Vol. 3, April 2025

THE CAPITOL ECONOMICS JOURNAL

Capitol Economics Journal

Volume III, Spring 2025

The CEJ is published annually and all publication policies and processes are conducted in accordance with international standards of academic ethics and integrity. The CEJ accepts and publishes original, undergraduate research papers relating to all fields of economics. Consideration for each year's edition opens at the beginning of the fall semester and closes one month prior to the end of the semester.

Read the journal online:

https://www.gwuues.org/publications

Contact: ues.gwu@gmail.com

George Washington University Undergraduate Economics Society
University Student Center
800 21st Street NW
Washington, D.C. 20052

| Contents

| Letter from the Editor

Dear Reader,

It is my great pleasure to present this year's edition of the Capitol Economics Journal! After several years of dormancy, I am incredibly proud to see its revival, providing a platform to showcase the dedication and hard work of our writers. This journal represents the culmination of tireless efforts from our authors, editors, and support team, and it stands as a testament to their commitment to economic research.

The Capitol Economics Journal offers undergraduate economics majors at George Washington University a rare opportunity to have their senior theses published, solidifying their research as a concrete contribution to the field. Over the past academic year, our writers have worked diligently to apply their coursework to original empirical research, refining their work through rigorous revisions guided by their peers and the esteemed faculty of the George Washington University Economics Department. This edition highlights the breadth and versatility of economic study, featuring research that spans public taxation policy, environmental considerations in South America, and the relationship between public transit accessibility and residential home values.

I would like to extend my deepest gratitude to the team that made this publication possible—the Executive Board of the Undergraduate Economics Society, our esteemed publishing organization, our dedicated team of associate editors, and of course, our talented writers. I am also profoundly grateful to the George Washington University Student Government Association for their financial support in bringing this journal to life. This endeavor would not have been possible without the dedication, enthusiasm, and perseverance of this incredible team.

Most importantly, I want to thank **you, the reader**, for taking the time to engage with the research presented in this journal. Your curiosity and willingness to explore these contributions to the field of economics make this publication truly meaningful.

Warm Regards,

Christopher M. Coulter
Editor-in-Chief
Capitol Economics Review

Breaking the Curve: An Analysis of the South American Environmental Kuznets' Curve Outlier

Mallory Kussman

George Washington University

Abstract

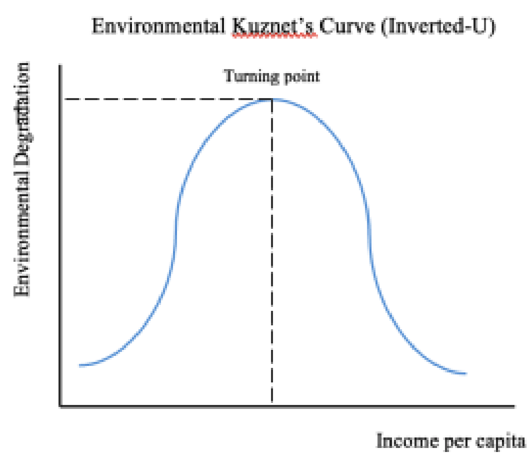
The Environmental Kuznets Curve (EKC) predicts an inverted-U relationship between environmental degradation and economic development. While the EKC suggests that middle- income countries should experience peak carbon emissions, South America's emissions remain below expected levels for their development status. This paper uses 32 years of data on the 12 sovereign South American nations to investigate whether industrialization or broader measures of development better explain the relationship between trade openness and carbon emissions in their unique economic trajectories. Using polynomial quadratic regression models analyzing trade, development, and industrialization, the study finds that industrialization is a stronger predictor of the relationship between trade and emissions than development, finally providing evidence supporting the Environmental Kuznets Curve's application to South America with the new determinant of industrialization. This analysis highlights the need to integrate regional economic contexts into international and environmental economic research. **JEL Codes:** Q56, F18, O13

1 Introduction

This paper examined whether development or industrialization was a better indicator of a strong relationship between carbon emissions and nominal trade for South American countries. The existing literature on this subject is based off two contradictory theories in the discipline of economic development. The pollution haven hypothesis, sometimes referred to as 'the race to the bottom theory', posited by Copeland and Taylor in their 1994 analysis of the impacts of the North American Free Trade Agreement (NAFTA) on greenhouse gas emissions which has since been expanded upon to cover all the world's regions, states that free trade increases world pollution (Copeland and Taylor 1994). As in a competitive global market, countries with weaker environmental regulations have a financial and logistical comparative advantage in more pollutant-intensive industries and production methods. This might encourage other countries to weaken environmental regulations to better compete, resulting in environmental degradation and increased pollution and greenhouse gas emissions.

Opposing this there is the pollution halo hypothesis, coined by Walter and Ugelow in 1979 and Pethig in 1976, which states that greater access to global markets reduce world pollution (Walter and Ugelow 1979) (Pethig 1976). They claim that as multinational corporations bring cleaner technologies, better environmental practices, and institutional focuses on environmental protectionism, they raise local standards. Through the transfer of knowledge and personnel, higher expectations, and social pressure within and toward the corporation, Foreign Direct Investment (FDI) can create a "halo" effect wherein stricter environmental practices spread to the less-developed countries hosting them, eventually resulting in reduced pollution and stricter policies (Porter and Van der Linde 1995).

Bridging these two theories, the Environmental Kuznets' Curve theory suggests the existence of an inverted-U shaped relationship between environmental degradation, shown through various measures of pollution, and development, shown here through income per capita (Kuznets 1955). According to this theory, as a nation's economy develops and income per capita grows, pollution initially grows as the nation's populace consumes more resources and this phase usually overlaps with the process of industrialization which is often lacking in regulation in early stages. However, after achieving a certain level of income per capita and middle-class growth, environmental quality tends to improve as people demand cleaner air, water, and have more time and bandwidth to focus on causes like environmental sustainability, now that meeting their basic needs for survival is less challenging. This leads to governments adopting stricter regulations and companies engaging in greater research and development of clean technology. Thus, this theory proports the accuracy of the pollution haven hypothesis in early development stages and the pollution halo hypothesis in later stages. Development then became the determinant of which theory will apply to what country and why this may change over time. With development added as a determinant, economic development researchers



were able to produce and reproduce papers showing the strong relationship between trade and development (Yasmeen et al. 2019) (Wang et al. 2019) (Azam et al. 2020) (Hassan et al. 2020) (Onwachukwu et al. 2021).

South America opposes the Environmental Kuznet's Curve's assertion that middle income countries are the highest polluting category, as their nominal and especially per capita carbon emissions are significantly below what the models predict for their income category. Elmarzougui et al.'s regional research on the pollution haven hypothesis found that "the exceptions [to the hypothesis' proposition that increasing trade increases pollution] are CO_2 emissions in OECD and South American countries which tend to be respectively reduced and augmented by increases in trade openness" (Elmarzougui et al., 2016). There are many beliefs as to why South America stands as an outlier, as they don't fall in the high-income category like the OECD countries that can then be explained by the Environmental Kuznets' Curve. For example, abundant renewable energy resources, natural resource-based exports, and while it is generally believed that public demand for environmentally conscious policy and production increases with per capita income, the strong political culture and engagement found in many South American nations may have allowed them to bypass the Environmental Kuznets Curve's purported developmental prerequisite.

This paper seeks to prove that South America's outlier status from the Environmental Kuznets' Curve is due to the comparatively less-industrialized nature of their economies as opposed to nations with similar development levels. "As trade, services, data, people, and ideas internationalized, they didn't do so uniformly or consistently" (O'Neil 2022). Despite their relatively high income per capita and overall human development scores, the typical markers of a developed economy, i.e. a greater importance of trade in a nation's GDP, advanced infrastructure, high degrees of manufacturing and innovation, economic diversity, etc. are frequently missing in the South American context. Duran, Mussachio, and della Paolera found that after the abolishment of the continent's widespread import-substitution industrialization policies and rejoining of the global economy in the 1980s, the international context, internal policies, the countries' initial conditions upon independence, and the nature of the countries' export products discouraged industrialization (Duran, Mussacio, and della Paolera 2017). This works in collaboration with other explanations for South America's absence from the Environmental Kuznets' Curve, namely the focus of their economies on natural-resource exports, to provide evidence for the claim that in the South American context, industrialization is a better determinant of the nature of the relationship between changes in trade and pollution levels than development.

To prove that industrialization is a more accurate and nuanced indicator than development for the relationship described in the Environmental Kuznets Curve paper, this paper examines two quadratic polynomial regression models, analyzing industrialization and development levels respectively as a determinant of the relationship between trade, measured as a percentage of GDP, and carbon emissions, measured in metric tons, to determine that the 12 sovereign South American economies, analyzed over a 32-year period from 1990-2022, are more closely tied to industrialization than development level. The rest of the paper is organized as follows: the study's model methodology and data sources are presented in the next section, followed by results and robustness checks, and finally the study's conclusions.

2 Quadratic Polynomial Regression Model with Interaction Effects

In the spirit of Elmarzougui et al. (2016) and Frankel and Rose (2005), this paper estimates the long run relationship between CO_2 emissions, nominal trade values, and development level with the following empirical specification:

$$Y_i = \beta_{0i} + \beta_1 T_1 + \beta_2 X_i + \beta_3 Z_i + \beta_4 Z_i^2 + \beta_5 (X \cdot Z)_i + \epsilon_i \tag{1}$$

Listed below are the descriptions of the terms in this equation:

- Y_i: Greenhouse Gas Emissions (measured in metric tons of CO_2 emissions per capita) in Country i
- X_i: Trade Openness (measured as the sum of imports and exports as a percentage of GDP) in Country i
- Z_i: Development Level (measured by the UN Human Development Report where each country receives a rating between 0 being the least developed and 1 being the most developed) in Country i
- T_i: Year (standardized) in Country i
- β_{0i}: An intercept term in Country i
- β_1: Coefficient capturing the inherent significance of the passage of time on emissions
- β_2: Coefficient capturing the direct effect of trade on emissions
- β_3: Coefficient capturing the effect of development level on emissions
- β_4: Coefficient capturing the effect of non-linear development on emissions
- β_5: Interaction term capturing how the effect of trade on emissions differs between developed and less developed countries

- ϵ_i: Error term representing unobserved factors affecting emissions in Country i

Similarly, this paper assumes the long run relationship between CO_2 emissions, nominal trade values, and industrialization level can be estimated by the following empirical specification:

$$Y_i = \theta_{0i} + \theta_1 T_i + \theta_2 X_i + \theta_3 D_i + \theta_4 D_i^2 + \theta_5 (X \cdot D)_i + \epsilon_i \tag{2}$$

Listed below are the descriptions of the terms in this equation:

- Y_i: Greenhouse Gas Emissions (measured in metric tons of CO_2 emissions per capita) in Country i
- T_i: Year (standardized) in Country i
- X_i: Trade Openness (measured as the sum of imports and exports as a percentage of GDP) in Country i
- D_i: Industrialization Level (measured measured by the United Nations' Industrial Development Organization's Industrialization Intensity Index where each country receives a rating between 0 being the least industrialized and 1 being the most industrialized) in Country i
- θ_{0i}: An intercept term in Country i
- θ_1: Coefficient capturing the inherent significance of the passage of time on emissions
- θ_2: Coefficient capturing the direct effect of trade on emissions
- θ_3: Coefficient capturing the effect of industrialization level on emissions
- θ_4: Coefficient capturing the effect of non-linear industrialization on emissions
- θ_5: Interaction term capturing how the effect of trade on emissions differs between industrialized and less industrialized countries
- ϵ_i: Error term representing unobserved factors affecting emissions in Country i

As exemplified by Elmarzougui et al. (2016) and Frankel and Rose (2005), this paper uses a polynomial quadratic regression model as a method of cross-country time series analysis with a squared development and industrialization term respectively in order to account for "non-linearities which may arise from non-homotheticities in production or consumption to support the EKC" (Elmarzougui et al. 2016). As they stated, the Environmental Kuznets' Curve hypothesis would be confirmed by a positive coefficient on the base industrialization and development terms and a negative coefficient on the squared industrialization and development terms.

Moving beyond Elmarzougui et al. and Frankel and Rose's models, an interaction term was added to account for the determination of the impact of the relationship between trade and development or industrialization respectively. Given that positive

β_3 or θ_3 would suggest that increased trade leads to a larger increase in emissions in developed/industrialized countries compared to less developed/industrialized countries, a comparison between the β_3 and θ_3 will be used to show which indicator, development level or industrialization level is ultimately a better predictive variable for whether increased trade openness will increase greenhouse gas emissions.

While the models above provide a robust framework for analyzing the relationship between carbon emissions, trade openness, and development or industrialization levels, one must acknowledge the possibility of omitted variable bias and the lack of key control variables, as pointed out by the referee. Factors such as urbanization levels, public transit availability, and the percentage of energy derived from renewable sources are likely to play a significant role in shaping greenhouse gas emissions but are not included in these specifications. Use of extraneous variables were limited due to the "one or the other" nature of the investigation. As the intent is to determine if industrialization or development is the better indicator, excluding these control variables that might be more correlated with one than the other and thus alter the coefficients this paper intends to analyze. Future extensions of this analysis could address these limitations by finding a way to incorporate these control variables, thereby providing a more comprehensive understanding of the determinants of carbon emissions and their relationship with trade.

3 Description and Visualization of Data (Carbon Emissions, Trade, Development Levels, and Industrialization Levels)

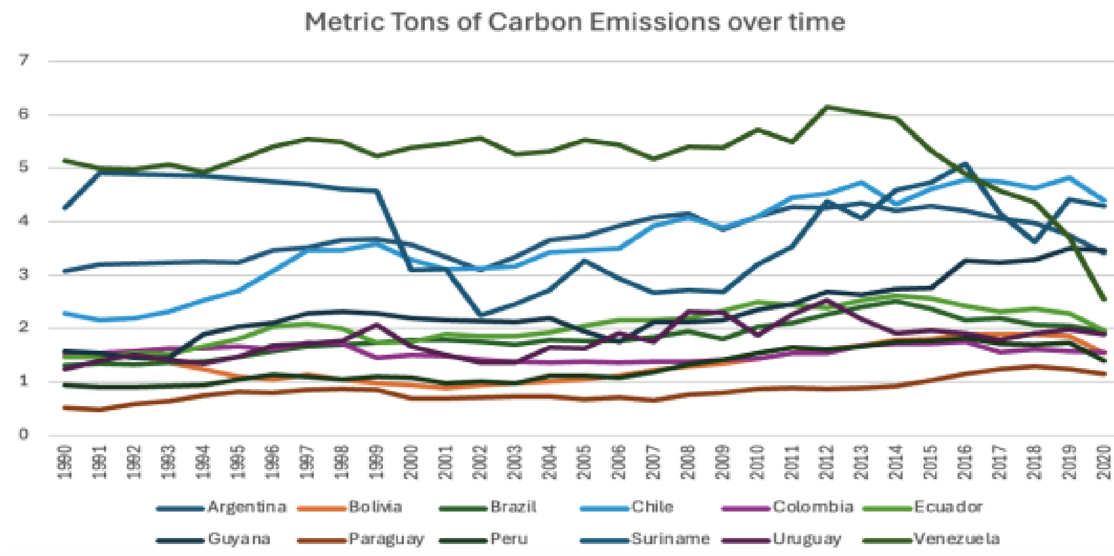
Variable Descriptions and Sources			
Variable (Model Nomenclature)	Variable	Measurement	Source
Y	Greenhouse gas emissions	Metric tons of CO_2 emissions per capita by country	World Bank Development Indicator Database
X	Trade openness	Sum of imports and exports as a percentage of the country's GDP	World Bank Development Indicator Database
Z	Development level	The UN Human Development Index gives each country a score between 0 and 1, the score is the geometric mean of normalized indices for each of three indicators: life expectancy at birth (health), mean of years of schooling for adults aged 25 years and more and expected years of schooling for children of school entering age (education), and gross national income per capita (standard of living).	UN Human Development Report
D	Industrialization level	The UN Industrial Development Organization's Industrialization intensity index is measured by the simple average of the share of Manufacturing Value Added (MVA) in GDP and the share of medium and high-technology (MHT) activities in MVA. The former captures the role of manufacturing in the economy and the latter the technological complexity of manufacturing.	UN Industrial Development Organization's Competitive Industrial Performance Index Dataset

There are 12 sovereign nation states in South America: Argentina, Bolivia, Brazil, Chile, Colombia, Ecuador, Guyana, Paraguay, Peru, Suriname, Uruguay, and Venezuela. With an exclusion of the dependent territories of the Falkland Islands (UK), French Guiana (France), and South Georgia and the South Sandwich Islands (UK) from this analysis due to the inaccessibility of their data as a separate statistic from their administrating state's data.

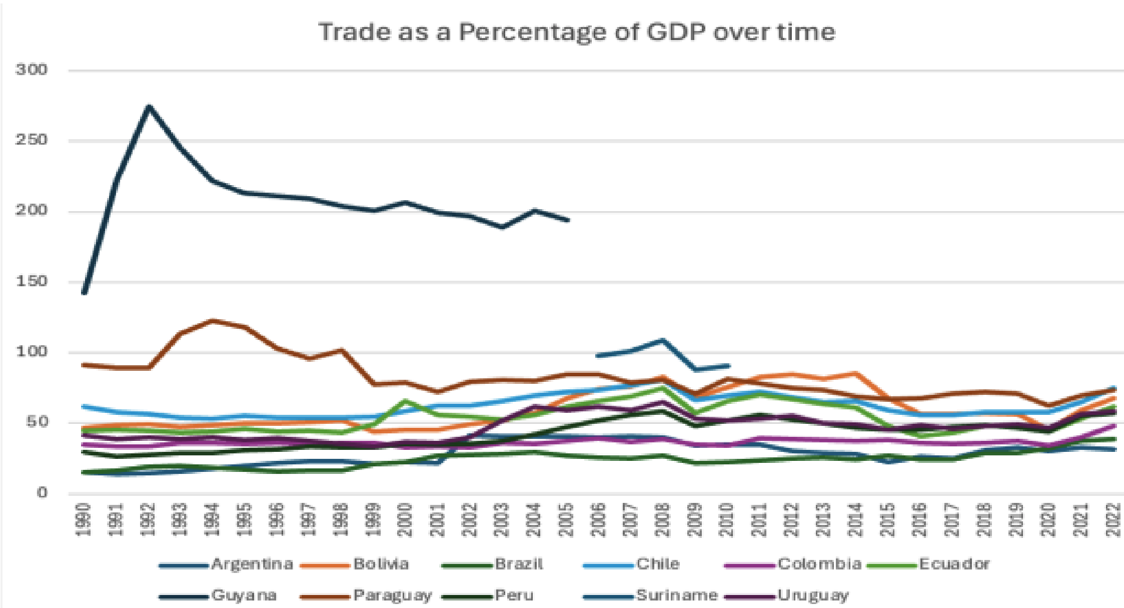
The World Bank's World Development Indicators Databank was used in gathering data on greenhouse gas emissions and trade openness. Selecting the 12 sovereign states of South America as the countries, CO_2 emissions (metric tons per capita) and trade (series, and all 50 years for which data might be available. Downloading the data in table form displayed that data on CO_2 emissions per capita largely wasn't available prior to 1990. In attempt to adhere the model more closely to previous literature, exploration of using kilotons of CO_2 emissions, HFC gas emissions (thousand metric tons of CO_2 equivalent), and Nitrous oxide emissions (thousand metric tons of CO_2 equivalent) as the environmental degradation metric, but they have unreliable data, so a returned to the initial metric tons of CO_2 per capita metric. Suriname only reported trade as a percentage of GDP data to the World Bank in five of the specified years, Venezuela hasn't reported trade as a percentage of GDP for the past nine years, and Guyana hasn't reported trade as a percentage of GDP since 2005. However, given that trade as a percentage of GDP was measured as early as 1974, there are still sufficient observations.

In gathering data on industrialization level, the United Nations Industrial Development Organization's Data Browser webpage was used. Further navigation to the Competitive Industrial Performance Index dataset enabled the download of annual data on Industrial Intensity Index Scores for each country, further use of a pivot table ensured it fit the dataset and final analysis. However, data on Guyana was unavailable in this dataset. Similarly, in gathering development level data, the UN Human Development Reports website which offered links to their dataset on a year-by-year or country history basis was chosen. Opting then to download country-level histories of HDI scores, use a pivot table to ensure it fit the dataset, and analyze it.

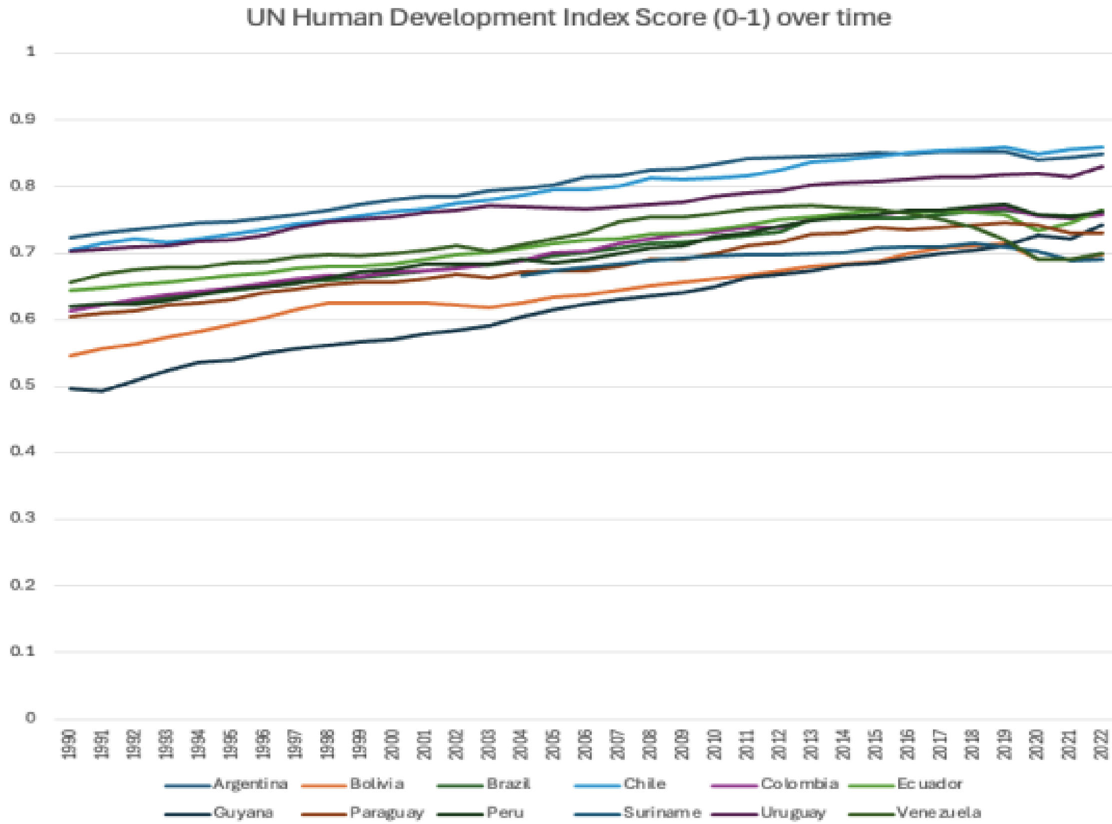
Metric tons of carbon emissions over time can be visualized below. This graph shows most South American nations having low and consistent levels of carbon emissions over the 30-year period from 1990-2020. That excludes Brazil who, in emitting over 5 metric tons of Carbon Dioxide per year until 2014, qualified for the majority of the period as a moderate polluter. As well as Suriname who, while maintaining their low polluter status not yet reaching 5 metric tons in a year, has had a far more turbulent emission history. Finally, it is worth noting that while Chile lacks any concerning spikes or drops to denote instability in the emissions history, it is now approaching the 5 metric ton mark of becoming a moderate polluter.



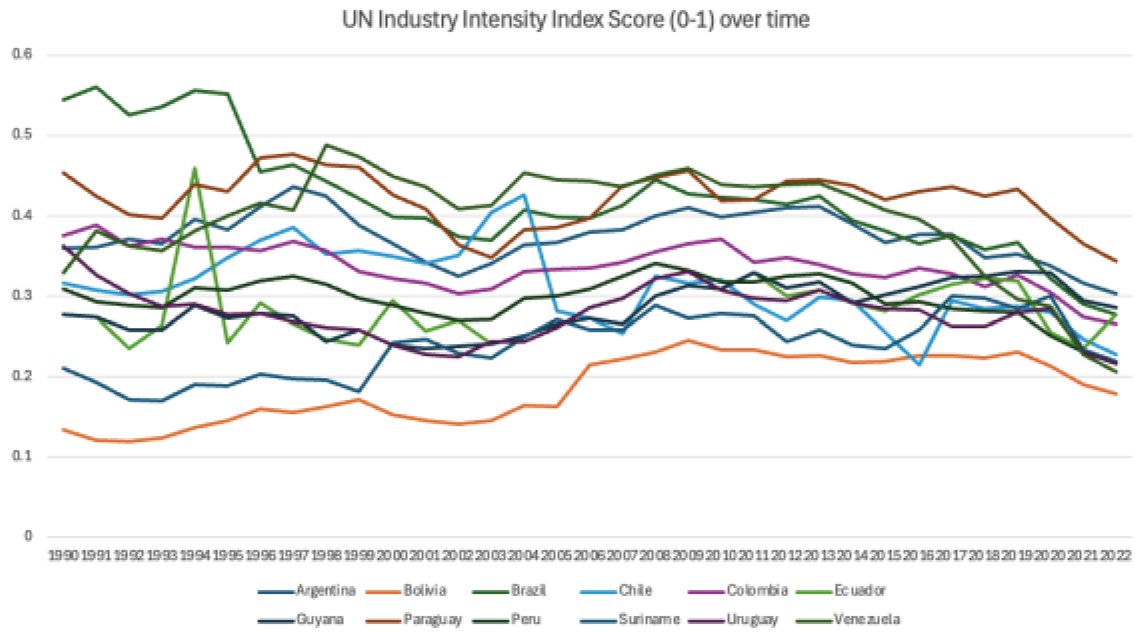
Trade as a percentage of GDP over time can be visualized below. This graph shows Suriname, Venezuela, and Guyana's previously mentioned inconsistencies in reporting. One can also see that Suriname is an immense outlier during the early 90s, with trade accounting for over 100% of its GDP, as in addition to massive amounts of their country's income coming from exporting gold and oil, but also relying almost completely on imports for food, manufactured goods, and clothing. While Paraguay also exceeds 100% in the mid 90s due to reliance on imports, the other South American nations remain comfortably consistent with their low percentage of GDP relating to trade.



Human development over time can be visualized below as a score between 0 and 1. This graph shows South America's infamous high human development levels. With the low development category encompassing every score below 0.55, only Suriname has fallen into the lowest category in our 32-year period. The rest of the South American nations have maintained consistently medium and high development scores throughout the period.



Finally, industrial intensity over time can be visualized also as a score between 0 and 1. This graph shows South America's shockingly low levels of industrialization, especially considering their impressive human development scores. One thing in particular to note is that the graph's upper most line reflects a score of 0.6 when the index is in fact out of 1. Brazil, having once had a somewhat robust manufacturing sector, lost the majority of its economy's focus on industry with the dissolution of import-substitution industrialization policies in the 1980s and 1990s. Just as many countries that show minor gains in industrial intensity, also show minor, or major in Brazil, Paraguay, and Chile's cases, losses in industrial intensity. This goes to bolster the premise that low industrialization levels are more consistent with strong relationship carbon emissions and trade than high development levels.



4 Results

Review of the Trade-Industrialization Model:

$$Y_i = -44.488579 + 0.021T_i + 0.027X_i + 21.085D_i - 18.535D_i^2 - 0.095(X \cdot D)_i + \epsilon_i \quad (3)$$

Starting with the trade-industrialization model, the positive coefficient of 0.021 for the year variable T_i suggests that, on average, emissions increase slightly over time, with an increase of 0.021 metric tons of CO_2 per capita per year. This reflects the compounding effect of economic growth, population growth, and increased energy consumption over time, even as some countries adopt greener technologies. It suggests that global trends, such as industrial expansion and rising living standards, continue to drive emissions upward. This falls in line with the expectations of the pollution haven hypothesis and the initial assumptions of the Environmental Kuznets Curve theory.

The coefficient of 0.027 for trade openness indicates that each one percent increase in trade (as a percentage of GDP) is associated with an increase of 0.027 metric tons of CO_2 per capita. This suggests that trade liberalization and globalization tend to contribute to increased emissions, likely due to increased production, transportation emissions from shipping and logistics, and the Energy-intensive production processes outsourced to trade-heavy economies that the pollution haven hypothesis predicts.

In examining the industrialization coefficient, one can see a significant positive relationship: for each one-unit increase in the industrialization index (a shift from the least to most industrialized), emissions increase by 21.085 metric tons of CO_2 per capita. This dramatic increase reflects the resource-intensive nature of industrial activities, such as manufacturing, construction, and energy generation, which are heavily reliant on fossil fuels. This result also supports the basis in this research that countries at higher levels of industrialization are likely to emit significantly more than less industrialized countries.

The interaction term captures how the effect of trade openness on emissions changes with industrialization. The negative coefficient of -0.095 suggests that trade openness has a smaller impact on emissions as a country becomes more industrialized. This coefficient reflects how industrialization moderates the emissions-trade relationship, highlighting the nuanced interaction between these variables.

Additionally, this model boasts three coefficient term p-values that meet the requirements for a significance code of 0; this promotes the importance of the three variables: trade, industrialization, and the trade-industrialization intercept term as highly influential in determining the level of carbon emissions in South America. Similarly, both industrialization and the trade-industrialization interaction term score a significance code of zero on their F-statistic p-value, suggesting that these terms significantly explain variability in the dependent variable of carbon emissions.

Review of the Trade-Development Model:

$$Y_i = 117.685 - 0.059T_i + 0.012X_i - 13.721Z_i + 21.129Z_i^2 - 0.009(X \cdot Z)_i + \epsilon_i \quad (4)$$

In the trade-development model, the negative coefficient for the year variable suggests that, on average, CO_2 emissions per capita decrease slightly over time, at a rate of 0.059 metric tons per year. This finding contradicts the findings of the industrialization model and the predictions of the pollution haven hypothesis, while supporting the expectations of the pollution halo hypothesis. If true, the decline may reflect a gradual shift toward greener technologies, improvements in energy efficiency, and potentially reduced dependence on carbon-intensive activities in South America.

The positive coefficient for trade openness suggests that each one increase in trade (as a percentage of GDP) is associated with an increase of 0.012 metric tons of carbon emissions per capita. The small magnitude of this coefficient implies that trade openness has a relatively modest direct impact on emissions, at least in comparison to other terms, in this model. The increase could be attributed to trade-driven industrial production, energy use, and transportation emissions. However, the effect is smaller than in the industrialization model, suggesting that trade's emissions impact is less pronounced in the context of development.

The negative coefficient for development indicates that as a country becomes more developed, emissions decrease. Specifically, development's coefficient of -13.721 indicates that as the development level increases, emissions initially decrease. For each one-unit increase in the development level (a shift from a low score to a higher score on the development scale), emissions drop by 13.721 metric tons of carbon dioxide per capita. This aligns with the pollution halo hypothesis' notion that more developed countries may transition to cleaner industries, adopt green technologies, and implement stricter environmental regulations.

The interaction term between trade openness and development level has a negative coefficient, indicating that the effect of trade openness on emissions diminishes as countries become more developed. Here one may note that the research question number one has been answered: the coefficient of the trade-industrialization interaction term (-0.095) is more significant than that of the trade-development interaction term (0.009) by a magnitude greater than ten, suggesting that industrialization is a better predictor of the relationship between trade and emissions in South America than the existing literature's proposition of development levels.

This model attained only one coefficient term p-value of zero. This one variable was the year adjustment variable, implying that the variable doing the most to help predict carbon emissions in South America is simply passage of time, this is especially poignant when you consider that in this model, only development-squared of trade, development, development- squared, and the trade-development interaction term received significance code below 1. However, both development and the year term score a significance code of zero on their F- statistic p-value, suggesting first that these terms significantly explain variability in the dependent variable of carbon emissions and secondly that these terms may be linked in this model meaning that development's impact on carbon emissions may be seen as an effect inherent in South America's development across the 32-year time span studied.

Overall, the trade-development model provides valuable insights into how emissions evolve as countries develop. It highlights the nuanced relationship between trade, development, and emissions, emphasizing that the impacts of trade openness depend heavily on a country's development stage. However, the weaker interaction term and lower significance levels compared to the trade-industrialization model suggest that industrialization remains the more dominant factor driving emissions in South America.

Results in Respect to the Environmental Kuznets' Curve Hypothesis:

Now to discuss the non-linear effects of development and industrialization. This paper's analysis of the trade-development model concurs with previous literature insofar as it does not adhere to the Environmental Kuznets Curve. As previously mentioned, Elmarzougui and Frankel and Rose state that the Environmental Kuznets' Curve hypothesis would be confirmed by a positive coefficient on the base industrialization and development terms and a negative coefficient on the squared industrialization and development terms.

The development coefficient was negative and the development-squared coefficient of positive 21.129 on the squared development term suggests a strong, non-linear U-shaped relationship between development and emissions. This non-linear relationship may reflect that highly developed countries have higher per capita consumption and energy demands, which could lead to increased emissions despite advancements in technology and regulations. The squared term's positive coefficient indicates a (non-inverted) U-shaped relationship where emissions first decrease as trade and development initially increase then decrease as the variables reach the EKC turning point. This is the exact opposite relationship of what the Environmental Kuznets' Curve hypothesis predicts. Additionally finding that the relationship between trade and emissions weakens as development levels increase, suggesting that more developed countries may trade with lower emissions impacts.

In the trade-industrialization model, the negative coefficient of -18.535 on the squared industrialization term D_{i2} implies a non-linear, inverted U-shaped relationship between industrialization and emissions. This suggests that emissions increase as countries industrialize, but after reaching the EKC turning point, the emissions growth rate slows or even reverses. This could reflect a transition to more efficient, cleaner technologies or environmental regulations as industrialization advances. Finally, the emissions impact of trade diminishes as countries become more industrialized, suggesting that industrialized, even more so than developed, countries may be more capable of managing the environmental impacts of trade. This supports the expectations of all three critical pieces of literature upon which this research was based: the pollution haven hypothesis, the pollution haven hypothesis, and the Environmental Kuznets Curve. While all of these theories have been researched extensively and largely proven, this industrialization-trade model applies the theories to the South American outlier in a way not thoroughly expressed in previous development determinant-focused literature.

5 Robustness Checks

Trade-Development Model:

This paper utilizes several statistical checks to ensure the robustness and validity of the trade-development model, the variables, as well as the errors. First, a t-test was chosen to ensure all individual terms in this more complicated quadratic regression model are statistically meaningful. It was employed at the level of $\alpha = 0.025$ and using the significance requirement of $t^* > 1.96$ for infinite degrees of freedom, the development model's $|t^*| = 0.375$ for the trade-development interaction term does not meet the necessary requirements to reject the null hypothesis $H_0 : \beta_5 = 0; H_A : \beta_5 \neq 0$. Secondly, an f-test was chosen for its better capability for comparison's purposes i.e. its determination of if adding multiple variables significantly improves the model's function. Then applying an f test at $\alpha = 0.01$ and using the significance requirement of $f^* > 3.017$ for 5 degrees of freedom in the numerator and infinite degrees of freedom in the denominator, the development's model $f^* = 37.85$, the model overall does meet the requirements necessary to reject the null hypothesis $H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$; H_A: at least one $\beta_i \neq 0$. Finally, in examination the model's p-value of 0.00000000000000022, one can clearly see this is vastly lower than any reasonable standard for α (0.1, 0.05, 0.025, 0.001) and shows confidence in the model due to abnormal distribution of errors.

Trade-Industrialization Model:

This paper similarly employs several statistical checks to ensure the robustness and validity of the trade-industrialization model, the variables, as well as the errors. First, employing a t-test at $\alpha = 0.025$ and using the significance requirement of $t^* > 1.96$ for infinite degrees of freedom, the industrialization model's $|t^*| = 3.521$ for the trade-industrialization interaction term does meet the necessary requirements to reject the null hypothesis $H_0 : \theta_5 = 0, H_A : \theta_5 \neq 0$. Then applying an f-test at $\alpha = 0.01$ and using the significance requirement of $f^* > 3.017$ for 5 degrees of freedom in the numerator and infinite degrees of freedom in the denominator, the development's model $f^* = 9.795$, the model overall does meet the requirements necessary to reject the null hypothesis $H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$; H_A: at least one $\beta_i \neq 0$. Finally, in examining the model's p-value of 0.00000001011, one can clearly see this is vastly lower than any reasonable standard for α (0.1, 0.05, 0.025, 0.001) and shows confidence in the model due to abnormal distribution of errors.

Simplified Multivariate Linear Regression Model:

As a final robustness check, this paper employs a simplified version of the previous models that excludes terms like the squared industrialization and development metrics in attempt to directly compare the interaction terms without any distortions in their coefficients caused by multicollinearity. Further, this avoids the issues associated with comparing coefficients across models that may be influenced by underlying data variation or difference in standard error, thus producing a more direct and mathematically sound model for use of coefficient comparison alone.

$$Y_i = \Lambda_{1i} + \Lambda_2(X \cdot Z)_i + \Lambda_3(X \cdot D)_i + \epsilon_i \tag{5}$$

Listed below are the descriptions of the terms in this equation:

- Y_i: Greenhouse Gas Emissions (measured in metric tons of CO_2 emissions per capita) in Country i
- X_i: Trade Openness (measured as the sum of imports and exports as a percentage of GDP) in Country i
- Z_i: Development level (measured by the UN Human Development Report where each country receives a rating between 0 being the least developed and 1 being the most developed) in Country i
- D_i: Industrialization Level (measured measured by the United Nations' Industrial Development Organization's Industrialization Intensity Index where each country receives a rating between 0 being the least industrialized and 1 being the most industrialized) in Country i
- Λ_{1i}: An intercept term in Country i
- Λ_{2i}: Interaction term capturing how the effect of trade on emissions differs between developed and less developed countries
- Λ_{3i}: Interaction term capturing how the effect of trade on emissions differs between industrialized and less industrialized countries

$$Y_i = -8.20643 + 0.01345(X \cdot Z)_i - 0.11769(X \cdot D)_i + \epsilon_i \tag{6}$$

The final model allows us to directly compare the coefficients. The positive coefficient of 0.01345 on the trade-development interaction term indicates that trade openness has a slightly increasing effect on emissions as countries become more developed. Whereas the negative coefficient of -0.11769 on the trade-industrialization interaction term indicates that the impact of trade openness on emissions decreases as industrialization increases. Both these findings affirm our previous models.

The coefficient on the trade-industrialization interaction term is greater than the coefficient on the trade-development interaction term by a magnitude greater than ten. This suggests that industrialization plays a much more significant role in moderating the relationship between trade openness and emissions, reaffirming the previous models' assertions.

6 Conclusion

This paper has shown that in the South American context, industrialization is a more accurate predictor of the relationship between trade and carbon emissions than development. The unique historical and economic context of South America offers insight into inapplicability of theoretical models like the Environmental Kuznets' Curve and its traditional determinant of development, which suggests that environmental degradation follows an inverted U-shaped path as economies develop/industrialize. South American economies, shaped by legacies of colonization and dependency on resource extraction, have developed along different trajectories compared to other middle-income regions, relying heavily on primary exports rather than high-emission industrial sectors. Additionally, strong political culture enables greater economic and human development score growth for their income category as compared to the rest of the world.

South American countries pursued a model of import-substitution industrialization (ISI) through much of the 20th century. This model declined in the 1980s in favor of more neoliberal trade policies, resulting in greater integration into the global economy but with limited industrial growth. As countries moved away from ISI, their economic structures became characterized by renewed emphasis on exporting natural resources, often leaving manufacturing and more advanced industrial sectors underdeveloped. The region's abundant renewable resources, political engagement, and public awareness around environmental concerns have also played a role in maintaining relatively low per capita emissions, further distinguishing South America from other regions at similar income levels and preventing the continent's connection to the traditional Environmental Kuznets' Curve. While South American nations are largely unique in this high development – low industrialization setup, further research might explore this relationship's applicability to less developed nations than those seen in South America.

In the trade-industrialization model presented here, emissions initially increase with trade as industrial activities expand. However, an inverted-U relationship becomes evident, aligning more closely with the Environmental Kuznets Curve hypothesis and suggesting that, as industrialization reaches higher levels, emissions may level off or decline. This reflects a gradual shift towards cleaner practices and technologies, suggesting that industrial maturity, more than general human development, brings about reductions in emissions intensity as production processes improve. In contrast, the trade-development model does not show the same relationship, reinforcing the view that South America's low-emission profile is more closely tied to its limited industrialization than to broader developmental metrics.

This analysis established industrialization as a more significant determinant for the relationship between trade and carbon emissions than was available in previous literature and successfully adhered the continent to a new form of the Environmental Kuznets Curve. This shows that historical and structural differences in economies like those found in South America ought to be considered in policymaking. Their unique economic path requires a tailored approach to understanding their environmental impacts. Going forward, policies aiming to reduce carbon emissions may benefit from focusing on developing of the region's renewable energy potential. Further, this analysis shows policy makers that industrialization need not be seen as the only path to goals such as human development, environmental sustainability, or income growth, as South American economies have achieved many of these goals while focusing on developing their agricultural and natural resource sectors, and self-sufficiency-focused production goals. This study highlights the importance of integrating historical economic context into environmental research, acknowledging the impact of trade on emissions is shaped by the depth and nature of industrial activity within each country.

References

[1] Afesorgbor, Sylvanus Kwaku, and Binyam Afewerk Demena. "Trade Openness and Environmental Emissions: Evidence from a Meta-Analysis." *Environmental and Resource Economics* 81, no. 2 (February 1, 2022): 287–321. https://doi.org/10.1007/s10640-021-00627-0.

[2] Azam, Muhammad, Liu Liu, and Najid Ahmad. "Impact of Institutional Quality on Environment and Energy Consumption: Evidence from Developing World." *Environment, Development and Sustainability* 23, no. 2 (February 1, 2021): 1646–67. https://doi.org/10.1007/s10668-020-00644-x.

[3] Badía, Daniel. "The Major Challenges Facing Latin America." MAPFRE, July 18, 2024. https://www.mapfre.com/en/insights/economy/challenges-latin-america/.

[4] Benavente, José Miguel. "Economic Diversification in Latin American Countries: A Way to Face Tough Times Ahead." In *Breaking the Oil Spell*. International Monetary Fund. Accessed October 1, 2024. https://www.elibrary.imf.org/display/book/9781513537863/ch007.xml.

[5] Copeland, Brian, and M. Scott Taylor. "North-South Trade and the Environment." *The Quarterly Journal of Economics* 109, no. 3 (1994): 755–87. https://econpapers.repec.org/article/oupqjecon/v_3a109_3ay_3a1994_3ai_3a3_3ap_3a755-787.htm.

[6] Economic Commission for Latin America and the Caribbean. "Poverty Rates in Latin America Remain Above Pre-Pandemic Levels in 2022, ECLAC Warns." Accessed October 1, 2024. https://www.cepal.org/en/pressreleases/poverty-rates-latin-america-remain-above-pre-pandemic-levels-2022-eclac-warns.

[7] Duran, Xavier, Aldo Musacchio, and Gerardo Della Paolera. "Industrial Growth in South America: Argentina, Brazil, Chile, and Colombia, 1890–2010." In *The Spread of Modern Industry to the Periphery since 1871*, edited by Kevin Hjortshøj O'Rourke and Jeffrey Gale Williamson, 1st ed., 318–42. Oxford University Press, 2017. https://doi.org/10.1093/acprof:oso/9780198753643.003.0013.

[8] Elmarzougui, Eskandar, Bruno Larue, and Lota D. Tamini. "Trade Openness, Domestic and Foreign Investments, and the Environment." *Modern Economy* 7, no. 5 (May 3, 2016): 591–605. https://doi.org/10.4236/me.2016.75065.

[9] Frankel, Jeffrey, and Andrew K. Rose. "Is Trade Good or Bad for the Environment? Sorting Out the Causality." *Working Paper Series*, September 2003. https://ideas.repec.org//p/ecl/harjfk/rwp03-038.html.

[10] Hassan, Mahmoud, Walid Oueslati, and Damien Rousselière. "Environmental Taxes, Reforms and Economic Growth: An Empirical Analysis of Panel Data." *Economic Systems* 44, no. 3 (September 1, 2020): 100806. https://doi.org/10.1016/j.ecosys.2020.100806.

[11] Herre, Bastian, Pablo Arriagada, and Max Roser. "The Human Development Index and Related Indices: What They Are and What We Can Learn from Them." *Our World in Data*, February 12, 2024. https://ourworldindata.org/human-development-index.

[12] Kuznets, Simon. "Economic Growth and Income Inequality." *American Economic Review* 45 (March 1955).

[13] Le, Thai-Ha, Youngho Chang, and Donghyun Park. "Trade Openness and Environmental Quality: International Evidence." *Energy Policy* 92 (May 1, 2016): 45–55. https://doi.org/10.1016/j.enpol.2016.01.030.

[14] Masera, Diego. "Industrialization in Latin America and the Caribbean: Challenges and Opportunities." *Industrial Analytics Platform*, April 21, 2022. https://iap.unido.org/articles/industrialization-latin-america-and-caribbean-challenges-and-opportunities.

[15] O'Neil, Shannon. "Why Latin America Lost at Globalization—and How It Can Win Now." *Council on Foreign Relations*, August 25, 2022. https://www.americasquarterly.org/article/why-latin-america-lost-at-globalization-and-how-it-can-win-now/.

[16] Pethig, Rüdiger. "Pollution, Welfare, and Environmental Policy in the Theory of Comparative Advantage." *Journal of Environmental Economics and Management* 2, no. 3 (1976): 160–69. https://econpapers.repec.org/article/eeejeeman/v_3a2_3ay_3a1976_3ai_3a3_3ap_3a160-169.htm.

[17] Porter, Michael E, and Claas Van Der Linde. "Toward a New Conception of the Environment-Competitiveness Relationship." *Journal of Economic Perspectives* 9, no. 4 (November 1, 1995): 97–118. https://doi.org/10.1257/jep.9.4.97.

[18] Walter, I, and JL Ugelow. "Environmental Policies in Developing Countries." *Ambio*, 1979, 102–9.

[19] Wang, Zhen, Cai Li, Qiaoling Liu, Beibei Niu, Sha Peng, Liangchun Deng, Ping Kang, and Xiaoling Zhang. "Pollution Haven Hypothesis of Domestic Trade in China: A Perspective of SO2 Emissions." *Science of The Total Environment* 663 (May 1, 2019): 198–205. https://doi.org/10.1016/j.scitotenv.2019.01.287.

The Impact of Metro Accessibility on Residential Property Values and Population Clustering: A Spatial Analysis of Washington, D.C.

Maryana Shnitser

George Washington University

Abstract

The Washington, DC Metro system is the third largest heavy rail system in the US, but only 30% of DC residents live within walking distance of a Metro station (Sustainable DC, n.d.). This study examines the relationship between proximity to metro stations and residential property values using a hedonic pricing model implemented by Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR) models. The results indicate that property values tend to decrease as distance from Metro stations increases, with OLS revealing a consistent overall positive relationship between proximity and property value. The GWR model, however, allows spatial variability, demonstrating that the effect of Metro proximity on property values varies across neighborhoods. A Getis-Ord Gi^* analysis of population clustering identifies spatial clusters where residents are densely concentrated near Metro stations, emphasizing the role of transit access in influencing urban density and property demand. These results show the role of transit access in shaping urban real estate markets and provide empirical evidence for targeted transit-oriented development (TOD) strategies aimed at increasing property values and accessibility in Washington, DC. **JEL Codes:** R41, R30, C21

1 Introduction

As urbanization continues to reshape cities, effective public transportation is critical to improving connectivity and accessibility. Public transit systems allow residents to easily access employment centers, recreational activities, and essential services. Public transit also reduces traffic congestion and urban sprawl (Choi et al., 2021; Yang et al., 2024). This study estimates the relationship between proximity to metro stations and residential property values in Washington, DC. Additionally, the study explores population data to determine whether there is significant population clustering within half a mile of metro stops. Understanding how public transportation affects both property values (price) and population concentration (quantity) is critical for evaluating the welfare impact of public transit and the spatial patterns of urban growth.

Residential housing is a competitive market, and people's willingness to pay for properties near metro stations reveals preferences. Past studies have consistently concluded that there is a positive relationship between proximity to public transit systems and property values. This existing research has predominantly focused on generalized impacts of metro proximity without accounting for spatial variations or spatial clustering. Overlooking critical spatial variations often results in high levels of spatial autocorrelation in regression results, which distorts finding. A study by Benjamin and Sirmans (1996) that focused on Washington, D.C. found that distance to metro stops significantly influenced rental prices, with each one-tenth mile increase in distance from the station resulting in a decrease in rent of about 2.50%. This study aims to provide more recent estimates for Washington, DC that also consider local variations in coefficient values.

This paper uses an Hedonic Pricing Model (HPM) to account for variations in housing characteristics, neighborhood characteristics, and distances to metro stations. It uses an Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR) to quantify the impact of metro proximity on property values and compare whether the error caused by spatial autocorrelation is reduced when taking into account spatial variations with a GWR. This builds on the methods presented in a study by Zhang et al. (2021), which analyzed the impact of proximity to transit centers on property values in the Stamford and Hartford areas in Connecticut and found that GWR had better modeling performance than OLS.

A study by Tan et al. (2019) used remote sensing data to analyze land use changes before and after new metro lines were constructed. The study compared high resolution satellite imagery and used remote sensing analysis to detect land use changes around metro stations (vacant land to building and buildings to vacant land). Population changes were analyzed before and after the opening of metro stations using Difference-in-Differences (DiD) models. The study found that the population density increased in central stations and suburban stations experienced land revitalization. This study incorporates population clustering analysis as a proxy for infrastructure development patterns to understand how metro accessibility affects population concentration, using a Getis-Ord Gi analysis to identify statistically significant clusters of high or low population densities.

Understanding these relationships has implications for future planning in Washington, DC. and other cities as they look to invest in public transit infrastructure. Transit infrastructure is expensive, and planners generally expect that investments will not only enhance mobility but also increase property values in surrounding areas. This increase in property values generates higher tax revenues that benefit the city in the long run (American Public Transportation Association, 2015). Using a GWR to map local coefficient values and understanding population clustering allows urban planners to make more informed decisions about resource allocation and utility maximization (Dziauddin et al., 2014; AlQahtani and Anjomani, 2021). These insights can guide the development of new transit lines or stations and improvements to current service that maximize economic and community benefits.

2 Study Area and Data

Washington, D.C. (District of Columbia) is the capital of the United States (US). It is located on the east coast of the US, bordered by Maryland (MD) to the north and the east and by Virginia (VA) to the south and the west. The city is served by the Metro Rail and the Metro Bus system, which are operated by the Washington Metropolitan Area Transit Authority (WMATA). The Circulator is another bus service that was operated by the DC Department of Transportation (DDOT) in public-private partnership with RATPDev since 2005. A phase out of the Circulator began on October 1, 2024, with an end of service date of December 31, 2024. WMATA and DDOT are working to design and implement a new 'Better Bus Network' that will expand Metro Bus service to areas previously served by the Circulator. The 'Better Bus Network' is expected to be completed in July 2025 (District Department of Transportation, 2024). Because of these service changes, this study only analyzes the impact of proximity to Metro Rail stations in Washington, DC. There are 40 Metro Rail stations across the 68.35 square miles of DC.

This study uses residential property values from Redfin Data Center for properties sold between October 2023 and October 2024. Incorrect data samples with null fields are removed. If there are multiple samples within the same apartment building, one unit is randomly selected as a representative sample for that building because the GWR cannot be run with duplicate longitude and latitude fields. In total, this study uses 4,073 properties across DC. The property dataset contains information on the number of bedrooms, number of bathrooms, square footage, lot size, type of home, and age of home. This study uses median income, educational attainment, and poverty rate by census tract, obtained from the US Census Bureau's American

Community Survey (ACS) of economic and social characteristics. These variables capture key socio-economic characteristics that are relevant to property values.

Median income represents the average economic status of residents within each census tract. Higher median incomes are often an indicator of greater purchasing power, which can increase demand for housing and drive up property values (Brigham, 1965). Poverty rate captures socio-economic conditions that may negatively affect property values. Higher poverty rates are often associated with reduced demand for housing, lower investment in neighborhood infrastructure, and potentially higher crime rates, which may deter prospective buyers (Ware, 2014). This study measures educational attainment as the percentage of residents with at least a Bachelor's degree in each census tract. This variable acts as an indicator of knowledge-based economic and social capital. Higher levels of educational attainment are often associated with higher wages, better school districts, and more stable property markets, all of which can contribute to higher property values (Brigham, 1965). This study uses a national park shapefile from OpenDataDC and the Euclidean Distance tool in ArcGIS Pro to calculate the distance to parks variable. This variable is included to control for neighborhood amenities that are unrelated to metro proximity. Proximity to parks can increase property values, especially in urban areas where access to green space may be limited, by making properties more desirable for the recreational opportunities and aesthetic values that they provide (Anderson et al., 2006). These variables were selected as key indicators of factors influencing property values, while acknowledging that they are not exhaustive. Variables such as crime rates and proximity to primary schools were excluded to reduce potential collinearity and focus on distinct socio-economic factors tied to property values.

Shapefiles metro lines and stations are publicly available on OpenDataDC. This study uses the Euclidean Distance tool in ArcGIS Pro to calculate the distance to the nearest metro stop. Euclidean Distance calculates the shortest straight line distance to the nearest point, which usually does not reflect the actual travel distance in urban areas as street networks and buildings alter the path that a person can take to reach a metro station. In this study, waterbodies, national parks, and military bases are erased from the DC shapefile when calculating the distance to metro stations because it is assumed that a person could not cross these land features to access a metro station. A street network dataset of DC was not available, but would yield better insights by incorporating more accurate travel distances and times into the analysis. Figure 1 shows the location of the properties used in this analysis in DC, Figure 2 provides summary statistics of explanatory variables, and Figure 3 provides reference maps for DC wards and quadrants that will help contextualize results presented later in the paper.

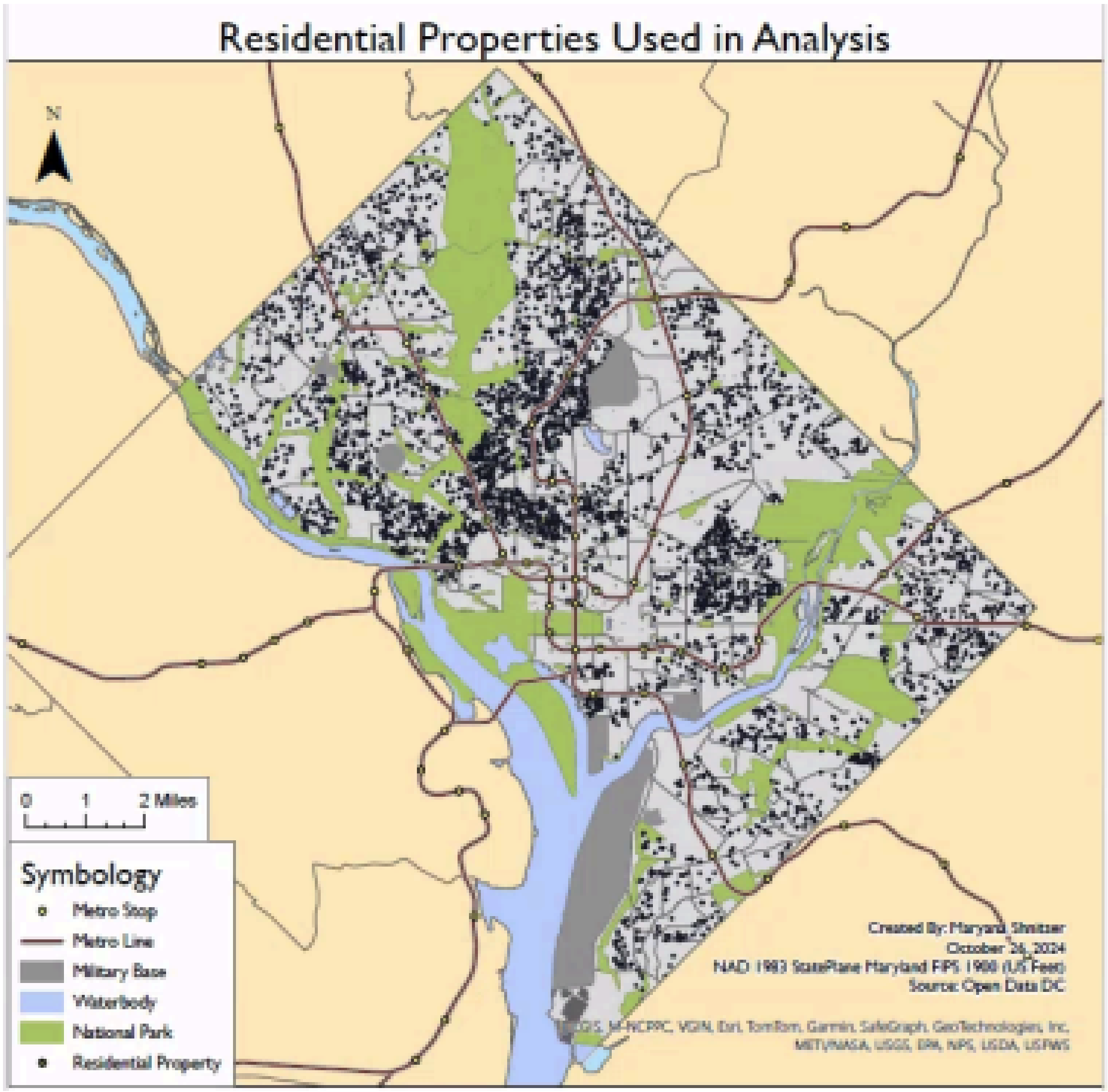


Figure 1: Map of Property Samples in DC

Variables	Min	Mean	Max	Std. deviation
House Characteristics				
Bedroom	0	3.11	14	1.44
Bathroom	1	2.54	18	1.34
Property Type	1	2.03	5	0.77
Size (ft sq)	299	1,981	246,315.00	4,027.46
Lot size (sq ft)	0	2,445	155,905	4021.08
Year Built	1	3.61	3.61	1.76
Neighborhood Characteristics				
Median Income	0	83,108.53	167,031	29774.61
Educational Attainment (%)	12	47.31	96.23	20.16
Poverty Rate (%)	0	7.93	52.8	9.98
Distance to Parks (mi)	0	0.102	0.502	0.07
Distance to Transit				
Distance to Metro Stations (mi)	0	0.725	2.49	0.451

Figure 2: Summary Statistics of Explanatory Variables

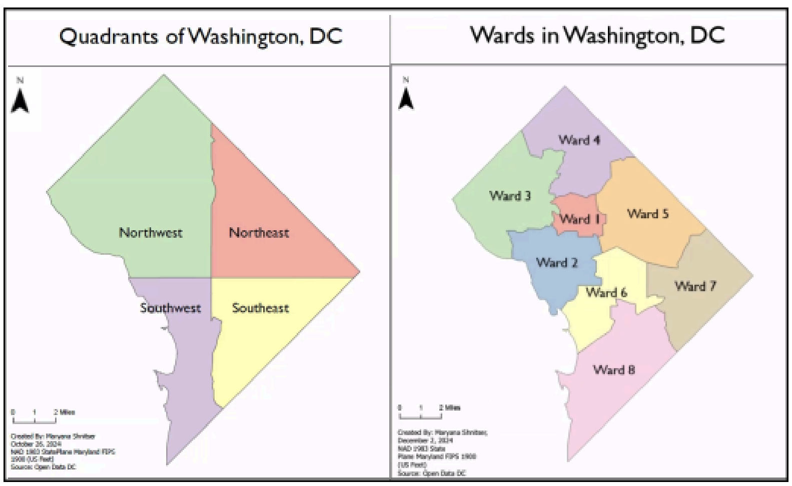


Figure 3: The map on the left shows the quadrants of DC and the map on the right shows the eight wards of DC

3 Methods

Hedonic pricing models are commonly used in evaluating residential property values because houses have different characteristics, and they are purchased based on these characteristics and their prices. Therefore, its value can be attributed to these factors. The following model separates these effects to isolate and estimate the impact of metro proximity on residential property values while controlling for property-specific and neighborhood traits (Zhang et al., 2021; Bajari and Kahn, 2003):

$$P = f(N, H, T, e) \tag{1}$$

Listed below are the descriptions of the terms in this equation:

- P: Price of the Property
- N: Combination of Neighborhood Traits (median income, rate, educational attainment, poverty rate, and distance to parks)
- H: Combination of House Characteristics (number of bedrooms, number of bathrooms, square footage, lot size, type of home, and age of house)
- T: Distance to a Metro Station
- e: Error Term

The property values are log transformed to have a normal distribution. The multiple regression is specified as follows:

$$\begin{aligned}\log(\text{Property\_Value}) = {} & \beta_0 + \beta_1(\text{Bedroom}) + \beta_2(\text{Bathroom}) + \beta_3(\text{Square\_Footage}) \\ & + \beta_4(\text{Lot\_Size}) + \beta_5(\text{Type\_of\_Home}) + \beta_6(\text{Age\_of\_Home}) \\ & + \beta_7(\text{Median\_Income}) + \beta_8(\text{Distance\_to\_Park}) + \beta_9(\text{Poverty\_Rate}) \\ & + \beta_{10}(\text{Education}) + \beta_{11}(\text{Distance\_to\_Metro}) + \epsilon \end{aligned} \tag{2}$$

OLS is a commonly used regression technique that models the relationship between a single dependent variable and related explanatory variables. However, OLS produces a single set of coefficients across a study area, even though many relationships are not spatially homogeneous. The index used to measure spatial autocorrelation in OLS results is Global Moran's Index (hereafter, Moran's I). Real estate data is inherently spatial data, as every unit has a location. Therefore, the following GWR model is implemented to reduce spatial autocorrelation:

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{p} x_{ij}\beta_j(u_i, v_i) + \epsilon_i \tag{3}$$

Listed below are the descriptions of the terms in this equation:

- (u_i, v_i): Coordinates of the i^{th} point
- $\beta_0(u_i, v_i)$: Location of specific intercept
- $\beta_j(u_i, v_i)$: Location specific intercept at point i
- x_{ij}: j^{th} variable related to $\beta_j(u_i, v_i)$
- p: Number of local parameters to be estimated
- e: Error Term

GWR is an extension of OLS that recognizes spatial variations and gives observed data near location $\beta_j(u_i, v_i)$ more influence than those further away, allowing coefficients to vary with location (Brunsdon, Charlton, and Fotheringham, 1998).

Additionally, this paper uses a Getis-Ord Gi^* hot-spot analysis to identify clusters of high or low values within a dataset and explores the distribution of local spatial clustering groups. It identifies hot spots (high-value clusters) and cold spots (low-value clusters) in spatial variables by looking at the feature within the context of neighboring features. The local sum for a feature and its neighbors is compared proportionally to the sum of all features. When the local sum is very different from the expected local sum and too large to result from chance, the result is a statistically significant z-score. A larger z-score indicates more intense clustering of high values, and a smaller z-score indicates clustering of low values. The following equation calculates the Gi^* statistic:

$$G_i^* = \sum_{j=1}^{n} w_{i,j}x_j - \frac{X\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{n\sum_{j=1}^{n} w_{i,j} - (\sum_{j=1}^{n} w_{i,j})2}{n-1}}} \tag{4}$$

Listed below are the descriptions of the terms in this equation:

- G_i^*: Standardized statistics for location i
- x_j: Value of the population at location j
- $w_{i,j}$: Spatial weight between location i and j
- n: Number of features
- X: Mean of attributes across all locations
- S: Standard deviation across the study area

Hotspots represent clusters of high population density, and coldspots are areas of low population density (Ord and Getis, 1995).

4 Results

Variables	Coefficient	Robust Std. Error	Robust t-statistic	VIF
Intercept	11.921167	0.047637	250.250551**	
House Characteristics				
Bedroom	0.103302	0.013696	7.542702**	2.784506
Bathroom	0.262666	0.013638	19.259472**	2.620354
Property Type	0.032986	0.015431	2.137641**	1.957077
Size (ft sq)	0.000003	0.000004	0.653146	1.366584
Lot size (sq ft)	0.000008	0.000002	4.702016**	1.775484
Year Built	-0.020846	0.004021	-5.18408**	1.11706
Neighborhood Characteristics				
Median Income	0.000007	0.00	19.241827**	1.832538
Educational Attainment (%)	0.004765	0.000466	10.231204**	1.864643
Poverty Rate (%)	-0.005938	0.000928	-6.401092**	1.521921
Distance to Parks	-0.000146	0.000023	-6.235928**	1.176817
Distance to Transit				
Distance to Metro Stations	-0.073255	0.016129	-4.541792**	1.154427
Number of observations: 4,073				
R-squared value: 0.681138				
Adjusted R-squared value: 0.680275				
Akaike's Information Criterion (AICc): 4684.816791				

Figure 4: Results from OLS Regression. Statistically significant results ($p < 0.05$) are indicated by: **

The results from global OLS regression are shown in Figure 3. Almost all variables are statistically significant at the 5% level. The negative coefficient of 'Distance to Metro Station' variable indicates a positive relationship with proximity to metro stops because value decreases with each increase in distance from metro stops. For each one-mile increase in distance from a metro station, property values decrease by about 7.33%, all else being equal. This is smaller than previous studies specific to DC, such as the Benjamin and Sirmans (1996) paper that estimated a 2.50% decrease for each one-tenth mile, or 25% for each one mile increase. Figure 4 shows results from the Moran's I that this study uses to analyze spatial autocorrelation by checking whether the residuals exhibit a random spatial pattern: The results indicate that there is very

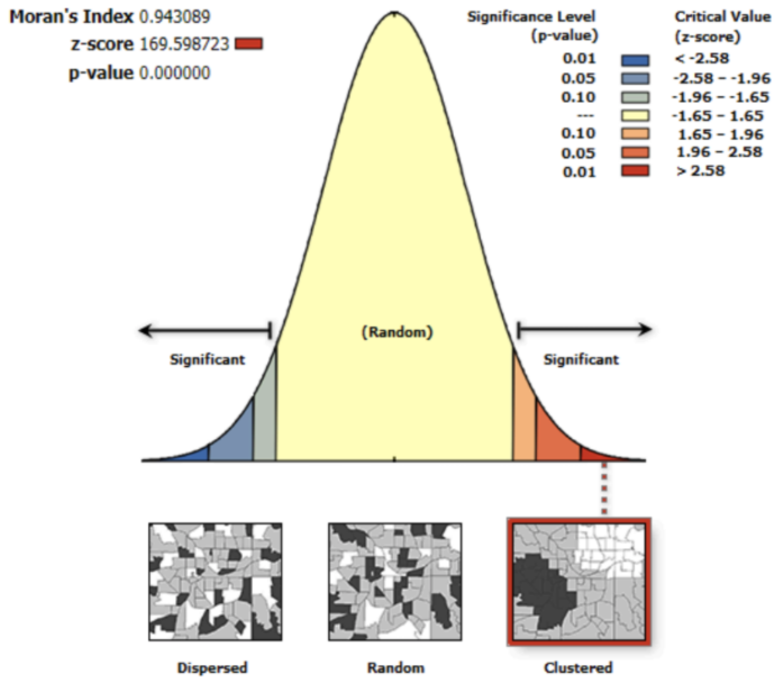


Figure 5: Results from Global Moran's I, a test of spatial autocorrelation

strong positive spatial autocorrelation. The z-score of 169.60 shows that there is a less than 1% likelihood that the clustered pattern is a result of random chance.

The GWR mostly confirms the results of the OLS model while also describing non-stationary spatial relationships. Because this study is focused on the impacts of proximity to public transit, Figure 6 only shows the local coefficients of the explanatory variable 'Distance to Metro Stop' from the GWR. The GWR produces a map of continuous coefficient estimates that allows for a more nuanced understanding of how local conditions affect property values. The GWR shows a substantial improvement in model fit with an adjusted R^2 of 0.8165, higher than that of the OLS model, 0.680275. The increase in the R^2 value indicates that accounting for spatial heterogeneity significantly improves the model's ability to capture patterns in data. The AICc value also decreased from 4684.82 to 2858.71. This provides strong evidence that the GWR is better suited for accounting spatial variability in property values when compared to the OLS model and that failing to account for this variability leaves a lot of the spatial structure in the observations unexplained.

These coefficients, ranging from 0.79 to -0.50, indicate the presence of spatial vari-

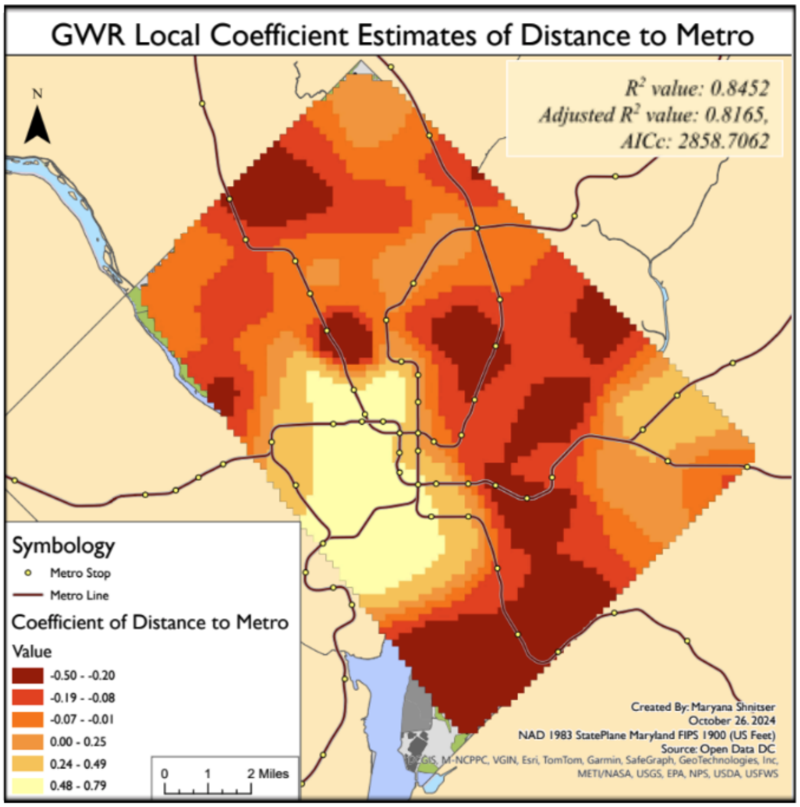


Figure 6: Map showing local coefficient of the 'distance to metro stop' variable on residential property values

ation, with some areas having positive relationships. The GWR estimates coefficient values at all locations, not just those with observed data points which is usually an advantage of GWR. However, in the case of DC, this leads to unusual effects. Land use throughout the city must be considered to properly interpret these results. A majority of areas that have high positive coefficients, that indicate that property values increase as distance to metro stops increases, are in and around the National Mall. Much of this land is dominated by government buildings, museums, and national parks, which are owned and operated by the federal government and the National Park Service (NPS). The other notable area of positive coefficient values is located by the northeast corner of DC. This area includes the National Arboretum, which is another plot of land where traditional market dynamics may not apply. Both of these areas are commercial and tourist hubs that do not have many residential property data points, so typical dynamics may not apply even though coefficients are estimated across the entire study area.

The positive coefficient extends into Downtown neighborhoods and Foggy Bottom. These results may be affected by the proximity of these residential properties to federally used land, which lacks residential property samples. Another contributing factor could be the presence of the George Washington University, where proximity to campus may outweigh the appeal of metro access. Students, who are the primary renters in this area, may be willing to pay a price premium for proximity to the university campus rather than transit accessibility. Additionally, this trend might reflect the neighborhoods' access to other high-value amenities such as the State Department, the Kennedy Center, and the White House. Residents in these areas may value proximity to these centers over transit convenience. High walkability scores in and the abundance of surrounding restaurants, stores, and cultural attractions in a compact geographic area may also reduce reliance on public transit. Negative coefficients may be seen further from the city center because there may be a greater reliance on public transit in these areas to reach central areas. While these are potential explanations for the positive coefficients in these areas, further research is needed to fully understand the factors contributing to this effect in these outlying neighborhoods.

Despite these exceptions, there are areas with negative coefficient values across most residential areas of DC that are outside of central zones. Living in these areas allows residents to pay relatively lower rents or mortgage costs while benefiting from accessible metro stations, providing a balance of affordability and connectivity. This suggests that properties closer to metro stations have higher values because they provide convenient access to public transportation, a valuable amenity for many urban residents. Proximity to metro stations allows residents to easily commute without relying on personal vehicles. For example, this access allows people to reach jobs in the Central Business District of DC, as well as in nearby "Edge Cities" that serve as major commercial and employment hubs, such as Bethesda and Silver Spring in Maryland and Tysons in Virginia. This accessibility increases demand for housing in these areas as it attracts residents who prefer to live in urban areas. The results of the GWR support that this preference is correlated with higher property values in many areas. This aligns with previous studies that conclude that the improvement of accessibility to employment and other amenities provided by transit access add premiums on residential property values but with spatial variations over geographic areas, producing positive effects in some areas, but negative effects in others (Dziauddin, Powe, and Alvanides 2014; Du and Mulley, 2006; Du and Mulley, 2012; Yang et al., 2024; Zhang et al., 2021).

Visualizing these coefficients is important to identify areas where investments in transit might yield the most significant impact on property values, guiding urban planning and development strategies. It is also helpful in identifying underserved areas to create targeted interventions that address people's needs for access to services and to transportation (Yang et al., 2024). A notable area with high negative coefficient values is in ward eight, located in the lower southeast corner of DC. This may be because ward eight has historically had limited access to amenities and services when compared to other areas of the city. Public transportation tends to be the main form of transportation for many residents to access jobs and essential services in other parts of the city (Census Reporter, n.d.). This means that proximity to metro stations significantly reduces commute times and makes these properties more desirable, leading to the large impact of metro proximity on property values (Dziauddin, Powe, and Alvanides 2014). Ward eight also has limited access to public transportation, with one metro station serving the entire area. With the information provided by the GWR, this is one of the places that could be identified as an area that could benefit from improvements to service to make public transportation services equitable and accessible to all residents, including low-income and minority populations. While the OLS regression gave a constant coefficient for the entire study area, visualizing how coefficient values vary across space with a GWR suggests that the influence of metro proximity on residential property value is contextual and requires spatial considerations such as land use and existing socio-economic differences.

Analyzing population clustering alongside property values can also reveal preference for proximity to metro stations. When metro stops open, there tends to be an increase in residential property development to meet the demand for housing near public transit stations (Tan et a., 2019). While this study does not specifically examine changes in residential property availability, examining areas of high population clustering in relation to metro stops indicates demand for metro accessible residential properties (AlQahtani and Anjomani, 2021). It also shows that there may have been an increase in property development to accommodate this demand in these areas (Champagne, Dubé, and Barla, 2022). Figure 7 shows areas of significant population clustering by census block, as determined by the Getis-ord Gi^* analysis, and Figure 8 overlays a 0.5 mile buffer to provide a visual representation of how this relates to metro proximity:

Areas of negative clustering are mostly seen in southwest DC, which is where the National Mall and other federal buildings are located. Areas of high population clustering are mostly located along metro lines that are closer to downtown in northwest DC. This suggests a greater reliance on public transit in these neighborhoods, while areas in northeast DC and the northwest corner may be more car dependent.

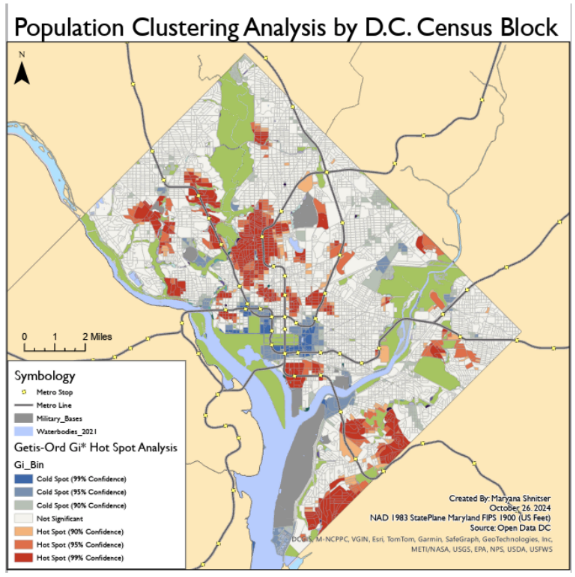


Figure 7: Getis-Org Gi^* index showing population clustering in DC

When comparing these results to the GWR coefficients, the analysis reveals that areas with the highest negative coefficient values for metro distance, where property values decline sharply as distance from metro stops increases, do not show significant population clustering. This effect may be attributed to limited housing availability near these metro stations. Here, increasing the housing supply near metro stations might help meet demand for housing near transit stations.

Understanding areas with high population density can offer insights into the dynamics of transit-oriented development (TOD). TOD is the development of higher density housing, retail, and commercial spaces around transit hubs. It promotes walkable communities and reduces dependency on personal vehicles (Zamir et al., 2024). In the context of Washington, DC, TOD can lead to increased property values by creating more desirable places to live because there is a synergetic relationship between rail proximity and walkability (Duncan, 2011). As property values rise, ridership on public transit often increases as well, generating revenue that can help offset the costs associated with building and improving transportation systems (Cervero, 2007). This helps understand the relationship between transit access and housing market dynamics that can inform the strategic planning of transit stops and property development (Dziauddin, Powe, and Alvanides 2014).

By addressing both general trends and spatial variability, this study provides insights for urban planning and policymaking. The results of this analysis highlight the importance of considering spatial relationships when examining the relationships

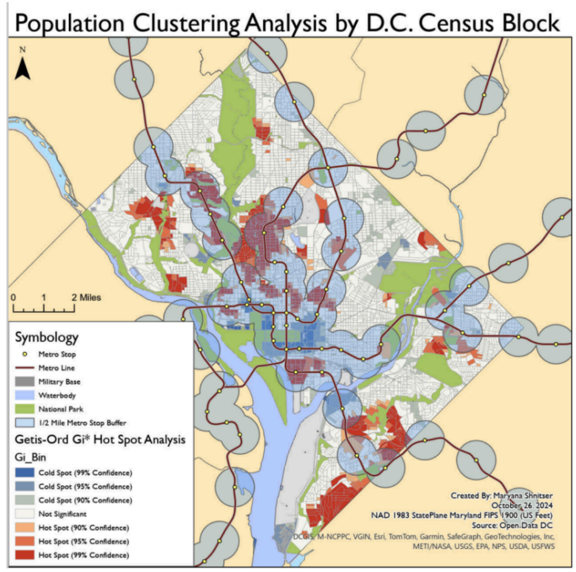


Figure 8: Getis-Ord Gi^* index showing population clustering in DC overlaid with a buffer showing areas within a 0.5 mile radius of metro stops

between transit access, property values, and population density. While the OLS coefficient value aligns with previous studies that conclude that there is a positive relationship between metro proximity and property values, the GWR reveals that this effect varies across space because of spatial heterogeneity in the explanatory variables. The GWR confirms that proximity to metro stations has a significant, spatially variable effect on property values in Washington, DC, emphasizing the need to incorporate spatial considerations into analyses of property values and TOD. Because of these relationships, urban planners should account for localized dynamics when designing strategies to maximize the economic and social benefits of transit investments.

5 Conclusion

This study assesses the impact of metro proximity on residential property values in Washington, DC using both a hedonic pricing model implemented by OLS and GWR models. In total, this study includes 4,073 properties. The findings indicate that there is a positive relationship between proximity to public transit and property values. OLS results provide a baseline that indicates that as distance from a metro station increases by one mile, property values decrease by approximately 7.33%. The GWR model reveals substantial spatial heterogeneity in this relationship, demonstrating that the effect of metro accessibility on property values varies significantly across neighborhoods. This is because spatial data has spatial dependence, and the GWR takes local spatial relationships into consideration. The GWR produces better data fitting with lower AICc and adjusted R^2 values compared to the OLS.

In residential areas where metro stations or properties near metro stations are scarce, proximity to transit plays a disproportionately strong role in shaping property values, likely because of added accessibility benefits for residents. In contrast, some areas near the National Mall and downtown DC show positive GWR coefficients, reflecting a decreased reliance on transit in favor of other high-value amenities and unique local characteristics. Further, a Getis-Ord Gi^* hot-spot analysis that identifies clusters of high population density. The results emphasize the concentration of dense, transit-oriented communities around particular metro stops in DC, while other regions with high property value sensitivity to metro proximity remain lower density, suggesting potential for targeted transit-oriented development.

These results may provide future guidance for transit development in Washington DC and insights to guide initiatives focused on improving housing accessibility and transit options in high-demand, transit-dependent areas. The results also emphasize the need for spatially differentiated planning strategies and investment, particularly in underserved areas. These plans can increase property values and improve access to essential urban amenities, ultimately generating revenue for the city. As DC pursues sustainable development, enhancing public transit access is crucial to reducing car dependency and urban sprawl.

Future research may benefit from incorporating precise travel networks that better represent distances to metro stations. It could also incorporate additional explanatory variables that provide more detailed neighborhood characteristics. Additionally, expanding this analysis to metro accessible areas in Maryland and Virginia to analyze the impact of metro proximity and population clustering on suburban areas that tend to have lower population density would create a better regional understanding of urban dynamics.

References

[1] Acuña, Rhea. 2023. "End of the Line: The Impact of New Suburban Rail Stations on Housing Prices." *Journal of Transport and Land Use* 16 (1): 67–86. https://doi.org/10.5198/jtlu.2023.2199.

[2] AlQuhtani, Saad, and Ardeshir Anjomani. 2021. "Do Rail Transit Stations Affect the Population Density Changes around Them? The Case of Dallas-Fort Worth Metropolitan Area." *Sustainability* 13 (6): 3355. https://doi.org/10.3390/su13063355.

[3] American Public Transportation Association. 2015. *Value Capture for Public Transportation Projects.* ATPA. https://www.apta.com/wp-content/uploads/Resources/resources/reportsandpublications/Documents/APTA-Value-Capture-2015.pdf.

[4] Anderson, Soren T., and Sarah E. West. 2006. "Open Space, Residential Property Values, and Spatial Context." *Regional Science and Urban Economics* 36 (6): 773–89. https://doi.org/10.1016/j.regsciurbeco.2006.03.007.

[5] Benjamin, John, and Stacy Sirmans. 1996. "Review of Mass Transportation, Apartment Rent and Property Values." *The Journal of Real Estate Research* 12 (1): 1–8. https://www.jstor.org/stable/pdf/44152427.

[6] Brigham, Eugene F. 1965. "The Determinants of Residential Land Values." *Land Economics* 41 (4): 325. https://doi.org/10.2307/3144665.

[7] Brunsdon, Chris, A. Stewart Fotheringham, and Martin E. Charlton. 1996. "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity." *Geographical Analysis* 28 (4): 281–98. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x.

[8] District Department of Transportation. 2024. *Bus Priority Program.* Accessed October 30, 2024. https://buspriority.ddot.dc.gov/.

[9] Census Reporter. "Census Profile: Ward 8, DC." Accessed October 30, 2024. http://censusreporter.org/profiles/61000US11008-ward-8-dc/.

[10] Cervero, Robert. 2007. "Transit-Oriented Development's Ridership Bonus: A Product of Self-Selection and Public Policies." *Environment and Planning A: Economy and Space* 39 (9): 2068–85. https://doi.org/10.1068/a38377.

[11] Cervero, Robert, and Michael Duncan. 2002. "Transit's Value-Added Effects: Light and Commuter Rail Services and Commercial Land Values." *Transportation Research Record: Journal of the Transportation Research Board* 1805, no. 1 (January 2002): 8–15. https://doi.org/10.3141/1805-17.

[12] Champagne, Marie-Pier, Jean Dubé, and Philippe Barla. 2022. "Build It and They Will Come: How Does a New Public Transit Station Influence Building Construction?" *Journal of Transport Geography* 100 (April): 103320. https://doi.org/10.1016/j.jtrangeo.2022.103320.

[13] Choi, Kwangyul, Han John Park, and Jim Dewald. 2021. "The Impact of Mixes of Transportation Options on Residential Property Values: Synergistic Effects of Walkability." *Cities* 111 (April): 103080. https://doi.org/10.1016/j.cities.2020.103080.

[14] Damm, David, Steven R. Lerman, Eva LernerLam, and Jeffrey Young. 1980. "Response of Urban Real Estate Values in Anticipation of the Washington Metro." *Journal of Transport Economics and Policy* 14 (3): 315–36. https://doi.org/10.2307/20052588.

[15] Duncan, Michael. 2021. "The Impact of Transit-Oriented Development on Housing Prices in San Diego, CA." *Urban Studies* 48 (1): 101-127. https://doi.org/10.1177/0042098009359958.

[16] Dziauddin, Mohd Faris, Neil Powe, and Seraphim Alvanides. 2014. "Estimating the Effects of Light Rail Transit (LRT) System on Residential Property Values Using Geographically Weighted Regression (GWR)." *Applied Spatial Analysis and Policy* 8 (1): 1–25. https://doi.org/10.1007/s12061-014-9117-z.

[17] Nasri, Arefeh, and Lei Zhang. 2014. "The Analysis of Transit-Oriented Development (TOD) in Washington, D.C. And Baltimore Metropolitan Areas." *Transport Policy* 32 (March): 172–79. https://doi.org/10.1016/j.tranpol.2013.12.009.

[18] Pan, Qisheng, Haixiao Pan, Ming Zhang, and Baohua Zhong. 2014. "Effects of Rail Transit on Residential Property Values." *Transportation Research Record: Journal of the Transportation Research Board* 2453 (1): 118–27. https://doi.org/10.3141/2453-15.

[19] Ryan, Sherry. 1999. "Property Values and Transportation Facilities: Finding the Transportation-Land Use Connection." *Journal of Planning Literature* 13 (4): 412–27. https://doi.org/10.1177/08854129922092487.

[20] Zhang, Bo, Weidong Li, Nicholas Lownes, and Chuanrong Zhang. 2021. "Estimating the Impacts of Proximity to Public Transportation on Residential Property Values: An Empirical Analysis for Hartford and Stamford Areas, Connecticut." *ISPRS International Journal of Geo-Information* 10 (2): 44. https://doi.org/10.3390/ijgi10020044.

Assessing the Tax Incentives for Electric Vehicles: Effectiveness and Evidence from the United States Based on Model and State Variation

Tsung-Han (Henry) Tsai

George Washington University

Abstract

This study examines the impact of electric vehicle (EV) tax credits on the sales and adoption of EVs. I analyze the quarterly sales data of EV models from 2020 to 2024 that qualify for tax incentives following the Inflation Reduction Act (IRA), relative to their non-EV counterparts. The introduction of the IRA is associated with a significant increase in EV sales, ranging from 80% to 90%. I also computed the elasticity of EV registrations for each unit of government tax credit using the maximum tax credit provided by each state. I find that a $1000 tax rebate is associated with an increase in the number of EV registrations by 15.7%. Collectively, these results indicate a potential increase in consumer responsiveness towards tax credits over time. **JEL Codes:** H23, L62, Q58

1 Introduction

The Biden administration's 2022 Inflation Reduction Act (IRA) introduced the Clean Vehicle Tax Credit, replacing the previous tax benefits associated with the Energy Improvement and Extension Act of 2008. This act offers a tax credit of up to $7,500 for new electric vehicle (EV) buyers, depending on the vehicle's model, the components used, and the assembly location (Buckberg, 2023). The goal is to promote a more sustainable transportation system by making EVs and plug-in hybrid electric vehicles (PHEVs) more financially accessible. Some state governments also complement federal incentives with their tax credit programs, further encouraging EV adoptions. Examining the impact of these tax credits on EV sales is critical for evaluating the policy's effectiveness. This evaluation can contribute to the extensive body of EV diffusion literature and deepen our understanding of fiscal incentives' effects on consumer behavior, providing essential information for policymakers to improve future EV policies.

Econometric methods can be used to determine if tax incentives increase EV sales by isolating the causal impact of tax credits. The first model in this study examines vehicle variation, using tax credits for new EV buyers as the independent variable. The "policy shock" is the tax credit reform linked to the IRA that began on January 1, 2023. Quarterly new EV sales are the dependent variable, treated as the natural log of the number of new vehicles sold. Employing a difference-in-differences (DiD) design, the treatment group will include nine popular EV models previously ineligible or only partially eligible for tax credits but became eligible after the IRA (IRS, 2022). The control group includes ten popular ICEVs similar to the EV models in the treatment group but not eligible for the IRA tax credit. To demonstrate the effectiveness of the EV tax credit, the difference between the treatment and control groups must exhibit a statistically significant difference after the treatment in 2023.

The second model examines the variations in maximum EV tax credits across states and their effect on EV adoption. The independent variable is the maximum EV tax credit available to new EV buyers at the time of purchase, while the dependent variable is the natural log of each individual state's EV registration count. This simple Ordinary Least Squares (OLS) regression then controls for the maximum EV registration fee and unemployment rate to control for the financial disincentives for purchasing EVs and the potential impact of consumer confidence.

This study falls within the economic discipline of public finance and examines the influence of tax incentives on the EV and PHEV markets. DeShazo, Sheldon, and Carson (2017) used a case study in California and claimed that a \$2,500 tax incentive for EVs would result in a 7% increase in EV sales. Jenn, Springel, and Gopal (2018) did a similar analysis of the influence of various monetary incentives on EV adoption nationwide, showing that each \$1,000 in tax rebate may raise EV sales by 2.6%. Other studies have focused on different rebate schemes. Gallagher and Muehlegger (2011) investigated the efficacy of income tax credits, deductions, and sales tax exemptions. They find that sales tax exemptions resulted in a more than sevenfold increase in hybrid vehicle sales compared to income tax credits. Cole et al. (2023) studied IRA-specific provisions and forecasted the market share of EV sales by 2030, controlling for indirect expenses such as maintenance, home charging station installation, and gasoline costs. They projected that the impact of charging station subsidies and tax rebates might result in an 18 percentage point increase in EV sales compared to a scenario in which the IRA was not implemented. Overall, this body of research highlights the importance of subsidies in accelerating EV adoption and the role of tax incentives in changing customer behaviors.

The DiD design of this study, focusing on the variations between EV and ICEV models, contributes to the existing literature by analyzing the impact of the IRA's EV tax incentives on EV sales at the model level. While much of the prior research focused on household, state-level, or city-level data concerning EV purchases, there has been limited attention on the comparative sales of EV models against their non-EV counterparts. Model-specific variation is increasingly more relevant, given that the Inflation Reduction Act (IRA) imposes more stringent requirements on model-specific manufacturing, such as the location of final assembly, the types of components utilized, and whether the production involves components or raw materials from foreign entities of concern, such as China and Russia (Buckberg, 2023).

This paper begins in Section 2 with an overview of the current EV landscape in the United States, including various government programs in place to encourage EV purchases, as well as the challenges associated with assessing their efficacy. Section 3 outlines the data utilized in this research, including the sales of different EV and ICEV models, state-level tax credits, and other control variables. Section 4 discusses the DiD design and the simple OLS regression used for analyzing state-level variations. Section 5 presents the results of the analysis, including a DiD coefficient of 0.9320, 0.8469, and 0.8649, along with an EV Tax Credit coefficient of 0.0002036, 0.0001953, and 0.000157, with the latter controlled for other variables. The DiD coefficients are all statistically significant, but the EV tax credit coefficients are not. Despite this, their size is quite large in comparison to the mainstream results in the existing literature, and when combined with the DiD analysis results, it appears that tax credits are extremely effective at encouraging EV adoption among American consumers. Section 6 examines the study's robustness, emphasizing both the limitations of the dataset available and alternative interpretations. Section 7 concludes the analysis by examining the policy implications of the findings.

2 U.S. EV Policies: Past, Present, Progresses and Challenges

Prior to 2023, when the IRA's Clean Vehicle Tax Credit took effect, EV manufacturers were subject to a quota system. Once an EV manufacturer sells 200,000 eligible vehicles in the U.S., the tax credit begins to phase out, eventually falling to zero (Buckberg, 2023). Tesla hit the 200,000 cap in 2018 and lost it completely in 2019 (Shepardson, 2019). GM met its credit limit in 2018 and lost credit eligibility in 2020 (Shepardson, 2019). With the complete phase-out of these tax credits, it became evident to automakers and policymakers that this was a policy that punished success. Top EV manufacturers, such as Tesla and General Motors, which created some of the earliest models that consumers purchased in large numbers, were the first to lose their credits. As competition in the EV market intensified, particularly from China and, to a lesser extent, Europe, in addition to lobbying by automakers who justified their need for continued support on the grounds of "zero-carbon transportation" and "keeping green technology and manufacturing at home," policymakers in Washington soon concluded that more action was necessary (Buckberg, 2023).

As President Biden signed the IRA into law in 2022, with its Clean Vehicle Tax Credit provision taking effect in 2023, some of the top-selling EVs that had previously exhausted their tax credits for hitting the 200,000 limits—such as Tesla's Model 3 and Model Y, as well as GM's Cadillac Lyriq, Chevrolet Bolt, and Bolt EUV—became eligible for tax credits once again. However, the IRA introduced several new conditions this time around. The pre-IRA credit offered a $7,500 nonrefundable tax credit to all new EV purchases, provided the automaker had not exceeded its vehicle sales limit. This initial version imposed no additional restrictions on the vehicle's price, the buyer's income, the country of assembly, or the origin of other components. The IRA's credit, however, imposes a number of new restrictions, including a mandate that the car be assembled in North America, along with caps on price and buyer income, and specific requirements for the origin of battery components and critical minerals. Consequently, instead of a single $7,500 tax credit, the IRA now offers two separate $3,750 credits: one based on the origin of battery components and the other on critical mineral materials. The restrictive price and income limits also aim to prevent high-income individuals and those who can already afford EVs from benefiting, as these tax credits are unlikely to influence their purchasing decisions (Buckberg, 2023).

In an effort to reduce the reliance on "foreign entities of concern," like China and

Russia, for critical components and materials, Sen. Joe Manchin (D-WV) pushed a last-minute arrangement in 2022 to include conditions that would promote a more "secure" supply chain within North America or among ally nations. According to the International Energy Agency, China currently controls a large portion of the processing of rare earth elements and battery materials. It produces approximately 90% of the rare earth elements used in magnets, 50% to 70% of cobalt, and 35% of nickel. IRA's rigorous standards substantially reduce the number of vehicles that qualify. As a result, models such as the Mercedes EQS and Lucid Air are no longer eligible, and the Nissan Leaf and Rivian temporarily lost their eligibility before they made the necessary adjustments. In essence, the IRA created clear winners and losers as it went into effect, all to increase EV sales while preserving "green manufacturing" in the United States (Buckberg, 2023).

Though the federal government offers benefits for EVs, such as a maximum tax credit of $7,500 and subsidies for infrastructure improvements like charging stations, many states have also implemented their own tax credit programs to encourage EV adoption (Cole et al., 2023). Based on data from 2022 and 2023, California, Delaware, and Maine offer up to $7,500 in tax credits for eligible EV buyers, while Colorado and Oregon provide up to $5,000 (Jaros and Hoffer, 2023). Conversely, some states, including Arizona and Texas, do not offer any tax rebates (Jaros and Hoffer, 2023). Certain states even impose extra fees, with Alabama and Wyoming charging a $200 annual EV registration fee, and Mississippi and Washington charging $150 (Jaros and Hoffer, 2023). States also offer various other incentives, including income tax credits and deductions, waivers of state sales tax, single-passenger access to high-occupancy vehicle (HOV) lanes, and exemptions from registration, emissions testing, and parking fees (Gallagher and Muehlegger, 2011). These incentives can be substantial, often amounting to several thousand dollars, and help offset the cost of EVs.

All of these benefits offset a large percentage of EV costs. Take the top-selling Tesla Model Y, which averaged around $59,000 and can reach $64,000 in 2022 when fully loaded (Carfigures.com, 2024). The federally given $7,500 IRA tax credit covers 12% of vehicle cost. California, Delaware, and Maine residents may even receive an additional $7,500, bringing the total tax credit to $15,000, 25% of the vehicle's cost. Federal and state governments have already subsidized 25% of the vehicle's cost in this case, which is substantial for any buyers eligible for these subsidies. HOV lanes, registration and parking fee exemptions, and other benefits may also be available. A case study on hybrid vehicles' tax credits also reveals that customers capture at least 80 cents per dollar of federal subsidy (Salle, 2007). In most cases, this combination of federal and state subsidies is effective at lowering the barrier of entry for prospective EV buyers, making these “cleaner” vehicles more appealing to a broader demographic.

According to the White House, the Biden administration’s investments have resulted in a threefold increase in EV sales and a more than 40 percent increase in publicly accessible charging stations since he assumed office (White House, 2023). There are now over three million electric vehicles on the road and more than 135,000 public EV chargers nationwide (White House, 2023). Nevertheless, it is insufficient to attribute this phenomenon only to the effects of the tax credit reforms associated with the IRA. A naive comparison of tax credits and vehicle sales can conflate the increase in both additional and non-additional EV purchases. In this context, the term ”additional” pertains to purchases made by consumers who opt for EVs due to the tax incentive, as opposed to purchasing ICEVs in the absence of such incentives (Sheldon and Dua, 2019). Purchases made by consumers who would have purchased EVs regardless of the tax incentives are ”non-additional” (Sheldon and Dua, 2019). Borenstein and David (2016) describe this categorization in their analysis, demonstrating that tax incentives might not be the primary driver for new EV acquisitions, as evidenced by the significantly higher likelihood of high-income households purchasing EVs.

There is some evidence suggesting that EV tax credits are generally ineffective, with non-additional purchases constituting the majority of new EV acquisitions, as demonstrated by Hoekstra, Puller, and West (2017). They found that households eligible for federal tax incentives were no more likely to purchase a new vehicle within 6–9 months following the program’s inception than those not eligible. However, mainstream scholarship presents a different view. Studies by Sheldon and Dua (2019), DeShazo, Sheldon, and Carson (2017), Jenn, Springel, and Gopal (2018), and Gallagher and Muehlegger (2011) all indicate statistically significant and sizable increases in EV sales and adoptions following the introduction of different federal or state-level incentives. These studies employ varying research designs, from time and entity-fixed effects to DiD, based on the different data available. In Sections 3 and 4, I will discuss the open-source data available for my research and explain why I chose DiD as the preferred research design to assess changes in consumer behavior due to the EV tax credits.

3 Data on EV and ICEV Models and State-Level Tax Credit

The first dataset is composed of panel data featuring nine top-selling EVs eligible for IRA's Clean Vehicle Tax Credit and ten top-selling ICEV models, serving as counterparts to the EVs in the treatment group, along with their quarterly sales figures from Q1 2020 through Q4 2023. Table 1.1 in the appendix shows a complete list of the EV models selected for the treatment group and their non-electric counterparts selected for the control group. Quarterly EV sales data for EVs were sourced from Statista, which obtained the original figures from Cox Automotive, a company specializing in automotive industry data. Statista provides data on quarterly EV sales up until Q1 2023 (Cox Enterprises, 2023). The latest figures, from Q2 2023 to Q4 2023, are available in Cox Automotive's "Electric Vehicle Sales" quarterly reports (Cox Enterprises, 2023).

Among the 10 non-electric vehicles chosen for the control group, six are the best-selling SUV models in the U.S. for 2023, based on Statista's ranking (GoodCarBadCar.net, 2023). This selection aligns with the prevalence of SUVs in the treatment group, except for the Tesla Model 3, Tesla Model X, Ford F-150 Lightning, and Rivian R1. The Tesla Model 3, being a sedan, is matched with the Toyota Camry, the top-selling sedan (Reynolds, 2023). The Tesla Model X, a large-size SUV, is paired with the best-selling large-size SUV, the Chevrolet Tahoe (Good Car Bad Car, 2022). For pick-up trucks, the Ford F-150 Lightning and Rivian R1 T are paired with their best-selling non-EV counterparts, the Ford F-Series and Chevrolet Silverado, respectively (Hearst, 2024). The sales data for non-EVmodels of the control group are accessible on Carfigures.com, which compiles vehicle sales data into a dataset and has proven to be quite reliable after verifying the numbers against the manufacturers' quarterly reports (Carfigures.com, 2024).

Sales	Observations	Mean	Std. dev.	Max	Min
Common Sample	269	54606.23	51537.21	352677	3
Control Group (ICEVs)	154	81580.88	50217.75	352677	14213
Treatment Group (EVs)	115	18483.64	23671.69	105158	3

Table 1.2: Summary Statistics on Model Sales for DiD Analysis

Table 1.2 shows that the common sample includes 269 observations with a mean sales figure of 54,606.23 and a large standard deviation of 51,537.21. The control group has a higher mean sales figure of 81,580.88 compared to the treatment group's mean of 18,483.64, suggesting that ICEVs have higher sales numbers than EVs on average. The large standard deviations for both groups indicate a wide distribution of sales figures around the mean. Table 1.2a in the appendix presents the summary statistics for the same dataset on sales, displayed in logarithmic form for further reference.

Each point represents the sales of a certain vehicle model for a particular quarter in Figure 1, the corresponding scatterplot of Table 1.2. The blue dots represent ICEV sales, which are generally higher and show a broader distribution across the quarters. There are a few outliers indicating quarters with exceptionally high sales, which explains the high standard deviation seen in Table 1.2. The green dots represent EV sales, which are less scattered and generally lower than ICEV sales, indicating a more consistent but lower volume of sales. Refer to Figure 1a in the appendix to view the scatter plot of sales presented in logarithmic form.

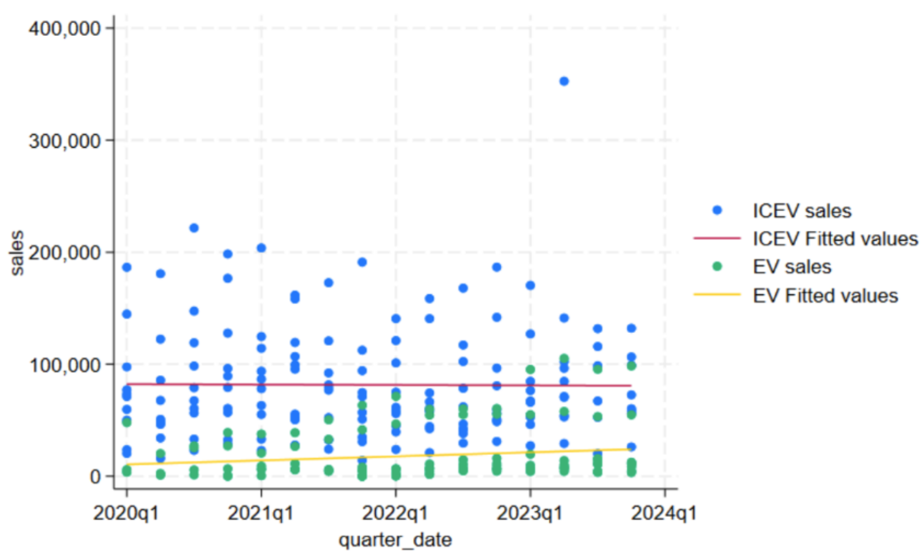


Figure 1: EV (Treatment) and ICEV (Control) Model Sales from 2020Q1 to 2023Q4

An exceptional outlier to note is the blue dot towards the top right, which represents the sale of the Ford F-Series, the most popular pick-up truck in the United States, with 352,677 units sold during the second quarter of 2023. I chose not to omit this outlier from my analysis because one of the top-selling EVs is the Ford F-150 Lightning, an EV substitute for the Ford F-Series. Ford F-150 Lightning is essentially a mirror image of the F-Series, with the primary difference being that it is an EV and qualifies for tax credits. Excluding it from the analysis would overlook an important substitution effect between the two models.

Another important consideration when analyzing the data is the presence of seasonal fluctuations in vehicle sales. Specifically, sales tend to be higher in Quarter 3, Quarter 4, and the first quarter of the subsequent year, compared to Quarter 2. This trend is observable in the case of EVs, exemplified by the Tesla Model Y, the top-selling EV in the United States. In the second quarter of 2021, sales were recorded at 38,877 units, followed by steady increases in the third (50,430), fourth (63,386), and first quarter of 2022 (71,358). However, sales then declined to 59,822 in the second quarter of 2022. A similar pattern is evident for ICEVs, using the Toyota RAV4, the top-selling ICEV, as an example.In the second quarter of 2022, sales were 66,493, then increased to 102,456 in the third quarter, remained relatively high at 96,600 in the fourth quarter, and continued at 84,704 in the first quarter of 2023. Factors such as seasonal promotions by car dealerships during key holiday periods (e.g., Thanksgiving, Christmas) or at the start of the year (e.g., New Year sales) may contribute to these fluctuations. Additionally, increases in disposable income—such as year-end bonuses—may boost consumer confidence as well, thereby prompting decisions to replace older vehicles. These seasonal variations have important implications for the research method employed in this analysis—DiD, which will be discussed in greater detail in Section 4.

The second dataset consists of cross-sectional data encapsulating various metrics such as the maximum EV tax credit, EV registration fee, EV registration count, and the unemployment rate across all 50 U.S. states. The information regarding the maximum EV tax credit and registration fees for individual states was obtained from a think tank known as the Tax Foundation, with its website's latest update in September 2023, drawing its figures from the U.S. News & World Report and State Statutes (Jaros and Hoffer, 2023).

State-level EV registration counts were sourced from the Department of Energy's Alternative Fuels Data Center, with the available data extending up to 2022 (U.S. Department of Energy, 2023). Therefore, the dataset and the associated regression analysis will utilize the 2022 EV registration data alongside the state-level tax credit information last updated in 2023. This approach, though not ideal, is appropriate as there have not been significant changes in the tax credit schemes from 2022 to 2023 for most states. The state-level unemployment rates were provided by Statista, which bases its figures on the Bureau of Labor Statistics' unemployment data for 2022 (Bureau of Labor Statistics, 2023).

Variable	Observations	Mean	Std. dev.	Max	Min
EV Tax Credit	50	1448	2267.44	7500	0
EV Registration Fee	50	79.10	71.16	213.7	0
EV Registration Count 2022	50	48728.2	128737.4	903620	640
Unemployment (%)	50	3.4	0.74	5.4	2.1

Table 2: Summary Statistics of State-Level Tax Credit and EV Registration

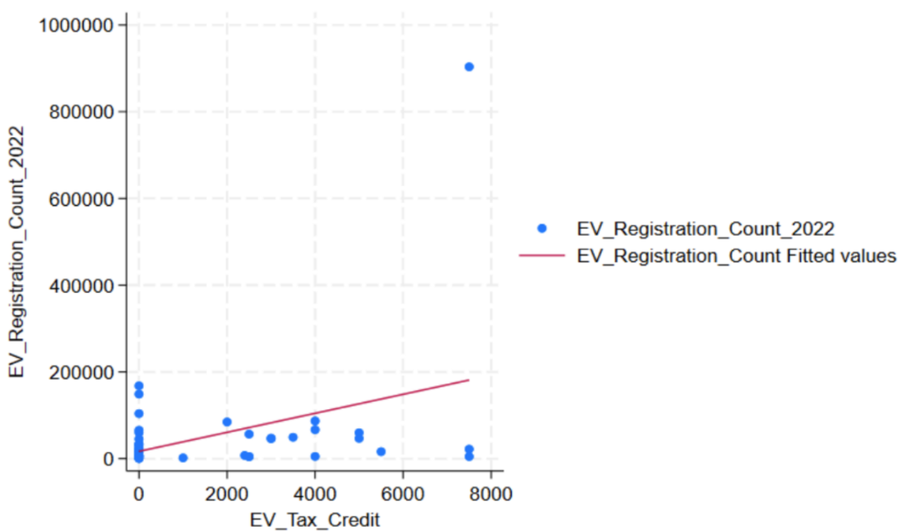


Figure 2: Tax Credit and EV Registration of 50 States

Table 2 presents summary statistics for four different variables across 50 states. The average state EV Tax Credit is $1,448, with a substantial standard deviation of $2,267.44, indicating significant variation among states. The average EV Registration Fee is around $79.10, and there is considerable variation (standard deviation of $71.16). For the EV Registration Count in 2022, the average across the states is 4,872.8, with a very high standard deviation of 12,873.74. The Unemployment rate (%) averages at around 3.4%, with a narrower standard deviation of 0.74%, suggesting less variability across states.

Figure 2 illustrates the relationship between the maximum EV Tax Credit and EV Registration Counts in 2022 across the 50 states. With each point representing a state, the plot indicates a positive relationship between tax credits and EV registration, as seen by the upward trend of the fitted line. However, there are states with high

EV registration counts that deviate significantly from the fitted line, suggesting the presence of outlier states like California, which has 903620 EVs registered as of 2022. I chose not to omit California because it is one of the states most actively involved in promoting EV adoption, as reflected in its leading number of EV sales and adoptions. Excluding California would mean omitting a significant portion of the vehicles already used in my analysis. More specifically, a substantial share of the vehicles analyzed in the first DiD design likely occurred in California.

Overall, the data highlights that there is a positive correlation between tax credits and EV registration, though the high standard deviations in tax credits and registration fees suggest significant disparities in state policies. The overall characteristics remain consistent even when the EV registration count is expressed in logarithmic form, as shown in Figure 2a in the appendix.

4 Empirical Models: DiD on Model Variation and State-Level Variation

The first model I intend to apply is a conventional DiD equation, as illustrated by the equation below. Table 3 describes each of the variables used in the following regression.

$$\ln(Sales_i) = \beta_0 + \beta_1(After_i) + \beta_2(Treatment_i) + \beta_3(After_i \cdot Treatment_i) + \epsilon_i \quad (1)$$

Quarter_Date	Identifies a particular quarter (Q1, Q2, Q3, or Q4) of a given year. "2023q1" would denote the first quarter (January to March) of the year 2023.
Sales	The total quantity of vehicles sold within a specific quarter.
Treatment	A binary variable used to categorize vehicle models based on their type. If a vehicle is an EV, the treatment variable is assigned a value of 1.
After	A binary variable marking the time periods after the assigned treatment has been assigned. The first cutoff for treatment was when the tax rebate provision of the IRA came into effect in 2023q1, which expanded the eligibility for EV tax credits. Multiple cutoff periods will be tested in this study, specifically 2023q2 and 2023q3. If a period falls after the cutoff, the "after" variable is assigned a value of 1.

Table 3: Variable Table for the DiD Regression

In this DiD regression, β_3, the coefficient for the interaction term $(After_i{\cdot}Treatment_i)$, represents the change in the percentage of average sales for the treatment group in the post-treatment period relative to the counterfactual scenario where the new EV tax credit was not implemented. If β_3 is not statistically significantly different from 0, it suggests that the treatment (in this case, EV subsidies under the IRA) did not have an effect on EV sales.

The first reason I chose a DiD design is the seasonal fluctuations in vehicle sales, as detailed in Section 3. The data indicates that sales are typically higher in Quarter 3, Quarter 4, and the first quarter of the following year, compared to Quarter 2. This pattern is observed for both EVs and ICEVs. Initially, I considered using a simple Regression Discontinuity (RD) design to analyze EV sales before and after the implementation of the IRA's clean vehicle tax credit in Q1 2023. However, seasonal fluctuations—driven by factors such as holiday promotions (e.g., Thanksgiving, Christmas), New Year sales, and increases in disposable income (e.g., year-end bonuses)—pose a challenge. These factors create potential simultaneous shocks that compromise the robustness of using RD to analyze this policy treatment. Specifically, it becomes difficult to isolate the effect of the new EV tax credit provisions of the IRA, effective January 2023, as the sole "shock" influencing EV sales. Any concurrent external shocks impacting EV sales would undermine the validity of the RD design. By adopting a DiD approach, comparing EV sales with a control group of ICEVs, I can better account for these seasonal effects and mitigate the issue of simultaneous shocks.

This model also draws inspiration from existing scholarship, specifically the DiD design used by Sheldon and Dua (2019) to evaluate the impact of California's "Replace Your Ride" program on households' vehicle purchasing decisions in 2015. This program offered subsidies for EV and hybrid buyers under a certain income threshold. The treatment group consisted of car buyers below the income threshold who were eligible for the EV and hybrid tax incentives. The control group consisted of individuals marginally above this income bracket, hence disqualified from the tax incentives. In their DiD analysis, the first "difference" compared the sales of EVs and hybrids among the income groups ineligible and eligible for the tax rebates. The second "difference" compared these sales before and after the beginning of the subsidy program in May 2015. Their findings indicated a statistically significant increase in EV and hybrid sales for new vehicle buyers in the LA metropolitan area, with sales figures climbing by approximately 50% and 70%, respectively, following the program's implementation.

Sheldon and Dua (2019) relied on a limited dataset of vehicle sales figures and household income data to identify rebate eligibility, lacking more detailed information. My circumstances are similar, with my dataset limited to sales numbers for vehicle models, vehicles' tax rebate eligibility, and the start date of the IRA's Tax Credit policy in January 2023. Following the research design of Sheldon and Dua (2019), I also employ a DiD model. Nonetheless, the IRA's tax credit criteria depend not on income like California's "Replace Your Ride," but on specific characteristics of the vehicle model and its manufacturing process. These include the final assembly's location, the components used, and whether the assembly involved components or raw materials from foreign entities of concern, notably China and Russia. Therefore, instead of assessing sales variation based on income brackets, my analysis will pivot on variations among individual EV models and their respective subsidy qualifications.

Furthermore, compared to Sheldon and Dua (2019), who marked the introduction of California's "Replace Your Ride" program in June 2015 as a singular "shock" treatment and gave almost no time for the public to respond to the new policy, my approach will implement multiple shock treatments subsequent to the introduction of the IRA's tax credit provision in January 2023 and their corresponding regression. This method allows for a more nuanced analysis that acknowledges the time consumers may require to respond to policy changes. While Salle (2007) indicated that consumers can respond strategically to new policies, as evidenced by his model-specific DiD analysis of Toyota's Prius and its non-hybrid sedan counterpart after tax rebate came into effect following the Energy Policy Act of 2005—which revealed that buyers adjusted their purchasing timing to maximize subsidy benefits—such immediate responses may not be directly applicable to the context of the IRA. The immediate behavioral adjustments seen in Salle's (2007) research highlight the buyers' strategic response to tax credits, yet the complex model-specific eligibility of the IRA necessitates a more cautious expectation regarding consumer behavior.

Taking into consideration that some buyers may delay their vehicle purchases due to confusion or uncertainty—especially since many EVs lost their original tax credits either partially or entirely for not meeting specific manufacturing requirements—I will introduce multiple shock treatments across quarters such as 2023q1, 2023q2, and 2023q3 and run multiple regressions (Domonoske, 2024). This approach accommodates the potential hesitance of consumers who are more cautious about purchasing immediately after the policy's enactment, ensuring that my model captures the varied consumer reactions and the possible deferral of vehicle purchase decisions in the wake of the IRA's new regulations.

The second model that I plan to implement is a simple regression model with controls, drawing from the approach used by Gallagher and Muehlegger (2008) in their study of how state-level incentives impact hybrid vehicle sales.

$$\begin{aligned}\ln(EV\_Registration\_Count\_2022_i) = \beta_0 &+ \beta_1(EV\_Tax\_Credit_i) \\ &+ \beta_2(EV\_Registration\_Fee) + \beta_3(Unemployment_i) \\ &+ \epsilon_i \end{aligned} \tag{2}$$

In this simple OLS regression model utilizing cross-sectional data at the state level, i denotes individual states. The model specifies the natural log of EV registration count in 2022 for each state i as the dependent variable, with the state's EV tax credit as the independent variable. Additionally, the model includes the EV registration fee and the unemployment rate as control variables to adjust for their potential confounding effects.

Gallagher and Muehlegger (2008) used state-level, quarterly hybrid sales data spanning from 2000 to 2006 and analyzed a comprehensive array of tax incentives, such as state sales tax waivers, income tax credits, and other non-tax incentives. By employing entity/state fixed effects and time fixed effects methods, their study accommodated both flexible national adoption patterns and time-invariant state preferences. Their findings indicated that sales tax waivers were particularly effective, associated with more than a tenfold increase in HEV sales, in stark contrast to the relatively modest impact of income tax credits.

Following the research design of Gallagher and Muehlegger (2008), the ideal approach would involve regressing EV sales figures against the total EV subsidies offered by individual states, before applying state/entity and time fixed effects to adjust for nationwide adoption trends and time-variant state-specific preferences, and subsequently estimating the coefficient for EV tax credit. Nonetheless, due to limited access to data, the most comprehensive variation in state-level EV tax credits that I was able to gather consists solely of a cross-sectional dataset including variables such as the maximum EV tax credit and EV registration fee, updated last in 2023, alongside EV registration counts from 2022. Given the cross-sectional nature of the data, the inclusion of state and time fixed effects in my analysis is not feasible.

Nevertheless, such limitation does not prevent this data from providing valuable insights regarding the interplay between EV tax credits and state-level EV registration counts. By conducting a straightforward regression analysis and incorporating control variables, this study can still explain the dynamics of this relationship. Similar to Gallagher and Muehlegger (2008), who factored in additional tax incentives like sales tax waivers and income tax credits, this research incorporates the EV registration fee, a possible disincentive to EV purchases, and unemployment rates, reflective of consumer confidence/sentiment as demonstrated in Malovana et al. (2021). Such controls enable examining these variables' collective impact on the causal relationship between EV tax credits and EV registration numbers. This work also extends the analysis of Gallagher and Muehlegger (2008) by including proxies for consumer confidence and EV registration fees, which their state-fixed effects may partially but not wholly capture.

5 Results and Analysis: An Effective Policy?

β_3	
2023Q1 • Robust Std. Err. • T-Stat • P-Value	0.9320112 • (0.1416205) • 6.58 • 0.000
2023Q2 • Robust Std. Err. • T-Stat • P-Value	0.8469667 • (0.1526432) • 5.55 • 0.000
2023Q3 • Robust Std. Err. • T-Stat • P-Value	0.8649349 • (0.2165618) • 3.99 • 0.001

Table 4: Regression Table for DiD Coefficient at Different Cutoff Points

β_3, the coefficient associated with the interaction term ($After_i \cdot Treatment_i$), represents the percentage difference in average sales for the treatment group during the post-treatment period compared to a counterfactual where the new EV tax credit policy was not enacted. This coefficient is valued at 0.9320, 0.8469, and 0.8649 for the treatment periods of 2023Q1, 2023Q2, and 2023Q3, respectively. This suggests that, on average, the introduction of the IRA's Clean Vehicle Tax Credit is associated with an increase in EV sales by 93.20% in 2023Q1, 84.69% in 2023Q2, and 86.49% in 2023Q3, contingent on consumer reactions to the policy in each quarter. All P-values are below 0.05, meaning that all of these coefficients are statistically significant at the 5% significance level.

This result is consistent with the findings of Sheldon and Dua (2019), who showed a statistically significant increase in EV sales of about 50%. In fact, my coefficients, which average between 80 and 90 percent, are positive, similar to Sheldon and Dua's estimates, but 30 to 40 percentage points higher. In my case, the 2023 IRA tax credit is almost too effective. One straightforward explanation for the discrepancy is that the 2023 federal tax credit examined in this study is more effective than the tax credit provided by California's "Replace Your Ride" program in 2015.

The hypothesis that the IRA is more effective is not unreasonable, considering that the California state government offered only $4,500 in tax credits for each EV under the "Replace Your Ride" program in 2015, while the federal government offered between $3,750 and $7,500 in tax credits for eligible EV models in 2023. The infrastructure supporting EVs, such as the density of charging station networks, has also become more prevalent in the United States in 2023 compared to 2015, making the purchase and maintenance of EVs more accessible. The implicit assumption is that when indirect, non-cash incentives become more readily accessible, making owning and operating an EV easier, consumers will also be more responsive to EV tax credits, creating a positive feedback loop. This is possible when analyzing Cole et al. (2023), who found that government spending on charging stations is significantly more effective than tax rebates, and Salle's (2007) finding that consumers have the tendency to adjust the time of their EV purchases to maximize benefits they could get.

An alternative explanation is that the increase in EV sales has been overestimated due to excessive noise in the dataset, suggesting that the causal effect of the tax credit is larger than it should be due to high data variability, which has "obscured" the smaller causal effect. This is possible given that Sheldon and Dua (2019) examined a very specific segment of the population in their DiD analysis of California's "Replace Your Ride" tax credit. They examined a narrow subset of new vehicle sales in the Los Angeles Metro Area, specifically transactions involving trade-ins, where eligibility for the tax rebate required 1) replacing an older vehicle with an EV or hybrid, 2) residing in the South Coast Air Quality Management District, and 3) earning less than an income threshold. In contrast, the dataset I compiled and used includes new EV and ICEV sales across the United States, but there is no particular "tax credit" value for each EV transaction, instead relying on a presumed "shock" event as a cutoff point for the DiD study. This limitation in a DiD research design can only allow me to focus on the broader effect of a "shock" and may capture more than just the causal effects of tax rebates on EV purchases, such as a cyclical economic expansion that may have boosted consumer confidence or other reforms of EV tax benefits on the state or local levels, indicating the need for a more precise dataset to determine the actual effect of subsidies on EV sales.

Regressors	(1)	(2)	(3)
EV Tax Credit (β_1) • Robust Std. Err. • T-Stat • P-Value	0.0002036 • (0.0001089) • 1.87 • 0.068	0.0001953 • (0.000123) • 1.59 • 0.119	0.000157 • (0.0001105) • 1.42 • 0.162
EV Registration Fee (β_2) • Robust Std. Err. • T-Stat • P-Value		-0.0007694 • (0.0035024) • -0.22 • 0.827	-0.0002661 • (0.0035269) • -0.08 • 0.940
Unemployment (β_3) • Robust Std. Err. • T-Stat • P-Value			0.6425174 • (0.2520432) • 2.55 • 0.014
Intercept (β_0) • Robust Std. Err. • T-Stat • P-Value	9.382096 • (0.2463144) • 38.09 • 0.000	0.6425174 • (0.2520432) • 21.87 • 0.000	0.6425174 • (0.2520432) • 7.47 • 0.000
R^2	0.0964	0.0976	0.1969

Table 5: Regression Table for State-Level Analysis

β_1	(1)	(2)	(3)
F-Stat	3.49	2.52	2.02
Prod > F	0.0677	0.1192	0.1622

Table 6: F-Test on β_1

Column (1) of Table 5 presents the outcome of regressing the natural log of EV registration counts against the maximum EV subsidies provided by individual states. With a β_1 value of 0.0002036, the interpretation is that an increase of 1 dollar in the maximum EV tax credit is associated with an average rise in EV registration counts by approximately 0.02036%. Column (2) repeats this regression while controlling for the maximum EV registration fee, a method states use to generate tax revenue that also acts as a disincentive for potential EV buyers. With an adjusted β_1 of 0.0001953, the state-level data suggests that for each 1-dollar rise in the maximum EV tax credit, there is an associated increase of about 0.01953% in EV registration counts, controlling for the maximum EV registration fee. Column (3) introduces another control, the unemployment rate, which is used as a proxy to gauge consumer confidence/sentiment and can significantly influence EV registrations. With an adjusted β_1 of 0.000157, this analysis suggests that, on average, a 1-dollar increase in a state's maximum EV tax credit is associated with a 0.0157% increase in EV registration counts, holding unemployment rate and the maximum EV registration fee constant. Table 6 presents the outcomes of conducting an F-test on each β_1 coefficient, none of which exhibit statistical significance at the 5% significance level.

Even though none of these coefficients show statistical significance, their comparison with other figures in existing literature is not entirely without merit. The elasticity estimate of DeShazo, Sheldon, and Carson (2017) suggests that a tax credit of $2,500 is associated with a 7% increase in EV adoption. When applying this scale to the coefficient of 0.015%, which controls for EV registration fees and unemployment rates, it results in an estimated 39% increase—nearly six times their estimate. Similarly, Jenn, Springel, and Gopal (2018) found that $1,000 of tax credit correlates with a 2.6% increase in EV sales; scaling this to my coefficient of 0.0157% yields approximately 15.7%, almost six times their figure. While this elasticity is significantly higher than those generally found in the literature, it aligns with the general understanding that tax credits are effective. However, similar to the results from the previous DiD estimate, it appears almost too effective in my case.

Although I cannot completely dismiss the possibility that American consumers may have become generally more responsive to the EV tax credit in 2022 compared to earlier periods—as previously suggested to explain the large DiD coefficient—it is more likely that my regression severely overestimates the actual elasticity of EV demand per unit of government tax credit. This overestimate is primarily a result of the limited access to data, specifically having access to only each state's maximum tax credit for EVs as of 2023. Ideally, access to detailed transaction-level data that specifies the amount of tax credit given for each new EV registered in a state would allow for a more accurate assessment of the tax credit's causal effect. Moreover, not all vehicles registered in a state are eligible for the full maximum tax credit, which my dataset failed to capture. Therefore, a more robust analysis would require panel data covering multiple EV sales and corresponding credits received across different states over extended time periods. This would enable me to employ entity and time-fixed effects as used by Gallagher and Muehlegger (2008) and ideally control for state-level consumer sentiment and gas prices, similar to the strategy employed by Jenn, Springel, and Gopal (2018). However, this study is limited by the availability of only cross-sectional data.

6 Robustness Check: Parallel Counterfactual Trend and Consumer Confidence

The assumption of a parallel counterfactual trend is fundamental to any DiD research design. In order to establish the validity of the prior estimates of β_3, it is necessary for the treatment group, representing EV sales figures, to demonstrate a trend that is parallel to that of the control group, comprising non-EV sales figures, in the absence of the new Clean Vehicle Tax Credit introduced by the IRA in 2023. The question of whether the counterfactual trend would have been parallel can never be known. Nevertheless, examining parallel pre-trends can be used as a tool to assess this assumption. Figure 3 presents the observed means of the control and treatment groups, specifically highlighting their trends before and after the beginning of treatment in the first quarter of 2023. From a visual perspective, the graph fails to provide evidence for parallel pre-trends, as the observed mean of the EV group exhibits notable fluctuations. A similar pattern can be observed when examining other graphical diagnostics for parallel trends, specifically those with the shock treatment assigned to the second and third quarters of 2023. For more information, please consult Figures 3b and 3c in the appendix.

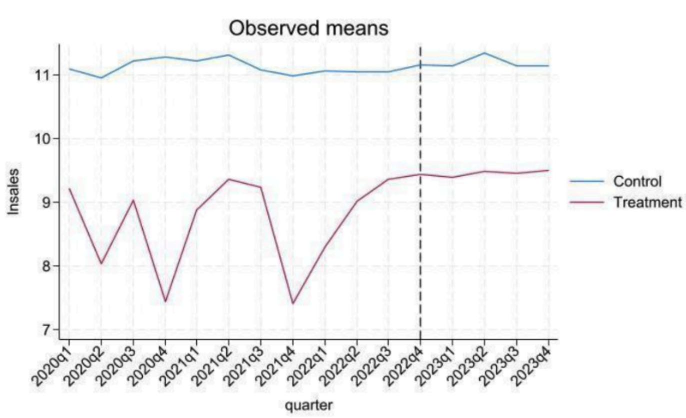


Figure 3: Graphical Diagnostic for Parallel Trends - 2023Q1 as Treatment

H0: Linear trends are parallel	2023Q1	2023Q2	2023Q3
F-Stat	6.76	10.69	13.30
Prob > F	0.0181	0.0043	0.0018

Table 7: Parallel Trends Test (Pretreatment Time Period)

The results of the parallel trends test presented in Table 7 confirm the visual intuition. The null hypothesis posits that linear trends are parallel. Yet, low P-values of 0.0181, 0.0043, and 0.0018 for the treatment periods of 2023Q1, Q2, and Q3, respectively, clearly indicate that the null hypothesis has been rejected in favor of the alternative hypothesis that linear trends are not parallel, with 95% confidence. However, at a significance level of 99%, the null hypothesis for 2023Q1 is not rejected, suggesting that the parallel trends condition holds "weakly" for this quarter. Despite this, in the majority of other scenarios, the parallel trends test has shown that the parallel counterfactual trends assumption does not hold, and there is little that can be done to remedy this issue with the limited data available for this study.

Furthermore, examining parallel pre-trends alone is not sufficient to confirm the robustness of this DiD analysis. Another critical element inherent in any DiD approach is ensuring that the treatment being analyzed—in this context, the new EV tax credit provisions of the IRA, effective from January 2023—is the sole "shock" affecting the sales figures of EVs and ICEVs. The presence of any concurrent external shock that differentially impacts the sales trajectories of EVs and ICEVs would compromise the robustness of the estimated β_3 discussed previously.

Could there be any such external shock capable of affecting either the trajectories of EV or ICEV sales? One potential factor is the variation in consumer sentiment from late 2022 to the end of 2023. According to the University of Michigan's quarterly consumer sentiment index, consumer sentiment climbed from 58.8 to 64.6, then dropped to 62.3, climbed to 69.6, and finally dropped again to 64.9 between the 2022Q4 and 2023Q4 (University of Michigan, 2023). In the first DiD regression presented in Table 4, where the "after" treatment period begins in 2023Q1, a simultaneous rise in consumer sentiment from 58.6 to 64.6 could imply that the demand for both ICEVs and EVs might have increased, potentially at differing rates. This variation in consumer sentiment could challenge the assumption the IRA is the only shock of this period by introducing a simultaneous event that also serves as an external shock, thereby compromising the robustness of the estimated β_3 from the previous analysis. Once again, there are no simple remedies for this problem aside from running another DiD regression that aims to reduce the influence of any simultaneous shock events as much as possible, likely involving eliminating the occurrence of any concurrent external shocks by analyzing a separate period or event that is more isolated.

In the second OLS regression model, which examines state-level EV registration counts, EV tax credits, and the control variables of EV registration fee and unemployment, a major concern for robustness is the omission of other important variables. For example, Gallagher and Muehlegger (2008) factored in the availability of charging stations and discovered its statistically significant causal effect on the consumers' inclination to buy EVs. This consideration is crucial when analyzing states in the southern region, which typically have a lower density of charging stations compared to coastal states, which tend to have higher charging station density hence higher EV registration counts (United States Department of Energy, 2024). Furthermore, nationwide shifts in attitudes towards EVs—for reasons unrelated to EV subsidies or consumer sentiment, such as the introduction of new Tesla models or a revived optimism for sustainable transportation—can also influence EV adoption. Addressing these issues is unlikely with the limited cross-sectional data available to me. However, should I gain access to panel data that spans multiple time periods and covers a range of states, I could employ entity/state fixed effects to control for characteristics that do not change over time within states, and time fixed effects to account for trends affecting all states equally, thereby simultaneously addressing both of these issues.

7 Conclusion

In this study, two primary models were utilized to assess the impact of EV tax credits on electric vehicle adoption. The initial model is a Difference-in-Differences (DiD) approach that designates EVs as the treatment group and ICEVs as the control group. It leverages the Inflation Reduction Act's (IRA) new Clean Vehicle Tax Credit provision as a pivotal "shock" event and explores multiple treatment periods to accommodate potential delayed adjustments in consumer behavior. On average, the enactment of the IRA's Clean Vehicle Tax Credit was associated with an increase in EV sales by 93.20% with 2023Q1 as treatment, 84.69% with 2023Q2, and 86.49% with 2023Q3, each depending on consumer reactions to the policy in different time periods. With all P-values less than 0.05, these findings are statistically significant at the 5% significance threshold.

The second model is an OLS regression model that uses EV registration counts as dependent variables and the maximum EV tax credit of individual states as independent variables. The model controls for maximum EV registration fee, a disincentive for EV purchases, and unemployment, a proxy used to reflect consumer sentiment. With an adjusted coefficient of 0.000157, this analysis suggests that, on average, a 1-dollar increase in a state's maximum EV tax credit is linked to a 0.0157% increase in EV registration counts, holding the unemployment rate and the maximum EV registration fee constant. However, this coefficient failed to pass the F-Test and is therefore not statistically significant at the 5% significance level.

Despite mixed results, it remains clear that EV tax credits effectively boost EV sales and adoptions, consistent with most mainstream literature findings. However, the size of their causal effect appears excessively large in my analysis. One possible explanation for this discrepancy is that American consumers may be more responsive to the federal tax credit associated with the IRA compared to past tax incentives. This hypothesis is possible based on the findings from Cole et al. (2023) and Salle (2007), who suggest that investments in charging stations and other infrastructure can further stimulate EV demand. Additionally, consumers also have the tendency to time their EV purchases to maximize the benefits they receive. When combining the two, the reasonable implication is that consumer demand for EVs may vary over time and be influenced by the level of infrastructural development, leading to differing responses to tax credit policies implemented at different times.

An alternative explanation for the exceptionally large consumer response to the EV tax credit could be that my regression significantly overestimates the actual elasticity of EV demand per unit of government tax credit. This overestimation is primarily due to my limited access to data, specifically having only the maximum tax credit information for EVs in each state as of 2023. As demonstrated in cases like the IRA and various state-level subsidy schemes such as California's "Replace Your Ride" program, not all EVs qualify for the full maximum tax credit, a detail my data fails to capture. Resolving this issue is challenging, as I only have access to a cross-sectional dataset rather than panel data that covers multiple EV sales and their corresponding credits received. Having a panel dataset would allow me to employ more econometric tools to isolate tax credit's casual impact on EV registration counts.

Given that the DiD analysis in this study does reveal a statistically significant, positive impact of the EV tax credit on EV sales, a logical next step is to investigate whether the elasticity of demand for EV per unit of government tax credit really does vary over time. Furthermore, besides making some EVs previously ineligible for any tax credit eligible for the $7500 tax credit, IRA also introduced more stringent manufacturing and component requirements that made some models lose their eligibility. Following the withdrawal of subsidies, it is naturally relevant to question whether their sales figures have indeed been declining. In April 2023, the Biden administration announced its list of vehicles eligible for the tax credit, causing the market to feel uneasy again as several models previously qualified for the IRA tax credit temporarily lost their eligibility or were forced to adjust their production and assembly of EV and battery components to North America (Domonoske, 2024). Analyzing the effects of EVs losing their eligibility and how firms have adapted their investment strategies in the EV sector is vital for understanding the implications for the U.S. Green Transition.

References

[1] "2022 U.S Large SUV Sales Figures (With Rankings)." 2022. GCBC. https://www.goodcarbadcar.net/2022-us-large-suv-sales-figures/.

[2] Borenstein, Severin, and Lucas W. Davis. 2016. "The Distributional Effects of US Clean Energy Tax Credits." *Tax Policy and the Economy* 30 (1): 191–234. https://doi.org/10.1086/685597.

[3] Bureau of Labor Statistics. n.d. "Annual Unemployment Rate by State U.S. 2022." Statista. Accessed March 21, 2024. https://www.statista.com/statistics/223675/state-unemployment-rate-in-the-us/.

[4] Clinton, Bentley C., and Daniel C. Steinberg. 2019. "Providing the Spark: Impact of Financial Incentives on Battery Electric Vehicle Adoption." *Journal of Environmental Economics and Management* 98 (November): 102255. https://doi.org/10.1016/j.jeem.2019.102255.

[5] Cole, Cassandra, Michael Droste, Christopher Knittel, Shanjun Li, and James H. Stock. 2023. "Policies for Electrifying the Light-Duty Vehicle Fleet in the United States." *AEA Papers and Proceedings* 113 (May): 316–22. https://doi.org/10.1257/pandp.20231063.

[6] Cox Enterprises. 2023a. "Electric Vehicle Sales Report Q2 2023." https://www.coxautoinc.com/wp-content/uploads/2023/07/Q2-2023-Kelley-Blue-Book-Electric-Vehicle-Sales-Report.pdf.

[7] ———. 2023b. "Electric Vehicle Sales Report Q3 2023." https://www.coxautoinc.com/wp-content/uploads/2023/10/Q3-2023-Kelley-Blue-Book-Electric-Vehicle-Sales-Report.pdf.

[8] ———. 2024. "Electric Vehicle Sales Report Q4 2023." https://www.coxautoinc.com/wp-content/uploads/2024/01/Q4-2023-Kelley-Blue-Book-Electric-Vehicle-Sales-Report.pdf.

[9] ———. n.d. "U.S. BEV Quarterly Sales by Model." Statista. Accessed February 22, 2024. https://www.statista.com/statistics/1284477/united-states-bev-quarterly-sales-by-model/.

[10] DeShazo, J.R., Tamara L. Sheldon, and Richard T. Carson. 2017. "Designing Policy Incentives for Cleaner Technologies: Lessons from California's Plug-in

Electric Vehicle Rebate Program." *Journal of Environmental Economics and Management* 84 (July): 18–43. https://doi.org/10.1016/j.jeem.2017.01.002.

[11] Domonoske, Camila. 2024. "The $7,500 Tax Credit for Electric Cars Has Some Big Changes in 2024. What to Know." *NPR*, February 21, 2024. https://www.npr.org/2023/12/28/1219158071/ev-electric-vehicles-tax-credit-car-shopping-tesla-ford-vw-gm.

[12] Buckberg, Elaine. 2023. "Clean Vehicle Tax Credit: The New Industrial Policy and Its Impact." *Stanford Institute for Economic Policy Research*. August 2023. https://siepr.stanford.edu/publications/policy-brief/clean-vehicle-tax-credit-new-industrial-policy-and-its-impact.

[13] "Electric Vehicle Registrations by State." 2023. *Alternative Fuels Data Center: Maps and Data*. July 2023. https://afdc.energy.gov/data/10962.

[14] "Federal Tax Credits for Pre-Owned Plug-in Electric and Fuel Cell Vehicles." n.d. Accessed March 21, 2024. https://www.fueleconomy.gov/feg/taxused.shtml.

[15] Gallagher, Kelly Sims, and Erich Muehlegger. 2011. "Giving Green to Get Green? Incentives and Consumer Adoption of Hybrid Vehicle Technology." *Journal of Environmental Economics and Management* 61 (1): 1–15. https://doi.org/10.1016/j.jeem.2010.05.004.

[16] Hoekstra, Mark, Steven L. Puller, and Jeremy West. 2017. "Cash for Corollas: When Stimulus Reduces Spending." *American Economic Journal: Applied Economics* 9 (3): 1–35. https://doi.org/10.1257/app.20150172.

[17] IRS. n.d. "Manufacturers and Models for New Qualified Clean Vehicles Purchased in 2022 and Before — Internal Revenue Service." Accessed February 22, 2024. https://www.irs.gov/credits-deductions/manufacturers-and-models-for-new-qualified-clean-vehicles-purchased-in-2022-and-before.

[18] Muehlegger, Erich, and David S. Rapson. 2022. "Subsidizing Low- and Middle-Income Adoption of Electric Vehicles: Quasi-Experimental Evidence from California." *Journal of Public Economics* 216 (December): 104752. https://doi.org/10.1016/j.jpubeco.2022.104752.

[19] The White House. 2023. "FACT SHEET: Biden-Harris Administration Announces New Private and Public Sector Investments for Affordable Electric Vehicles." *The White House Briefing Room*. April 17, 2023. https://www.whitehouse.gov/briefing-room/statements-releases/2023/04/17/fact-sheet-biden-harris-administration-announces-new-private-and-public-sector-investments-for-affordable-electric-vehicles/.

[20] Wu, Xi, Jing Gong, Brad N Greenwood, and Yiping Song. n.d. "No Longer Riding Dirty: The Effect of Electric Vehicle Subsidies on the Diffusion of Emerging Technologies in Automobile Markets."

Appendix

Data:

The excel file continuing the quarterly EV and ICEV sales figures can be accessed via this link. The other file containing state-level tax credit information can be accessed here. The sources of some of the original data and numbers has been described in both this appendix and the data section of the actual paper.

Treatment Group	Control Group
Tesla Model Y	Toyota RAV4
Tesla Model 3	Toyota Camry
Chevy Bolt EV/EUV	Honda CR-V
Tesla Model X	Chevrolet Tahoe
Ford Mustang Mach-E	Nissan Rogue
Ford F-150 Lightning	Ford F-Series
Volkswagen ID.4	Chevrolet Equinox
Rivian R1T	Chevrolet Silverado
Hyundai Ioniq 5	Jeep Grand Cherokee
	Hyundai Tucson

Table 1: Table 1.1: EV and ICEV Model Selection for the Control and Treatment Group

The treatment group comprises 9 EV models that are eligible for the EV tax credit, selected from Statista's top-selling EV data based on Cox Automotive's figures and Cox Automotive's quarterly reports from Q2 2023 to Q4 2023 (Cox Enterprises, 2023). Their tax credit eligibility before and after the IRA can be verified on websites of the Internal Revenue Service (IRS) and Department of Energy's (DoG) Office of Energy Efficiency and Renewable Energy (IRS, 2022).

For the control group, 6 ICEV models chosen are among the best-selling SUV models in the U.S. for 2023, based on Statista's ranking (GoodCarBadCar.net, 2023). This selection aligns with the prevalence of SUVs in the treatment group, except for the Tesla Model 3, Tesla Model X, Ford F-150 Lightning, and Rivian R1. The Tesla Model 3, being a sedan, is matched with Toyota Camry, the top-selling sedan (Reynolds, 2023). The Tesla Model X, a large-size SUV, is paired with the best-selling large-size SUV, the Chevrolet Tahoe (Good Car Bad Car, 2022). For pick-up trucks, the Ford F-150 Lightning and Rivian R1 T are paired with their best-selling non-EV counterparts, Ford F-Series and Chevrolet Silverado, respectively (Hearst, 2024). All quarterly sales data for non-EVmodels of the control group are accessible on Carfigures.com (Carfigures.com, 2024). For analytical convenience, after initial data cleaning and restructuring, specifically assigning dummy variables to the first 9 EV models as 0-8 and the last 10 ICEV models as 9-18, the dataset is prepared for further analysis.

Sales	Observations	Mean	Std. dev.	Max	Min
Common Sample	269	10.20656	1.588462	12.77331	1.098612
Control Group (ICEVs)	154	11.13663	0.5993772	12.77331	9.561913
Treatment Group (EVs)	115	8.961079	1.648094	11.56322	1.098612

Table 1.2a: Summary Statistics on Model ln(*sales*) for DiD Analysis

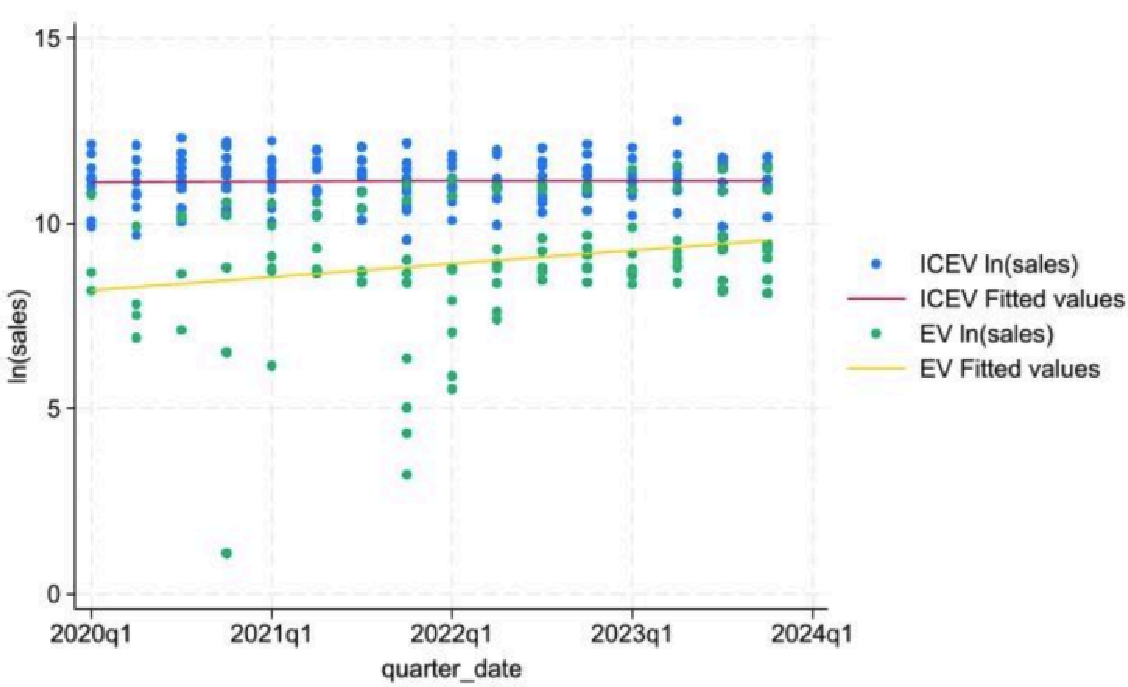


Figure 1a: EV (Treatment) and ICEV (Control) Model ln(*sales*) from 2020Q1 to 2023Q4

Data Collection for the State-Level Tax Credit Regression Model

The second dataset consists of cross-sectional data encapsulating various metrics such as the maximum EV tax credit, EV registration fee, EV registration count, and the unemployment rate across all 50 U.S. states. The information regarding the maximum EV tax credit and registration fees for individual states was obtained from a think tank known as the Tax Foundation, with its website's latest update in September 2023, drawing its figures from the U.S. News & World Report and State Statutes (Jaros and Hoffer, 2023).

State-level EV registration counts were sourced from the Department of Energy's Alternative Fuels Data Center, with the available data extending up to 2022 (United States Department of Energy, 2023). Therefore, the dataset and the associated regression analysis will utilize the 2022 EV registration data alongside the state-level tax credit information last updated in 2023. This approach, though not ideal, is appropriate as there have not been significant changes in the tax credit schemes from 2022 to 2023 for most states. The state-level unemployment rates were provided by Statista, which bases its figures on the Bureau of Labor Statistics' unemployment data for 2022 (Bureau of Labor Statistics, 2023).

Figure 2a: Tax Credit and ln (EV Registration Counts) of 50 States

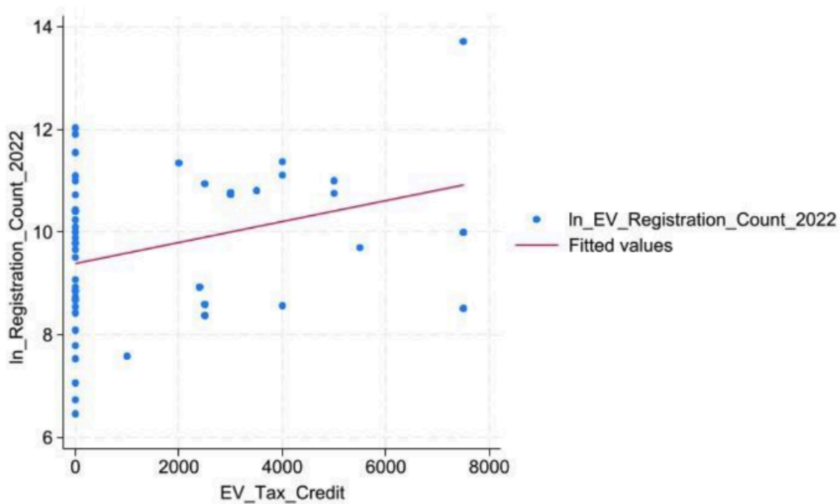


Figure 2a: Tax Credit and ln($EV\_Registration\_Counts$) of 50 States

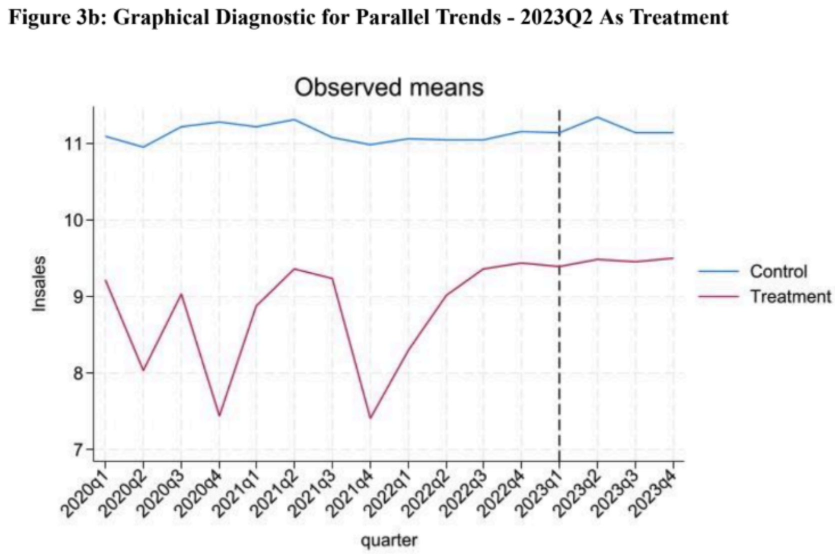


Figure 3b: Graphical Diagnostic for Parallel Trends - 2023Q2 as Treatment

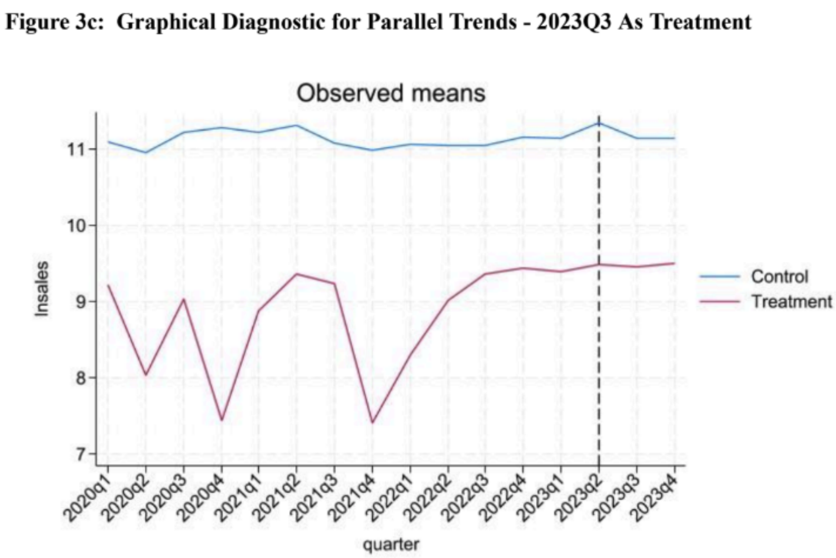


Figure 3c: Graphical Diagnostic for Parallel Trends - 2023Q3 as Treatment

Programs and Files

Stata18 is the only software program used to conduct data analysis throughout this project. The do file with all the commands can be accessed here. When importing the excel files into Stata, make sure to use the "Import" function then select "Excel spreadsheet." Remember to click the "use first row as variable" label.