(0.814)	(0.117)	(0.0828)	(0.386)	(0.0407)	(0.147)	(0.235)	(0.0608)
F_11	20.14***	0.257***	2.032***	-6.219***	0.187***	1.532***	2.003***	-0.158***
	(2.582)	(0.0562)	(0.242)	(0.839)	(0.0282)	(0.181)	(0.235)	(0.0495)
F_12	22.78***	0.806***	1.805***	-4.237***	0.0977***	1.559***	2.011***	-0.0676
	(3.235)	(0.116)	(0.267)	(0.733)	(0.0246)	(0.215)	(0.269)	(0.0598)
F_13	19.56***	0.880***	1.733***	-5.577***	0.218***	1.342***	1.567***	-0.189***
	(2.227)	(0.0933)	(0.171)	(0.782)	(0.0287)	(0.162)	(0.191)	(0.0625)
F_14	17.48***	1.101***	1.428***	-5.000***	0.365***	1.288***	1.579***	-0.0565
	(2.802)	(0.154)	(0.275)	(0.886)	(0.0559)	(0.200)	(0.236)	(0.0480)
F_15	12.21***	0.261***	1.462***	-4.385***	0.210***	1.335***	1.455***	-0.0732*
	(1.489)	(0.0635)	(0.196)	(0.602)	(0.0291)	(0.182)	(0.201)	(0.0426)

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Event Study Results (Page 2 of 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	co	no	no2	so2	o3	pm2.5	pm10	nh3
F_16	8.572***	-0.148	0.573***	-1.188***	0.149***	0.717***	0.763***	-0.0982***
	(1.354)	(0.105)	(0.0875)	(0.338)	(0.0248)	(0.0984)	(0.0998)	(0.0342)
F_17	-7.631***	-0.219***	-1.224***	0.436	-0.00395	-0.671***	-1.013***	-0.220***
	(1.983)	(0.0836)	(0.276)	(0.404)	(0.0277)	(0.164)	(0.219)	(0.0488)
F_18	8.966***	-0.383***	1.748***	-4.641***	0.439***	1.054***	1.184***	0.427***
	(2.171)	(0.102)	(0.251)	(0.658)	(0.0563)	(0.171)	(0.202)	(0.0736)
F_19	2.380	-0.641***	0.458***	-0.260	0.220***	1.919***	2.021***	0.0969***
	(1.790)	(0.153)	(0.143)	(0.688)	(0.0256)	(0.312)	(0.339)	(0.0359)
F_20	8.595***	-0.0179	1.376***	-2.635***	0.139***	1.006***	1.241***	0.0986***
	(1.813)	(0.148)	(0.197)	(0.486)	(0.0254)	(0.170)	(0.207)	(0.0276)
L_1	6.384***	0.164***	0.445***	-2.100**	-0.0436	0.239***	0.192*	-0.0724***
	(0.775)	(0.0484)	(0.0912)	(0.925)	(0.0524)	(0.0855)	(0.101)	(0.0205)
L_2	5.582***	0.233***	0.510***	-0.184	0.0620	0.273**	0.273**	-0.0212
	(0.841)	(0.0537)	(0.101)	(0.838)	(0.0426)	(0.110)	(0.138)	(0.0176)
L_3	4.175***	0.157**	0.163**	-1.148	-0.0690	0.0357	-0.0851	-0.0335*
	(0.810)	(0.0680)	(0.0813)	(0.918)	(0.0472)	(0.136)	(0.178)	(0.0182)
L_4	2.122**	-0.0254	0.136	0.413	-0.0470	0.408***	0.390**	-0.0636***
	(1.062)	(0.107)	(0.123)	(0.909)	(0.0464)	(0.151)	(0.186)	(0.0229)
L_5	6.261***	-0.0651	0.193***	1.519	-0.0794*	-0.000748	-0.0199	-0.0343***
	(0.763)	(0.0673)	(0.0676)	(0.962)	(0.0436)	(0.107)	(0.130)	(0.0111)
L_6	5.690***	-0.246***	0.272***	0.851	0.0881*	-0.0284	-0.0404	-0.0790***
	(0.871)	(0.0940)	(0.0687)	(0.843)	(0.0466)	(0.158)	(0.189)	(0.0215)
L_7	3.606***	-0.0579	-0.0655	1.988*	-0.0829	-0.472*	-0.854**	-0.0443**
	(1.246)	(0.0918)	(0.170)	(1.104)	(0.0561)	(0.261)	(0.374)	(0.0202)
L_8	-0.482	-0.0342	-0.694***	2.682***	-0.202***	-0.386**	-0.649***	-0.103***
	(1.332)	(0.102)	(0.214)	(1.032)	(0.0555)	(0.183)	(0.233)	(0.0260)
L_9	13.68***	0.663***	0.962***	0.640	-0.00682	1.111***	1.135***	-0.0174
	(1.528)	(0.0837)	(0.162)	(0.795)	(0.0390)	(0.138)	(0.155)	(0.0204)
L_10	13.02***	0.540***	0.703***	0.601	-0.0574	0.379***	0.403***	-0.0265**
	(1.158)	(0.0631)	(0.0831)	(0.785)	(0.0390)	(0.0954)	(0.105)	(0.0124)

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Event Study Results (Page 3 of 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	co	no	no2	so2	o3	pm2.5	pm10	nh3
L_11	16.40***	0.691***	1.584***	0.973	0.0715***	1.128***	1.212***	0.0998***
	(1.961)	(0.116)	(0.187)	(0.597)	(0.0256)	(0.137)	(0.153)	(0.0224)
L_12	2.704	-0.230	-0.112	0.996	-0.106***	0.189	0.189	-0.113***
	(1.938)	(0.164)	(0.190)	(0.801)	(0.0327)	(0.147)	(0.165)	(0.0190)
L_13	5.013**	0.512***	-0.275	2.031***	-0.155***	0.208	0.218	-0.0834***
	(1.988)	(0.154)	(0.229)	(0.768)	(0.0399)	(0.151)	(0.192)	(0.0296)
L_14	9.405***	0.160	0.977***	0.583	0.0709***	0.695***	0.929***	-0.0384**
	(1.218)	(0.113)	(0.135)	(0.664)	(0.0226)	(0.0924)	(0.109)	(0.0164)
L_15	7.399***	0.327***	0.376***	0.611	-0.143***	0.107	0.104	-0.0698***
	(1.174)	(0.119)	(0.108)	(0.521)	(0.0346)	(0.108)	(0.123)	(0.0195)
L_16	4.186***	0.333***	0.243*	0.651	-0.0910***	0.222**	0.288***	-0.0189
	(1.085)	(0.0819)	(0.139)	(0.475)	(0.0312)	(0.0879)	(0.104)	(0.0118)
L_17	7.588***	0.769***	0.740***	-0.0920	0.0425	0.454***	0.481***	0.0470***
	(1.196)	(0.119)	(0.110)	(0.241)	(0.0262)	(0.0647)	(0.0769)	(0.0123)
L_18	5.408***	0.475***	0.166	0.304*	-0.131***	0.0331	0.0452	-0.0113
	(0.705)	(0.0704)	(0.106)	(0.166)	(0.0283)	(0.0513)	(0.0543)	(0.00785)
L_19	6.818***	0.562***	0.245**	1.007***	0.0397***	0.476***	0.500***	-0.0486***
	(0.881)	(0.0743)	(0.0950)	(0.245)	(0.0128)	(0.0380)	(0.0464)	(0.0183)
L_20	1.644**	0.0653	0.149**	0.0880	-0.0645***	0.174***	0.207***	-0.0482***
	(0.725)	(0.0513)	(0.0754)	(0.141)	(0.0113)	(0.0577)	(0.0683)	(0.00878)
Constant	298.8***	3.078***	14.30***	48.39***	2.341***	6.781***	8.926***	0.966***
	(0.646)	(0.0340)	(0.0483)	(0.363)	(0.0208)	(0.0654)	(0.0747)	(0.0182)
Stacked Observations	46580	46580	46580	46580	46580	46580	46580	46580

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Event Study Joint Significance Tests

Table 10: Joint Significance Tests of Pre-treatment Window Coefficients

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	(CO)	(NO)	(NO_2)	(O_3)	(SO_2)	($PM_{2.5}$)	(PM_{10})	(NH_3)
F Test Value	211.06	212.76	142.93	200.99	101.81	136.78	158.81	128.17
P Value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.000

Event Study Interaction Term Plots

Below are the plots for each of the eight pollutant gases for a pre-treatment window of 20 weeks and a post-treatment window of 20 weeks.

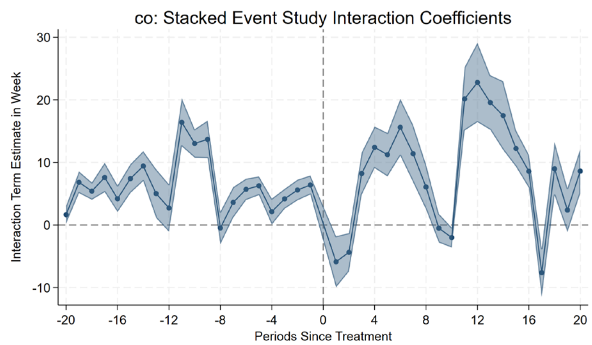


Figure 1: CO Event Study Plot

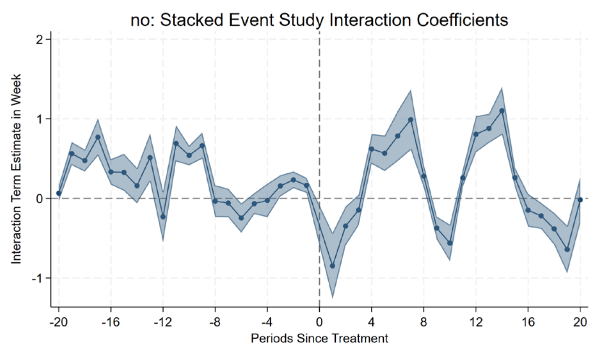


Figure 2: NO Event Study Plot

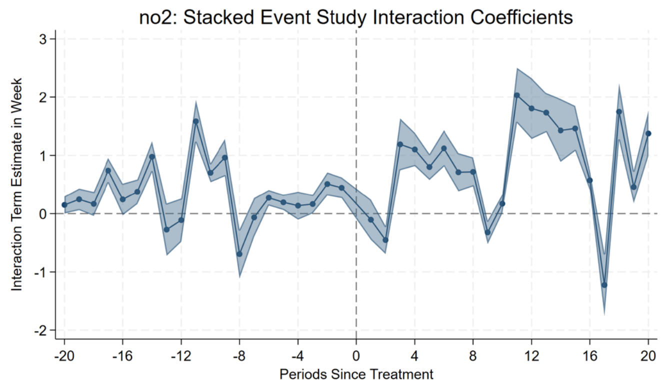


Figure 3: NO_2 Event Study Plot

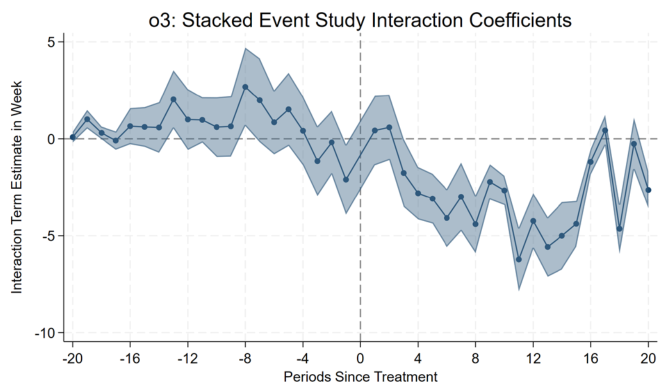


Figure 4: O_3 Event Study Plot

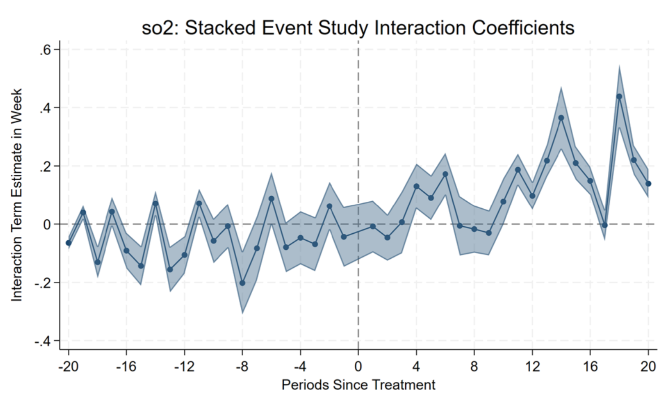


Figure 5: SO_2 Event Study Plot

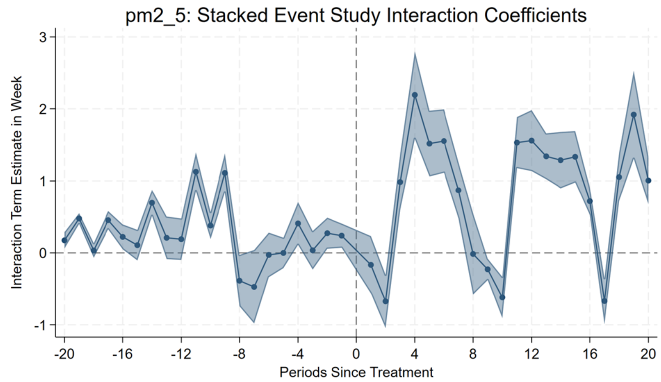


Figure 6: $PM_{2.5}$ Event Study Plot

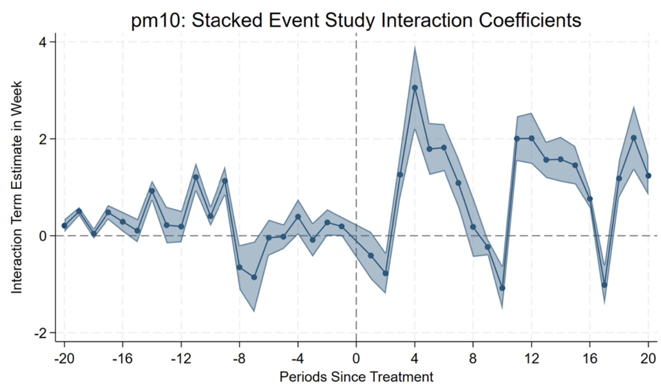


Figure 7: PM_{10} Event Study Plot

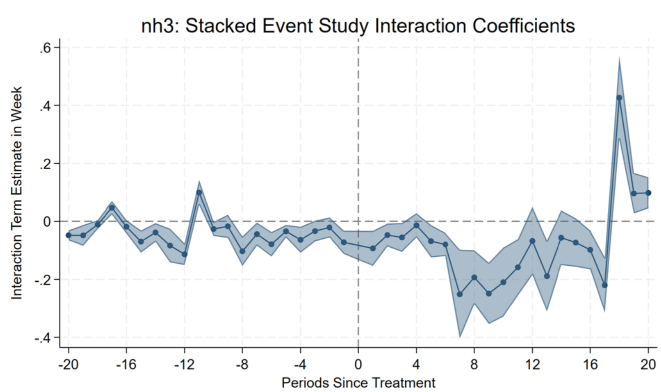


Figure 8: NH_3 Event Study Plot

0.1 Intersections Plots by City

Treated intersections at Green Light cities are plotted with red dots whereas control intersections are plotted with a light blue center dot. All intersections include a 1-mile radius surrounding them.

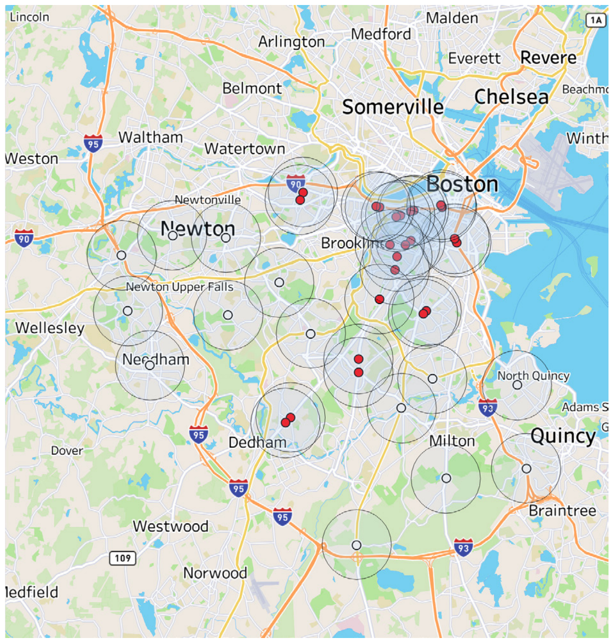


Figure 9: Boston Treated Intersections. Note for many of Boston's treated intersections, treatment simultaneously began with a nearby intersection. Therefore, we can see on the map that there are many pairs of treated intersections with very close geographic proximity to one another.

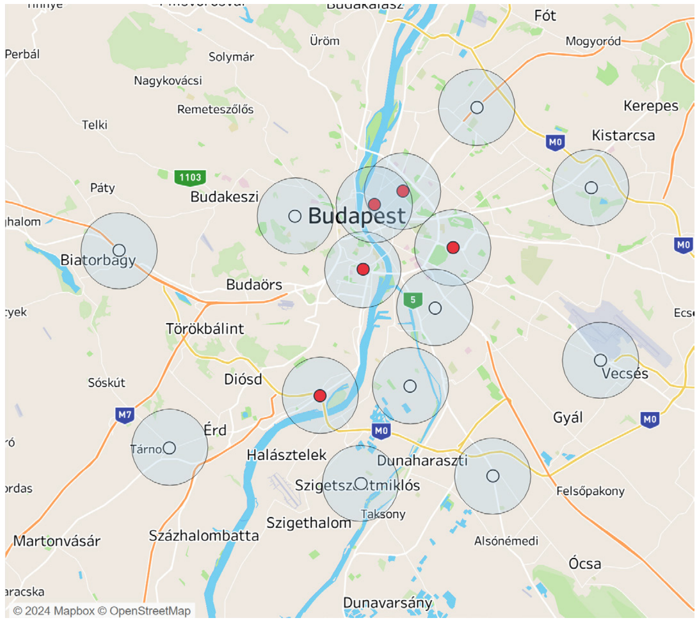


Figure 10: Budapest Treated Intersections

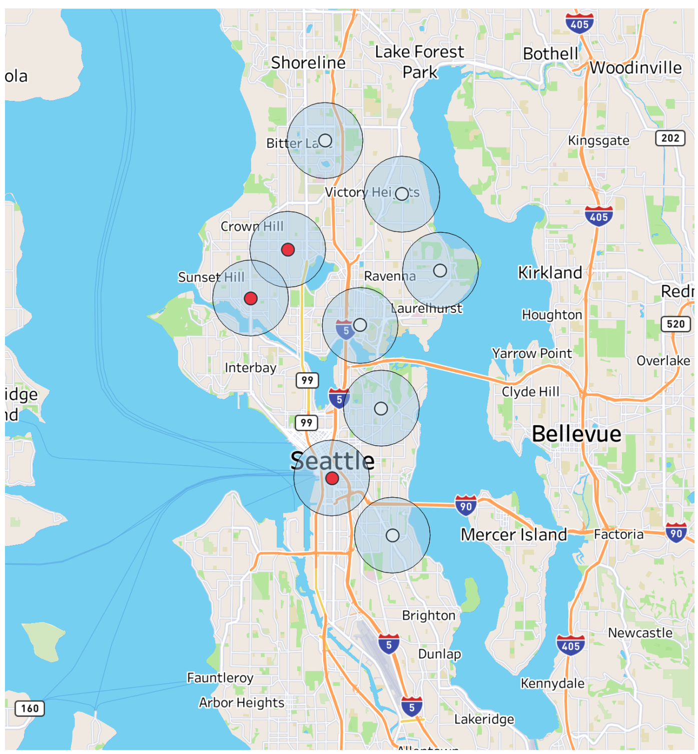


Figure 11: Seattle Treated Intersections

Motherhood and the Market: Paid Maternity Leave Policy Impacts on Women's Wages and Labor Force Participation *

Mackenzie Bivin

Abstract

Gender equalities within the labor market can largely be attributed to the decrease in women's labor force participation and wages upon motherhood. Paid maternity leave is one possible method to ease the transition to having a child while still allowing the mother to participate within the labor market, making it a possible aid in decreasing inequality. By leveraging heterogeneity in maternity leave policies that exist across states, this paper estimates impacts of paid maternity leave (PML) policies on the labor force outcomes of women with children from ages zero to six. I find that PML policies increase labor force participation for mothers the first two years after a child's birth, causing women, particularly those with no college education, to return to work earlier. Additionally, I find an increase in overall wages for women with access to PML due to an increased presence of people within the labor force but find no increase for wages conditional on a woman being in the labor force. This result likely occurs because the people entering the labor force earlier tend to be at the bottom of the income distribution, which is validated by observing college education levels. This decreases the estimated impact of PML policy.

*I would like to thank Nathaniel Hendren and Ishaana Talesara for their advice throughout this project.

Introduction

Inequalities exist across genders within the labor market. Many of these inequalities can be attributed to the child penalty: although career trajectories for women and men start off the same, after a woman has her first child, the earnings and career growth for each sex deviates with women falling behind. The gap widens within the first years of motherhood before leveling off to about a 25% difference in employment (Child Penalty Atlas) and a 31% difference in wages (Kleven et al.). Within the United States, it is estimated that 84% of the gender gap can be explained via the child penalty (Child Penalty Atlas). As such, to determine how to decrease the impact of the child penalty, one can examine specific policies to determine the impact on labor force outcomes of mothers in attempt to shrink the child penalty gap.

As this gap in labor market outcomes widens in the transition to parenthood, I turn to policies that may come into effect at this time. One such policy is maternity leave. Maternity leave provides workers time off of work to recover and take care of and connect with their newborn child. While paid maternity leave policies are widely popular in other top competitive countries (Earle et al.), the United States only provides unpaid job protected leave through the Family Medical and Leave Act of 1993 (FMLA). Any further policies relating to maternity leave are provided by the states. Heterogeneity exists across states in maternity leave options, with some providing only FMLA protections, some expanding FMLA protections to small businesses, and others providing their own PML programs. In fact, PML policies are on the rise with 7 states enacting laws to implement such polices in the last five years. Income increases from PML policies could strengthen people's labor force attachments and provide them with the ability to work and earn more. High take-up of PML and increases in labor force attachment internationally motivates me to investigate the impacts of PML policies on women's labor force outcomes in the United States.

In this paper, I leverage existing differences in PML policies across states to estimate the impacts of having access to such policies on various labor force related outcomes. I rely on a difference-in-difference estimation, comparing take-up, labor force participation, and wages and log wage of women with children birthed pre and post state-offered PML with women with young children in states that do not have any maternity leave provided by the state outside of the FMLA. Specifically, I look at the impact of maternity leave take-up the year the mother has a child. I also investigate the consequences on labor force outcomes in the long term, looking at the influence it plays on women with children from ages zero to six. This paper serves as an avenue to investigate whether PML policies reduce the negative effects of having a child on women's labor force outcomes.

I find that maternity leave take-up increases when women have access to PML policies compared to when they do not. This increase is four percentage points for women who worked more than 13 weeks the previous year and have a child younger than one.

Labor force participation increases for mothers with children ages zero to two years who have access to PML programs compared to mothers of children of the same age range who do not. Specifically, labor force participation increases by 2.3 percentage points for mothers of children less than one year old, 2.6 percentage points for mothers of one-year old children, and 4.7 percentage points for mothers of two-year old children. For mothers of children three years or older (with the exception of mothers of five-year old children), having access to PML programs has no significant impact on labor force participation. This suggests that PML policies increase labor force attachment, and women who have access to PML programs at the time they have a child are more likely either to remain at work or return to work when their child is one to two years old, instead of returning to work when the child is older.

Regressing wages on treatment shows that PML policies increase wages of women with children ages zero to four. These increases are estimated to be about $25-45, but with roughly $20 confidence intervals, smaller impacts cannot be ruled out. When regressing log wages on treatment, and therefore only looking at the impact of PML policies on wages conditional on women being in the labor force, I find only an increase in log wages for women with children who are less than one year old. This result could be due to the wage increase provided to women on maternity leave by PML programs. These differences in regressions suggest that increases in wage stems from PML programs providing women with a stronger attachment to the labor force. Additionally, this lack of change in log wage can be partially attributed to labor force participation increasing via PML programs for people with no college education who typically make less money, increasing the presence of lower wages and bringing down any possible wage increase achieved by more educated women with higher wage potential and/or earnings.

In this paper, I first provide a background on how PML policies have evolved in America and existing literature on these policies. Then I describe my data and empirical framework for measuring long term impacts of PML policies on women's labor force outcomes., I present the results of my analysis, draw conclusions, and propose potential areas for further research.

Background

The United States is the only nation of globally competitive economies not to offer paid maternity leave and is one of two not to offer paid paternity leave (Earle et al. 2011). Instead, the United States federal government offers job-protected leave for up to twelve weeks within a year-long period via the FMLA. However, the FMLA restricts who has access to this protected leave when having a child. To qualify for the FMLA, one must have worked 1250 hours in the past year at a firm with 50 or more employees and live no more than 75 miles from the worksite. As such, people who work at smaller firms do not qualify for federally protected leave. These constraints lead to only about half of all workers being eligible under the FMLA. (Rossin-Slater et al. 2011) Nevertheless, the FMLA increased coverage and take-up with no statistically significant negative impacts on women's employment or wages. (Waldfogel 1999).

Beyond the FMLA, any protection for women in the workplace having children is provided by individual states. Eleven states either offer programs in which parents are provided paid leave for the birth or adoption of a child or have passed a paid leave law with a guaranteed future implementation date. However, only three of these states have afforded their residents a PML program for more than the past ten years, which will be my focus as I estimate long term impacts.

California has provided paid leave through the California Paid Family Leave (PFL) Program since 2004. This program offers benefit payments for up to 8 weeks. These payments are 60 to 70 percent of weekly wages earned within the five to 18 month period before a mother's claim start date, up to a maximum of $1,620 per week through 2024. To qualify for coverage, a mother must have birthed a child into her family within the past 12 months, paid into State Disability Insurance in the past five to 18 months, and not taken the eight weeks of leave in the past year.

New Jersey implemented its New Jersey Family Leave Insurance Program in 2009. The program provides up to 66 percent of weekly wages paid over twelve weeks of leave (an increase from the original six weeks of leave prior to 2020). To qualify, one must earn above a threshold for wages ($14,200) the previous year. Funding for this program comes from payroll reductions. This program applies to parents who not only birthed the child but adopted the child into their family as well.

Rhode Island introduced Temporary Caregiver Insurance in 2013. Under this policy, workers are provided with up to 6 weeks of paid leave during a year-long period. Funding comes from payroll deductions. The payment, ranging from $130 to $1,043, is equal to a person's highest quarter (3 months) of earnings during the previous year multiplied by 4.62 percent. To qualify, a woman must have made a baseline amount of money within the past year (in 2024 this number was $ 16,800).

The policies within California, New Jersey, and Rhode Island only provide maternity leave and do not provide job protection. However, these policies can be used alongside the FMLA or other state job protection policies so that a woman can have paid leave while also having a job to return to.

Previous research has found that California's PFL program increased take-up of maternity leave and length-of maternity leave, with the average maternity leave length shifting from three weeks to six or seven weeks (Rossin-Slater et. al 2011).

Various studies have been conducted to investigate the relationship between maternity leave policies and labor force participation. Within Great Britian, increased maternity rights were associated with more woman returning to work earlier (Gregg et al. 2003) and with increased labor force attachment for women (Burgess et al. 2008). Baum and Ruhm found that California's PFL program led to higher work and employment probabilities for mothers nine to twelve months after giving birth, increasing job continuity among those with relatively weak labor force attachments. In California, after the implementation of the state's PFL program, labor force participation increased for young women relative to young men and older women (Das and Polachek 2014). In a study on long term impacts of PML policies on labor force participation, labor force detachment for mothers decreased significantly for the first five years of motherhood (Jones and Wilcher 2020).

In terms of equality, the results are mixed. Rossin-Slater et al. recorded larger maternity leave take-up growth from less advantaged groups, specifically those who were less educated, unmarried, and nonwhite, while Jones and Wilcher found that long term improvements to labor force participation were driven by those with higher education and greater social privilege.

Research that investigates the impact of PML policies on other labor force components tends to be short-term focused and less clear. Stanczyk found following the first year of motherhood, women with access to a PML program had 4.1 percent higher incomes and a 10.2 percent lower risk of poverty than women who did not have access to such programs. Baum and Ruhm estimated an increased number of hours worked but did not find a statistically significant impact on wages. Rossin-Slater et al. found an increase in hours worked for mothers of one to three year olds but did not provide information on wages. Thus, there is merit in the investigation of PML programs and the impact on women's wages.

Data

I obtained my data from the Current Population Survey (CPS) via the Integrated Public Use Microdata Series (IMPUS). The CPS collects monthly data from approximately 60,000 households, representing about 100,000 people within the United States. These surveys, often conducted by field workers or over the phone, collect a wide range of data including demographics, labor force participation, industry information, and household information. The households are randomly selected after being stratified by regions. Households are surveyed for four consecutive months, not surveyed for the next eight months, and then interviewed again for another four months before ending their participation in the CPS.

I included data over a 20-year span, from 1999-2019. I limited the data to women aged between 22 and 50 as this span gave me a significant amount of data for women who would typically be having children (through birth or adoption) and participating in the labor force as well. The control data consists of women from states in which, over the specified period, the state provides no additional maternity leave policies to supplement the FMLA. Accordingly, the women in these states only have access to unpaid protected leave. These control group states are Ohio, South Carolina, Texas, and Utah. The state is provided for the year the survey is taken, so I am unable to identify and account for women who moved into the state after having a baby and therefore participated in a different maternity leave policy.

From this data, I was able to determine if a woman was on maternity leave by using the data variable WHYABSNT: a variable that details why a person had not gone to work within the week prior to the survey being conducted. When attempting to understand maternity leave take-up, it is important to look at both the impact of PML on maternity leave for all women as well as women who qualify for PML. In all three states that offer PML, women must have worked or earned a baseline level the previous year. Therefore, use of the variable WKSWORK2 is appropriate, which details a range of the number of weeks worked in the previous year. Although this will not account for all possibilities (for example, the eligibility of a woman who had a baby in the previous year would be determined by two years prior), this approach still offers a better estimate of the impacts of the policies on women who would qualify for PML. However, this data is only available through the CPS's Annual Social and Economic Supplement (ASEC), which is just taken annually in March. Accordingly, the dataset for this regression will be much smaller compared to all women with a child less than one year old. For wages, I utilized the variable EARNWEEK2, which represented the weekly earnings of the surveyed person. Only about 25 percent of total people surveyed provided their weekly earnings.

I was able to identify the ages of children through the variable YNGCH, which indicates the age of the youngest child in the household. The data, however, did not provide the ages of all of the children in a household. This method of looking at child age only allowed me to estimate impacts of labor force outcomes of the mother based on the age of the youngest child. Although this approach does not lead to an understanding of how many children a mother has and how that may impact her ability to participate in the labor force, valuable insights into the impact on labor force nevertheless can be gained. Specifically, because the youngest child typically requires the most supervision and care, this child's presence is likely to dominate the impact of the mother's labor force outcomes.

To create the treatment variable, I utilized the survey year, the age of the youngest child, and the state of the mother's residence. If the mother was in a state in which a PML program was enacted, I then performed a calculation to see if the youngest child was born in a year after the PML policy was effective by subtracting the age of the child from the year the data was collected to determine if the child was born when the PML program was in effect. If so, then the treatment was in effect as well represented by a 1 within the dataset.

Empirical Method

I utilize different policies among states and different implementation periods to estimate the impact of PML policies on various labor market outcomes. Specifically, this is conducted through a difference-in-difference method. I compare an outcome variable (maternity leave take-up, labor force participation, or wages) between those who have access to PML programs and those who only have access to the FMLA. As I want to understand the impact of PML on certain labor market outcomes over the long term, I run individual regressions for women with children of each individual age. The only exception to this approach occurs when evaluating maternity leave take-up. As take-up is traditionally utilized by the mother close to the birth of the child, the only regression time I will be running for this variable of interest is for children under the age of one.

The regression is as follows:

$$y_{it} = \beta_0 + \beta_1(\text{Treatment}_{it}) + +\beta_2 X_i + \gamma_t + \delta_i + \epsilon_{it}$$

This regression will be run independently for each age a from 0 to 6. Here, y_{it} will represent either maternity leave take-up, labor force participation, or wage, dependent on the specific regression being run. The treated variable is a binary variable that represents if mother i who has a child of age a in year y had access to paid maternity leave when the child was born (meaning did mother i have the child in California after 2004, New Jersey after 2009, or Rhode Island after 2013). X_{it} is a vector of control variables, including mother's age, education level, marital status, and race. γ_t represents time fixed effects, and δ_i represents state fixed effects. β_1 is our coefficient of interest.

Results

Maternity Leave Take-Up

First, I examine maternity leave take-up. In table 1, I investigate two regressions. Regression one considers the impact of having access to PML when giving birth on the likelihood of a woman being on maternity leave if she has a child that is less than one years old. Here, I see a one percentage point increase on the likelihood of a woman being on maternity leave when she has access to PML. As the constant coefficient of maternity leave take-up is at about 8 percent, this is a substantial growth caused by access to a PML program. However, in looking at the confidence intervals, I cannot rule out no effect of PML on maternity leave take-up.

Regression two is similar to regression one except it only includes women who qualify for maternity leave. Regression two estimates a significant but much higher increase of 4.75 percentage points via the treatment. As this regression only includes the ASEC data, the sample is much smaller, and the results noisier. The confidence interval ranges from a one percentage point increase to an eight percentage point increase. This result strongly suggests that there is an increase in maternity leave take-up among mothers with newborns that qualify.

Table 1: Difference-in-Difference Regressions on Maternity Leave Take-up

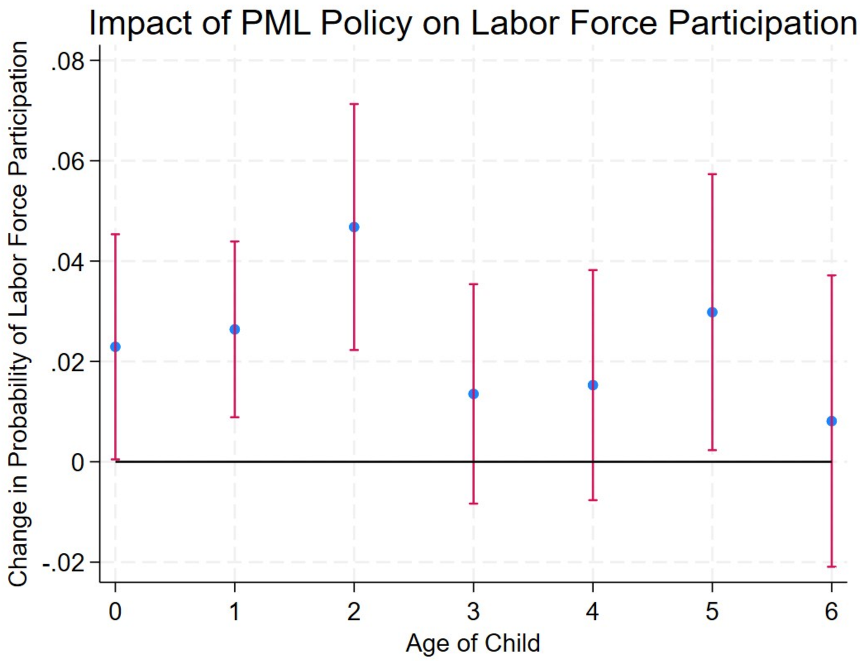
	(1)	(2)
Treatment	0.0103* (.00553)	0.0475** (.01782)
Confidence Intervals	[-.0012, .0218]	[.0103, .0847]
Year fixed effects?	Yes	Yes
State Fixed Effects?	Yes	Yes
Worked in Last Year?		Yes
R^2	0.02	0.02
Number of observation	87032	6320

In regressions (1)-(2), a mother being on maternity leave is being regressed on treatment, state and fixed effects, as well as control variables including education, age, marital status, and race.
*** = p < 0.01, ** = p < 0.05, * = p < 0.1

From this result, I draw the conclusion that women with newborns in states that offer PML, who are in the work force the year prior to having a child and thus qualify for PML, are more likely to go on maternity leave than women who did not have access to PML but also are in the work force a year prior to having a child. The same cannot be said for all women with newborns.

Labor Force Participation

As shown in figure one and table two, by regressing labor force participation on treatment, state and year fixed effects, and control variables, I found increases in labor force participation for mothers with state provided access to PML compared to mothers who did not have state provided access to PML when mothers had young children, specifically with children ranging from ages zero to two. When women have a newborn baby, I find a 2.3 percentage point increase in women's labor force participation in the year she gives birth where state provided PML is offered compared to when she lives in a state with no state provided PML. This is significant, albeit with a large confidence interval that cannot rule out an increase of only 0.04 percentage points. Consequently, I can confidently state there is a nonnegative impact of PML programs on labor force participation of women with children of age zero, though I am less confident in the actual quantified impact. I see a 2.6 percentage point increase for women who have a child of age one and a 4.7 percentage point increase for women who have a child age two. The Child Penalty Atlas provides that the United States child penalty employment gap is 25 percent, and so PML policies can decrease this gap by roughly 8 percent (2.3 percentage points) when women have a child of age one and roughly 19 percent (4.7 percentage points) when women have a child of age 2. Both of these values are statistically significant. After a child reaches the age of three, I can no longer say that there is an impact of PML policy on women's labor force participation, with the exception of children who are five years old. It is important to note that as the child's age increases, there are less instances within the data of mothers who have children that same age that were born when PML policy was enacted in the state (for example in California in 2011, women with one year olds are treated but women with four year olds are not as they gave birth before the policy was put in place). So, the regressions become noisier.



The coefficients on labor force participation for mothers who had state-provided access to PML compared to mothers who did not have state-provided access to PML are plotted. Each coefficient is taken from a child-age specific regression that includes controls as well as state and year fixed effects. The bars represent the 95% confidence intervals.

Figure 1: Impact of PML Policy on Labor Force Participation

Table 2: Difference in Differences Regressions on Labor Force Participation

	age 0	age 1	age 2	age 3	age 4	age 5	age 6
Treated	.02292**	.02639***	.04679***	.01353	0.01528	.02982**	.00811
	(.0108)	(.0084)	(.0117)	(.0105)	(.0110)	(.0131)	(.0139)
R^2	0.07	0.07	0.06	0.06	0.06	0.07	0.06
Number of observations	87032	97257	87960	77508	69356	64744	60707

In the above regressions, labor force participation is regressed on the treatment, state and fixed effects, as well as control variables including education, age, marital status, and race. Each regression is ran on mothers that have a youngest child of the specified age.

*** = significant at the 1% level, ** = significant at the 5% level, * = significant at the 10% level

These results suggest that paid maternity leave policies increase labor force attachment by shifting women who would return to the labor force once their child is slightly older (three to four years old) to now return to the labor force slightly earlier or stay in the labor force when they have children. However, as the child gets older, a PML program does not appear to impact women's probability of labor force participation.

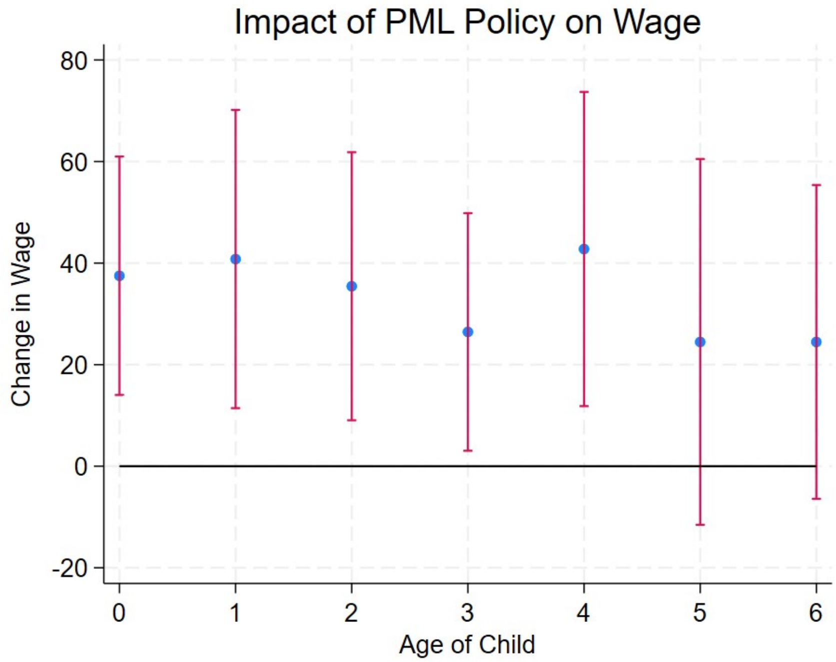
Wages and Log Wages

I regressed wages on PML policy exposure, fixed state and year effects, and control variables. The results are detailed in table 3 and graph 2. Access to PML for a woman when she has a child leads to an average increase of $37.50 weekly earnings when the child is age 0, $40.79 weekly earnings when the child is age 1, $35.44 weekly earnings when the child is age 2, $26.44 weekly earnings when the child is age 3, and $42.77 weekly earnings when the child is age 4 compared to women with children at these ages who did not benefit from a PML program when they gave birth to their children. I do not find any impact of exposure to PML programs for mothers with children five or older. These coefficients have quite large confidence intervals. For example, for the impact for mothers with children at age one, the confidence interval spans over a $60 spread. As such, it is possible that with this wide confidence interval, the impact of PML policies on wage could plausibly be much smaller, in the $5-10 range.

In addition to looking at the impact of PML on wages for all women within groups that report a wage, I focus specifically on the women who are within the labor force by regressing log wages on exposure to PML programs with only the people who report positive weekly earnings. This further decreases the size of the dataset. However, by looking at log wage, I can estimate the impact of PML policies on percent changes in wages for women in the labor force. The results can be seen in figure 3 and table 4. Here, the only year in which log wage increase for women who had access to PML programs when they gave birth compared to women who did not is when the child is age 0. Access to PML programs increases a woman in the labor force's weekly earnings by 4.97 percent. The confidence interval ranges from 0.7 percent increase to a 9.2 percent increase, which is a large range. This increase could partially stem from the increase in wage women receive if they take PML, meaning they receive a portion of their paycheck while on leave. This will impact both women who go on maternity leave and remain in the labor force due to PML as well as women who were already in the labor force and will take maternity leave regardless of whether it will be paid.

The discrepancy between results of the impact of PML policy on wages and log wages suggests that the impact is related to the women dropping out of the labor force as the wage regression includes people not in the labor force while the log wage regression does not. This difference might be explained by the impact of PML policy on labor force participation. There was an increase in labor force participation for the first two years of motherhood for women with access to PML programs compared to those without access. Therefore, this shift of increased employment will impact the wage regressions because more women benefitting from PML programs will be earning non-zero amounts than women without, and this will shift the treatment effects on wages upward. However, once I condition the regression only to women within the labor force, there is no effect, meaning the increase captured in the regression of wages comes from the increase of women in the labor force.

Another area I can investigate regarding this difference is the composition of those who are increasingly entering or remaining in the labor force after having access to a PML program upon giving birth. A specific concern is that the people who have increased labor force retention with access to a PML program are those with lower wages than average. The increase in lower wage jobs will shift the composition of the market to a larger number of people with lower wages, leading to an underestimation of the impact of PML policies on log wages. I was able to explore this possibility by revisiting labor force participation and running separate regressions based on education attainment.



The coefficients on wages for mothers who had state-provided access to PML compared to mothers who did not have state-provided access to PML are plotted. Each coefficient is taken from a child-age specific regression that includes controls as well as state and year fixed effects. The bars represent the 95% confidence intervals.

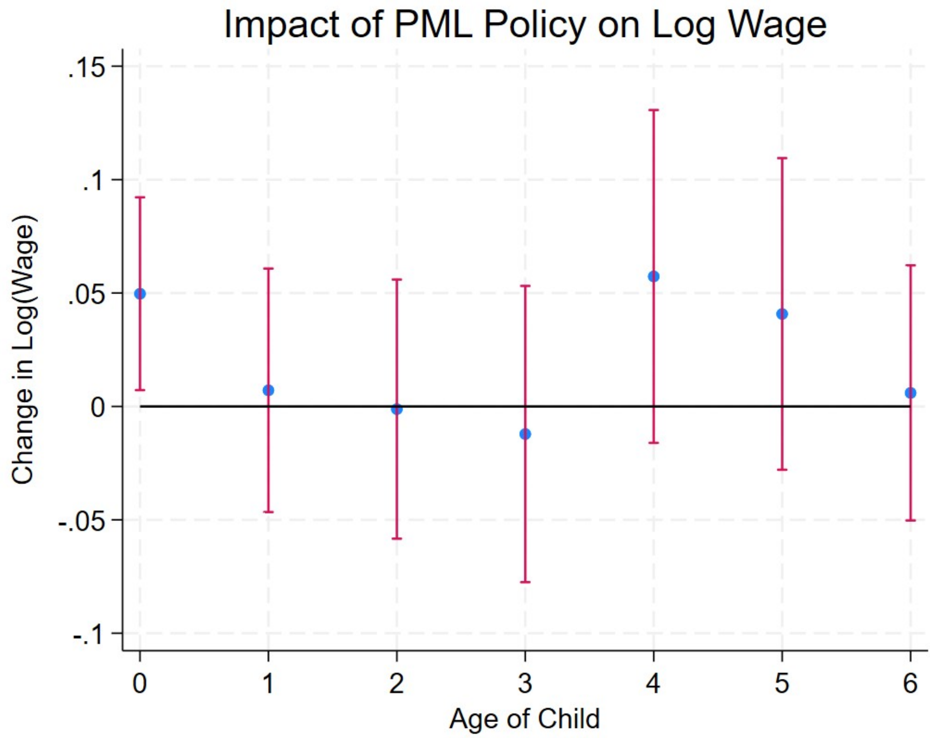
Figure 2: Impact of PML Policy on Wages

Table 3: Difference in Differences Regressions on Wages

	age 0	age 1	age 2	age 3	age 4	age 5	age 6
Treated	37.501***	40.794***	35.439**	26.439**	42.767***	24.481	24.473
	(11.25)	(14.08)	(12.65)	(11.22)	(14.83)	(17.27)	(14.81)
R^2	0.14	0.13	0.13	0.12	0.12	0.11	0.12
Number of observations	19936	23200	21498	19486	17698	16719	15884

In the above regressions, wages are regressed on the treatment, state and fixed effects, as well as control variables including education, age, marital status, and race. Each regression is ran on mothers that have a youngest child of the specified age who report their weekly earnings.

*** = $p < 0.01$, ** = $p < 0.05$, * = $p < 0.1$



The coefficients on wages for mothers who had state-provided access to PML compared to mothers who did not have state-provided access to PML are plotted. Each coefficient is taken from a child-age specific regression that includes controls as well as state and year fixed effects. The bars represent the 95% confidence intervals.

Figure 3: Impact of PML Policy on Log Wages

Table 4: Difference in Differences Regressions on Labor Force Participation

	age 0	age 1	age 2	age 3	age 4	age 5	age 6
Treated	.0497**	.0071	-.0012	-.0122	.0573	.0408	.0060
	(.0204)	(.0257)	(0.0273)	(.0313)	(.0352)	(.0329)	(.0270)
R^2	0.31	0.29	0.29	0.26	0.25	0.24	0.25
Number of observations	9785	11696	11155	10137	9445	8980	8717

In the above regressions, log wages are regressed on the treatment, state and fixed effects, as well as control variables including education, age, marital status, and race. Each regression is ran on mothers that have a youngest child of the specified age who report weekly earnings above $0, meaning they are employed and currently working or on paid leave.
*** = p < 0.01, ** = p < 0.05, * = 0.1

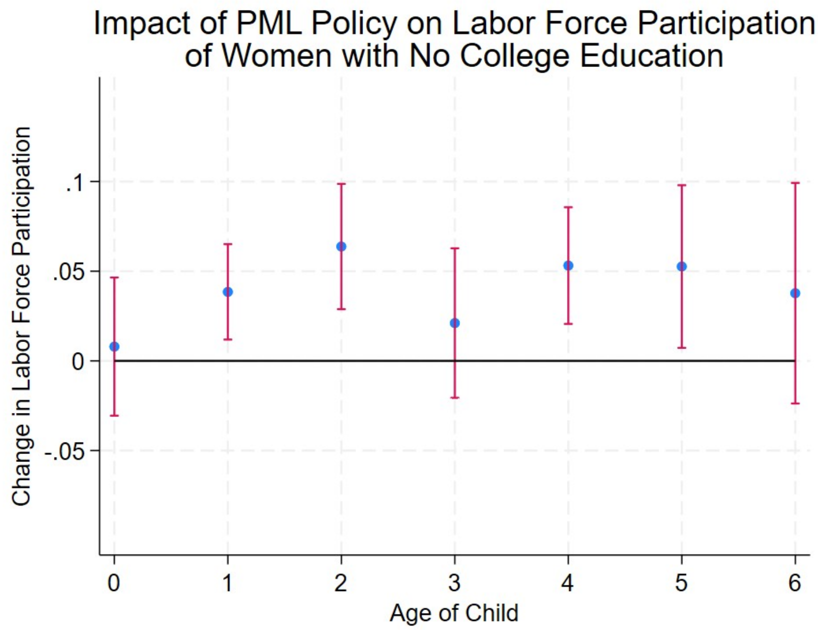
Educational attainment can be used as a proxy for potential wages. In table 3, I provide a breakdown on the mean and standard deviation for different levels of educational attainment. The higher the education attainment level, the higher the average income. I then use this proxy to contemplate how women of varying educational attainment levels are impacted differently by PML policies coming into effect, namely the composition of women in the work force and their potential incomes shift.

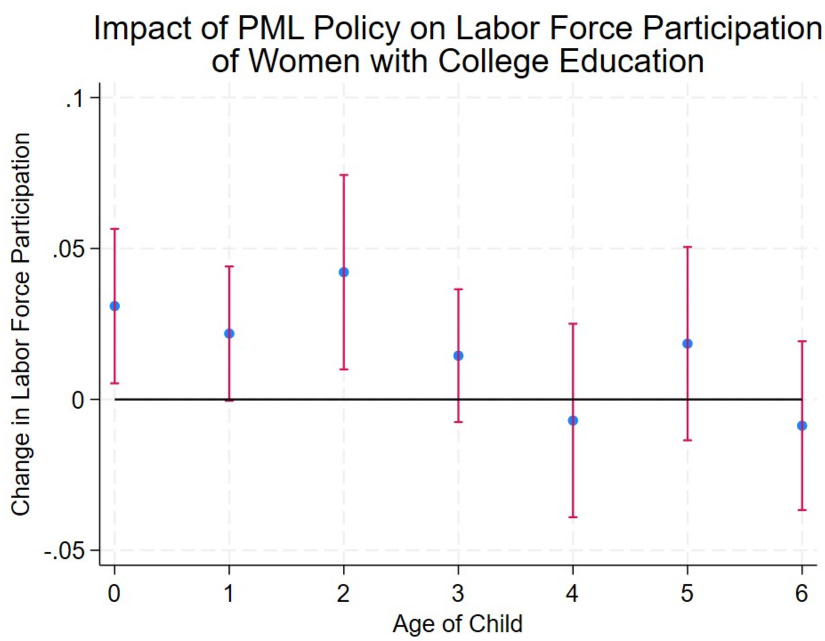
To understand how the labor force participation of women attaining different levels of education was impacted by PML policies, I ran two regressions of the same design as the previous methods but only include women with no college education in one and women with at least some college experience in the other. The results are shown in figures 4 and 5. Figure 5 shows that access to a PML program did not have a significant impact on the labor force participation of women with some level of college education with the exception of women with a two-year old child. On the other hand, in figure 4, there exists some positive significant impacts of access to PML programs on the labor market participation of women without any college education who have children ages zero to five, except for women with a three-year old child. Accordingly, this overall shift in labor force participation that occurs with the opportunity to participate in a PML program appears primarily with women of lower educational attainment. Thus, the women who increase their labor force participation as a result of access to a PML program are going to be disproportionally of lower educational attainment and lower potential income. As such, as the above theory suggested, the composition of people in the labor market will change to include more lower-waged people, which will shift the effect of the treatment on log wage down.

Table 5: Wages of Women Across Educational Attainment

	Less than HS	HS	Some College	Bachelor or Higher
Number of Mothers	6940	16876	21399	24700
Mean Wage	322.52	451.47	550.63	1022.343
Standard Deviation	167.25	284.97	362.66	657.88

This table looks at average wages of women broken up into different educational levels for women who have a child age six or younger. Women are divided into four categories based on their educational attainment: less than HS, which represents women who did not get their high school degree; HS, which represents women who received their high school diploma; Some College, which represents women who completed some college or received an associate degree; and Bachelor+, which represents women who received a bachelor degree or higher.





The coefficients on labor force participation for mothers with specific educational attainment who had state-provided access to PML compared to mothers with that same educational attainment who did not have state-provided access to PML are plotted. The top graph includes women with no college education, while the bottom graph includes women with college education. Each coefficient is taken from a child-age specific regression that includes controls as well as state and year fixed effects. The bars represent the 95% confidence intervals.

Figure 4: Impact of PML Policy on Labor Force Participation for Women Different Educational Attainment

Conclusion

In this paper, I utilize differences in maternity leave policies across states to estimate impacts of PML programs on maternity leave take-up, labor force participation, and wages. First, I discover a small increase in maternity leave take-up for women who have access to PML programs, but this result is only significant when my regression focuses on women who have worked within the past year and therefore qualify for maternity leave.

I see an increase in labor force participation for women who are covered under a PML policy when they have children zero to two years old. This signifies a shift in labor force attachment, in which women who have access to a PML program are more likely to either remain in the labor force when they have a baby or return earlier to work, specifically when the baby is one or two years old. As this same effect is not seen when the children reach three years of age and older, this signifies a shift in returning to the labor market when one's child is one or two years old instead of returning to the market when one's child is three or four. When breaking down the regressions by educational attainment as a proxy for income potential, it is revealed that the greatest impacts of PML policies on women's labor force participation are focused on women who did not receive any college education.

There is an increase in overall wages from access to PML programs for women with children from age zero to four. However, when looking at the impact of PML policies on log wage, I see only an impact on year zero, which is likely significant due to people being paid while on maternity leave. This difference across regressions can be explained by the theory that the change in wage is primarily driven by the shift from unemployment to employment in early years of motherhood. In addition, as women without college education are increasing their labor force participation under PML policies and they typically earn lower wages than college educated women, the corresponding shift in wage composition of the market likely leads to an underestimation of the impact of PML policies on log wages.

There exist further areas of research that would be beneficial to undertake. I would like to investigate more robustly into the impacts of PML policies on log wages, possibly separating by quantiles of wage. Additionally, I would like to go back and utilize maternity leave take-up as a first stage effect to understand the impact of paid maternity leave on labor force outcomes, not just the impact of the policy. Furthermore, I am interested in understanding the effect of the number of children a woman has on labor force outcomes, something I was unable to consider here due to data restrictions. Looking at how paid family leave policies impact both women and men and estimating movements of the child penalty could provide interesting insight into the benefits provided. Additionally, I would be fascinated to compare the benefits of PML policies to other policies aimed at decreasing gender inequality and evaluate the advantages and drawbacks of these policies to understand how PML policies compare in achieving gender equality while developing a more comprehensive idea of the value of large scale PML policy implementation.

References

Boyle Stanczyk, A., (2019). Does Paid Family Leave improve household economic security following birth? Evidence from California, 93(2), 262-304

Baum, C.L., & Ruhm, C.J. (2016). The effects of paid family leave in California on labor market outcomes. Journal of Policy Analysis and Management, (35)2, 333-356. https://doi.org/10.1002/pam.21894

Burgess, S., Gregg, P., Propper, C., & Washbrook, E. (2008). Maternity rights and mothers' return to work. Labour Economics, 15(2), 168-201.

California Employment Development Department. "Paid Family Leave." https://edd.ca.gov/en/disability/paid-family-leave/

Child Penalty Atlas. "Child Penalty in Employment." Accessed December 11, 2024. https://childpenaltyatlas.org/[2]3

Das, Tirthatanmoy and Solomom W. Polachek. "Unanticipated Effects of California's Paid Family Leave Program." IZA Discussion Paper No. 8023, March 2014

Earle A, Mokomane Z, Heymann J. International perspectives on work-family policies: lessons from the world's most competitive economies. Future Child. 2011 Fall;21(2):191-210. doi: 10.1353/foc.2011.0014. PMID: 22013634.

Gregg, P., Gutiérrez-Domènech, M., & Waldfogel, J. (2003). The Employment of Married Mothers in Great Britain: 1974-2000. Centre for Economic Performance, London School of Economics and Political Science.

IPUMS CPS. "Current Population Survey Data for Social, Economic and Health Research." Integrated Public Use Microdata Series, University of Minnesota. Accessed December 11, 2024. https://cps.ipums.org/cps/

Jones, Kelly and Britni Wilcher. "Reducing maternal labor market detachment: A role for paid family leave." Washington Center for Equitable Growth Working Paper, March 2020.

Kleven, Henrik J., Camille Landais, and Jakob E. Søgaard. "Child Penalties Across Countries: Evidence and Explanations." VoxEU.org, Centre for Economic Policy Research, December 10, 2020. https://cepr.org/voxeu/columns/child-penalties-across-countries-evidence-and-explanations.

New Jersey Department of Labor and Workforce Development. "Family Leave Insurance." https://www.nj.gov/labor/myleavebenefits/worker/fli/

Rhode Island Department of Labor and Training. "Temporary Disability / Caregiver Insurance." https://dlt.ri.gov/individuals/temporary-disability-caregiver-insurance

Rossin-Slater, M., Ruhm, C. J., & Waldfogel, J. (2011). The Effects of California's Paid Family Leave Program on Mothers' Leave-Taking and Subsequent Labor Market Outcomes. NBER Working Paper No. 17715. National Bureau of Economic Research.

Waldfogel, J. (1999). The Impact of the Family and Medical Leave Act. Journal of Policy Analysis and Management, 18(2), 281-302.

The Effect of Quotas on Wages: The Homegrown Player Rule in the Premier League*

Kate Ellison

Abstract

Quotas have long been used to increase representation of certain groups, especially in high-ranking positions. Relying on classic models of labor supply and demand, quotas may also increase wages for workers helped by the quota. In this paper, I look at whether a foreign player restriction in the English Premier League impacts player wages and transfer fees. I find that this quota has no effect on the average of domestic players in the Premier League, although it does increase domestic player wages at clubs bound by the constraint of the quota. The quota also decreases domestic player transfer fees.

Introduction

The question of how quotas impact worker employment and compensation is of key importance from a policy standpoint. Quotas are often touted as a tool to improve workplace diversity and place minorities in high-ranking positions, but more research is needed on whether they merely improve representation, or if they also improve outcomes for workers included in the quota. There is a large body of literature exploring how employment and wages respond to policy changes such as minimum wages, parental leave, anti-discrimination laws, and many others (Card and Krueger, 1994; Antecol et al., 2018; Agan and Starr, 2018). Much of this work exploits a difference-in-differences design to compare the effect before and after a policy intervention. However, research exploring the wage effects of quotas specifically is much sparser. Existing literature has found some evidence for a modest increase in wages for women and workers with disabilities whose positions are protected or created due to quotas (Bertrand et al., 2019; Szerman, 2022). However, the Premier League quota provides a somewhat different context, as it was designed to protect the declining representation of a former majority group, rather than catapult underrepresented groups into higher positions. Therefore, the Homegrown Player Rule provides a compelling example of how quotas can impact wages in this somewhat more unusual case.

In this paper, I ask the question of whether quotas increase worker wages. The example I explore is the 2010 adoption of the Homegrown Player Rule in the Premier League, England's

*I would like to thank Nina Roussille, Simon Jäger, Sara Ellison, Glenn Ellison, Lisa Ho, and Viola Corradini for their advice throughout this project. Additionally, I am very grateful to Jori Veng Pinje, not only for his generous provision of the wage data, but also for his helpful guidance. Finally, I would like to thank Neha Pant for her help in data scraping and Mackenzie Bivin for her help in developing my intuition and for her moral support throughout.

top-tier professional soccer league, following the example of Kleven et al. (2013), which exploits data availability in professional soccer to answer more broadly applicable economic questions.

Implemented the start of the 2010-2011 season, the Homegrown Player Rule requires all Premier League clubs to register a minimum of eight homegrown players. If they cannot register eight homegrown players, they are penalized one spot on their roster per homegrown player they are short. Considering the twenty-five person roster, this amounts to a restriction of a maximum of seventeen non-homegrown players per roster. In order for a player to qualify as homegrown, they must have been trained at an English or Welsh soccer club for at least three years between the ages of fifteen and twenty-one. The Homegrown Player Rule places no restrictions based on nationality. Henceforth, however, I will be considering the impact of this rule on domestic as opposed to homegrown players for reasons outlined in my Empirical Methods section.

Player transactions in professional soccer occur in two different forms. First, clubs pay wages to their contracted players. When a player signs for a club, they are usually given a contract that lasts for a few years, paying a certain fixed wage across that period. Players will then re-negotiate new contracts with their club before their previous contract runs out, with new contracts often leading to higher wages. Second, two clubs carry out transactions in order for a player to move from one club to the other. This typically means the buying club will pay a negotiated fee to the selling club, though it is not uncommon for free transfers to occur, meaning no transfer fee is paid.

As my primary interest, I use a triple differences (DDD) empirical strategy to analyze the impact of the Homegrown Player Rule on domestic players' wages. Finding a null result on overall wages for domestic players, I then consider potentially differential impacts for clubs for which the Homegrown Player Rule was binding versus nonbinding, again using a DDD design. I find that the Homegrown Player Rule increased wages for domestic players at clubs constrained by the Homegrown Player Rule quota. To supplement this analysis, I again use a DDD design to look at the impact of this rule on player transfer fees.

I will first build a simple theoretical framework to model the potential effects of a quota like the Homegrown Player Rule. Then, I describe my data and empirical methods, and discuss the necessary assumptions. Finally, I present the results of my analysis, draw conclusions, and propose potential areas for further research.

Theoretical Framework

In order to model the effects of a quota, I propose a simple framework to demonstrate how foreign and domestic soccer players' wages and transfer fees could be affected by a restriction on the number of foreign players. For simplicity, I assume that foreign and domestic players have the same labor supply, that is: $L_{s,dom} = L_{s,for} = L_s$. Then, I suppose that a team has demand $L_{d,dom}$ for domestic players and $L_{d,for}$ for foreign players. I also assume that, before the Homegrown Player Rule, the team is constrained by its roster limit, meaning it would prefer to employ more than twenty-five total foreign and domestic players, but it cannot (though I later relax this assumption). Furthermore, the team will be constrained by the Homegrown Player Rule, meaning it currently employs more foreign players than the quota

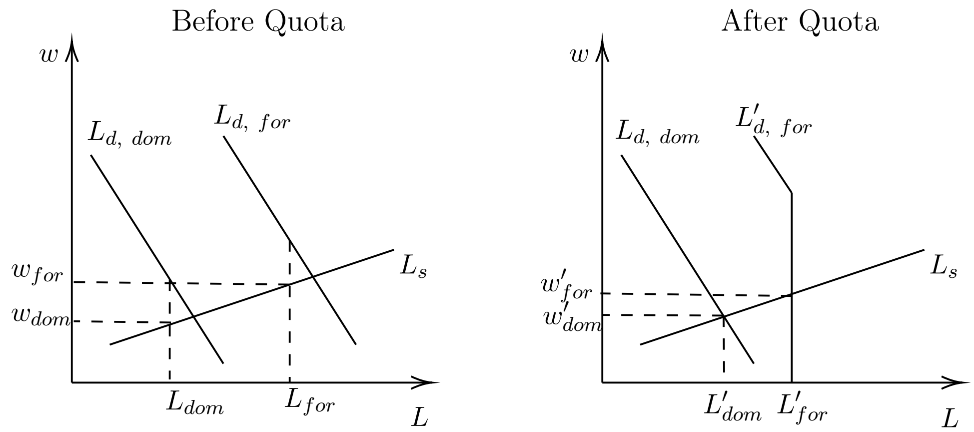


Figure 1: Demand for Foreign and Domestic Workers after a Quota

will allow. Then, its situation is modeled on the left of Figure 1.

In this case, the team chooses the optimal number of foreign and domestic players by equating the marginal benefit of hiring one more foreign or domestic player. However, the team is unable to hire up until the marginal benefit is zero, meaning they pay below the marginal product of labor for both foreign and domestic labor. Wages are instead determined by the labor supply curve, as indicated in Figure 1.

After the Homegrown Player Rule, $L_{d,for}$ becomes completely inelastic at the limit of seventeen foreign players. The team hires at $L'_{for} = 17 < L_{for}$. Because the team was previously constrained in the number of domestic players, it can and will want to now hire more, so $L'_{dom} > L_{dom}$. In Figure 1, the team is shown hiring up to the intersection of the domestic labor supply and demand curves, though this may not be the case (the team would not hire beyond that intersection). As we see in Figure 1, the model predicts that $\Delta w = w_{dom} - w_{for}$ will increase after the quota, meaning relative wages of domestic players will increase.

If we relax the assumption that the team's L_{dom} is constrained pre-quota, then even if w_{dom} is unchanged, Δw will increase due to the decrease in w_{for}. If the team's L_{for} is not constrained pre-quota, then we have the case where the quota is not binding. While this simple theoretical model predicts no effect on unconstrained teams because they are already hiring at their optimum, there may nevertheless be some effects. For example, their best domestic players might be tempted away by higher wages at previously constrained clubs, leading them to substitute towards the now-in-lower-demand foreign players.

To consider the impact of the quota on transfer fees, we assume that teams will only pay transfer fees for players whose marginal product of labor is greater than the wages they demand. Then, transfer fees are determined by some function of the player's marginal product to the two teams, along with the teams' bargaining power. Under these assumptions, average transfer fees for foreign players will weakly increase as a result of the quota because the average difference between foreign players' marginal products and wages demanded increases. On the other hand, transfer fees for domestic players will weakly decrease because the average difference between their marginal products and wages will decrease.

Both of these model's predictions are borne out in the data, as shown below. Domestic players' transfer fees decrease across the board in the Premier League after the Homegrown Player Rule's implementation. Domestic wages do not significantly increase overall, but wages for domestic players do increase at clubs where the Homegrown Player Rule quota is binding.

Data

In order to analyze the empirical effect of the Homegrown Player Rule, I use three separate datasets. Though professional soccer players' wages are often discussed in the media, they are not required to be published by clubs, and thus it is difficult to find reliable wage data. In order to observe impact on player wages, I obtained a dataset taken from the video game Football Manager.[1] Football Manager wages are compiled from a variety of sources, but are not provided directly from clubs or players, so they should be seen as estimates. However, the accuracy of Football Manager data has proven to be relatively high when validated against external sources (Veng Pinje, 2013).

The wages dataset contains data from the 2005-2006 season to the 2011-2012 season in the top five European leagues (Premier League in England, La Liga in Spain, Ligue 1 in France, Serie A in Italy, and Bundesliga in Germany). These five leagues are widely considered to be the highest-level leagues in the world. Each observation in the wage dataset corresponds to a player-season, containing data for each player in one of the top five leagues in a particular season. The variables included are player name and demographics, club and league details, and wages. For each player, gross weekly wages are given in Pounds sterling, including any guaranteed bonuses in players' contracts. It should be noted, however, that conditional bonuses are also common in professional soccer contracts, although those are not included in the Football Manager dataset. I do not have a separate dataset on player bonuses, so this is a potential area for bias in the wage data.

The player demographics include nationality, and some players have multiple nationalities. For those players with multiple nationalities, I consider them to be domestic players in any country corresponding with one of their nationalities, although they would almost certainly not qualify as homegrown in all of those countries. Additionally, each player has a unique ID number within the dataset, allowing me to track players throughout the time period of 2005-2012. Summary statistics on wages in the Premier League and the other top four European leagues are given in Table 1.

In contrast to wages, transfer fees are usually published by clubs. I was able to download a dataset on transfer fees which had been scraped from Transfermarkt.us, a website that tracks many different variables related to soccer players, including their transfer fees. I was unable to find specifics on how Transfermarkt determines and verifies transfer fees, although it is widely regarded as a trusted source for soccer data.

This dataset contains data on all transfers to and from teams in any of the top five European leagues from 1992-2022, though I restricted my analysis to 2005-2015 in order to reduce the impact of other confounding events, such as Brexit. The transfers dataset contains one observation per transfer event, meaning a transfer of a player between two clubs. For each transfer, the dataset records the player's name, both clubs' names, the

[1] I am very grateful to Jori Veng Pinje for providing me with these data.

direction of the transfer (in vs out), the type of transfer (loan vs permanent move), some player characteristics, and the transfer fee paid by the buying club. For permanent transfers, the transfer fee is typically published by at least one of the involved clubs, though if both clubs agree, the fee may be undisclosed. Reasons for not disclosing a transfer fee could include concern about fan reaction, jealousy from other players, and specific club preferences, all of which could potentially bias the reported transfer fees. Overall, around 15% of permanent transfers' fees are unreported.

However, the transfer dataset does not include data on player nationalities. In order to compare transfer fees of foreign versus domestic players, I needed player nationalities. Because the wages dataset only covers the period from 2005-2012, not all players in the transfer dataset are covered by the wage dataset, so I could not use nationality data from the wage dataset. I was able to scrape nationality data from FBRef.com, which lists the nationality of all top professional players, male and female, from the 1950s onwards. This is a simple cross-sectional dataset with one observation per player, and each observation listing just the player's name and nationality. According to FBRef, this dataset was compiled from players' FIFA-registered nationalities. FIFA is the main governing body for global football. As a result, each player is listed with only one nationality, which is thought of as their "main" nationality.

To link these datasets, I match player names from the nationality dataset with observations in the transfer dataset. Many players show up multiple times in the transfer dataset, representing multiple transfers so the player's nationality was applied to all observations. One complication in merging the datasets is the non-uniqueness of names in the nationality dataset. For each duplicate name, if all duplicates have the same nationality, I keep the name in the dataset. If duplicate names have different nationalities, then I drop them, because I will not be able to later make the distinction of whether that player is domestic or not in their league.

In total, around 3% of players in the nationality dataset were removed because of other players of their same name from different countries. This figure seems remarkably high, but it is largely driven by many soccer players' preferences to go by a single name as opposed to a first and last name. It should also be noted that the majority of the players in the nationality dataset never played in the top five European leagues and would not be included in final analysis regardless. When merging nationality and transfer data, of the roughly 100,000 transfers in the transfer dataset, 18% of them were unable to be matched with a player in the nationality dataset, largely driven by small differences in spelling, accents, punctuation, and nicknames. There could be cause for some concern here if names that are unmatched are correlated with certain nationalities.

Finally, in order to remove duplicate transfer records (from a player moving from one club in the top 5 leagues to another), I kept only transfers in the "in" direction (roughly 50% of transfers). Table 2 presents summary statistics for the transfers that are able to be used in regression (i.e. they have known transfer fees).

Empirical Methods

Triple Differences

I use a triple differences identification strategy to estimate the impact of the Homegrown Player Rule on wages and transfer fees. This consists of comparing an outcome variable (player wage or transfer fee) between domestic and foreign players, before and after the Homegrown Player Rule, between treated and untreated groups. Without considering for which clubs the Homegrown Player Rule is binding, the Premier League is the only treated league, while the other four top European leagues serve as control leagues. I selected these four European leagues as controls mostly due to similarity in quality and popularity, but it is important to note as well that the Premier League is unique in its global popularity and financial wherewithal. Therefore, these four leagues cannot be seen as perfect controls for the Premier League in absence of a Homegrown Player Rule.

In my regressions comparing Premier League clubs that are constrained and unconstrained by the Homegrown Player Rule, the treated group represents players at clubs at which the Homegrown Player Rule was binding in the 2008-2009 season (the last control season). The control group represents players at clubs that were not constrained by the Homegrown Player rule in 2008-2009.

Due to my triple differences strategy, the regression has many interaction terms, with β_7 being the coefficient of interest.

$$\begin{aligned}
y_{it} = \beta_0 &+ \beta_1(\text{Treated}_{it}) + \beta_2(\text{Post}_t) + \beta_3(\text{Domestic}_{it}) \\
&+ \beta_4(\text{Treated}_{it} \times \text{Post}_t) \\
&+ \beta_5(\text{Treated}_{it} \times \text{Domestic}_{it}) \\
&+ \beta_6(\text{Post}_t \times \text{Domestic}_{it}) \\
&+ \beta_7(\text{Treated}_{it} \times \text{Post}_t \times \text{Domestic}_{it}) + \beta X_{it} + \gamma_t + \delta_i + \epsilon_{it}
\end{aligned}$$

In this regression, y_{it} represents an outcome of either player wage, transfer fee, log wage, or log transfer fee, depending on the regression specification. Treated_{it} represents whether an individual i is in the treatment group (as described above) at time t, Post_t represents whether the observation takes place before or after 2009, and Domestic_{it} represents whether or not player i is counted as a domestic player in their league at time t. X_{it} represents a vector of control variables, including player age, age^2, and league and club fixed effects. γ_t represents time fixed effects, and δ_i represents player fixed effects.

Note that, in this regression, I am comparing *domestic* and *foreign* as opposed to *homegrown* and *non-homegrown* players. This is in order to better compare between treatment and control leagues, because in leagues other than the Premier League, there does not exist a homegrown designation. I found it too arduous a data collection task to determine which players in each leagues had been trained in that country for at least three years before they turned twenty-one. This is certainly an area that warrants further research, as the effect on homegrown players is plausibly different from the effect on domestic players, who may or may not qualify as homegrown. When classifying players as domestic, I count players of both Welsh and English nationality as UK domestic players.

Additionally, I classify observations from the 2009-2010 season onward as post-treatment due to concern over anticipatory effects of the Homegrown Player Rule. These effects can be seen in Figure 3 where there is a significant dip in foreign player transfer fees in the Premier League in 2009. The Homegrown Player Rule was announced in 2009, but did not go into effect until a year later at the start of the 2010-2011 season. Therefore, it is reasonable that clubs would have made transfer and contract decisions in 2009 with the future rule in mind. The result of this regression can be interpreted as the bonus (either in wages or transfer fees) given to domestic players over foreign players as a result of this exogenous quota. Of course, this empirical strategy is only valid if the assumptions underpinning a triple differences design are satisfied.

For the constrained treatment group, clubs are included if they had fewer than eight domestic players in the 2008-09 season. The histogram of domestic players per club in 2008-09 is plotted in Figure 7. There are four clubs (out of twenty) in the treatment group: Arsenal, Chelsea, Liverpool, and Sunderland.

I would have also liked to include a quadruple differences regression comparing domestic vs foreign players at constrained vs unconstrained clubs in the Premier League vs other four leagues before vs after the Homegrown Player Rule. However, the key issue with this regression is the definition of constrained vs unconstrained in leagues other than the Premier League. In order to be constrained in the Premier League, a club must have fewer than eight domestic players in the 2009-10 season. When this same definition is applied to the other four leagues, there is only one club (out of 98) that fits this definition, as the Premier League is the most cosmopolitan league. Using this one club as a control group would not be very meaningful, meaning the results of this quadruple differences regression would not be intuitively interpretable.

Parallel Trends

In a triple differences empirical method, we require one parallel trends assumption, which can be written as:

$$\begin{gathered}
(E[y_{0,it}|\text{Treated}_{it}=1,\text{Domestic}_{it}=1,\text{Post}_t=1]-E[y_{0,it}|\text{Treated}_{it}=1,\text{Domestic}_{it}=1,\text{Post}_t=0]) \\
-(E[y_{0,it}|\text{Treated}_{it}=1,\text{Domestic}_{it}=0,\text{Post}_t=1]-E[y_{0,it}|\text{Treated}_{it}=1,\text{Domestic}_{it}=0,\text{Post}_t=0]) \\
= \\
(E[y_{0,it}|\text{Treated}_{it}=0,\text{Domestic}_{it}=1,\text{Post}_t=1]-E[y_{0,it}|\text{Treated}_{it}=0,\text{Domestic}_{it}=1,\text{Post}_t=0]) \\
-(E[y_{0,it}|\text{Treated}_{it}=0,\text{Domestic}_{it}=0,\text{Post}_t=1]-E[y_{0,it}|\text{Treated}_{it}=0,\text{Domestic}_{it}=0,\text{Post}_t=0])
\end{gathered}$$

as shown in Olden and Møen (2022). In this equation, $y_{0,it}$ represents the outcome of wages or transfer fees if treatment had never occurred. In words, this refers to a counterfactual world in which the Homegrown Player Rule was not implemented. Then, the difference-in-differences of foreign versus domestic and pre- versus post-2009 outcomes in the treatment group should equal the the difference in differences of foreign versus domestic and pre- versus post-2009 outcomes in the control group. In order to assess the validity of these assumptions, I look at the pre-trends for foreign versus domestic players in the Premier League and control leagues.

In Figure 2, the average wages of foreign and domestic players are plotted in the Premier League and the four control leagues. We see that the pre-trends seem roughly parallel up through the 2008-2009 season, the season before the Homegrown Player Rule was announced.

The only year that suggest some cause for concern is 2006, when the trends in the treatment group are somewhat divergent to trends in the control group.

Similarly, I plot the same averages for transfer fees in Figure 3. The parallel trends assumption is somewhat less compelling for transfer fees. The gap between foreign and domestic players' wages appears to widen before 2009 in the control leagues, while this does not occur in the Premier League. Interestingly, in the Premier League, domestic players' fees are higher than foreign players' fees before 2010, but lower after. In contrast, in the four control leagues, foreign players' transfer fees are consistently higher than those of domestic players.

Another cause for concern in these graphs is the significant but temporary drop in foreign players' transfer fees in the control leagues from 2010-2012. Overall, the volatility of transfer fees across the period from 2005-2015 could suggest that there are other uncontrolled factors impacting transfer fees, like financial regulations, taxation, or selection into publication of transfer fees.

Spillover Effects

Another key assumption to a triple differences empirical method is that the treatment and control groups are independent, and that any treatment applied to the treatment group has no effect on the control group. In the case of the Homegrown Player Rule, we might be concerned that the rule would have spillover effects into the four control leagues. For example, if clubs in the control leagues employed many English players, and their wages increased as a result of better outside options and more demand for English players, this would increase wages on average for foreign players in the control leagues, biasing the DDD estimate.

However, in this case, I do not believe that spillover effects should cause great concern. There is very little mobility of English players to the control leagues. From the 2005-06 season to the 2011-12 season, there are only 33 player-season observations of English or Welsh domestic players employed by control league clubs, compared to a total of 1458 English or Welsh domestic player-season observations total. Additionally, using the transfer data, there were only eight transfers of English or Welsh domestic players to clubs in the control leagues from 2005-2015. Due to these small numbers, I do not believe that spillover effects will significantly disrupt the independence of the treatment and control groups. Full statistics for the origin clubs of UK domestic players transferring to Premier League clubs are given in Table 3.

Nevertheless, I also provide results from a simple differences-in-differences regression comparing UK domestic players and all other players before and after the Homegrown Player Rule in Table 8. Of course, this regression loses the dimension of the Premier League vs control leagues, so it does not account for the general trend of globalization and increased hiring of foreign players across leagues. However, its results may be less threatened by potential spillover effects. The regression equation is given by:

$$y_{it} = \alpha_0 + \alpha_1(\text{Post}_t) + \alpha_2(\text{Domestic}_{it}) + \alpha_3(\text{Post}_t \times \text{Domestic}_{it}) + \beta X_{it} + \gamma_t + \delta_i + \epsilon_{it}$$

where α_3 is the coefficient of interest, and all variables are defined as in the DDD regression.

Results

First, I look at evidence for whether the Homegrown Player Rule increased representation of domestic players in the Premier League. Results are presented in Table 4. I find that, in absolute terms, the number of domestic players at Premier League clubs does not increase after the Homegrown Player Rule. This should not be too surprising, though, considering that the Homegrown Player Rule was intended to slow the trend of increasing numbers of foreign players in the Premier League, not completely reverse it. When compared to the four control leagues, there still is no relative increase in the number of domestic players in the Premier League. However, when I compare constrained and unconstrained clubs within the Premier League, I find that the quota does indeed increase the number of domestic players at constrained clubs.

In Table 5, I present the results for the triple differences regression of the effect of the Homegrown Player Rule on wages for all players in the Premier League. In regressions (1) - (4), log weekly wages are regressed on the full set of interaction terms in the regression given above, as well as age controls and certain fixed effects, indicated for each regression. I find that there is no effect of the Homegrown Player Rule on domestic players' wages, given that the magnitudes of the coefficients of interest are small (between -4% and -8%) and none of them are statistically significant at the 5% level. In Figure 4, I plot the DDD coefficients by year on log wages for domestic Premier League players, based on Regression (4) from Table 5 including the full set of controls. There appears to be a slight downward trend in wages after the Homegrown Player Rule, but the only season with significantly lower wages than the 2008-2009 season is the 2011-2012 season.

In Table 6, I present the results from regressions measuring the effect of the Homegrown Player rule on wages of domestic players at constrained teams. For Regression (1), I run a simple difference-in-differences regression comparing domestic player wages at constrained vs unconstrained clubs before and after the Homegrown Player rule. I estimate that wages of domestic players at constrained clubs increases by 36% when compared to wages of domestic players at constrained clubs. Regressions (2)-(4) are triple differences regressions comparing wages of domestic vs foreign players at constrained vs unconstrained clubs before and after the homegrown player rule. In Regression (4), with all controls included, I find that domestic player wages at constrained clubs rose by 42% after the Homegrown Player rule. In Figure 5, I plot the coefficients on log wages of domestic players at constrained clubs using the same controls as in Regression (4) from Table 6. We see that wages for these domestic players increase significantly in the first treatment year and remain well above pre-treatment levels. It should be noted that while these effects are large and significant, they are also imprecisely measured, with standard errors around 10-15 percentage points.

To supplement my analysis of player wages, I use the same triple differences strategy to analyze the impact on transfer fees. Table 7 presents the results of these regressions. As with wages, I regress transfer fees on the full set of interaction terms from my regression equations. Regressions (1) - (4) are performed using transfer fees. However, Regression (5) provides the log regression of 1 + transfer fee in order to still include the modal transfer fee of 0. All regressions (1)-(4) show a large decrease in transfer fees of domestic players, between €1.6 million and €1.9 million. Regression (4) excludes the 99th percentile of transfer fees. Transfer fees have a right-skewed distribution, meaning that a few big-ticket transfers before or after the 2010 Homegrown Player Rule implementation could potentially bias estimates. However, we see in regression (4) that this does not appear to be the case, as the DDD coefficient does not change much. The main difference is that the standard error is lower because the standard deviation of transfer fees is lower. All of these estimates are significant at the 5% level, and the coefficient in regression (4) is significant at the 1% level. In Regression (5), which runs the DDD regression on log transfer fees, the effect is also very large, with the Homegrown Player Rule estimated to reduce transfer fees by 30%. In Figure 6, just as with wages, I plot the coefficients from the DDD regression of transfer fees. We see a steady decline in transfer fees after the implementation of the rule.

Conclusion

While I do not find an overall positive wage effect on domestic players, I do find that the Homegrown Player Rule increased wages for domestic players at clubs constrained by the rule. This result offers some promising evidence for the ability of quotas to not only increase representation, but also increase wages of workers protected by the quota. The zero effect on overall wages for domestic players in the Premier League, though, may be cause for concern. One potential explanation could be that the best domestic players simply moved from unconstrained to constrained clubs who were looking to replace their talented foreign players. Additionally, wages are sticky, especially in soccer where wages are given in multi-year contracts. My analysis of wages only considers wages up through 2012, which may not allow time for wages to fully adjust after the implementation of the rule.

The decrease in transfer fees, while also consistent with model predictions, should not be accepted unequivocally either. Selection into transfer fee publication is a potential concern, as well as the volatility of transfer fees across seasons.

In the case of both wages and transfer fees, it is important to recognize that neither are complete measures of player compensation, nor do they encompass all transactions that occur in the professional soccer market. Just to name a few: professional soccer players have sponsorship deals with companies; player and club agents receive sizable commissions from transfer fees, and those figures are rarely publicized; finally, differential tax rates for foreign and domestic players mean that net wages for foreign and domestic players may be significantly different from their gross wages, which are used in this analysis.

In the case of the Homegrown Player Rule, questions still remain. Potential areas for future research include looking at more recent data to estimate longer-term effects on wages, as well as developing an empirical strategy to distinguish the effect on wages and transfer fees of all homegrown players, not just domestic players.

More broadly, the effect of quotas on wages remains an area of interest for future empirical research. While representation is an important metric, at the end of the way, workers really care about wages. In addition to increasing representation, one of the express goals of quotas should be to improve compensation for workers covered by the quota. In order to assess the impact of quotas on workers, it is necessary to examine what happens to relative wages as a result of quotas.

References

Agan, Amanda and Sonja Starr, "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment," *Quarterly Journal of Economics*, 2018, *133* (191), 235.

Antecol, Heather, Kelly Bedard, and Jenna Stearns, "Equal but Inequitable: Who Benefits from Gender-Neutral Tenure Clock Stopping Policies?," *American Economic Review*, 2018, *108* (9), 2420–2441.

Bertrand, Marianne, Sandra E Black, Sissel Jensen, and Adriana Lleras-Muney, "Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway," *The Review of Economic Studies*, 2019, *86* (1), 191–239.

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

Kleven, Henrik Jacobsen, Camille Landais, and Emmanuel Saez, "Taxation and International Migration of Superstars: Evidence from the European Football Market," *American Economic Review*, 2013, *103* (5), 1892–1924.

Olden, Andreas and Jarle Møen, "The triple difference estimator," *The Econometrics Journal*, 2022, *25*, 531–553.

Pinje, Jori Veng, "Wage- and Tax-Induced Mobility of Footballers among 30 European Countries," 2013.

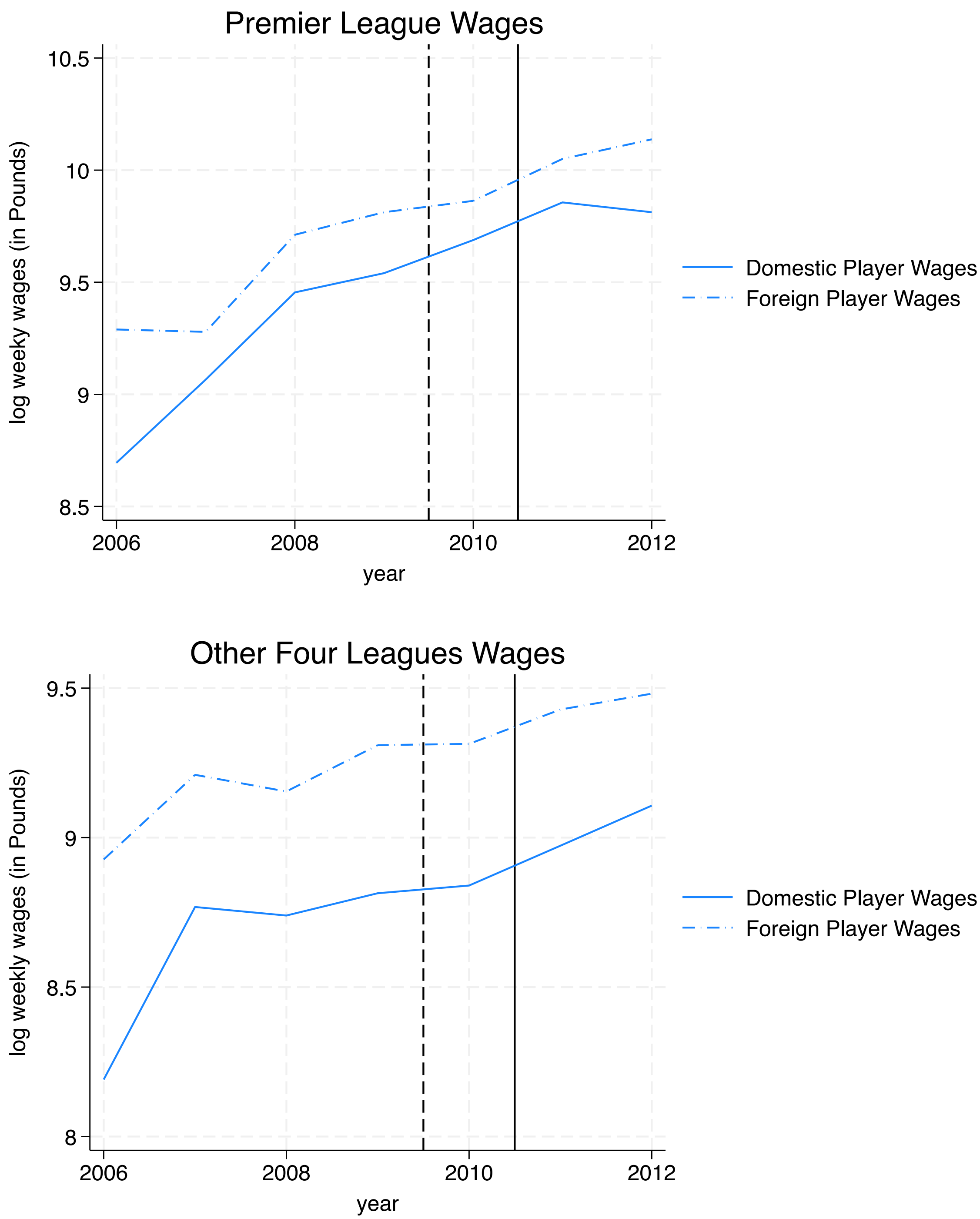
Szerman, Christiane, "The Labor Market Effects of Disability Hiring Quotas," 2022.

Figures and Tables

Table 1: Wages in the Top 5 European Leagues, 2005-2012

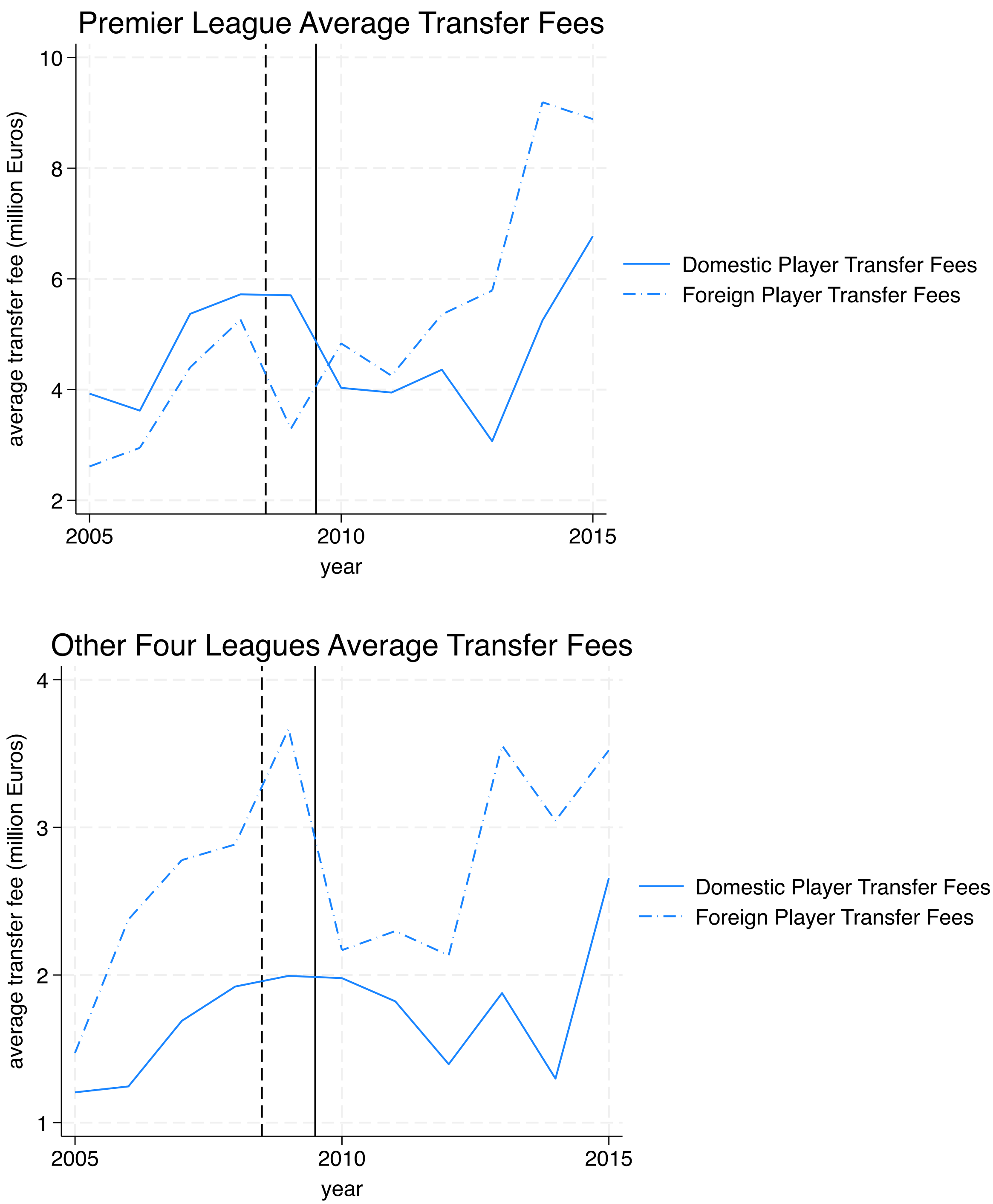
	Premier League		Four Control Leagues	
	Pre-2009	Post-2009	Pre-2009	Post-2009
Number on payroll	1833	1390	7153	5852
Mean	20.6	30.8	12.6	17.9
Standard Deviation	20.8	28.4	13.7	24.2
Minimum	0	0	0	0
Maximum	160	180	169	283

This includes all estimated weekly wages paid to players in the top 5 European leagues from 2005-2012. Mean, standard deviation, minimum, and maximum are given in thousands of Pounds.



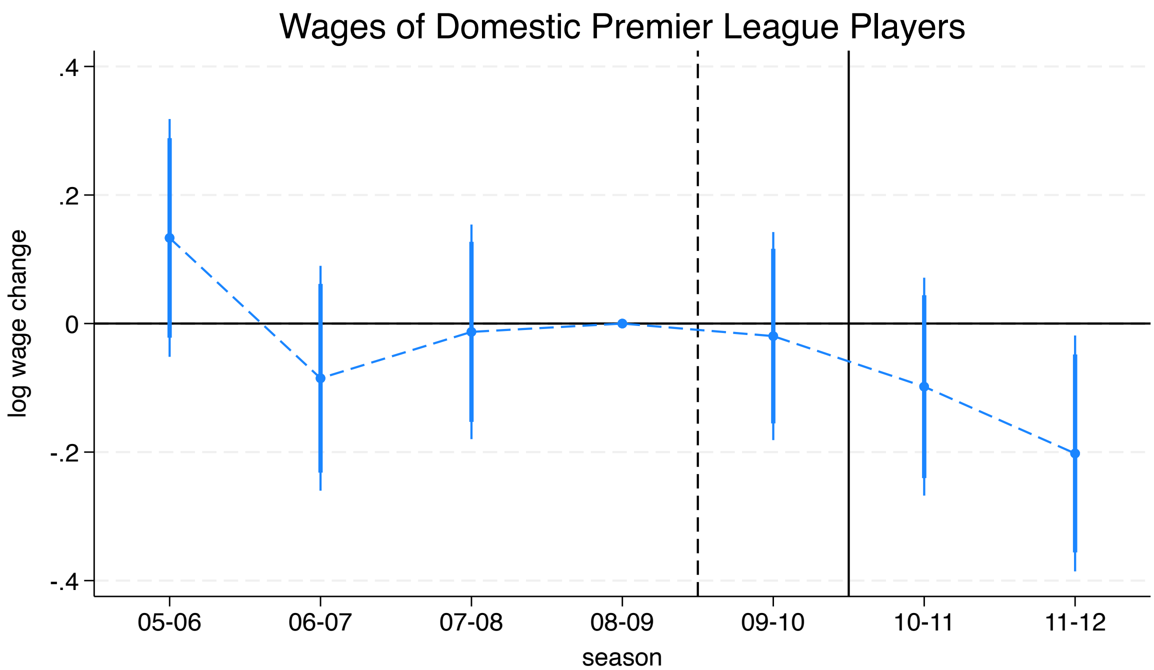
Log weekly wages (in GBP) of domestic and foreign players are plotted from 2006-2012. Each year on the x-axis corresponds to wages during the season ending in that year (i.e. 2009 refers to the 2008-2009 season).

Figure 2: Wages Trends Premier League and Control Leagues



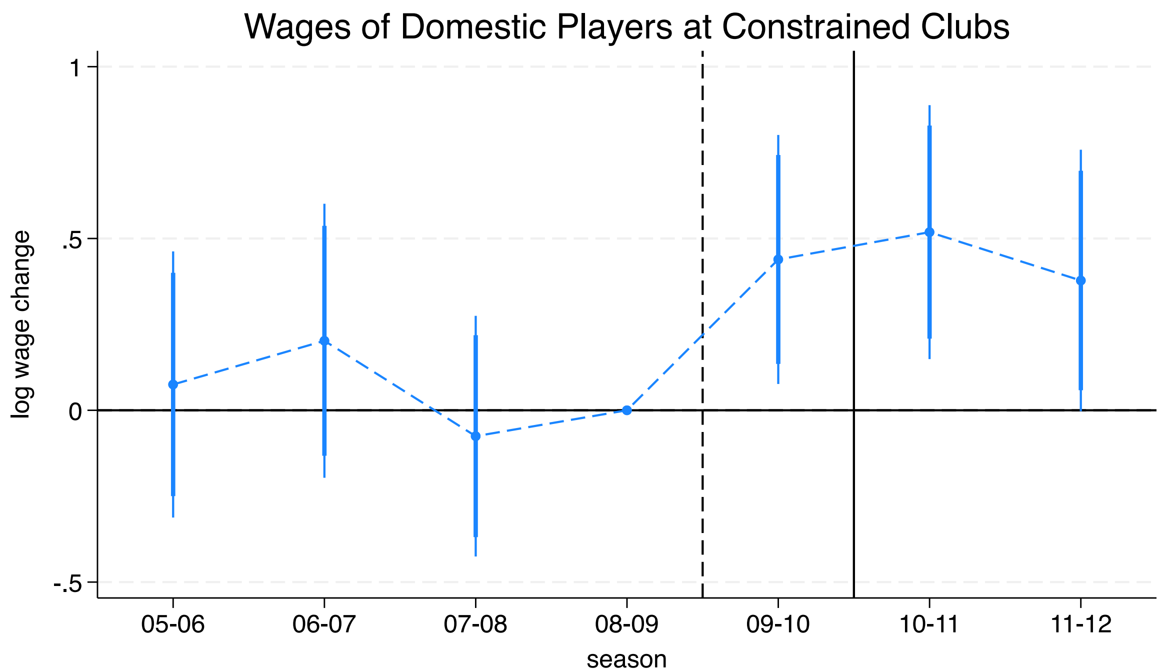
Average transfer fees of (in millions of Euros) of domestic and foreign players are plotted from 2005-2010. Each year on the x-axis corresponds to transfers occurring during that calendar year.

Figure 3: Transfer Fee Trends Premier League and Control Leagues



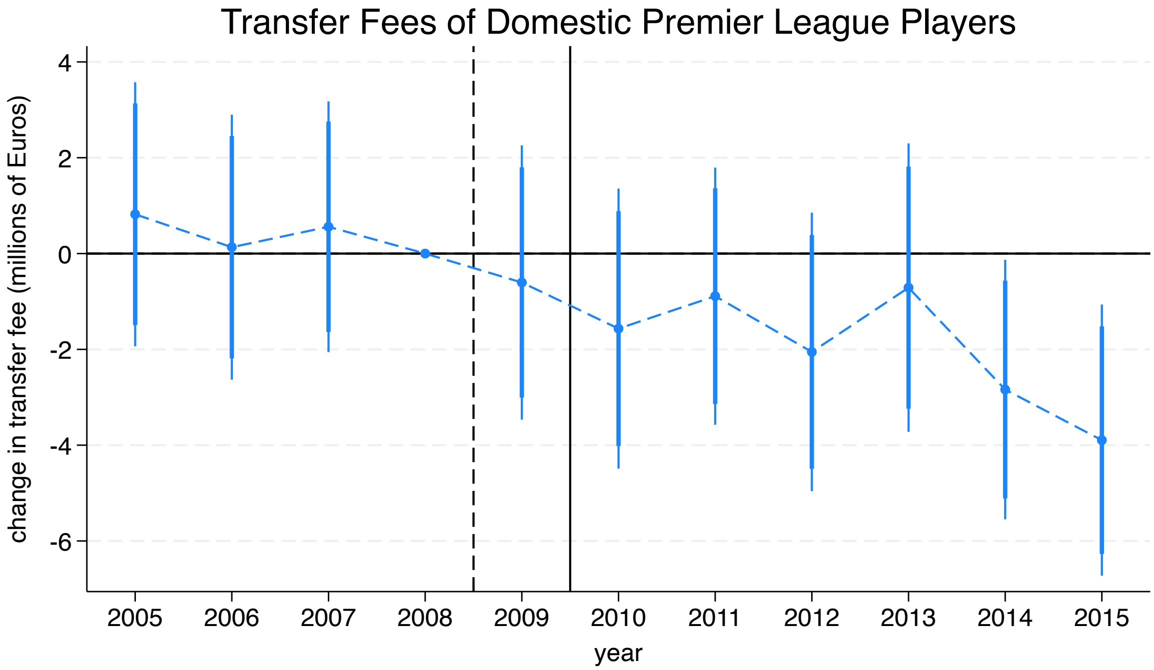
The coefficients on log wages for domestic Premier League players are plotted, compared to foreign Premier League players and control league players. These coefficients are derived from a regression that includes age controls, and league, club, and player fixed effects. The bars represent 90% and 95% confidence intervals.

Figure 4: Change in Premier League Domestic Player Wages



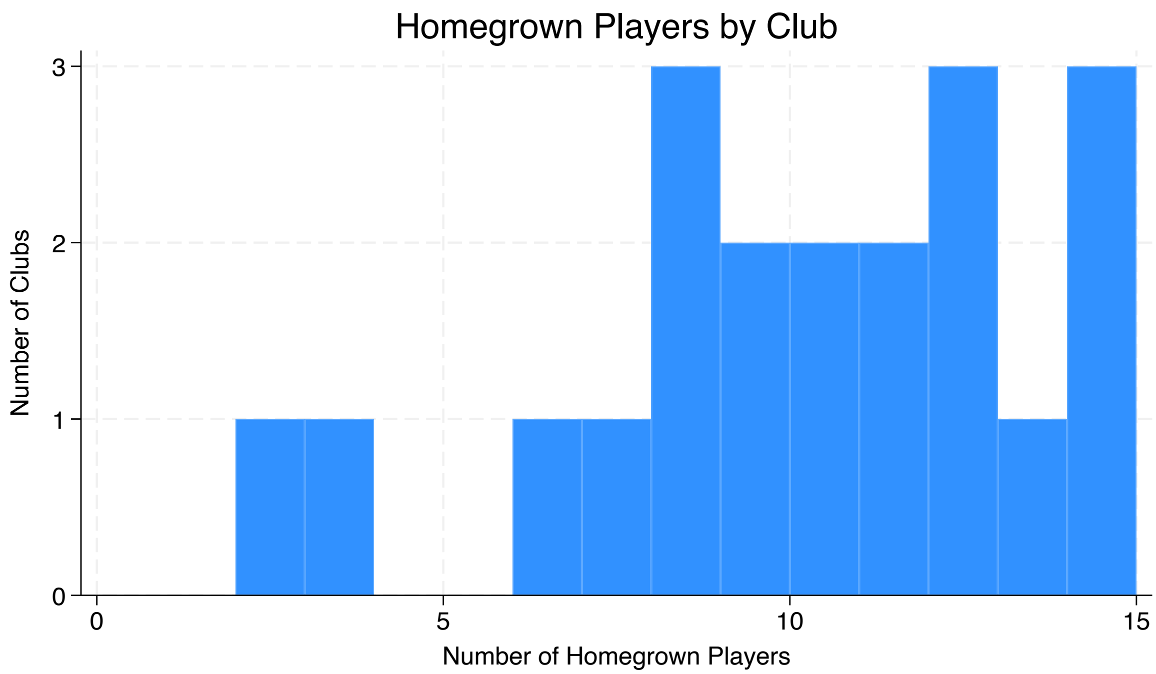
The coefficients on log wages for domestic players at constrained clubs are plotted, compared to foreign players and players at unconstrained clubs. These coefficients are derived from a regression that includes age controls, and league, club, and player fixed effects. The bars represent 90% and 95% confidence intervals.

Figure 5: Change in Premier League Constrained Domestic Player Wages



The coefficient on transfer fees for domestic Premier League players are plotted, compared to foreign Premier League players and control league players. These coefficients are derived from a regression that includes age controls, and league and club, and player fixed effects. The bars represent 90% and 95% confidence intervals.

Figure 6: Change in Premier League Domestic Transfer Fees



This histogram plots the number of UK domestic players at each Premier League club in the 2008-2009 season (the last control season). While domestic is not a perfect analog for homegrown only one player in the league that year would have qualified as homegrown but not domestic, so this is a pretty accurate representation of which clubs were constrained by the Homegrown Player rule.

Figure 7: Number of Domestic Players At Premier League Clubs in 2008-09

Table 2: Transfer Fees in the Top 5 European Leagues, 2005-2015

	Premier League		Four Control Leagues	
	Pre-2009	Post-2009	Pre-2009	Post-2009
Number of transfers	603	952	1947	3715
Mean # of transfers/year	151	136	487	531
Mean	4.1	5.7	2.0	2.4
Standard Deviation	6.1	8.9	3.6	5.8
Minimum	0	0	0	0
Maximum	43.9	76	35.5	101

This includes all transfers into the Premier League and the Control Leagues from 2005-2015 for which the transfer fee was known. Mean, standard deviation, minimum, and maximum are given in millions of Euros.

Table 3: Origins of UK Domestic Players Transferring to Premier League Clubs, 2005-2015

Club transferring from	Number	Percent
Other English club	417	95
English youth team	8	2
Club in other 4 European leagues	5	1
Club in other foreign league	8	2
Total	438	100

This includes all transfers into the Premier League and the Control Leagues from 2005-2015 for which the transfer fee was known. Other English club refers to any club in the English and Welsh football pyramid, including teams in the Premier League and lower league. English youth team refers to the youth team of any English club. A club in the other 4 European leagues refers to clubs in La Liga, Ligue 1, Bundesliga, or Serie A.

Table 4: Number of Domestic Players at Premier League Clubs

	(1)	(2)	(3)
Number of domestic players	-26.91*	-42.67	32.75**
	(14.85)	(38.71)	(14.83)
Regression	DID	DDD	DDD
Constrained?			Yes

Regression (1) is the DID regression of number of domestic vs foreign players before and after the HG player rule. Regression (2) is the DDD regression of number of domestic vs foreign players before and after the HG player rule in the Premier League vs the control leagues. Regression (3) is the DDD regression of the number of domestic vs foreign players before vs afteer the HG player rule at constrained vs unconstrained clubs
*** = significant at the 1% level, ** = significant at the 5% level, * = significant at the 10% level

Table 5: Triple Differences Regression of Wages

	(1)	(2)	(3)	(4)
$\text{Treated}_{it} \times \text{Post}_t \times \text{Domestic}_{it}$	-0.0517	-0.0703	-0.0665	-0.0643
	(0.0804)	(0.0562)	(0.0561)	(0.0547)
age	1.0626***	1.1206***	1.1151***	1.1067***
	(0.0241)	(0.0555)	(0.0553)	(0.0543)
age^2	-0.0180***	-0.0201***	-0.0201***	-0.0197***
	(0.0004)	(0.0006)	(0.0006)	(0.0006)
Year fixed effects?	Yes	Yes	Yes	Yes
Player fixed effects?		Yes	Yes	Yes
League fixed effects?			Yes	Yes
Club fixed effects?				Yes
R^2	0.34	0.82	0.82	0.83
Number of observations	16210	13714	13714	13714

In regressions (1)-(4) log weekly wages (in thousands of GBP) are regressed on all the terms from the above regression.
*** = significant at the 1% level, ** = significant at the 5% level, * = significant at the 10% level

Table 6: Constrained vs Unconstrained Regression of Wages

	(1)	(2)	(3)	(4)
$\text{Treated}_{it} \times \text{Post}_t \times \text{Domestic}_{it}$	0.3631***	0.8706***	0.4406***	0.4218***
	(0.1197)	(0.1986)	(0.1319)	(0.1342)
age	0.8797***	1.0618***	0.8863***	0.9042***
	(0.1604)	(0.0498)	(0.0973)	(0.0996)
age^2	-0.0184***	-0.0181***	-0.0172***	-0.0176***
	(0.0016)	(0.0009)	(0.0012)	(0.0012)
Regression	DID	DDD	DDD	DDD
Year fixed effects?	Yes	Yes	Yes	Yes
Player fixed effects?	Yes		Yes	Yes
Club fixed effects?	Yes			Yes
R^2	0.87	0.40	0.86	0.87
Number of observations	1192	3222	2658	2658

In regressions (1)-(4) log weekly wages (in thousands of GBP) are regressed on all the terms from the above regression.
*** = significant at the 1% level, ** = significant at the 5% level, * = significant at the 10% level

Table 7: Triple Differences Regression of Transfer Fees

	(1)	(2)	(3)	(4)	(5)
Transfer fee	Yes	Yes	Yes	Yes	
Log(1+TF)					Yes
$\text{Treated}_{it} \times \text{Post}_t \times \text{Domestic}_{it}$	-1.8603**	-1.8611**	-1.6307**	-1.7671***	-0.3140***
	(0.7903)	(0.7904)	(0.7143)	(0.5591)	(0.1044)
age	1.5884***	1.5586***	2.0003***	1.6259***	0.4095***
	(0.1354)	(0.1343)	(0.1437)	(0.1064)	(0.0219)
age^2	-0.0334***	-0.0330***	-0.0404***	-0.0331***	-0.0085***
	(0.0025)	(0.0025)	(0.0027)	(0.0020)	(0.0004)
Year fixed effects?	Yes	Yes	Yes	Yes	Yes
League fixed effects?		Yes	Yes	Yes	Yes
Club fixed effects?			Yes	Yes	Yes
Excluding 99th percentile?				Yes	
R^2	0.08	0.09	0.34	0.33	0.39
Number of observations	7217	7217	7217	7149	7216

In regressions (1)-(4) transfer fees or log transfer fees (in millions of Euros) are regressed on all the terms from the above regression.
*** = significant at the 1% level, ** = significant at the 5% level, * = significant at the 10% level

Table 8: Difference-in-Difference Regressions of Wages and Transfer Fees

	(1)	(2)	(3)
Log wage	Yes	Yes	
Transfer fee			Yes
$\text{Post}_t \times \text{Domestic}_{it}$	0.1084*	-0.0882**	-0.1957
	(0.0565)	(0.0370)	(0.7046)
age	1.1098***	1.1359***	1.6063***
	(0.0246)	(0.0554)	(0.1402)
age^2	-0.0189***	-0.0203***	-0.0338***
	(0.0005)	(0.0006)	(0.0026)
Year fixed effects?	Yes	Yes	Yes
Player fixed effects?		Yes	
R^2	0.29	0.81	0.04
Number of observations	16210	13714	7217

In regressions (1)-(4) log weekly wages (in thousands of GBP) or transfer fees (in millions of Euros) are regressed on the terms from the difference-in-differences regression.
*** = significant at the 1% level, ** = significant at the 5% level, * = significant at the 10% level

The Effect of Changes in Fuel Prices on Adoption of Farming Mechanization: A Cross-Country Analysis

by

Lukas Hanson-Puffer

Submitted to the Department of Economics
On May 10, 2024 in Partial Fulfillment of the
Requirements for the Degree of
Bachelor of Science in Economics

Abstract

Ever since the industrial revolution, farming worldwide has become increasingly mechanized, and has corresponded with dramatic increases in crop yields, essential for global food security. This paper seeks to investigate a possible relationship between changing local diesel pump prices as a determinant of increasing mechanization of agriculture in a country. By using different local responses to global crude oil prices as an instrument for local diesel prices and machinery imports per capita by year as a proxy for adoption of mechanization, we found a substantial and significant positive association between the change in predicted national diesel price and level of imports of farming machinery per capita. This relationship held similar when accounting for time lags and GNI per capita fluctuations and increased in magnitude when excluding net exporters of farming machinery. Our analysis suggests that this is because of the high correlation between cereal commodities and crude oil prices which increase farmer incomes and therefore increase purchases of farming machinery.

Thesis Supervisor: Benjamin Olken

Title: Professor of Economics

1 Introduction

Every country has a unique story of economic development and the number of strategies for development is only growing. Globalization has opened countless doors of connection and specialization, meaning that economic development today looks vastly different to what it did during the industrial revolution and will continue to evolve as our technology and knowledge bases advance further. Despite these differences, all societies throughout history have faced the same issue of how to obtain enough food to feed their population. The advent of agriculture was a key to the development of civilization since a society can only grow if one person can grow more food than one person consumes so that a labor force is free to do other jobs to maintain and advance the society.

Overall, the world's ability to grow more food for every person has increased dramatically since the start of the 20th century. In 1900, 72.1% of people employed were employed in the agricultural sector, thereby meaning that a person could theoretically produce enough food for 1.39 people (Grigg, 1975). This is in sharp contrast to in 2022, when only 26% of the global labor force being employed in agriculture, or each person growing food for 3.85 people (World Bank, 2024). This growth has enabled the global population to ballon from 1.6 billion in 1900 to over 8 billion today (World Bank, 2022).

One of the most effective ways to increase agricultural yield is through using machinery such as tractors and combine harvesters, which have helped contribute to wheat yield in the United Kingdom growing from 0.99 tons per hectare in 1850 (Broadberry et al., 2015) to over 8 tons per hectare in 2000 (FAO, 2023). Yield, however, varies widely by country based on the level of mechanization. The United Kingdom, for example, has the 5th highest wheat yield in the world and has one of the highest levels of mechanization at 2.31 horsepower per unit of agricultural land, compared with Somalia with the lowest wheat yield in the world at 0.4 tons per hectare and only 0.04 horsepower per unit of agricultural land (USDA, 2023). Therefore, by looking at the factors influencing the adoption of mechanized farming equipment, we can potentially advise policies to increase mechanization and improve yield. Given that population in developing countries is expected to grow 90% between 2017 and 2050, improving agricultural productivity will be crucial to ensuring food security (United Nations, 2022).

In this study, we conducted a quantitative analysis attempting investigate a possible relationship between changes in local energy prices and changes in the level of farming mechanization in a country. We ran multiple regressions of the change local diesel pump prices, using Brent Crude oil prices as an instrument to improve data availability, on the imports of farming machinery in a given country. This gave us country-specific coefficients, measuring the pass-through of global oil prices onto local diesel prices. We included controls for heterogeneity in GNI per capita growth and for if the country is a net exporter of machinery, as well as both time and country fixed effects.

Based on existing literature on price effects, we identified three hypotheses for what effect changes in diesel prices could have on the value of farming machinery imports in a country for a given year. First, lower operational costs could incentivize farmers to mechanize. Alternatively, higher diesel prices could make newer, more fuel-efficient equipment more attractive, thereby increasing imports as diesel prices increase. Finally, diesel prices could have an effect on farmer income, through impacting commodity grain prices, as well as on operational cost, which could impact the ability of farmers to afford mechanized equipment.

We first measured the feasibility of predicting diesel price by country from its relationship with the global Brent Crude oil price. We found that countries have elasticities that vary widely from -0.19 in Bahrain to 1.79 in Nigeria, but with most countries near the global average of 0.63. This means that for every logarithmic level increase in crude oil price per barrel, the average country will see a 0.63 level increase in diesel pump price per liter. The resulting predicted diesel prices track well with the actual diesel price values with R-squared valued largely concentrated between 0.6 and 1.

Then, using these predicted diesel prices, we found that a logarithmic unit increase in diesel price corresponds with a 0.552 logarithmic unit increase in machinery imports per capita, significant at the 1% significance level and found that time lag had insignificant impact on imports. This strong and significant result was still present when excluding net exporters of farming machinery and when incorporating GNI per capita growth with coefficients of 0.735 and 0.488, respectively. In the discussion section, we found that this positive relationship was most likely caused by the close association between diesel prices and commodity grain prices with regression coefficients between 0.75 and 0.83 for cereal crops. The attractiveness of fuel-efficient machinery was found to likely be less of a factor given the extremely long payback period required for the purchase of new machinery.

2 Theory and background

The mechanization of agriculture has changed form over time as technology improves and farming's role in society was redefined by industrialization. In this section we discuss historical mechanization, its potential impact on economic development measured in the Romer model, and previous studies conducted investigating the causes and effects of mechanization.

2.1 History of farming mechanization

The advent of agriculture 12,000 years ago, during the Neolithic Revolution, is widely credited with causing humans to abandon our previous nomadic, hunter-gatherer lifestyles (National Geographic Society, 2024). The first farming tool, however, was not invented until 7,000 years later in Ancient Sumeria when forked sticks were used as an early form of plough (Biering, 2024) Over the ensuing millennia, farming tools improved in efficiency, but were still predominantly hand-operated and labor intensive. This dramatically changed during the industrial revolution, starting with Jethro Tull's horse-drawn seed drill in 1701, before progressing to motorized equipment in the 20th century (McNeil, 2003).

During the latter half of the 20th century, the adoption of agriculture mechanization grew dramatically, but asymmetrically worldwide, greatly increasing agricultural yields as population exploded, especially in Asia. Using horsepower per hectare (hp/ha) of agricultural land as a measure of mechanization, Japan has seen an over six-fold growth from 3hp/ha in 1968 to 18.9hp/ha in 2011. This trend is similar to Thailand's 1100%, the Philippines' 1067%, and India's 635% growth in mechanization over the same time period (Bautista et al., 2017). Despite the rapid growth seen in South-East Asia, Africa has shown much slower adoption of mechanized farming equipment where only 18% of farms use some sort of tractor-powered machinery and only 35% have access to animal-powered equipment (Kirui, 2019).

2.2 Impact of mechanization on development

The primary advantage of mechanized farming is the potential to vastly increase agricultural yield – the weight of crops able to be grown per unit area. Looking at increasing mechanization in China, every 1% increase in mechanization has been shown to increase the yield of grain crops by 1.59% (Peng et al., 2022). Increasing mechanization during the industrial revolution and again during the green revolution of the 1960s has enabled global population to explode to over 8 million today despite growing at an average 1.3% in the 20th

century compared to an average 0.08% in a pre-industrial world (United Nations, 2022). The world population is expected to surpass 10 billion by 2058 with Africa projected to double in population, thereby making increasing agricultural efficiency key to both food security and increasing farmer incomes in Africa over the next 30 years (Banafa, 2019). This is especially relevant as regions will likely see diminished natural resources such as water and soil fertility due to climate change and environmental degradation (Emami et al., 2018).

The macroeconomic impact of farming mechanization can be shown through applying the Romer model to the agricultural sector. We see in equations (1) that agricultural output (Y) is defined by a function of technology level (A), level of capital (K), and labor force (L). Since the quantity of arable land in a country is largely fixed,[1] capital here refers to elements such as farming infrastructure and livestock. In the Romer model, the rate of technology growth ($\dot{A}$) is endogenous, the function for which is defined in equation (2).

$$(1) \quad Y = A * F(K, L)$$

$$(2) \quad \dot{A} = G(A, K, L)$$

Using the Cobb-Douglas production, we can now redefine the general form to the equations. In equation (3), α and β represent the proportion of output attributable to capital and labor respectively, which varies by country. In equation (4), $\gamma, \delta,$ and ϵ represent the proportion of change in technology level attributable to the current technology level, level of capital, and state of the labor force, while ω represents the efficiency of a country's innovation process. Through expanding the equation, we see in equation (6) the final output growth function for a country's agricultural sector.

$$(3) \quad Y = A * K^{\alpha} L^{\beta}$$

$$(4) \quad \dot{A} = \omega A^{\gamma} K^{\delta} L^{\epsilon}$$

$$(5) \quad \dot{Y} = A\left[\dot{A} * K^{\alpha} L^{\beta} + \frac{\delta}{\delta t}\left(K^{\alpha} L^{\beta}\right)\right]$$

$$(6) \quad \dot{Y} = \omega A^{\gamma} K^{\alpha} L^{\beta}\left(\frac{\dot{A}}{A} + \alpha\frac{\dot{K}}{K} + \beta\frac{\dot{L}}{L}\right)$$

[1] Future work could examine economic models with more fixed effects in addition to the area of arable land per country.

The mechanization of farming has a significant impact on this equation, primarily through improving the level of technology in the sector by increasing the capital stock in the form of increasing the stock of farming machinery. Elements that impact the ability of the country of mechanizing include the current capital stock (K) through impacting someone's ability to purchase machinery, the labor force (L) being trained in operating machinery, and the ability of a country to diffuse new technologies to the population (ω). The price of diesel would theoretically affect the profitability of investment in machinery which would subsequently affect K.

2.3 Existing literature

Multiple studies have been conducted on the mechanization of agriculture, mostly focusing on the effects of mechanization on the economy and wider society. Gollin (2002) found that low agricultural productivity acts as a delay on economic development. These findings were corroborated by Caunedo (2022) upon running a randomized control trial in multiple Indian provinces with the addition of measuring productivity and employment. Finally, in examining the green revolution, Foster and Rosenzweig (1996) found that the increasing agricultural yields had positive impacts on health and school enrollment in India. Matsuyama (1991) showed that increasing productivity in agriculture spills into other sectors, leading to the movement of labor to more productive sectors, and in-turn causing long-term economic growth.

More recently, studies have sought to identify factors that help drive mechanization in developing countries. Gollin (2014) found that the agricultural productivity gap between countries was primarily caused by differences in technology levels, quality of national institutions, infrastructure and access to markets, and human capital. Upon examining seven factors influencing mechanization, Ghosh (2010) found that the two most impactful were the size of landholding and access to irrigation sources. This paper seeks to build on these studies by conducting a regression analysis to investigate a possible relationship between energy prices and the mechanization of agriculture.

2.4 Mechanization cause hypotheses

Given the impact that mechanization can have on economic development, this paper seeks to conduct a quantitative analysis of the role of energy prices on adoption of this technology. We have identified three primary hypotheses for what the impact could be and why. Firstly, lower energy prices could represent lower cost of operation for mechanized farming equipment, thereby incentivizing farmers to purchase equipment. Alternatively, this effect could be offset since when diesel prices rise, fuel economy becomes more valuable, thereby incentivizing the purchase of new vehicles when fuel costs are high. Finally, diesel prices could influence farmers' economic situation in other ways such as affecting commodity prices of their crops which could impact their willingness and ability to purchase equipment beyond considering operational costs. Following the analysis, these three hypotheses will be discussed as to their potential influence on the results.

3 Data

This study primarily uses two datasets, looking to measure the local diesel price per liter in each country and the corresponding level of farming mechanization for a given year. In this section, we will give an overview of the datasets used as well as validating the combined dataset and explaining limitations of the data used.

3.1 Diesel prices

According to the United States Department of Agriculture (USDA), diesel fuel was responsible for 73% of on-site energy consumption in domestic agriculture, primarily used to power farming machinery such as tractors. This proportion would likely be even higher in developing countries due to lower availability of electricity. Therefore, this study uses national diesel pump prices per liter to measure the cost of energy for farmers. The World Bank[2] contains biennial data for median diesel pump prices in US dollar equivalent per liter for 227 countries and regions between 1991 and 2016 averaging just over 10 readings per country. For the time period, diesel pump prices averaged \$0.83 per liter with a standard deviation of \$0.46. There is also a sizeable range in prices from the negligible \$0.008 per liter in OPEC-member Venezuela, both in 2014 and 2017, to a high of \$3 per liter in conflict-strewn Eritrea those same years.

3.2 Mechanization of farming

There were multiple avenues that we explored when considering how to measure the switch in the agricultural sector from unmechanized hand tools such as scythes to machinery such as tractors and harvesters. We could measure mechanization (1) directly through World Bank's dataset on number of tractors per capita, (2) through its effects by looking at agricultural yield of cereal crops per area of arable land, (3) through mechanization's effects

[2] Data is collated by the German Agency for International Cooperation (GIZ)

on labor by using national employment in agriculture, but eventually decided upon (4) through measuring the value of imports of farming machinery.

While the direct metric would be preferable, the dataset only contains data until 2000, greatly diminishing the potential usefulness of the study. The decision to use import data from the United Nations Comtrade database was due to a combination of data availability and being a good proxy for the level of mechanization in a country. From the dataset, there are over 5,000 readings for the value of total farming machinery imports for 220 countries and regions between 1980 and 2022.[3] During the time period covered by the data, the general trend is that imports per capita are increasing over time at a rate of $0.67 per year, ranging from a low of $9.48 in 2000 to a high of $41.95 in 2007. Imports per capita vary widely by country with a most recent value average of $34.93 ranging a low of $0.07 in Albania to a high of $260.07 in Lithuania.

3.3 Combined dataset validation

After merging the datasets of predicted diesel price and farming machinery imports, we have 3,259 datapoints across 123 countries between 1987 and 2022. The considerable number of readings and countries represented suggests that this data is sufficiently comprehensive for our cross-country comparison study. However, given that around 37% of 195 UN countries are not represented, we must verify whether our subset of countries is a representative cross-section of the world in terms of geographic location and economic level.

As shown in ***Table 1***, the seven geographic regions defined by the World Bank are all represented, ranging from 33% in North America to 69% in Europe & Central Asia. While there is a discrepancy here, the regions with 50% or less representation (North America and East Asia & Pacific) will not have a substantial effect on the validity of the study. This is because multiple countries in East Asia & Pacific are small islands such as Tuvalu where there is little potential for mechanization due to lack of arable land. Additionally North America only has two total countries so will have negligible impact on the full dataset.

In terms of income representation, all segments apart from Low income have 60-70% representation, suggesting that the dataset is sufficiently representative of the world. Only 11 countries classified as Low income being included suggests that the under-representation

[3] Farming machinery was defined by SITC trade codes 721 - Agricultural machinery (excluding tractors) and parts thereof, 7223 – Track-laying tractors, and 7224 – Wheeled tractors.

could hinder the usefulness of the study results. However, it could be argued that these countries would be the least useful in determining the causes of farming mechanization since countries such as Malawi have persistently low levels of mechanization. Therefore, the slightly lower representation of Low income countries will have a smaller effect on the validity of the study.

3.4 Data limitations

As shown previously, the combined dataset of diesel prices and value of farming machinery imports is sufficiently comprehensive and representative to provide useful empirical results. Despite these strengths, there are some limitations to this dataset including (1) aggregation at the national level, (2) aggregation at the national level, and (3) import data mostly neglecting the secondhand market.

The first weakness of the data is the fact that diesel prices are aggregated the annual level. Looking just at the United States, diesel prices per gallon vary widely within years such as in 2008 when prices peaked in the middle of July at $4.764 before falling to less than half that price by the end of the year at $2.327 (US Energy Information Agency, 2023). Volatility in early 2022 was even more stark with prices jumping $1.64, a 45% increase, in just three months between January and March. Missing these short-term shocks in diesel prices could therefore limit the usefulness of the study as people's behavior would likely adapt faster than over the course of a year. This limitation is somewhat mitigated since while short-term consumption decisions would likely adapt quickly, the decision to invest in capital equipment would be predicated on longer-term trends.

The second weakness is also through aggregation, but this time aggregating diesel prices on a national level rather than regionally means that the dataset loses detail. Local discrepancies are by far the most pronounced in large countries such as the United States where average diesel prices in Hawaii are 56% higher than that in Oklahoma (American Automobile Association, 2024). However, this trend is also true in smaller countries such as the United Kingdom show significant discrepancies with diesel prices in Belfast being 12% lower than the county of Dumfriesshire, less than 50 miles away (Fleet News, 2024). It is reasonable to assume that farmer purchasing decisions would be based on local prices instead of national, which is not shown in the data. However, this limitation is also mitigated since local diesel prices are largely determined by factors such as land and labor costs of operating a filling station, which are factors that do not significantly change relative to other regions over time.

The third data limitation concerns the dataset of farming machinery imports and how imported goods are typically new, which would neglect a large secondhand market for farming machinery. Used vehicles would likely have a cheaper purchase price, but lower fuel efficiency due to improved technology leading to efficiency gains in new vehicles with average fuel economy for new having increased 29% between 2004 and 2019 (US Environmental Protection Agency, 2023). The lower purchase price of used vehicles would likely be more attractive to farmers buying their first piece of machinery, transactions that would be overlooked in the data. Although secondhand transactions are not measured directly, imports are a good proxy for this since farmers that are buying new equipment from abroad would likely sell their used equipment to another farmer domestically, thereby making the secondhand market closely tied to the import data used.

4 Using variation in crude oil passthrough across countries to predict local diesel prices

Since local diesel prices are substantially influenced by the global price of crude oil, combined with there being large gaps in the national diesel price data with less than 35% of country-year entries containing a reading for diesel prices between 1991 and 2016, we can see if crude oil prices can help us estimate diesel prices by country. Despite being a large influence on price, countries vary in how elastic local prices are based on changes in global prices due to factors such as transportation costs and taxes. Therefore, we can look to measure each country's unique price passthrough from global crude prices to local diesel prices. Crude oil prices are available aggregated annually from the Energy Information Agency (EIA) between 1987 and 2022 for WTI and Brent crude prices. Brent crude was chosen for this specification due to its more widespread use over WTI. As shown in ***Figure 1***, oil prices were much lower in the 1990s, bottoming out in 1998 at under $13 per barrel, but rose sharply in the early 2000s with a spike first during the Financial Crisis in 2008 before reaching a maximum of almost $112 per barrel in 2012. Prices then fell again in 2014 before another spike in 2022 following supply chain disruption and Russia's invasion of Ukraine.

4.1 Empirical setting

To account for the exogenous variable of oil prices and to improve data availability, we can evaluate global crude oil prices as an instrument for local diesel prices by measuring how well diesel price correlates with Brent crude oil prices. We can use an annual average for crude oil prices as an instrument for local diesel pump prices as shown in Equation (7). In the equation, ln (P_{it}) represents the natural logarithm of predicted diesel pump price for country i at year t, and ln (C_t) represents the natural logarithm of the price of Brent crude oil at year t. Coefficients α_i and β_i represent slope and intercept values for a regression of Brent crude oil price on local diesel pump price.

$$(7) \quad \ln(P_{it}) = \alpha_i \ln(C_t) + \beta_i$$

In this specification, each country's coefficients, α_i and β_i, represents how sensitive local diesel prices are to fluctuations in the global crude oil price and a constant. We can then predict the price of diesel P_{it} for a given country each year, thereby allowing us to increase the number of readings from 2,747 to 7,665 between 1987 and 2022 after excluding countries with less than five actual diesel price readings to ensure accuracy.

4.2 Results

Figure 2 shows the distribution of calculated country-specific coefficients (α_i), averaging 0.59 and ranging from -0.197 in Bahrain to a high of 1.788 in Nigeria with the majority of values falling between 0.5 and 1. Applying the general form of equation (1) with the China in 2012 as an example, we see below that the predicted diesel price for that country-year pair was $1.22 per liter, comparable to the actual average diesel price of $1.28 measured by the World Bank.

$$\ln(P_{China,2012}) = \alpha_{China} \ln(C_{2012}) + \beta_{China}$$

$$\ln(P_{China,2012}) = 0.774 * 4.715 - 3.447$$

$$P_{China,2012} = \$1.22 \; per \; liter$$

The histogram in ***Figure 3*** shows a range of correlation coefficients for the local diesel price on global crude prices. Most countries have a correlation coefficient clustered between 0.6 and 1, with a mean of 0.75. As shown in ***Figure 4***, Chile tracks the best with global crude as $r^2 = 0.97$, whereas Netherlands represents the 50th percentile at $r^2 = 0.81$ and Belize represents the lowest value with $r^2 = 0.0003$. We can therefore see that our predicted diesel pump price tracks well with the actual values with only 11 countries having a correlation coefficient of less than 0.5, 10 of whom derive the majority of their export value from petroleum products. The exception is Sudan where oil only represents 8% of exports, but who's economy would likely have been significantly impacted by its split from South Sudan in 2011. The high correlation coefficients shown in ***Figure 3*** suggest that Brent crude oil price can be effectively used to estimate the local diesel price by country. By using crude prices, we therefore can analyze 2,054 price readings between 1987 and 2022 across 81 countries while maintaining reasonable reliability for the dataset.

5 Effects of diesel prices on mechanization

Using crude oil prices as an instrument for local diesel prices, we can investigate the impact of prices on the mechanization of farming. In this section, we will evaluate import data as a proxy for mechanization, merge the datasets of predicted diesel prices and import data, and measure the relationship between machinery imports and diesel prices using an OLS regression.

5.1 Mechanization data justification

To show how machinery imports are a good indicator of mechanization, we can actually look at a country's exports. Given that exports can be seen as an indicator of domestic production, we can look at a country's trade data to identify which goods are imported vs available locally. As shown in ***Figures 5 & 6***, we compare trade data of the developing Democratic Republic of the Congo where farming is largely small-scale and by hand to the similarly economically sized but developed Slovenia where farms are large, industrial enterprises. We see from the figures that the DRC does not export any machinery[4] but imports $2.21 billion per year due to a lack of local production. This contrasts with Slovenia being a net exporter of machinery, exporting $4.91 billion annually. This example illustrates the need of developing countries to import machinery if they are to mechanize their agriculture industry, thereby making a change in farming machinery imports a useful proxy to measure increasing farming mechanization.

5.2 Empirical setting

To investigate the effects of diesel prices on imports of mechanized farming equipment, we can run an OLS regression on the combined panel dataset as shown in

---

[4] Machinery described here refers to the category of 'Machinery, mechanical appliances, and parts.'

Equation (8). In the equation, ΔM_{it} represents the change value of imports per capita between time t and time $t-1$ in country i, Δp_{it} represents the change in predicted diesel price between time t and time $t-1$ in country i, calculated in the previous section, and α and β are the regression coefficients. We can also investigate the effect of time lags to account for factors such as changes in behavior of farmers taking time, looking to be assured that a change in fuel prices is not a short-term shock. Equation (9) shows the OLS regression equation with time lags for one, three, and five years also using the predicted local diesel prices.

$$(8) \quad \Delta \ln(M_{it}) = \alpha \Delta \ln(p_{it}) + \beta + \gamma_t + \epsilon_i$$

$$(9) \quad \Delta \ln(M_{it}) = \alpha_1 \Delta \ln(p_{it}) + \alpha_2 \Delta \ln\left(p_{i(t-1)}\right) + \alpha_3 \Delta \ln\left(p_{i(t-3)}\right) + \alpha_4 \Delta \ln\left(p_{i(t-5)}\right) + \beta + \gamma_t + \epsilon_i$$

In the regressions, we used country fixed effects (ϵ_i) to control for unobserved heterogeneity between countries that are unlikely to change over time such as cultural differences and national institutions. We also include time fixed effects (γ_t) to control for factors that change over time but have effects that span globally such as economic cycles and lowered barriers to international trade. Standard errors are clustered by country to account for the fact that readings within countries are likely correlated.

In addition to the regression shown in equation (8), we ran further regressions, firstly including the natural logarithm of GNI per capita growth as a variable, shown in equation (10) as g_t, where α_1 and α_2 are the regression coefficients. The idea is to control for domestic economic shocks unrelated to oil price shocks and not captured in time fixed effects due to their effects only being felt by a single country. An example such as GNI per capita falling 5.5% in Haiti following the devastating earthquake in 2010 (World Bank, 2022). We also ran the regression without countries that were net exporters of farming machinery since these countries would likely see opposite effects of changes in oil price to those countries increasing mechanization, which could affect the validity of the results.

$$(10) \quad \Delta \ln(M_{it}) = \alpha_1 \Delta \ln(p_{it}) + \alpha_2 \Delta \ln(g_{it}) + \beta + \gamma_t + \epsilon_i$$

5.3 Results

Running the regression described in equation (8), we see in ***Table 2***, regression (1) that the coefficient on the change in natural logarithm of diesel prices (α) is 0.552 and is statistically significant at the 1% significance level ($Z = 3.33$) with a constant value of -0.0946. This would mean that a unit increase in the logarithm of diesel prices would correspond with a 0.552 unit increase in the value of farming machinery imports per capita as shown in equation (11). The strong positive association found goes against the hypothesis of lower fuel prices incentivizing farmers to buy machinery due to lower operating costs, thereby suggesting that there are other factors at play when informing the decision of purchasing mechanized farming equipment.

$$(11) \quad \Delta \ln(M_{it}) = 0.552 * \Delta \ln(p_{it}) - 0.0946 + \gamma_t + \epsilon_i$$

Regressions (2-4) in ***Table 2*** show the regression coefficients for the one, three, and five-year time-lag values of predicted diesel prices, described as $\alpha_1, \alpha_2, \alpha_3$, and α_4 in equation (9). As shown in the table, the coefficient on the change in diesel prices at time t (α_1) remains statistically significant at no less than the 5% significance level for all regressions, but the magnitude of the coefficient decreases slightly from 0.552 to 0.447. In regressions (2-4) of ***Table 2***, none of the coefficients on the change in diesel price variables are individually statistically significant. Upon running a joint test on coefficients α_2, α_3, and α_4 to investigate the possibility of joint significance, we find that $p > F = 0.777$, and is therefore statistically insignificant, even when considered together. We have therefore not included the time-lag variables in further regression tests in the study.

The regression results in ***Table 3*** show the coefficients when controlling for net exporters in regression (1). Regression (1) shows that the value of α in equation (8) is still statistically significant at the 1% significance level, with the value now calculated as 0.735 ($Z = 3.77$). This means that the positive association found in ***Table 2*** is even stronger when excluding net exporters of farming machinery with every logarithmic unit increase of the predicted diesel price, machinery imports per capita increase by 0.735 logarithmic units. This result therefore strengthens the conclusion that higher diesel prices are associated with increasing imports of farming machinery per capita.

The second regression in ***Table 3*** is based on the specification in equation (10), including GNI per capita growth as an endogenous variable not accounted for in either time of country fixed effects. We see from regression (2) results that the coefficient on predicted diesel prices (α_1) is 0.448 ($Z = 2.31$) and is still statistically significant at the 5% significance level. The coefficient on GNI per capita was found to be 0.338 but is not statistically significant at the 5% significance level $Z = 1.86$.

Based on the results in ***Tables 2 & 3***, we found that there is a statistically significant, positive relationship between the year-on-year change in predicted diesel price in a country and the value of farming machinery imports per capita. When excluding net exporters of mechanized farming equipment, we see that both the magnitude and significance of this relationship increases. We chose to exclude net exporters since the impact felt by these countries from changes in diesel price would have the opposite effect because an increase in global imports of farming machinery would represent income rather than an expense to these countries. This income effect from significant exports could interfere with the results when measuring the impact of diesel prices alone. The testing of time lags and GNI per capita as endogenous variables yielded insufficient statistical evidence to be further considered.

6 Falsification test

Given the unexpected results of the regression detailed in the previous section, we decided to run a falsification test to determine whether or not the results were a result of an issue with the empirical strategy. If the results are statistically significant in a comparable manner to the regression on farming machinery, it points to an error in empirical strategy. However, if the results are not statistically significant, it would suggest a real positive relationship between diesel prices and the purchase of mechanized farming equipment.

6.1 Empirical setting

To conduct this test, we will switch the dependent variable of farming machinery imports with textile[5] imports and then with precious metal imports[6] as shown in equation (12) where M_{it} represents imports per capita, otherwise using the same regression specification as before. These products were chosen since they should have very little covariance with diesel prices and consequently should show no association under the null hypothesis using the same regression specification as previously. Time and country fixed effects were used again to control for heterogeneity over geography and over time.

$$(12) \quad \Delta \ln(M_{it}) = \alpha_1 \Delta \ln(p_{it}) + \beta + \gamma_t + \epsilon_i$$

6.2 Results

Table 4 shows the OLS regression results for the falsification test with regression (1) measuring the effect on textile imports, and regression (2) measuring the effect on the value

[5] Defined as Harmonized System codes HS50-63 by the World Customs Organization (WCO)
[6] Defined as Harmonized System code HS71 by the World Customs Organization (WCO)

of precious metal imports with the same regression specifications as in the previous model. The coefficient of change in diesel prices is not statistically significant at the 5% significance level for either textile or precious metal imports with $Z = 0.53$ and $Z = 1.64$ respectively, considerably below the critical value of $Z = 1.96$. It can be therefore suggested that there is insufficient evidence for a relationship between changes in national diesel prices and the value of either textile or precious metal imports in a given year.

The results of the falsification test lend more credibility to the positive association found in the regression of diesel prices on farming machinery imports and significantly reduces the likelihood of the relationship being because of an error in econometric analysis. These findings therefore suggest an alternative, exogenous factor explaining why farmers appear to purchase more machinery as operating costs of such machinery increase.

7 Discussion

In theory, we would expect to have found a negative relationship between the change in national diesel prices and the value of farming machinery imports, however, based on the analysis conducted in the previous sections, we have found the opposite. Therefore, there must be an outside factor to explain the phenomenon of the presumed rising operational cost of equipment being associated with increased purchase of that equipment. In this section we will discuss two other hypotheses alluded to in Section 2 in the relationship between diesel prices and cereal crop prices, and the benefits of potentially improving fuel efficiency of equipment.

7.1 Relationship between oil prices and cereal crop prices

Looking at diesel prices, we explored one half of the financial equation for farmers in the form of costs, but revenue is equally important in farmer profits and subsequently the financial means to purchase machinery. What if diesel prices were also related to the revenue that farmers brought in, which could have the potential to offset higher operational costs? To measure this, we can look at the association between oil prices and prices for crops commonly grown by farmers.

By far the most abundant crops are cereal crops with production at 3.1 billion tons in 2021, dominated by maize (corn), rice, and wheat, which represent 92% of all cereal crops grown (Food and Agriculture Organization, 2022). Given the prevalence, low value density, and easy storage of these grains mean that prices are highly commodified in the global market. We can therefore use monthly data from the Federal Reserve Economic Data (International Monetary Fund, 2024) to measure the correlation between global Brent Crude prices and global grain prices for each of the three primary cereal crops.

As shown in ***Figure 7***, we see a strong, positive correlation between the global price of oil and the price of maize, rice, and wheat with R-values of 0.83, 0.75, and 0.79 respectively, well above the critical value of 0.097. Similar to crude oil, prices dipped to a low in the late 1990s before rising sharply in the 2000s and again in the early 2010s where maize reached over 270% of 1990 prices at $311 per ton in 2011. Prices then fell with oil prices in 2014 before aggressive price growth in 2022. Despite oil prices generally being more volatile, the only notable anomaly is the low price of rice in 2022 despite high prices for oil, maize, and wheat. This can likely be explained by large rice exporters such as India, representing 34% of global rice exports, (Observatory of Economic Complexity, 2022) maintaining economic ties with Russia following the invasion of Ukraine, thereby keeping oil prices lower than the global average.

This high correlation shows that, even though high oil prices may lead to high diesel pump prices within a country, this could be outweighed by farmers being able to sell grain at higher prices. However, similar to our findings in heterogeneity of pass through from crude oil prices to local diesel prices, we don't know the pass through from global to local grain prices. In countries with significant pass through, higher prices outweighing costs would mean that farmer profits would actually increase when diesel prices rise, thereby explaining how farmers could be more inclined to purchase machinery at this time.

7.2 Fuel efficiency of equipment

Since we used machinery import data as a proxy for mechanizing, we likely captured primarily the sale and purchase of new equipment rather than that on the secondhand market. As discussed previously, vehicle fuel economy has improved by 29% in just 15 years (Environmental Protection Agency, 2023) so purchasing new equipment would theoretically offset up to 23% of fuel costs. We see this trend in the automotive industry where a $1 increase in gasoline prices corresponds with the market share of the top quartile of vehicles by fuel economy increasing by 21.1% (Busse et al., 2011).

For tractors specifically, we can see how much farmers could potentially save by buying new, more efficient equipment to evaluate whether this is likely to happen in our data when fuel prices rise. Typically fuel costs account for 69% of tractor operating costs and 33%

of total costs (Edwards, 2015). Buying used machinery would lower ownership costs at the expense of higher fuel and maintenance costs. Cumulative maintenance costs increase exponentially over time to roughly 70% of vehicle purchase price after 10,000 hours of use (about 25 years). This would mean that buying a new tractor would save approximately $7 per hour of operation in maintenance and $6 per hour in fuel costs. At these saving rates, it would take over 20 years to offset the higher upfront cost of new machinery and even a $1 per liter increase in diesel price[7] would only cut the payback period down to 13.5 years, which is almost certainly unaffordable for farmers. Therefore, it is unlikely that fluctuations in diesel price substantially changed the distribution of new versus used equipment purchased.

8 Conclusion

As the preceding analysis has shown, we can draw meaningful conclusions regarding the use of crude oil prices as an instrument to predict national-level diesel pump prices, and regarding the association between these diesel prices and farming machinery imports in a country.

Our first conclusion in the study is that using Brent Crude oil prices can predict the local diesel prices in a country with reasonable accuracy. Countries vary in how sensitive their local diesel prices are to changes in the global oil price, but these seem to be relatively consistent over time with most countries situated near the global average of 0.63 with a 95% confidence interval of (0.58, 0.69). These country-specific elasticity coefficients tracked extremely well with the actual diesel prices reported by the World Bank as shown by the average 0.72 correlation coefficient with a 95% confidence interval of (0.69, 0.76).

We were also able to conclude that there is a positive and statistically significant association between changes in local diesel prices and national imports of farming machinery for a given year. The base regression found the coefficient on logarithmic change in diesel prices to be 0.552 and was largely unaffected by incorporating time lags in the regression model. The statistical significance of the relationship persists when accounting for national fluctuations in GNI per capita growth and the magnitude increases substantially to 0.735 when excluding countries that are net exporters of farming machinery. The falsification test finding no evidence of a statistically significant association of diesel prices and imports per

[7] This theoretical year-on-year increase is greater than every country-year pair in the dataset (maximum increase = $0.93 in Nigeria, 2021)

capita of textiles and precious metals further strengthens the conclusion of positive association in the regressions.

While we are not able to conclusively state the reason for the relationship between diesel prices and mechanization, the data suggests that the primary reason is that farmers benefit from increased income during times of high diesel prices due to the high correlation between grain commodity prices and crude oil prices. Since the majority of agriculture is dedicated towards highly commodified goods, both in crops and animal products. The high correlation coefficients of 0.83, 0.75, and 0.79 for maize, rice, and wheat respectively therefore strongly suggest that farmer incomes, and therefore ability to mechanize, are tied to global crude oil prices. We found that this correlation is more likely than farmers choosing to purchase new equipment for fuel efficiency purposes given the extremely long payback period required.

This study provided useful insights into the relationship between diesel price changes and the adoption of farming mechanization. Future research could build on this by performing a national-level case study investigating the policy determinants of mechanization similar to Ghosh (2010), in a country that has dramatically increased mechanization. South Korea would be a notable example, a country which has increased horsepower per unit of land from 0.0008 in 1961 to 9.98 in 2019 or the 3rd highest in the world (USDA, 2023). Work could also be done to further quantify how farmer incomes change with global crude oil prices to better identify when farmers are able to afford capital to increase agricultural yields.

Tables

Table 1: Countries represented in combined dataset by geographic region and income group

	Countries represented	Proportion of global total
East Asia & Pacific	14	47%
Europe & Central Asia	40	75%
Latin America & Caribbean	21	62%
Middle East & North Africa	13	62%
North America	1	50%
South Asia	5	63%
Sub-Saharan Africa	29	60%
High income	42	69%
Upper middle income	33	61%
Lower middle income	37	69%
Low income	11	42%

Table 2: OLS regression results of predicted diesel prices on farming machinery imports

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Change in log diesel prices	0.552***	0.527***	0.499***	0.447**
	(0.166)	(0.169)	(0.164)	(0.198)
Change (t-1)		-0.0592	-0.0325	-0.0305
		(0.173)	(0.178)	(0.153)
Change (t-3)			0.0377	0.0733
			(0.152)	(0.122)
Change (t-5)				-0.159
				(0.118)
Constant	-0.0946	-0.0696	-0.0673	-0.0351
	(0.0612)	(0.0766)	(0.0737)	(0.0802)
Observations	3,136	3,014	2,772	2,532
R-squared	0.069	0.064	0.072	0.081
Number of countries	122	121	120	119
Country FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: OLS regression results when (1) excluding net exporters and (2) accounting for GNI fluctuations

VARIABLES	(1) OLS	(2) OLS
Change in log diesel prices	0.735***	0.488**
	(0.195)	(0.211)
GNI per capita growth		0.338*
		(0.182)
Constant	-0.170*	-0.00967
	(0.0869)	(0.0720)
Observations	1,974	2,936
R-squared	0.061	0.082
Number of countries	79	120
Country FE	YES	YES
Time FE	YES	YES
Net exporters excluded	YES	NO

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

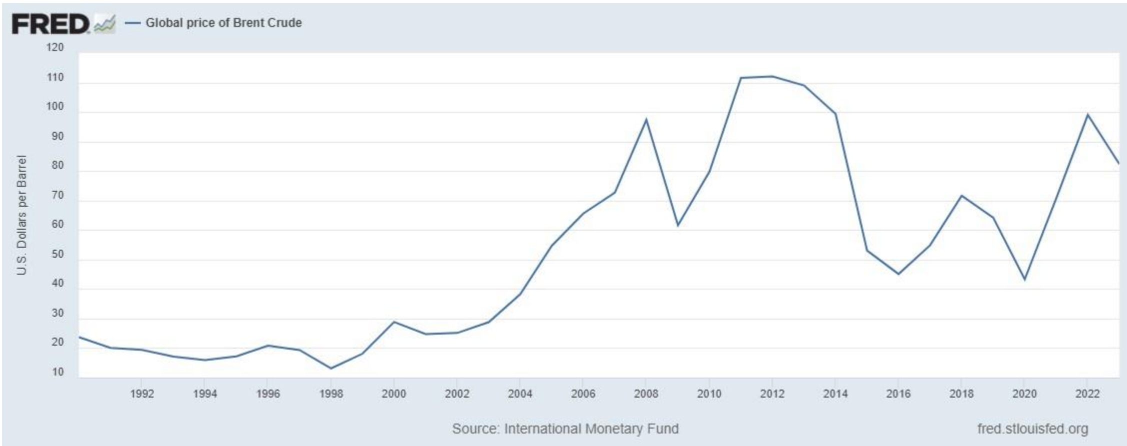
Table 4: OLS falsification test regression table of (1) Textiles and (2) Precious Metals

VARIABLES	(1) OLS	(2) OLS
Change in log diesel prices	0.0467	0.360
	(0.0879)	(0.220)
Constant	0.124***	0.148
	(0.0287)	(0.136)
Observations	2,596	2,518
R-squared	0.105	0.037
Number of countries	119	118
Country FE	YES	YES
Time FE	YES	YES

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figures

Figure 1: Annual mean price of Brent Crude oil 1987-2023



Retrieved from: Federal Reserve Economic Data at
https://fred.stlouisfed.org/series/POILBREUSDM

Figure 2: Histogram of country-specific coefficients

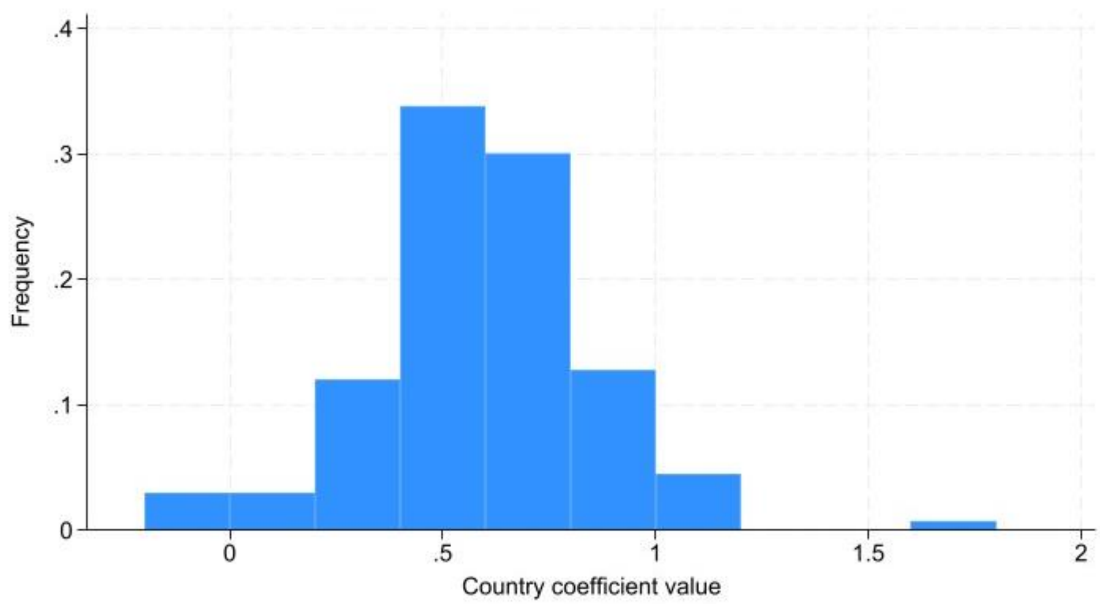


Figure 3: Histogram of R-squared values between crude oil prices and domestic diesel prices by country

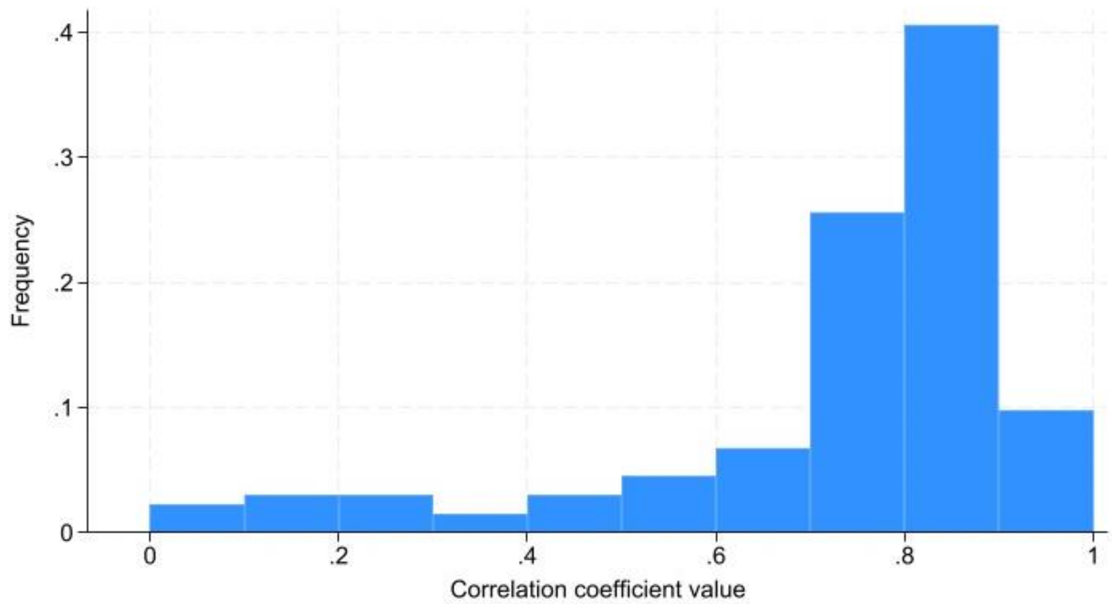


Figure 4: Examples of predicted diesel prices for three countries with varying levels of pass through from global crude oil prices (Chile = highest, Netherlands = median, Belize = lowest)

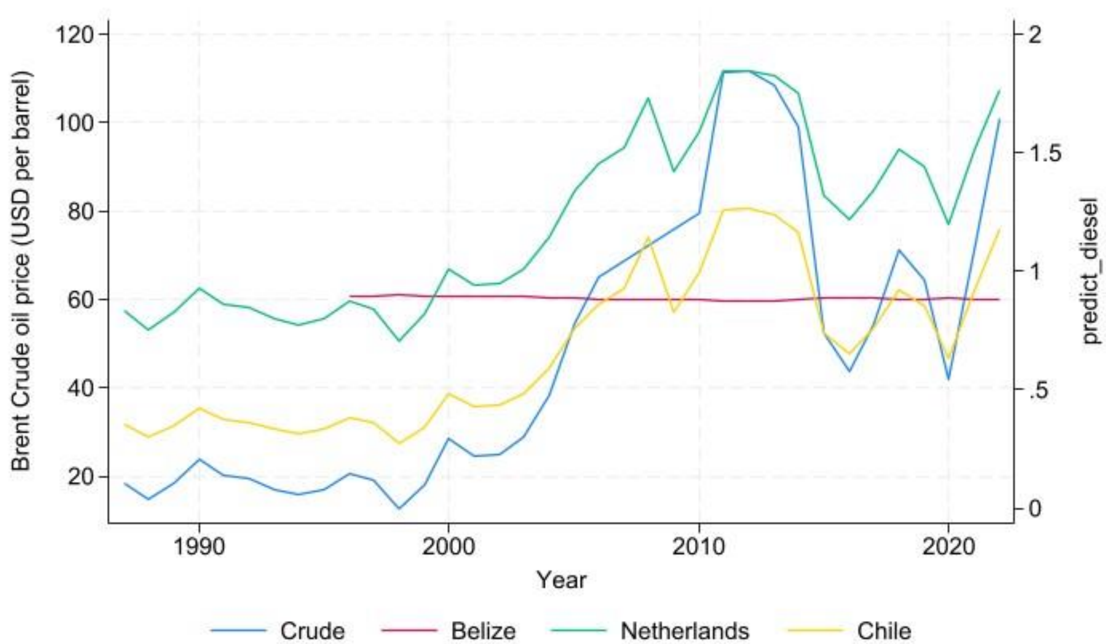
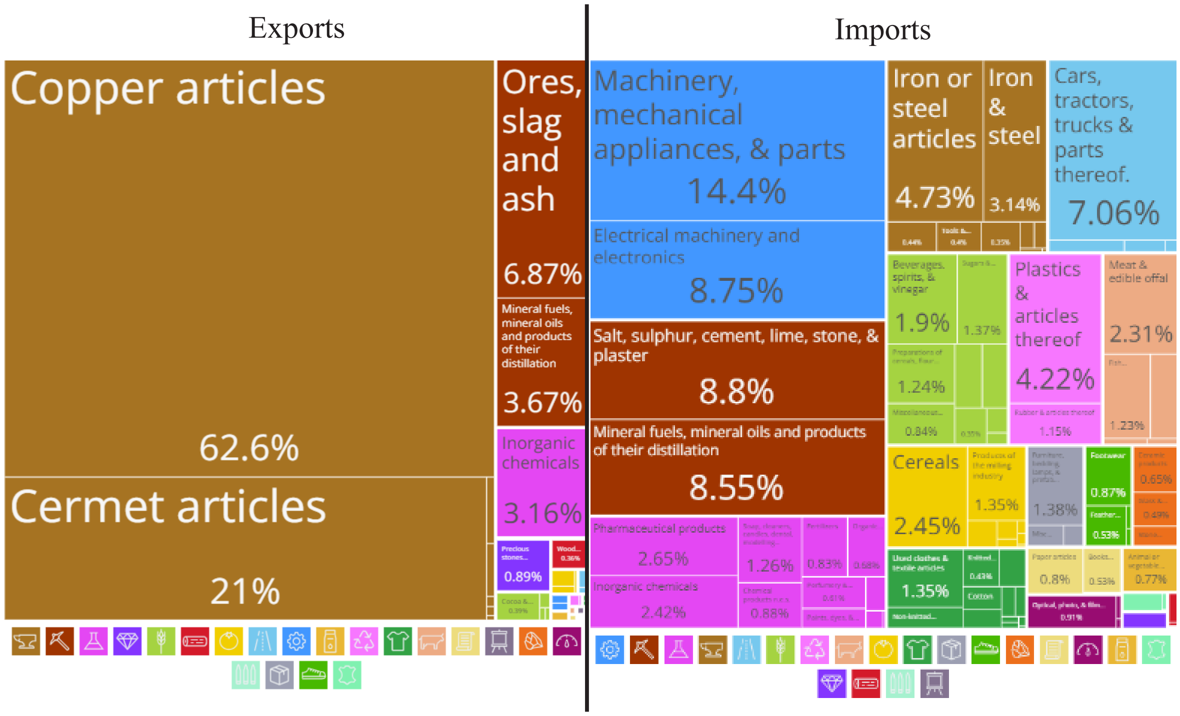
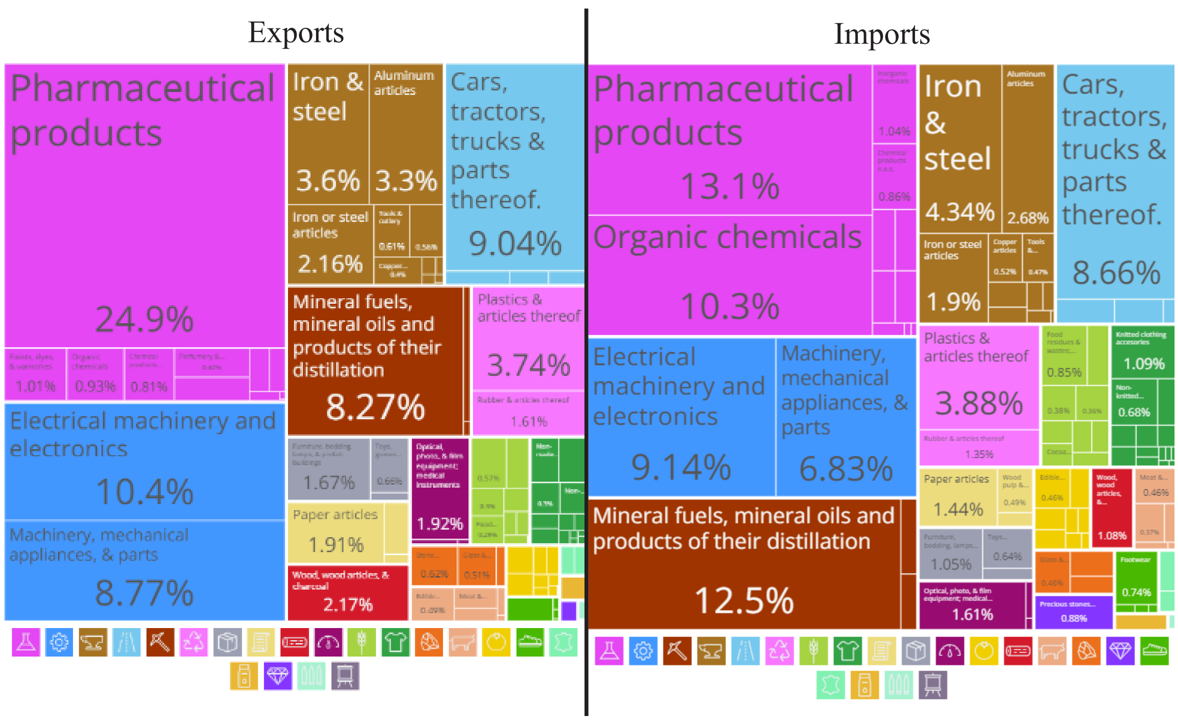


Figure 5: Trade breakdown for Democratic Republic of the Congo, 2022



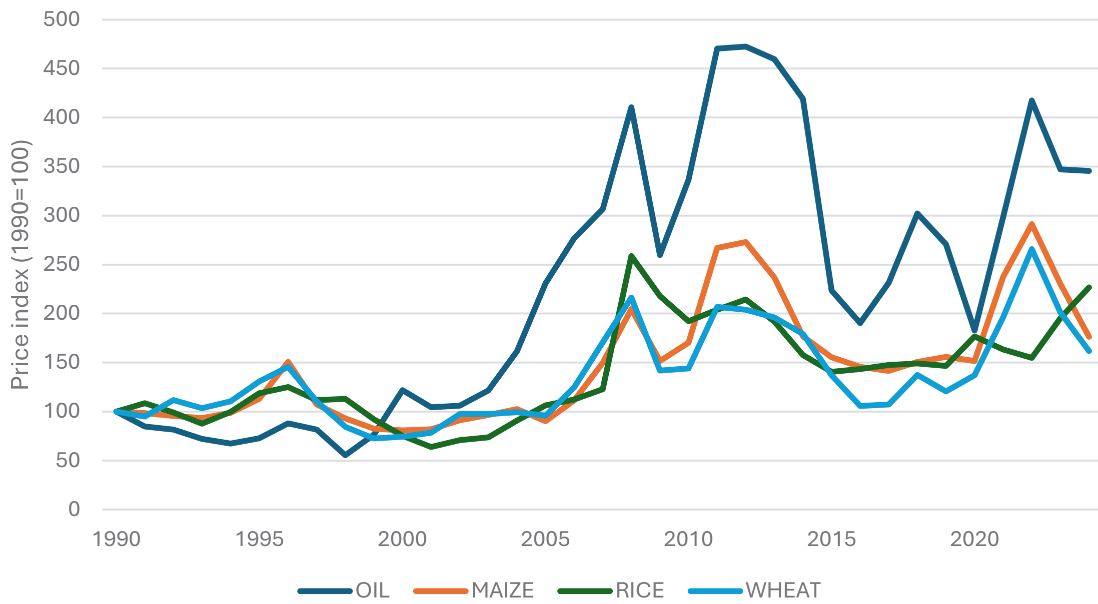
Retrieved from: Observatory of Economic Complexity at
https://oec.world/en/profile/country/cod

Figure 6: Trade breakdown for Slovenia, 2022



Retrieved from: Observatory of Economic Complexity at
https://oec.world/en/profile/country/slo

Figure 7: Annual global price index of Brent Crude oil and the three most abundant cereal crops (maize, corn, and wheat)



References

American Automobile Association. (2023). *AAA Gas Prices*. https://gasprices.aaa.com/state-gas-price-averages/

Banafa, B. (2019, April 9). *Mechanizing agriculture is key to food security*. Africa Renewal. https://www.un.org/africarenewal/magazine/april-2019-july-2019/mechanizing-agriculture-key-food-security

Bautista, E. G., Kim, J., Kim, Y., & Panganiban, M. E. (2017). Farmer's Perception on Farm mechanization and Land reformation in the Philippines. *Journal of the Korean Society of International Agricultue*, *29*(3), 242–250. https://doi.org/10.12719/KSIA.2017.29.3.242

Biering, D. (2024). *History of Agriculture Equipment: Important Developments and Examples*. https://www.tstar.com/blog/history-of-agriculture-equipment-important-developments-and-examples

Broadberry, S., Campbell, B., Klein, A., Overton, M., & Van Leeuwen, B. (2014). *British Economic Growth, 1270–1870* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781107707603

Busse, M. R., Knittel, C. R., & Zettelmeyer, F. (2011). *Pain at the Pump: The Effect of Gasoline Prices on New and Used Automobile Markets*. https://www.kellogg.northwestern.edu/~/media/Files/Faculty/Research/ArticlesBookChaptersWorkingPapers/gaspaper110506.ashx

Caunedo, J., & Kala, N. (2022). Mechanizing Agriculture. *Working Paper 29061*. http://www.nber.org/papers/w29061

Crops and livestock products. (2023). [..Csv]. Food and Agriculture Organization of the United Nations. https://www.fao.org/faostat/en/#data/QCL

Edwards, W. (2015, May). *Estimating Farm Machinery Costs | Ag Decision Maker*. Iowa State University Extension and Outreach. https://www.extension.iastate.edu/agdm/crops/html/a3-29.html

Emami, M., Almassi, M., Bakhoda, H., & Kalantari, I. (2018). Agricultural mechanization, a key to food security in developing countries: Strategy formulating for Iran. *Agriculture & Food Security*, *7*(1), 24. https://doi.org/10.1186/s40066-018-0176-2

Employment in Agriculture (SL.AGR.EMPL.ZS). (2024). [..Csv]. World Bank. https://data.worldbank.org/indicator/SL.AGR.EMPL.ZS

Fleet News. (2024, April). *Regional fuel prices*. https://www.fleetnews.co.uk/costs/fuel-prices/

Foster, A. D., & Rosenzweig, M. R. (2023). *Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution*.

Ghosh, B. K. (2010). Determinants of Farm Mechanisation in Modern Agriculture: A Case Study of Burdwan Districts of West Bengal. *International Journal of Agricultural Research*, *5*(12), 1107–1115. https://doi.org/10.3923/ijar.2010.1107.1115

GNI per capita, PPP. (2022). [dataset]. World Bank. https://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD

Gollin, D., Lagakos, D., & Waugh, M. E. (2014). The Agricultural Productivity Gap. *The Quarterly Journal of Economics*, *129*(2), 939–993. https://doi.org/10.1093/qje/qjt056

Gollin, D., Parente, S., & Rogerson, R. (2002). *The Role of Agriculture in Development*. *92*(2).

Grigg, D. B. (1975). The World's Agricultural Labour Force 1800-1970. *Geography*, *60*(3), 194–202.

IMF Primary Commodity Prices. (2022). [dataset]. https://www.imf.org/en/Research/commodity-prices

Kirui, O. (2019). The Agricultural Mechanization in Africa: Micro-Level Analysis of State Drivers and Effects. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3368103

Lavagne d'Ortigue, O. (2022). Agricultural production statistics 2000–2021. *Food and Agriculture Organization*.

McNeil, I. (Ed.). (2003). *An encyclopaedia of the history of technology* (Transferred to digital printing). Routledge.

Moscona, J. (n.d.). *Agricultural Development and Structural Change, Within and Across Countries*.

National Geographic Society. (2024, January). *The Development of Agriculture*. https://education.nationalgeographic.org/resource/development-agriculture

Peng, J., Zhao, Z., & Liu, D. (2022). Impact of Agricultural Mechanization on Agricultural Production, Income, and Mechanism: Evidence From Hubei Province, China. *Frontiers in Environmental Science*, *10*, 838686. https://doi.org/10.3389/fenvs.2022.838686

Pump price for diesel fuel (US$ per liter) (EP.PMP.DESL.CD). (2016). [.Csv]. World Bank. https://data.worldbank.org/indicator/EP.PMP.DESL.CD

TFP indices ad components for countries. (2023). [.Xlsx]. United States Department of Agriculture Economic Research Service. https://www.ers.usda.gov/data-products/international-agricultural-productivity

The 2023 EPA Automotive Trends Report: Greenhouse Gas Emissions, Fuel Economy, and Technology since 1975,. (2023). US Environmental Protection Agency. https://www.epa.gov/automotive-trends/download-automotive-trends-report

The Observatory of Economic Complexity. (2022). [dataset]. https://oec.world/en/

UN Comtrade Trade Data. (2022). [dataset]. United Nations Comtrade. https://comtradeplus.un.org/

United Nations. (2022). *World Population Prospects 2022* (3). United Nations, Department of Economic and Social Affairs, Population Division. https://population.un.org/wpp/Publications/

Weekly U.S. No 2 Diesel Retail Prices (Dollars per Gallon). (2024). [dataset]. US Energy Information Administration. https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMD_EPD2D_PTE_NUS_DPG&f=W

Shuffling the Deck: Did NYC's Lottery Admissions Reshape School Demographics?

Sonia Seliger

MIT Department of Economics

This paper investigates the impact of New York City's 2021 high school admissions policy reform, which replaced selective admissions at top-ranked schools with a lottery-based system. The study employs a difference-in-difference approach to analyze changes in racial and socioeconomic diversity across the city's public high schools, focusing on screened schools. While the findings suggest that the policy did not lead to significant demographic shifts, the results raise important questions about educational equity and systemic barriers.

1. INTRODUCTION

New York City, often referred to as a "melting pot", is one of the most diverse cities in the world. Yet, its public high school system, serving nearly 300,000 students, has historically been characterized by stark racial and socioeconomic segregation. Much of this segregation stems from the reliance on selective admissions criteria and geographic district priorities. In 2021, the city implemented a major reform to address these inequities, eliminating district and all other geographic priorities and requiring selective, top-ranked schools to adopt a lottery-based admissions system. Public perception was split: some viewed the change as a step towards reducing systemic inequalities, while others argued it compromised academic meritocracy and school quality.

This research examines the impact of the 2021 admissions policy change on the demographic composition of students attending New York City's most highly ranked screened schools. Specifically, it addresses the question: how did the 2021 transition to a lottery-style admissions process affect the racial and socioeconomic makeup of students in these schools? By focusing on equity over traditional academic outcomes, this study provides a novel approach to evaluating the effectiveness of the policy in promoting diversity and access to high-quality education. Given the deep-rooted patterns of segregation within the NYC public school system, understanding the implications of this policy change is essential for assessing its success in promoting educational equity.

2. INSTITUTIONAL BACKGROUND AND 2021 ADMISSIONS CHANGE

The NYC public high school admissions process is highly complex, determined by factors such as the order of program rankings on a student's application, seat availability, and the admissions method of each program. Students rank up to 12 programs, and are first considered for their top-choice program. If they do not receive an offer for that program, they are then considered for their second choice, and so on, until a match is made.

Prior to 2021, approximately half of NYC high schools had geographic priorities, such as district or borough residency requirements. For example, Manhattan's District 2 schools were known for prioritizing local students, creating disparities in access for those outside the district. After the 2021 reform, geographic priorities were largely removed, leaving no zoned schools in Manhattan, two in the Bronx, five in Staten Island, six in Brooklyn, and fifteen in Queens.

NYC high schools use three main admissions methods: screened programs, educational option (ed opt), and open admissions.

- **Screened programs**: Before 2021, these programs were highly selective, admitting students based on academic criteria such as grades, standardized test scores, and sometimes essays or interviews. After the reform, screened schools switched to a lottery system. Applicants are grouped into five tiers based on seventh-grade GPA, with Group 1 (highest GPA) given priority until seats are filled. While this method maintains some preference for high-performing students, it eliminates other screening criteria and, more importantly, geographic restrictions, allowing gifted students from any neighborhood, including lower-income areas, to compete for seats at top schools citywide.

- **Educational option (ed opt) programs**: Introduced in the 1970s, ed opt programs aim for a balanced academic mix by admitting students across high, middle, and low performance categories, with admissions partially determined by lottery.

- **Open admissions programs**: These programs do not use any screening criteria and rely solely on random selection. They often serve as a fallback option for students not admitted to more selective programs.

Some schools, especially larger ones, admit students through multiple programs (e.g. screened, ed opt, open, and zoned admissions). These schools still retain zoned priority in limited cases.

The topic of selective schools and issues of equity in access to education has been a significant topic of debate and research. Corcoran and Baker-Smith (2018) highlighted how geographic-based admissions preferences in New York City exacerbated socioeconomic divides by restricting access to high-quality schools for students outside of priority districts. These policies often mirrored residential segregation, as district lines closely followed patterns of racial and socioeconomic divides across neighborhoods. High-performing schools were frequently located in wealthier districts, predominantly serving affluent, predominantly white student populations. Meanwhile, students from lower-income areas - who were more likely to be Black, Hispanic, or from immigrant backgrounds - were systematically excluded from these schools due to geographic restrictions. This segregation reinforced disparities in access to educational resources, experienced teachers, and advanced coursework opportunities.

The 2021 reform aimed to reduce these inequities by removing district-based admissions, allowing all students, regardless of where they lived, to apply to any school in the city. By eliminating these geographic barriers, the policy sought to create a more equitable admissions process and expand access to high-quality schools for students from historically marginalized communities.

Academic opinions provide support for a lottery-based admissions system. Schwartz (2005) and Conley (2012) both advocate for lottery-based systems to reduce bias and competition, arguing that such systems create fairer opportunities while maintaining academic thresholds. For example, Schwartz's proposal for random selection among the top fifth of applicants mirrors some aspects of NYC's current five-group lottery system.

Other studies question the academic value of selective admissions. Abdulkadiroğlu, Angrist, and Pathak (2014) found little evidence that attending elite exam schools significantly improved student academic outcomes. My research takes a different focus, investigating whether the 2021 shift to lottery-based admissions impacted racial and socioeconomic diversity, particularly in NYC's most selective schools. This approach emphasizes equity and access rather than solely academic outcomes.

3. DATA

I obtained data on student racial and socioeconomic demographics from the NYC Public Schools InfoHub. This dataset includes aggregated information for each public high school from the 2019-2020 to 2022-2023 school years. For each school, the dataset provides details such as total enrollment and the percentage breakdown of students by race. Racial groups are categorized into six groups: Asian, Black, Hispanic, Native American, Native Hawaiian and Pacific Islander, and White.

The dataset also includes the Economic Need Index (ENI) for each school, which estimates the percentage of students facing economic hardship. The ENI is calculated as 1.0 for students eligible for public assistance from the NYC Human Resources Administration, those who have lived in temporary housing within the last four years, or students who entered NYC schools for the first time in the past four years and have a home language other than English. For all other students, the ENI is based on the percentage of families with school-age children in the student's census tract whose income is below the poverty level, as estimated by the American Community Survey. Schools with a higher ENI score generally serve a poorer student population. The ENI is used as a proxy for socioeconomic diversity, with a higher ENI indicating a greater concentration of students from lower-income families. By incorporating this measure, the ENI provides a way to assess the degree to which schools are serving economically diverse student populations, which is essential for understanding the broader impact of the policy change aimed at improving access to education for disadvantaged groups.

The initial dataset included 481 public schools, of which 62 charter schools were excluded because they follow a separate admissions process. My analysis focuses on schools that students could reasonably be indifferent with attending, particularly given the 2021 admissions process shift. To refine this focus, I removed schools that the "average" student would not realistically consider.

For instance, schools specializing in a specific field, such as *Food and Finance High School* or *High School for Violin and Dance* were excluded. While these schools often provide a balanced curriculum, their focus on niche subjects makes them less likely choices for students not interested in those fields. Furthermore, these schools were likely not significantly impacted by the policy changes, as students applying to them tend to do so because of a specific interest in the school's specialization. This means the 2021 admissions reform, which targeted broader accessibility and equity across general screened schools, likely had little effect on the applicant pools of such niche schools. Students interested in a specific subject, such as finance or the arts, would still prioritize applying to these schools regardless of the new admissions framework.

Additionally, I removed most audition-based schools focused on theater, art, or music, as well as single-gender schools, alternative and credit-recovery schools, international schools with a high percentage of English Language Learners, and schools that offer degrees or focus on career preparation (e.g. *High School for Fire and Life Safety* and *High School for Construction Trades*). These schools cater to specialized student populations and are less relevant to understanding the impact of the 2021 admissions change. A complete list of excluded schools can be found in the Appendix.

Table 1: Location and Admissions Method Dataset Makeup

	All Boroughs	**Manhattan**	**Bronx**	**Brooklyn**	**Queens**	**Staten Island**
All Admissions Types	190	42	48	50	40	10
Screened	39	21	6	8	4	0
Ed Opt	80	13	28	23	14	2
Open	20	3	5	6	5	1
Specialized	8	2	2	2	1	1
Multiple Programs	43	3	7	11	16	6

Table 1 summarizes the schools included in the final dataset, broken down by borough and admissions method. The table shows that Manhattan has the highest concentration of screened schools, reflecting its historical dominance in selective admissions. Meanwhile, other boroughs, particularly the Bronx and Staten Island, have a higher proportion of ed opt or multiple-program schools, indicating more varied approaches to admissions. This distribution highlights the geographical and programmatic diversity within the NYC school system and sets the stage for analyzing how the 2021 admissions change may have affected equity and diversity.

4. SUMMARY STATISTICS TABLES

Table 2: 2019-2020 Summary Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum	Number of Observations
Total Enrollment	893	1,029	140	6,040	190
Student % Asian	12.91%	15.25%	0%	82.30%	190
Student % Black	29.39%	23.74%	0.40%	88.40%	190
Student % Hispanic	41.33%	23.35%	3.50%	91.90%	190
Student % Native American	1.13%	1.16%	0%	8.00%	190
Student % Hawaiian/Pacific Islander	0.55%	0.53%	0%	3.50%	190
Student % White	12.89%	16.49%	0%	79.30%	190

Table 3: 2020-2021 Summary Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum	Number of Observations
Total Enrollment	892	1,036	142	5917	190
Student % Asian	13.09%	15.35%	0%	82.80%	190
Student % Black	28.70%	23.05%	0.70%	86.90%	190
Student % Hispanic	41.65%	23.13%	3.80%	93.10%	190
Student % Native American	1.14%	1.16%	0%	7.90%	190
Student % Hawaiian/Pacific Islander	0.58%	0.52%	0%	3.30%	190
Student % White	12.73%	15.97%	0%	76.00%	190

Table 4: 2021-2022 Summary Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum	Number of Observations
Total Enrollment	866	1,002	115	5957	190
Student % Asian	13.33%	15.37%	0%	84.00%	190
Student % Black	27.92%	22.37%	0.70%	86.20%	190
Student % Hispanic	42.41%	22.79%	3.80%	92.70%	190
Student % Native American	1.08%	1.14%	0%	8.30%	190
Student % Hawaiian/Pacific Islander	0.57%	0.51%	0%	3.20%	190
Student % White	12.24%	14.99%	0%	72.50%	190

Table 5: 2022-2023 Summary Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum	Number of Observations
Total Enrollment	852	988	110	5942	190
Student % Asian	13.26%	15.35%	0%	81.70%	190
Student % Black	27.68%	22.10%	0.80%	88.00%	190
Student % Hispanic	43.32%	22.36%	4.10%	92.70%	190
Student % Native American	1.09%	1.12%	0.00%	6.30%	190
Student % Hawaiian/Pacific Islander	0.57%	0.55%	0%	2.80%	190
Student % White	12.02%	14.36%	0.30%	66.90%	190

To start my analysis, I reviewed the summary statistics for my dataset. *Tables 2 through 5* above provide an overview of the racial breakdown for each school year. Among the 190 schools included, there were 169,745 students in the 2019-2020 school year, 169,429 in 2020-2021, 164,479 in 2021-2022, and 162,033 in 2022-2023. This decline in enrollment, particularly after 2021, reflects broader trends in New York City public schools, where overall student enrollment has been steadily decreasing in recent years and is not specific to the schools in this dataset.

The high standard deviations in the measures reflect the considerable variation in school size and racial composition. Some schools serve a few hundred students, while others have enrollments in the thousands. Similarly, racial demographics differ widely, with some schools serving predominantly one racial group and others being more diverse. This variability highlights the diverse nature of NYC public schools and emphasizes the importance of disaggregating the data to understand specific trends and changes over time.

5. POLICY CHANGE EFFECT ON RACIAL COMPOSITION

To measure the impact of the 2021 admissions change, I used a difference-in-difference approach. The treated group in my analysis is screened schools, which have historically been the most selective and "desired" schools among students. After the admissions change, anecdotally, screened schools "filled up" first and many of them, especially those in Manhattan, have had a significant change in their demographic composition. For the control group, I grouped all other school types together. The years before the policy change (2019-2020 and 2020-2021) are the pre-treatment period, and the years after the change (2021-2022 and 2022-2023) are the post-treatment period The outcome variables examined are the proportion of students from various racial groups and the average ENI at these schools.

To explore whether the racial composition of schools changed due to the 2021 admissions policy change, I first compared trends in racial demographics at screened and non-screened schools, illustrated in *Figure 1*. Prior to the 2021 policy change, the trends in racial composition appear parallel for both groups, satisfying the parallel trends assumption necessary for the difference-in-difference methodology.

One minor exception is the percentage of Hispanic students, which decreased slightly from 2019 to 2020 in screened schools but increased marginally in non-screened schools. However, this divergence is small and likely reflects random variation rather than a systemic difference. Post-policy, the trends for each racial group remain relatively consistent in non-screened schools, while screened schools show some slight shifts - for example, a gradual increase in the percentage of Asian students and a decrease in White students.

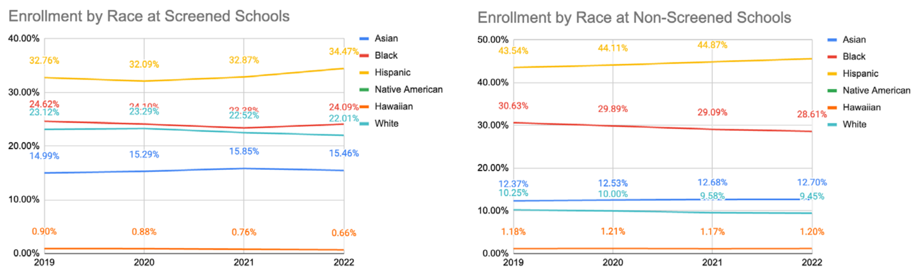


Figure 1: Enrollment by Race at Screened and Non-Screened Schools

6. EMPIRICAL ANALYSIS

To test these trends empirically, I conducted a difference-in-difference regression to assess whether the racial makeup of screened schools changed after the policy change. For example, in the case of Asian students, I used the following specification:

$$student\_pct\_asian_{it} = \beta_0 + \beta_1 screened_i + \beta_2 post\_policy_t + \beta_3 (screened_i \times post\_policy_t) + \epsilon_{it}$$

Here, $student\_pct\_asian_{it}$ represents the percentage of Asian students in school i during year t. The variable $screened_i$ is a dummy variable equal to 1 for screened schools and 0 otherwise, as screened schools are the "treated" in my study. The variable $post\_policy_t$ is a dummy variable equal to 1 for years after the 2021 policy change and 0 for years before. The interaction term $screened_i \times post\_policy_t$ captures the differential impact of the policy on screened schools. Errors are clustered at the school level to account for within-school correlations over time.

Table 6: Baseline Regression Results - Racial Composition

VARIABLES	Asian	Black	Hispanic	Native American	Native Hawaiian	White
screened x post policy	0.00275	0.00788	-0.00165	-0.00168**	-0.00049	-0.00334
	(0.0041)	(0.0051)	(0.0064)	(0.0008)	(0.0006)	(0.0071)
Constant	0.124***	0.303***	0.438***	0.012***	0.006***	0.101***
	(0.0128)	(0.0186)	(0.0187)	(0.0010)	(0.0004)	(0.0110)
Observations	760	760	760	760	760	760
R-squared	0.006	0.010	0.042	0.020	0.005	0.115

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This regression was repeated for each racial group: Asian, Black, Hispanic, Native American, Native Hawaiian and Pacific Islander, and White. *Table 6* displays the results. None of the coefficients for the interaction term $screened_i \times post\_policy_t$ were statistically significant, except for Native American students, whose enrollment in screened schools decreased by 0.168% relative to non-screened schools after the policy change. This effect was significant at the 5% level but is relatively small in magnitude. Although the decrease in Native American enrollment is statistically significant, it is important to note that Native American students represent a small fraction of the total student population (~1%). Therefore, the practical implications of this shift may be limited. Future research could explore potential contributing factors, such as changes in application patterns, targeted outreach, or community engagement efforts that may have influenced this trend.

7. POLICY CHANGE EFFECT ON ENI

While the policy change did not significantly alter the racial composition of screened schools, it is possible that it affected the socioeconomic makeup of these schools. To investigate this, I analyzed the ENI over the same time period. *Figure 2* shows ENI trends from 2019-2020 to 2022-2023 for both screened and non-screened schools. Prior to the 2021 policy change, the trends in ENI for these two groups are parallel, satisfying the parallel trends assumption. After the policy change, ENI increased for both groups, with a slightly larger increase in screened schools.

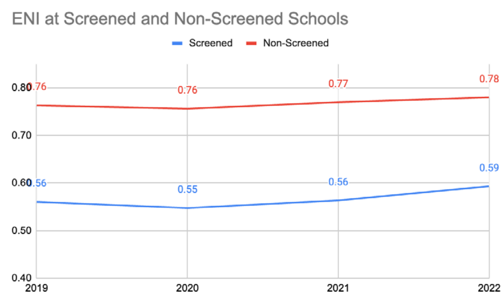


Figure 2: ENI at Screened and Non-Screened Schools

The graph suggests a post-policy increase in socioeconomic need among students at screened schools, but a regression analysis is necessary to confirm whether these changes are statistically significant and attributable to the policy shift. To test this, I repeated the baseline difference-in-difference regression, this time using ENI as the dependent variable. The specification is similar to the one used for racial composition:

$$ENI_{it} = \beta_0 + \beta_1 screened_i + \beta_2 post\_policy_t + \beta_3(screened_i \times post\_policy_t) + \epsilon_{it}$$

Here, ENI_{it} represents the ENI for school i during year t. The variable $screened_i$ is a dummy variable equal to 1 if the school is a screened school and 0 otherwise, while $post\_policy_t$ is a dummy equal to 1 for years after the policy change and 0 for years before. The interaction term $screened_i \times post\_policy_t$ captures the differential effect of the policy on ENI for screened schools. Errors are clustered at the school level to account for correlations over time within schools.

Table 7: Baseline Regression Results - ENI

VARIABLES	(1) ENI
screened	-0.206***
	(0.0334)
post policy	0.0185***
	(0.00175)
screened x post policy	0.00583
	(0.00757)
Constant	0.760***
	(0.0120)
Observations	760
R-squared	0.225

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The regression results indicate that, on average, before the policy change, screened schools had an ENI value 0.206 lower than non-screened schools. This finding is statistically significant at the 1% level, confirming that students at screened schools were wealthier than those at non-screened schools prior to the policy change. After the policy change, the ENI for screened schools increased by 0.0058 units relative to non-screened schools. However, this change is not statistically significant, suggesting that the policy did not meaningfully alter the socioeconomic composition of students at screened schools compared to non-screened schools. Interestingly, the regression also shows that the ENI increased by an average of 0.0185 units across all schools (screened and non-screened) after the policy change, and this result is statistically significant at the 1% level. This indicates that, on average, students across New York City public high schools became slightly poorer after the policy implementation.

8. INTERPRETATION AND BROADER IMPLICATIONS

These findings suggest that the 2021 policy change did not have a significant impact on the racial composition of students enrolled in screened schools. The minor changes observed in *Figure 1* appear to reflect broader trends or random variation rather than a systematic effect of the policy change. The figures provide helpful context for understanding the baseline racial distribution in screened versus non-screened schools. For example, screened schools consistently had a higher percentage of Asian and White students and a lower percentage of Black and Hispanic students compared to non-screened schools, both before and after the policy change. While these disparities remain, the lack of significant regression results indicates that the 2021 lottery-based admissions reform has yet to substantially shift these long-standing trends in student demographics.

One possible reason for the lack of significant change is implicit bias in the GPA-based grouping process. Students from historically underrepresented racial or socioeconomic backgrounds may have been disproportionately placed in lower GPA groups, limiting their chances of attending more selective schools. Even though the policy removed geographic restrictions, this GPA-based structure could have inadvertently perpetuated existing inequities, particularly if students from marginalized backgrounds face additional challenges in achieving higher GPAs.

Additionally, perceptions of declining school quality could play a role in shaping future policy debates. As some families may feel that the lottery-based system undermines the competitiveness of traditionally high-performing schools, there could be concerns about the long-term effects on school quality. These perceptions, whether rooted in real or perceived changes in academic rigor, could influence ongoing discussions about the role of selective admissions and the balance between equity and academic excellence in public schools.

The increase in ENI across all schools is noteworthy, even though the effect was not specific to screened schools. The rise could be indicative of broader socioeconomic shifts in New

York City, particularly in the wake of the pandemic. It is possible that wealthier families opted out of the public school system during this period, either by enrolling their children in private or religious schools or by moving to other areas for better schooling. This out-migration of higher-income students could contribute to the overall rise in ENI and align with anecdotal reports of families leaving urban areas during this period. Additionally, the economic impact of COVID-19 may have made some families poorer, further contributing to the rise in ENI as more low-income students enrolled in public schools. While the policy change did not have a significant effect on the ENI of screened schools specifically, the broader rise in poverty levels across the NYC public school system highlights the need for further research into enrollment patterns and the socioeconomic factors driving these changes.

9. NICHE RANKING BUCKETS

Following the change in admissions policy, students gained the ability to apply to selective and desirable schools that were previously restricted by geographic boundaries. With the introduction of the lottery system, I hypothesized that the most competitive schools “filled up” quickly with applicants who may have had different racial or socioeconomic backgrounds compared to the past. To explore this further, I categorized schools into five ranking “buckets” based on the 2025 Niche.com rankings for Best Public High Schools in New York. Of 1,283 schools:

- Bucket 1: Schools ranked 1-100
- Bucket 2: Schools ranked 101-200
- Bucket 3: Schools ranked 201-400
- Bucket 4: Schools ranked 401-700
- Bucket 5: Schools ranked 700+

While these rankings are for 2025, I argue that even if schools' rankings have shifted since 2021 due to the policy change, it's unlikely that any school would have experienced such a drastic change in its ranking as to move to a different bucket.

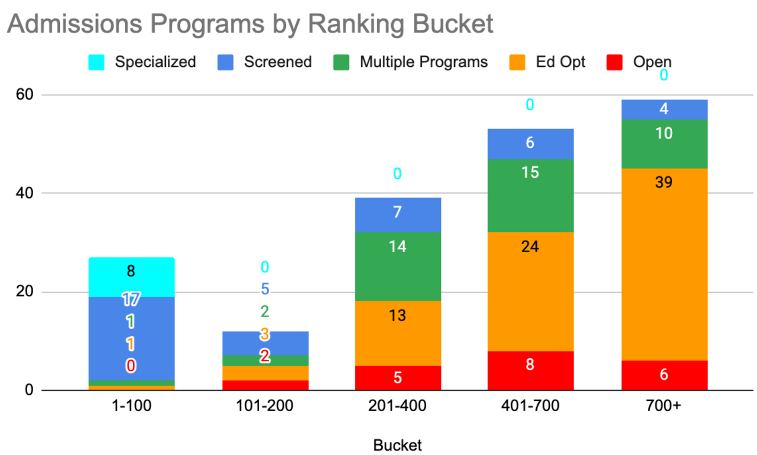


Figure 3: Admissions Programs Within Buckets

As shown in *Figure 3*, Bucket 1, which contains the most competitive and desirable schools, is made up mostly of specialized and screened schools. In contrast, as you move to lower-ranking buckets, there is a larger proportion of ed opt, open, and multiple admissions program schools.

10. POLICY CHANGE EFFECT ON RACIAL COMPOSITION WITHIN BUCKET 1

To determine whether the racial composition changed in the most desirable screened schools, I ran the same difference-in-difference regressions as before, but this time focusing only on schools in Bucket 1. *Table 8* displays the results of these regressions.

Table 8: Bucket 1 Regression Results - Racial Composition

VARIABLES	Asian	Black	Hispanic	Native American	Native Hawaiian	White
screened	-0.23484***	0.00918	0.05942	-0.00008	-0.00302**	0.16244**
	(0.0824)	(0.0321)	(0.0628)	(0.0017)	(0.0012)	(0.06)
post policy	0.0069	0.00645	0.0024	0.0005	0.0017	-0.0105
	(0.0101)	(0.0069)	(0.0059)	(0.0011)	(0.0011)	(0.0116)
screened x post policy	0.00181	-0.00213	0.01886	-0.00195	-0.0013	-0.01953
	(0.0119)	(0.0094)	(0.0114)	(0.0115)	(0.0013)	(0.0175)
Constant	0.476***	0.091***	0.147**	0.008***	0.0074***	0.225***
	(0.0128)	(0.0186)	(0.0564)	(0.0014)	(0.0009)	(0.0448)
Observations	108	108	108	108	108	108
R-squared	0.292	0.004	0.064	0.036	0.204	0.228

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

The coefficients show that, on average, screened schools in Bucket 1 have a 23.48 percentage point lower proportion of Asian students, a 0.3 percentage point lower proportion of Native Hawaiian and Pacific Islander students, and a 16.24 percentage point higher proportion of White students compared to non-screened schools. However, none of the coefficients on the interaction term - which captures the combined effect of being both a screened school and in the post-policy period - are statistically significant.

11. POLICY CHANGE EFFECT ON ENI WITHIN BUCKET 1

Next, I examined whether the ENI changed within Bucket 1 schools due to the policy change. *Figure 4* shows the ENI across all five buckets during the studied time period. The trend is clear: ENI increases as the school ranking decreases. Bucket 1 schools, ranked among the top 100 in New York, are significantly wealthier on average than schools ranked 101-200, which are wealthier than schools in lower-ranking buckets.

After the policy change, the ENI for Bucket 1 increased from 0.55 to 0.59. This indicates that the most selective and desirable schools saw a slight increase in the proportion of poorer students. This suggests that the policy may have succeeded in making these schools more accessible to students from lower-income backgrounds.

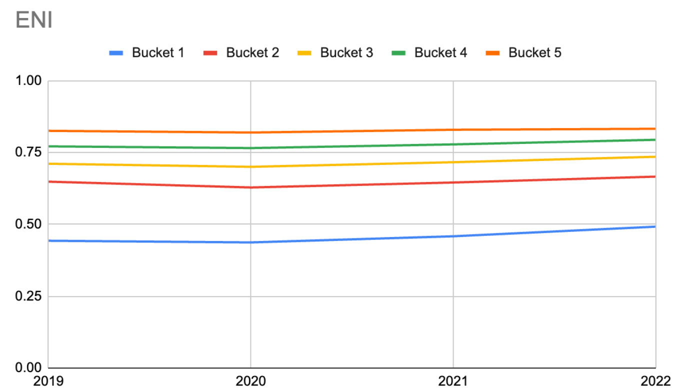


Figure 4: ENI Across Buckets

Finally, I ran the difference-in-difference regression again, using ENI as the dependent variable, but this time only for Bucket 1 schools. *Table 9* displays the result of this regression. The coefficient on screened schools (-0.0730) suggests that, on average, screened schools had a lower ENI compared to non-screened schools before the policy change, though this result is not statistically significant. The coefficient on post-policy (0.0239) is positive and statistically significant at the 1% level, indicating that the ENI increased by approximately 2.39 percentage points for non-screened schools after the policy change. The interaction term (0.0180) captures the additional change in ENI for screened schools post-policy, but it is not statistically significant. This suggests that while ENI increased overall, the effect did not differ significantly between screened and non-screened schools.

Table 9: Bucket 1 Regression Results - ENI

VARIABLES	(1) ENI
screened	-0.0730
	(0.0584)
post policy	0.0239***
	(0.00832)
screened x post policy	0.0180
	(0.0160)
Constant	0.486***
	(0.0489)
Observations	108
R-squared	0.070

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

12. DISCUSSION OF RESULTS

The regression results reveal that, contrary to expectations, the shift to a lottery-based admissions system did not significantly alter the racial or socioeconomic composition of students at screened schools. The 2021 NYC public high school admissions policy change, which removed district priority and expanded the selection pool for screened schools, was expected to influence both the distribution of students across schools and the demographic composition of enrolled students. However, my analysis found no statistically significant shifts in the racial and socioeconomic composition of students attending different types of schools. This raises important questions about the policy's effectiveness in promoting diversity and equity in school admissions.

A potential complication in interpreting these results is that some high-achieving students who previously aimed for screened schools may have opted for open or ed opt schools due to the perceived decline in school quality following the policy change. This "spillover" effect cannot be directly measured in this study, as the data available does not capture these nuanced changes in student choices across school types. Additionally, unobserved factors, such as individual student preferences, family priorities, and external influences like private tutoring or extracurricular activities, may have played a role in shaping the observed outcomes. However, I strongly believe that, in general, most students' choices remained similar to what they would have been prior to the change.

In particular, I argue that "higher-achieving" and ambitious students, particularly those in Groups 1 and 2 (the categories most likely to aim for screened schools), likely still ended up at screened schools, though possibly at a different one than they would have attended had the district priority not been removed. These students likely made strategic choices based on the revised admissions landscape, opting for schools that remained competitive or offered specific academic opportunities. The impact of the policy appears to have been a reshuffling of students within the screened schools category, rather than a large-scale shift away from these schools altogether.

On the other hand, "lower-achieving" and "less ambitious" students, who would have been more likely to attend an ed opt or open school prior to the admissions change, likely still ended up at one of these types of schools. These students may have had fewer choices due to the shift in the composition of applicants to more competitive schools, but it is reasonable to assume that they continued to be placed in schools within their academic range, in accordance with the general admissions criteria for these school types.

Thus, while the 2021 admissions policy change likely led to some redistribution of students among screened schools, the primary effect of the policy appears to be this reshuffling, rather than a significant reallocation of students to different school types (such as moving a large number of students from screened schools to open or ed opt schools). This finding suggests that the policy did not substantially alter the broader trends of student choice and distribution that existed prior to the change, especially among students with distinct academic profiles.

The analysis also highlights the importance of considering the role of school quality perceptions and how these perceptions influence student decisions. The potential decline in perceived school quality could have led some families to adjust their school preferences, but the general pattern of high-achieving students attending screened schools and lower-achieving students opting for ed opt or open schools appears to have remained relatively stable despite the policy change.

13. LIMITATIONS AND FURTHER RESEARCH

This study has limitations that should be acknowledged. Firstly, unobserved factors, such as the influence of school counselors, peer networks, and family dynamics, may have shaped student and family decisions in ways that the quantitative data cannot capture. Second, the exclusion of charter schools from the dataset could obscure broader trends in student distribution, as charter schools often serve as an alternative for families dissatisfied with traditional public schools. Lastly, the dataset does not include qualitative data on student and family perceptions of school quality or the specific reasons behind their school choices, which limits the ability to fully understand the mechanisms driving these outcomes.

Further research could explore these dynamics in greater depth by incorporating qualitative data, such as interviews or surveys, to capture perceptions of school quality and their role in shaping student's decisions. Additionally, future studies could include charter schools to provide a more comprehensive view of the broader educational landscape in NYC.

Understanding these nuances would help policymakers better anticipate the implications of structural changes in the admissions system and design more effective measures to promote equity in education.

14. CONCLUSION

This study finds that the 2021 NYC public high school admissions policy change, which replaced district priority with a lottery-based system, did not lead to significant shifts in the racial or socioeconomic composition of students at screened schools. These results suggest that the policy may not have achieved its intended goal of increasing diversity in top schools.

Given these findings, future research should explore how perceptions of school quality influence student decision-making, particularly when admissions processes change. Additionally, it would be valuable to examine the long-term impact of this policy shift on educational equity and access to high-quality schools, particularly for students from marginalized communities. Further research could delve deeper into the specific factors driving student choices and explore how these factors interact with structural changes in the admissions process.

From a policy perspective, the lack of significant demographic changes points to the need for more comprehensive reforms to promote greater equity in access to high-quality education. To better support students from diverse academic backgrounds, policymakers might consider additional measures, such as addressing disparities in access to resources, school quality, and opportunities. Future adjustments to the admissions policy could include targeting support for underrepresented students earlier in the education process, ensuring a more equitable distribution of resources, or implementing programs that specifically support students from marginalized communities in preparing for top schools.

In conclusion, while the 2021 policy change aimed to increase equity in school admissions, it has not significantly altered the demographic composition of NYC public high schools. Further adjustments to the admissions process, and a focus on improving all schools, are needed to achieve more equitable access to high-quality education for all students.

APPENDIX

References:

Conley, Dalton. Harvard by Lottery - the Chronicle of Higher Education, www.chronicle.com/article/Harvard-by-Lottery/131322. Accessed 25 Nov. 2024.

Corcoran, Sean Patrick, and E. Christine Baker-Smith. “Pathways to an elite education: Application, admission, and matriculation to New York City’s Specialized High Schools.” Education Finance and Policy, vol. 13, no. 2, Mar. 2018, pp. 256–279, https://doi.org/10.1162/edfp_a_00220.

“School Quality Report Citywide Data Archives.” School Quality Report Citywide Data, infohub.nyced.org/reports/students-and-schools/school-quality/school-quality-reports-and-resources/school-quality-report-citywide-data. Accessed 21 Oct. 2024.

Schwartz, Barry. Top Colleges Should Select Randomly from a Pool of ‘Good Enough,’ www.chronicle.com/article/top-colleges-should-select-randomly-from-a-pool-of-good-enough/. Accessed 25 Nov. 2024.

“The Elite Illusion: Achievement Effects at Boston and New York Exam Schools.” Econometrica, vol. 82, no. 1, 2014, pp. 137–196, https://doi.org/10.3982/ecta10266.

Excluded Schools:

1. 47 The American Sign Language and English Secondary School (02M047)
2. The Urban Assembly School for Emergency Management (02M135)
3. Stephen T. Mather Building Arts & Craftsmansh High School (02M139)
4. Manhattan Early College School for Advertising (02M280)
5. Urban Assembly Maker Academy (02M282)
6. Food and Finance High School (02M288)
7. High School of Hospitality Management (02M296)
8. Urban Assembly School of Design and Construction, The (02M300)
9. Urban Assembly Academy of Government and Law, The (02M305)
10. Lower Manhattan Arts Academy (02M308)
11. Urban Assembly School of Business for Young Women, the (02M316)
12. Gramercy Arts High School (02M374)
13. Manhattan Business Academy (02M392)
14. Business Of Sports School (02M393)
15. The High School For Language And Diplomacy (02M399)
16. High School for Environmental Studies (02M400)
17. Professional Performing Arts High School (02M408)
18. High School for Health Professions and Human Services (02M420)
19. Leadership and Public Service High School (02M425)
20. Manhattan Academy For Arts & Language (02M427)
21. International High School at Union Square (02M438)
22. Manhattan International High School (02M459)
23. High School of Economics and Finance (02M489)
24. Urban Assembly Gateway School for Technology (02M507)
25. Talent Unlimited High School (02M519)
26. Murry Bergtraum High School for Business Careers (02M520)
27. Repertory Company High School for Theatre Arts (02M531)
28. Union Square Academy for Health Sciences (02M533)
29. Manhattan Bridges High School (02M542)
30. New Design High School (02M543)
31. High School for Dual Language and Asian Studies (02M545)
32. Academy for Software Engineering (02M546)
33. Urban Assembly New York Harbor School (02M551)
34. Richard R. Green High School of Teaching (02M580)
35. The High School of Fashion Industries (02M600)
36. Chelsea Career and Technical Education High School (02M615)
37. Art and Design High School (02M630)
38. Life Sciences Secondary School (02M655)
39. Urban Assembly School for Media Studies, The (03M307)
40. The Urban Assembly School for Green Careers (03M402)
41. Wadleigh Secondary School for the Performing & Visual Arts (03M415)
42. Fiorello H. LaGuardia High School of Music & Art and Performing Arts (03M485)
43. High School for Law, Advocacy and Community Justice (03M492)
44. High School of Arts and Technology (03M494)
45. Special Music School (03M859)
46. Esperanza Preparatory Academy (04M372)
47. Young Women's Leadership School (04M610)
48. Eagle Academy for Young Men of Harlem (05M148)
49. The Urban Assembly School for Global Commerce (05M157)
50. Urban Assembly School for the Performing Arts (05M369)
51. Inwood Early College for Health and Information Technologies (06M211)
52. City College Academy of the Arts (06M293)
53. Community Health Academy of the Heights (06M346)
54. The College Academy (06M462)
55. High School for Media and Communications (06M463)
56. High School for Law and Public Service (06M467)
57. High School for Health Careers and Sciences (06M468)
58. Gregorio Luperon High School for Science and Mathematics (06M552)
59. The Laboratory School of Finance and Technology: X223 (07X223)
60. H.E.R.O. High (Health, Education, and Research Occupations High School) (07X259)
61. International Community High School (07X334)
62. Community School for Social Justice (07X427)

63. Bronx Design and Construction Academy (07X522)
64. Careers in Sports High School (07X548)
65. The Urban Assembly Bronx Academy of Letters (07X551)
66. Alfred E. Smith Career and Technical Education High School (07X600)
67. Health Opportunities High School (07X670)
68. Bronx Studio School for Writers and Artists (08X269)
69. Women's Academy of Excellence (08X282)
70. Renaissance High School for Musical Theater and the Arts (08X293)
71. Millennium Art Academy (08X312)
72. Archimedes Academy for Math, Science and Technology Applications (08X367)
73. Bronx Bridges High School (08X432)
74. Longwood Preparatory Academy (08X530)
75. School for Tourism and Hospitality (08X559)
76. Eagle Academy for Young Men (09X231)
77. Urban Assembly School for Applied Math and Science, The (09X241)
78. Bronx Early College Academy for Teaching & Learning (09X324)
79. Academy for Language and Technology (09X365)
80. Bronx International High School (09X403)
81. Bronx High School of Business (09X412)
82. Bronx High School for Medical Science (09X413)
83. Bronx School for Law, Government and Justice (09X505)
84. High School for Violin and Dance (09X543)
85. Claremont International HS (09X564)
86. Young Women's Leadership School of the Bronx (09X568)
87. Theatre Arts Production Company School (10X225)
88. The Marie Curie School for Medicine, Nursing, and Health Professions (10X237)
89. Bronx Academy for Software Engineering (BASE) (10X264)
90. Kingsbridge International High School (10X268)
91. Bronx School of Law and Finance (10X284)
92. International School for Liberal Arts (10X342)
93. High School for Teaching and the Professions (10X433)
94. Fordham High School for the Arts (10X437)
95. Bronx High School for Law and Community Service (10X439)
96. Celia Cruz Bronx High School of Music, The (10X442)
97. Crotona International High School (10X524)
98. Bronx Theatre High School (10X546)
99. High School for Energy and Technology (10X565)
100. Bronx Health Sciences High School (11X249)
101. Bronx High School for Writing and Communication Arts (11X253)
102. Academy for Scholarship and Entrepreneurship: A College Board School (11X270)
103. High School of Computers and Technology (11X275)
104. Bronx Academy of Health Careers (11X290)
105. Bronx High School for the Visual Arts (11X418)
106. High School of Language and Innovation (11X509)
107. New World High School (11X513)
108. High School for Contemporary Arts (11X544)
109. Bronx Aerospace High School (11X545)
110. Explorations Academy H.S. (12X251)
111. Pan American International High School at Monroe (12X388)
112. The Cinema School (12X478)
113. High School of World Cultures (12X550)
114. Dr. Susan S. McKinney Secondary School of the Arts (13K265)
115. Urban Assembly School of Music and Art (13K350)
116. Brooklyn Community Arts & Media High School (BCAM) (13K412)
117. Brooklyn International High School (13K439)
118. The Urban Assembly School for Law and Justice (13K483)
119. Urban Assembly Institute of Math and Science for Young Women (13K527)
120. Gotham Professional Arts Academy (13K594)
121. George Westinghouse Career and Technical Education High School (13K605)
122. City Polytechnic High School of Engineering, Architecture, and Technology (13K674)
123. PROGRESS High School for Professional Careers (14K474)
124. The High School for Enterprise, Business and Technology (14K478)
125. Williamsburg High School for Architecture and Design (14K558)
126. A-Tech High School (14K610)
127. Young Women's Leadership School of Brooklyn (14K614)
128. Digital Arts and Cinema Technology High School (15K429)
129. John Jay School for Law (15K462)
130. Cyberarts Studio Academy (15K463)
131. Cobble Hill School of American Studies (15K519)
132. Khalil Gibran International Academy (15K592)
133. Brooklyn High School of the Arts (15K656)
134. Brooklyn High School for Law and Technology (16K498)
135. The Brooklyn Academy of Global Finance (16K688)
136. Pathways in Technology Early College High School (P-Tech) (17K122)
137. Academy of Hospitality and Tourism (17K408)
138. International High School at Prospect Heights (17K524)
139. School for Human Rights, The (17K531)
140. High School for Public Service: Heroes of Tomorrow (17K546)
141. Brooklyn Academy of Science and the Environment (17K547)
142. Brooklyn School for Music & Theatre (17K548)
143. Clara Barton High School (17K600)
144. Academy for Health Careers (17K751)
145. Brooklyn Theatre Arts High School (18K567)
146. High School for Innovation in Advertising and Media (18K617)
147. Cultural Academy for the Arts and Sciences (18K629)
148. High School for Medical Professions (18K633)
149. Academy for Conservation and the Environment (18K637)
150. FDNY - Captain Vernon A. Richard High School for Fire and Life Safety (19K502)
151. High School for Civil Rights (19K504)
152. Performing Arts and Technology High School (19K507)
153. World Academy for Total Community Health High School (19K510)
154. Multicultural High School (19K583)
155. Transit Tech Career and Technical Education High School (19K615)
156. Academy of Innovative Technology (19K618)
157. W. H. Maxwell Career and Technical Education High School (19K660)
158. The Urban Assembly School for Collaborative Healthcare (19K764)
159. High School of Telecommunication Arts and Technology (20K485)
160. Urban Assembly School for Criminal Justice (20K609)
161. International High School at Lafayette (21K337)
162. Rachel Carson High School for Coastal Studies (21K344)
163. High School of Sports Management (21K348)
164. Life Academy High School for Film and Music (21K559)
165. William E. Grady Career and Technical Education High School (21K620)
166. Eagle Academy for Young Men II (23K644)
167. Teachers Preparatory High School (23K697)
168. International High School for Health Sciences (24Q236)
169. Academy of Finance and Enterprise (24Q264)
170. High School of Applied Communication (24Q267)
171. Pan American International High School (24Q296)
172. International High School at LaGuardia Community College (24Q530)
173. High School for Arts and Business (24Q550)
174. Aviation Career & Technical Education High School (24Q610)
175. Queens High School for Language Studies (25Q241)
176. Flushing International High School (25Q263)
177. East-West School of International Studies (25Q281)
178. World Journalism Preparatory: A College Board School (25Q285)
179. Business Technology Early College High School (26Q315)
180. Academy of Medical Technology: A College Board School (27Q309)
181. Rockaway Park High School for Environmental Sustainability (27Q324)
182. High School for Construction Trades, Engineering and Architecture (27Q650)
183. J.H.S. 157 Stephen A. Halsey (28Q157)
184. Hillside Arts & Letters Academy (28Q325)
185. Jamaica Gateway to the Sciences (28Q350)
186. Queens Gateway to Health Sciences Secondary School (28Q680)
187. High School for Law Enforcement and Public Safety (28Q690)
188. Young Women's Leadership School, Queens (28Q896)
189. Institute for Health Professions at Cambria Heights (29Q243)
190. Preparatory Academy for Writers: A College Board School (29Q283)
191. Benjamin Franklin High School for Finance & Information Technology (29Q313)
192. Cambria Heights Academy (29Q326)
193. Eagle Academy for Young Men III (29Q327)
194. Humanities & Arts Magnet High School (29Q498)
195. Energy Tech High School (30Q258)
196. Young Women's Leadership School, Astoria (30Q286)
197. Academy for Careers in Television and Film (30Q301)
198. Frank Sinatra School of the Arts High School (30Q501)
199. Information Technology High School (30Q502)
200. Newcomers High School (30Q555)
201. Academy of American Studies (30Q575)
202. The Eagle Academy for Young Men of Staten Island (31R028)

203. EBC High School for Public Service - Bushwick (32K545)
204. The Academy of Urban Planning and Engineering (32K552)
205. Achievement First Crown Heights Charter School (84K356)
206. Achievement First East New York Charter School (84K358)
207. Williamsburg Charter High School (84K473)
208. Leadership Preparatory Bedford Stuyvesant Charter School (84K517)
209. Kings Collegiate Charter School (84K608)
210. Achievement First Brownsville Charter School (84K626)
211. Brooklyn Ascend Charter School (84K652)
212. Northside Charter High School (84K693)
213. Brooklyn Prospect Charter School (84K707)
214. Excellence Girls Charter School (84K712)
215. Summit Academy Charter School (84K730)
216. Math, Engineering, and Science Academy Charter High School (84K733)
217. New Visions Charter High School for Advanced Math and Science III (84K738)
218. New Visions Charter High School for the Humanities III (84K739)
219. Coney Island Preparatory Public Charter School (84K744)
220. Unity Prep Charter School (84K757)
221. Leadership Preparatory Ocean Hill Charter School (84K775)
222. New Visions Charter High School for the Humanities
223. Coney Island Preparatory Public Charter School
224. Unity Prep Charter School
225. Leadership Preparatory Ocean Hill Charter School
226. Brooklyn LAB Charter School
227. Brooklyn Emerging Leaders Academy
228. Edmund W. Gordon Brooklyn Laboratory Charter School
229. Democracy Prep Endurance Charter School
230. Capital Preparatory (CP) Harlem Charter School
231. The Opportunity Charter School
232. Harlem Children's Zone Promise Academy 1 Charter S
233. Harlem Village Academy East Charter School
234. KIPP Infinity Charter School
235. Harlem Children's Zone Promise Academy II Charter
236. Democracy Prep Charter School
237. Success Academy Charter School - Harlem 1
238. New Heights Academy Charter School
239. DREAM Charter School
240. Success Academy Charter School - Harlem 3
241. Renaissance Charter High School for Innovation
242. Inwood Academy for Leadership Charter School
243. Democracy Prep Harlem Charter School
244. East Harlem Scholars Academy Charter School
245. Broome Street Academy Charter School
246. Harlem Prep Charter School
247. Harlem Village Academy West Charter School
248. New Visions Charter High School for Advanced Math
249. New Visions Charter High School for the Humanities
250. Renaissance Charter School
251. John W. Lavelle Preparatory Charter School
252. Lois and Richard Nicotra Early College Charter Sch
253. Bronx Lighthouse Charter School
254. New Visions Charter High School for Advanced Math
255. New Visions Charter High School for the Humanities
256. Hyde Leadership Charter School
257. International Leadership Charter High School
258. University Prep Charter High School
259. NYC Charter High School for Architecture, Engineer
260. Charter High School for Law and Social Justice
261. Metropolitan Lighthouse Charter School
262. American Dream Charter School
263. Dr. Richard Izquierdo Health and Science Charter S
264. The Equality Charter School
265. New Visions Charter High School for Advanced Math
266. New Visions Charter High School for the Humanities
267. South Bronx Community Charter School
268. Urban Assembly Charter School for Computer Science
269. AECI II: NYC Charter High School for Computer Engineering
270. Bronx Preparatory Charter School
271. KIPP Academy Charter School