Macalester Journal of Economics

Vol. 32 - Spring 2023

Macalester Journal of Economics

Volume 32 - Spring 2023

Contents

Foreword

The Macalester College chapter of Omicron Delta Epsilon, the international honors society in economics, proudly edits the Macalester Journal of Economics (MJE) every year. This year's editorial team – Cheikh Fall ('23), Valeska Fresquet Kohan ('23), Abbie Natkin ('23), Chiara Affatigato ('23) and Jonah Klein-Collins ('23)– have carefully selected eight papers on a variety of important topics. These papers are a sample of the research that our students produced in the last academic year. The sample includes a term paper written for a 200-level course on the economics of global food problems, four term papers written as part of a 300-level course on econometrics, and three capstone research papers written in two courses including behavioral economics and multinational corporations.

Anthropogenic climate change and its multitude of consequences are at the forefront of academic research and policy making around the globe. Unsurprisingly, a large share of this year's MJE papers are centered around this topic. The lead article is authored by Hufsa Ahmed ('24) and analyses the efficacy of state tax credits in raising solar panel adoption – a timely topic given the 2022 Inflation Reduction Act that revamped the Residential Clean Energy Credit. Within the US, nine states have offered varying personal tax credits to incentivize the adoption of solar panels. Using a two-way fixed effects model, among other empirical specifications, she finds no evidence that state tax credits significantly increase solar panel adoption irrespective of the size of the credit.

Other papers focused on anthropogenic climate change include the paper by Grace Generous ('24) and Mathilda Barr ('25). Grace Generous, for example, studies the effects of extreme heat days on the number of crimes committed in the context of Seattle, WA. The key finding is that unexpected and uncharacteristic extreme heat events can lead to a significant rise in the number of crimes committed during this period. This positive relationship appears particularly strong for crimes categorized as assault, larceny, and theft. In contrast to this study on crime, Mathilda Barr studies the impact of drought and climate change on the cattle cycle and ranch decision making patterns in the Southwestern United States. When faced with the 2022 extreme drought and poor prospects based on current climate predictions, many US ranchers decided to reduce their cattle head count causing a temporary price decline. Predictions of future price developments, however, point in the opposite direction due to a prolonged recovery in cattle heads (if at all) and significant feed grain supply tightening.

Aside from these three studies on environmental economics and climate change, the current issue of MJE brings forth two papers on the topic of international trade and Foreign Direct Investment (FDI), two papers on labor economics, and one study on the subject of behavioral economics. Each of these articles raises a fascinating research question and investigates a pressing issue: For example, does the engagement in international markets and accommodation of FDI stimulate economic growth? Or does this promotion of globalization expose countries to external shocks and create troublesome dependencies?

Zefan Qian ('23) investigates the former question and studies the relationship between FDI, exports, and economic growth in the context of East and Southeast Asia. To answer this question, he conducts a comprehensive time-series analysis that investigates the Granger causality between these three variables. The results are mixed and the relationships appear to change over time. In the more recent years, there is some evidence that suggests a growth promoting effect of FDI, which in turn stimulates further investment.

In contrast, Dino Weinstock ('23) and Patrick Fuchs ('23) investigate the latter question and study the changes in electricity prices faced by European countries in the aftermath of the Russian Invasion of Ukraine. The authors show that the increasing European dependence on Russian natural gas has its consequences. Using an event study design, the authors show that trade embargoes and the shutdown of natural gas pipelines raised electricity prices by more than €100/mWh in the most dependent countries. The results point to an important reality: reliance on international supply of critical commodities such as natural gas has its drawbacks, and these recent developments may intensify the discussions around global supply chain resiliencies and dependencies as well as energy supply diversification.

A different spin on yet another global topic – Economic Development – is presented in the paper by Abbie Natkin ('23) & Siri Hoff ('23). Determinants of economic development have been studied for centuries. In recent years, culture, and in particular, its effect on development through the channel of cooperation, has gained more attention in this strand of literature. In light of this research, the authors study the effect of campus culture on student cooperative behavior in public goods games. Specifically, the experiment evaluates whether cooperation declines when students from three different Minnesota liberal arts colleges play a public goods game with peers from another institution. The results provide some evidence that culture may shape cooperation as student behavior varied distinctly across the three schools.

The final set of papers contemplates topics in the realm of labor economics. The research addresses current and sensitive topics, such as the impact of refugees on labor market outcomes of natives. Mahmoud Majdi ('24) investigates this particular question in the context of refugee arrival and its effect on Norwegian employment. Using both an instrumental variables and fixed effects approaches, he finds inconclusive evidence regarding the refugee-employment relationship.

In another labor-oriented research project, Zak Yudhishthu ('24) studies land use regulations and local responses to labor demand shocks. Using a Bartik-type instrument (based on a location's exposure to national industry trends), the author shows that locations with greater levels of regulation experience stronger rent, but not permit, growth in response to a comparable labor demand shock. The author notes, however, these results are sensitive to the inclusion of control variables and deserving of further research.

On behalf of my colleagues in the Economics Department, I am delighted to present the research of these talented students. I am confident that you will find it enlightening and be impressed by the value of a liberal arts education.

Felix Friedt
Associate Professor of Economics

Using Effective Policies to Mitigate Climate Change: Do State Tax Credits Increase Residential Solar Panel Adoption?

Hufsa Ahmed
ECON 381: Introduction to Econometrics (Advisor: Amy Damon)

The world's supply of fossil fuels is not only quickly depleting, but the burning of fossil fuels is known to be the "primary cause of current climate change" (University of California Museum of Paleontology, 2023). Human activity and increased levels of carbon emissions from the burning of fossil fuels has contributed to record increases in global temperatures. With models predicting a 2 to 6 degrees Celsius increase in the earth's temperature over the next century, scientists estimate that only 8 years remain before the effects of climate change become irreversible (NASA, 2010; United Nations, 2019). As climate change becomes an increasingly visible phenomenon, there is a growing sense of urgency to make the transition from fossil fuels to more sustainable forms of energy in the United States. Just last year, in 2021, solar energy was recorded as the fastest growing source of electricity in the country (Pickerel, 2022). Solar power was once an expensive and inaccessible source of energy, but over the past 10 years, its price has decreased by 90% due to greater use and economies of scale. As of 2016, this has made the cost of solar energy comparable to traditional fossil fuel energy sources (Chrobak, 2021). Everyday Americans, however, are not widely adopting solar energy. Solar panel systems are a significant up-front investment with high purchase and installation costs totaling an average of $15,960 nationwide before incentives (Zagame, 2022). Furthermore, the benefits of a solar technology investment are mainly reaped in the long run. Although solar adopters will see their electricity bill reduced immediately, it can take anywhere from 5 to 15 years to break even on a financial investment into solar panels (Parkman, 2022). This duration can be reduced, and the financial appeal of solar panel adoption increased, through government-funded incentives.

Historically, tax incentives have been used by the government to promote socially desirable behavior. To date, 9 US states including Hawaii, Arizona, Massachusetts, Utah, Montana, New York, South Carolina, New Mexico, and Iowa have adopted their own versions of a residential solar tax credit with the goal of increasing solar panel use among their

Hufsa Ahmed is a junior with majors in Economics and Political Science. Correspondence concerning this article should be addressed to hahmed@macalester.edu.

residents (DSIRE, 2022). Most of these states adopted solar tax credits decades ago, but we have yet to widely explore whether these tax credits are leading to an increase in solar panel use. This paper will explore solar panel adoption rates among US states that do and do not have solar tax credit policies. Additionally, it will look at whether or not differences in state tax credit policies, specifically the dollar amount offered by each state's credit, affect solar panel adoption rates among states that do have tax credits. This analysis will answer the following question: do state tax credits increase residential solar panel adoption rates?

Studies on solar energy have long identified multiple characteristics and factors that affect adoption rates beyond the price tag associated with buying and installing solar panels. Surveys show that people with higher incomes, higher education, bigger homes, pro-climate ideologies, and higher household energy costs are more likely to have solar panels (Dastrup et al., 2012; Durham et al., 1988; Khurana et al., 2021; Sigrin et al., 2015). In neighborhoods where at least one other house already has solar panels, residents are more likely to get them as well due to increased exposure to solar energy as a viable energy alternative (Bollinger et al., 2020). Those who see panels as an addition to their home's resale value, which empirical studies show is at least a 3-4% increase in value, are also more likely to install solar panels (Dastrup et al., 2012).

Literature on the effect of government interventions, such as subsidies, on solar adoption gives insight into not only whether these policies are effective, but also in which circumstances they are most effective. Solar subsidies have been widely studied in many countries outside the United States. Recent research in Nepal, Chile, and India, for example, has shown increases in solar adoption when subsidies are used (Best & Nepal, 2022; Parsad et al., 2020; Walters et al., 2018). In the US, literature on subsidies being used to incentivize solar adoption is much less prevalent, but some papers written over the years have shown that they are an effective policy in the US as well (Crago & Chernyakhovskiy, 2017; De Groote & Verboven, 2019; Tibebu et al., 2021; Yokell, 1979). More specifically, studies emphasize that subsidies have the largest impact on adoption when they are implemented during the system installation process rather than at a future date as a delayed cash payment. This is because up-front financial support for installation costs decreases the risk of switching to solar energy (Bauner & Crago, 2015; De Groote & Verboven, 2019). Finally, as expected, the research also finds that the simpler the incentive is to receive, subsidy or otherwise, the more people adopt solar panels (Sarzynski et al., 2012).

Research looking into the efficacy of solar tax credits has been clustered within two time periods: the 1980s, when solar tax credits were first implemented, and the early 2010s, as solar energy became popular. Early research found that tax credits are effective in increasing solar adoption among residential households, but significantly more so when coupled with high prices of traditional forms of energy (Durham et al., 1988; Procter &

Tyner, 1984). Limited recent research affirms this, with a working paper that analyzes tax credits in three US states finding that they had a positive effect on residential solar adoption in areas with higher electricity costs (Caldas, 2020)]. An in-depth analysis of Hawaii's solar panel adoption rates finds that increasing electricity costs and declining solar installation costs have driven greater adoption rates. Tax credits have also boosted solar panel use in Hawaii, but have disproportionately benefited wealthy people (Coffman et al., 2016).

Synthesizing the literature, we know that government incentives that promote solar adoption, like subsidies and tax credits, have historically been effective, especially when traditional electricity prices are high and the financial risk of switching to solar energy is low. We also know that various demographic characteristics increase solar adoption rates. To best inform ongoing climate and energy policy conversations, additional up-to-date research is needed. There is a lack of recently published research on the impact of solar tax credits that takes into account major decreases in solar energy prices as of 2016. Additionally, although some research has been done on the effect of the implementation of a solar tax credit, no research has looked at the effect of a credit's size on residential solar adoption rates. Thus, my paper adds to existing literature by seeking to answer the question of how a state solar tax credit, and the magnitude of the credit offered, affect residential solar panel adoption. Analyzing both states that do and don't have a state tax credit, this paper conducts a nationwide analysis of all 50 states that does not currently exist.

Economic Theory and Intuition

Economic theory states that a rational decision-maker seeks to maximize their utility given a set budget constraint. In other words, consumers choose the "bundle" of goods that they believe provides the most benefit to them among all the bundles they can afford (Durham et al., 1988). When a consumer is choosing between traditional forms of energy and solar energy to run their household, we can assume that the perceived benefits of switching to solar come from both the long-term savings that they can accrue from using self-produced energy as well as intangible feelings of being socially responsible and knowing they are unaffected by electricity price changes (Durham et al., 1988). As a society, we have reached a point where solar energy and traditional electric energy are comparable in price, so the biggest deterrent from switching to solar energy is the risk that comes with the costly up-front investment into solar panels. From a behavioral standpoint, we know that rational people are typically risk averse and will make financial decisions that put them at ease. This is where economics says it is important for the government to step in and promote the socially desirable choice. By incentivizing solar energy use among its population, the government is helping address negative externalities of fossil fuel use, including greenhouse gas emissions and climate change. If the government provides a tax credit specifically, which assists in the up-front costs of adopting solar panels, a switch to solar energy becomes less risky, and the overall cost of using solar energy decreases for a consumer. Simple supply and demand logic indicates that a decrease in the price of a normal good results in an increase in the demand for that good. Thus, I expect that the relationship between the implementation of a solar tax credit and the number of solar panel installations is positive.

Data Description

The data for this analysis have been pulled from multiple sources. All solar panel data are from the Lawrence Berkeley National Laboratory's Electricity and Policy Markets division. Variables on demographic composition are from the American Community Survey completed by the US Census Bureau. The data on electricity prices come from the U.S. Energy Information Administration, the data on solar tax credit and net metering policies come from the Database of State Incentives for Renewables & Efficiency, and the data on government partisan alignment come from the National Conference of State Legislatures. This paper will use both a state-level and a tract-level data set to conduct analyses on the effect of tax credits on residential solar panel adoption rates.

State-Level Dataset

Table 1

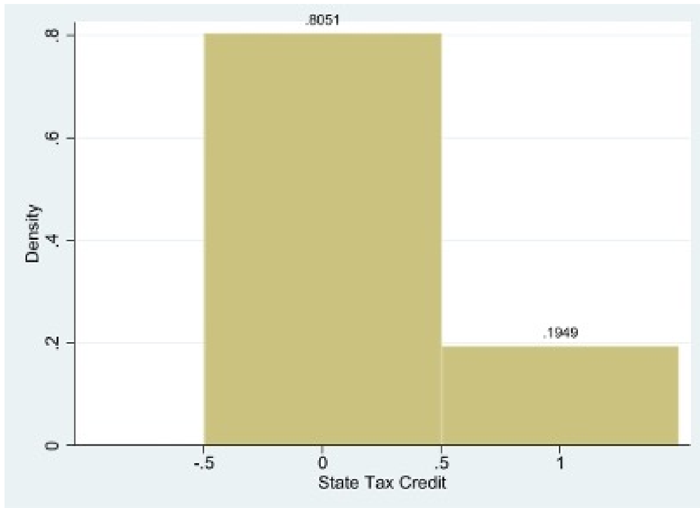
Descriptive Statistics

Variable	Obs.	Mean	S.D.	Min	Max
Number of Solar Panels	472	5691.72	19328.30	10	172812
State Tax Credit	472	0.195	0.397	0	1
Tax Credit Dollar Value	472	493.644	1347.153	0	6000
Population	472	7430000	7650000	623657	39557045
Owns a Home	472	1750000	1560000	177772	7502706
Has a Graduate Degree	472	609000	636000	57193	3779787
Median Income	472	29722.496	4584.299	20685	44936
Percent People of Color	472	24.429	13.27	4.575	77.824
Electricity Price	472	11.432	4.155	6.44	34
Statewide Net Metering Policy	472	0.869	0.388	0	1
Governor's Political Party	472	0.468	0.5	0	1

The first set of data provides the number of residential solar panel systems installed in all 50 states between the years 2010 and 2021. Table 1 provides a broad summary of this data set with 472 total observations. Given that only 9 states have solar tax credit policies, only 92 (19.5%) of the total observations come from a state with a credit in place (Figure 1).

Figure 1

Observations Split by State Tax Credit



Notes: 0 = no state tax credit and 1 = state tax credit

This data set has complete data across all 12 years for 32 of the 50 states. 11 of the remaining states have data for some years (Table 2), while the other 7 have no data at all. The states with no available data include Alabama, Alaska, Mississippi, North Dakota, South Dakota, Wyoming, and West Virginia.

Table 2

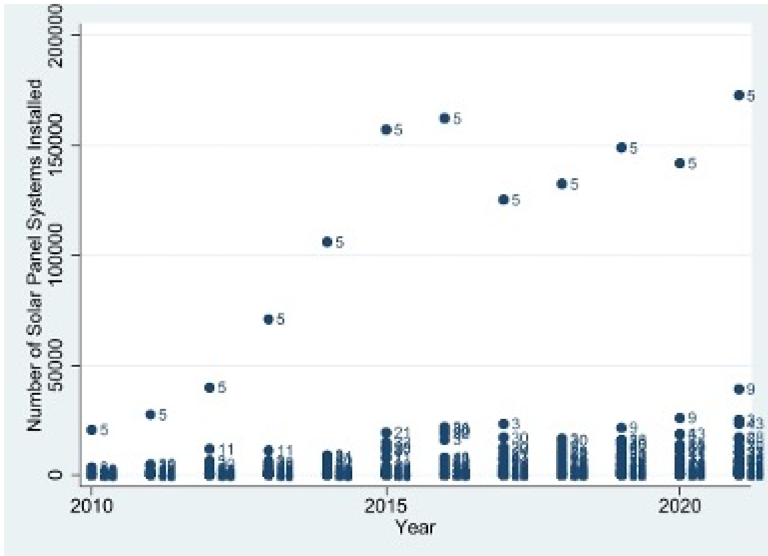
States with Partial Data

State	Years of Data	State	Years of Data
Arkansas	7	Kentucky	7
Delaware	11	Louisiana	11
Georgia	7	Nebraska	6
Indiana	9	Oklahoma	5
Iowa	9	South Carolina	9

Looking at important trends in the data, California, marked as state ID 5 in the figure below, is a large outlier in solar panel adoption rates among all states (Figure 2A). Removing California from the data, we more clearly see trends in solar panel installations in the remaining states (Figure 2B). While some states' solar panel adoption rates stay consistently low over time, others see sharp increases mid-decade or exponential growth over time.

Figure 2

Solar Panel Installation Rates Over Time



Notes: State 5 represents California

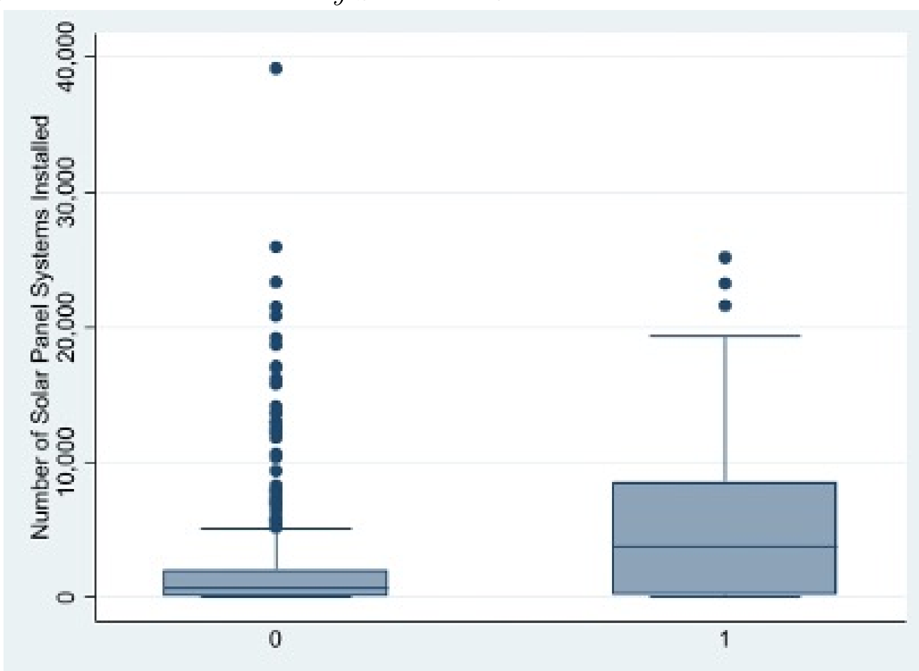
Figure 3

Solar Panel Installations by State & Over Time

Figure 3 splits observations between states with and without tax credits. Removing the biggest outlier, California, we see that the box on the right, comprised of observations with solar tax credits in place, has a greater median and a larger interquartile range than those without solar tax credits on the left. California aside, both boxes have outliers, many of which are not entirely visible in the figure. Thus, they are tabulated in Table 3, where we see a similar number of outliers between the two groups.

Figure 4

Annual Solar Panel Installations by State Tax Credit



Notes: 0 = no state tax credit and 1 = state tax credit

Table 3

Outliers in Annual Observations

States w/o any Tax Credits	# of Outliers	States w/o any Tax Credits	# of Outliers
Arkansas	7	Kentucky	7
Delaware	11	Louisiana	11
Georgia	7	Nebraska	6
Indiana	9	Oklahoma	5
Iowa	9	South Carolina	9

Notes: Not Including California.

Finally, among the 9 states that do have solar tax credits, New Mexico's credit has the highest dollar value, while Utah's 2021[1] tax credit has the lowest (Table 4).

Table 4

Descriptive Statistics for Tax Credit Dollar Values

Variable	Obs.	Mean	S.D.	Min	Max
Tax Credit Dollar Value	83	2807.229	1962.19	400	6000

Notes: Excludes 9 observations from South Carolina, which has a tax credit in place but does not assign a dollar value to it (the amount of the credit is instead based on a percentage of the total installation cost for a solar panel system).

Tract-Level Dataset

The second data set includes 32,959 observations for the states of Arizona, California, and Colorado for the years 2010 to 2020. Here, adoption numbers are counted by census tract, a unit of measurement smaller than a county. This data set will supplement the state-level data and provide additional information on the effect of state tax credits on solar panel system installations. Arizona, California, and Colorado have a combined total of 10,832 census tracts, and the data set has information for 6,263 of those tracts. Of the three states in this data set, only Arizona has adopted statewide tax credits. Thus, Arizona's 3,231 tract observations are the only ones from the data set that have been "treated" with a state tax credit (Table 5).

Table 5

Total Observations With and Without Tax Credits

Has a State Tax Credit	Freq.	Percent
No	29,728	90.20
Yes	3,231	9.80
Total	32,959	100.00

Table 6 shows a table of means summarizing this data set. The high number of outliers within this data set, due to the inclusion of California, makes it difficult to compare between states, so Table 7 helps quantify the distribution in solar adoption rates by state. Most notably, the median solar panel installation rate per tract across all three states is very similar despite Arizona being the only one with a solar tax credit. A closer comparison of each state's solar panel installations per year reveal no major trends over time (Figure 5).

[1] Utah provided a statewide solar tax credit with a flat dollar value until 2017 when they began to decrease the value of the tax credit annually until the policy ended in 2021

Table 6

Descriptive Statistics (Tract-Level)

Variable	Obs.	Mean	S.D.	Min	Max
Number of Solar Panels	32959	30.174	33.086	10	1050
State Tax Credit	32959	0.098	0.297	0	1
Tax Credit Dollar Value	32959	98.031	297.361	0	1000
Population	32959	5404.62	2281.995	620	39373
Owns a House	32959	1231.679	561.113	0	5815
Has a Graduate Degree	32959	489.602	440.886	0	5843
Median Income	32959	37574.763	15218.637	3832	135833
Percent People of Color	32959	31.172	18.879	0	97.865
Electricity Prices	32959	15.188	2.053	9.15	18
Statewide Net Metering Policy	32959	0.946	0.226	0	1
Governor's Political Party	32959	0.888	0.316	0	1

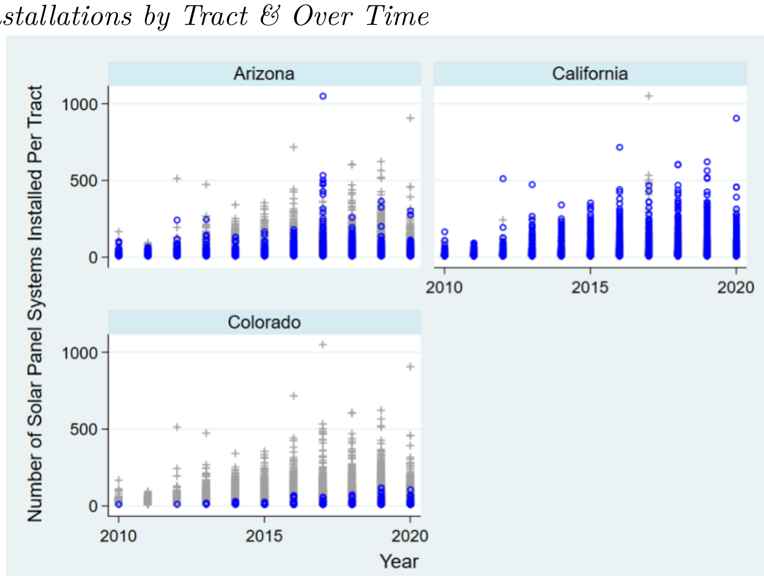
Table 7

Range of Annual Tract-Level Solar Panel Installations by State

State	Min	25^{th} Perc.	Median	75^{th} Perc.	Max
Arizona	10	13	20.000	34	1050
California	10	14	20.000	34	906
Colorado	10	12	16.000	24	119

Figure 5

Solar Panel Installations by Tract & Over Time



Empirical Strategy

Both the state-level and tract-level datasets I use are panel data, so I will split this section between cross-sectional regressions and panel regressions. My two variables of interest in the following regressions are *statetaxcredit*, a dummy variable that indicates if a state does or does not have a tax credit and *dollarcredit* , a continuous variable representing the size of a tax credit in a given state. In other words, this is the maximum dollars a household can receive from utilizing the credit. In my analysis, the null hypothesis is that $\beta_1 = 0$; I assume that there is no relationship between the presence of a state tax credit and a state's solar panel installation rates, or the size of a state's tax credit and the state's solar panel installation rates. Significant results would provide evidence that there actually is a relationship between solar panel installation rates and the independent variable of interest. Discussions on the shortcomings of this empirical strategy are included in the Limitations Section.

Cross-Sectional Regressions

I start with a basic state-level cross sectional regression looking at the variable *statetaxcredit* for the year 2020:

$$\begin{aligned} N_i = \beta_0 &+ \beta_1 statetaxcredit_i + \beta_2 population_i + \beta_3 poc_i + \beta_4 ownhouse_i \\ &+ \beta_5 income_i + \beta_6 graddegree_i + \beta_7 electricity_i + \beta_8 partisanship_i \\ &+ \beta_9 netmetering_i + \varepsilon_i \end{aligned} \tag{1}$$

where N_i is the number of solar panel systems installed. Controls make up the remainder of the regression, where population is state population, poc_i is the percent of the population that is nonwhite, $ownhouse_i$ is the number of residential houses that are owned, $income_i$ is the average income, $graddegree_i$ is the number of people with a graduate level degree, $electricity_i$ is the average electricity prices in cents per kilowatt of energy, $partisanship_i$ is a dummy variable representing the political party affiliation of the sitting governor, $netmetering_i$ is a dummy variable for whether or not the observation is from a state with a statewide net metering policy, and ε_i is the error term, all in state i.

One major threat to the internal validity of this analysis is endogeneity. Although I seek to find the effect of a state tax credit on solar panel installation rates, there is a reasonable argument to be made that states with naturally higher solar panel installation rates are more likely to adopt tax credit policies. To account for this potential reverse causality, I run an instrumental variable regression in equations (2) and (3) and use the percentage of Democratic legislators in a state's House of Representatives chamber as an excluded instrument. This instrument is correlated with the variable of interest, *statetaxcredit*, but does not explain N except through *statetaxcredit*. In this two-stage regression model, the first stage finds a predicted value for the state tax credit, $predictedstatetaxcredit_i$, while the second stage uses the predicted value to produce what should be an unbiased estimate of N, the number of solar panel systems installed in a given state.

$$\begin{aligned} predictedstatetaxcredit_i = {} & \alpha_0 + \alpha_1 percdemleg_i + \alpha_2 population_i + \alpha_3 poc_i \\ & + \alpha_4 ownhouse_i + \alpha_5 income_i + \alpha_6 graddegree_i \\ & + \alpha_7 electricity_i + \alpha_8 partisanship_i + \alpha_9 netmetering_i + \mu_i \end{aligned} \tag{2}$$

$$\begin{aligned} N_i = {} & \alpha_0 + \beta_1 predictedstatetaxcredit_i + \beta_2 population_i + \beta_3 poc_i + \beta_4 ownhouse_i \\ & + \beta_5 income_i + \beta_6 graddegree_i + \beta_7 electricity_i + \beta_8 partisanship_i \\ & + \beta_9 netmetering_i + \varepsilon_i \end{aligned} \tag{3}$$

where *percdemleg* is the percent of Democratic legislators in a state's House of Representatives in i state in the year 2020 and *predictedstatetaxcredit* is the predicted increase in probability that a state has a tax credit given a one unit increase in the percent of a House of Representatives' members who are Democrats. I utilize a linear probability model (LPM) framework to interpret the first stage regression whose outcome variable is binary. Although a logit or probit model might be more appropriate for a cross-sectional regression, because of a need for an ease of interpretation, an LPM works well. Finally, I use an F-test in the first-stage regression to see if the chosen excluded instrument is a strong predictor of the endogenous independent variable.

Panel Data Regressions

Using the complete panel data sets that follow observations over time, I can account for another large threat to internal validity: omitted variable bias. I use a between effects model in regression 4 (4.1 and 4.2 in the table) to further analyze my variable of interest, *statetaxcred*, at both the state and tract level. The lack of variation of *statetaxcred* within each state over the time period for which I have data prevents the use of the more ideal two-way fixed effects model which accounts for both time and entity invariant factors.

$$\begin{aligned} N_{it} = {} & \beta_1 statetaxcredit_{it} + \beta_2 population_{it} + \beta_3 poc_{it} + \beta_4 ownhouse_{it} \\ & + \beta_5 income_{it} + \beta_6 graddegree_{it} + \beta_7 electricity_{it} + \beta_8 partisanship_{it} \\ & + \beta_9 netmetering_{it} + \alpha_i + \varepsilon_{it} \end{aligned} \tag{4}$$

where α_i is the entity fixed effects and each variable is in state i or tract during t year.

Turning to my second variable of interest, *dollarcredit*, I now analyze the effect of different tax credit policies among states, which offer various dollar amounts in credit, on solar panel installation rates. I will use a two-way fixed effects regression model at the state level in regression (5) to account for time invariant and entity invariant variables, such as local cultural norms and the perceived amount of sun received by a state in a year, that would place bias on my results.

$$\begin{aligned} N_{it} = {} & \beta_1 statetaxcredit_{it} + \beta_2 population_{it} + \beta_3 poc_{it} + \beta_4 ownhouse_{it} \\ & + \beta_5 income_{it} + \beta_6 graddegree_{it} + \beta_7 electricity_{it} + \beta_8 partisanship_{it} \\ & + \beta_9 netmetering_{it} + \alpha_i + \gamma_t + \varepsilon_{it} \end{aligned} \tag{5}$$

where *dollarcredit* is the maximum dollar amount one can receive from a tax credit, γ_t is the time fixed effects, and α_i is the entity fixed effects in state i during year t. I will run the regression twice, once where Y_{it} represents N_{it}, the number of solar panels installed in state or tract i during year t (regression 5.1), and again with Y_{it} representing inverse hyperbolic sine, so that Y_{it} is now interpreted as the β_1 increase in solar panel installation rates resulting from a 1% increase in the dollar value of a tax credit, in state i during year t (regression 5.2). The variable of interest in this regression, *dollarcredit*, has many zero values because very few states have tax credits in place. Using inverse hyperbolic sine allows for a log transformation of the results with the zero values included.

Results

The regression results are split by the variables of interest in this paper: Table 8 lists all the regressions addressing the presence of a state tax credit, *statetaxcred*, while Table 9 analyzes the dollar value of each existing credit, *dollarcredit*. In Table 8, across all state-level regressions, *statetaxcred* is insignificant. However, the tract-level regression reports high significance, suggesting that a state with a tax credit expects to see a 77.56 unit decrease in the number of solar panel installations compared to states without a tax credit. This result is not robust, however: the *statetaxcred* coefficient in alternate regressions lacks significance when I change the controls (see Table 0.2 in the appendix). Altogether, the lack of consistent significance across all four regressions in this table indicates that the presence of a state tax credit does not lead to an increase in the use of solar panels by residential homes. Given the repeatedly high levels of significance across all but one of the control variables in the tract-level regression (4.2) and among the robustness checks in Table 0.2, I speculate that had the state-level dataset included more observations and variation, its results may also have been significant.

Additionally, the results of the two-stage least squares regression with the instrumental variable proved unhelpful and did not solve the endogeneity problem as expected. Not only were the results of the second-stage regression insignificant, as seen below, the first-stage regression provided an F-stat of only 1.54, showing that percent of Democrats in a state's House of Representatives is a weak predictor of *statetaxcred*.

Table 9, listing regressions for *dollarcredit*, again shows insignificant results. In regression 5.1, I had just enough variation in dollar credit values to run a two-way fixed effects model. Besides the fact that the results were insignificant, I have additional skepticism with the results because the entirety of the regression output is based on variations from just two states.

The robustness checks performed in Tables 10, 11, and 12 in the appendix reaffirm the insignificant results of the primary regressions included in this section.

Limitations

Two main concerns about the data used in this analysis prevent the results from being internally valid. The first is that the data is systematically unbalanced. The states with partial or no data are also the states that have the least solar infrastructure and solar usage (DSIRE, 2022). Not accounting for the places with the fewest solar panel in the country is problematic when looking for the true relationship between tax credit policies and solar panel installation rates. The second concern is omitted variable bias. Major variables unaccounted for in the regressions include the prices of solar panels and other existing state and local policies incentivizing residential solar adoption. These variables are currently found in the error term, placing a bias on each regression's variable of interest and violating the first Gauss-Markov condition stating that the expected value of the error is equal to zero.

Other Gauss-Markov conditions are violated in my regression analysis as well. California presents itself as an especially large outlier when observing any solar-related data. This skews the regression results and prevents outputs from being representative of the entire data set. Furthermore, neither of the independent variables of interest are independently distributed (they are not random), violating yet another condition for achieving the best linear unbiased estimator.

Building upon discussions above, a few more minor concerns around the internal validity of this paper's results exist. Encompassing states solar tax credit policies that vary in structure and detail is difficult using just a handful of continuous and binary variables. This leads to both a misspecification of the functional form and error-in-variables bias. For example, South Carolina's tax credit policy does not specify a dollar amount for its tax credit, and instead chooses to provide a tax credit for a set percentage of a household's solar installation costs. This policy cannot be encompassed by the variables in the regressions, and thus, this detail about South Carolina's policy was left out of the analysis. Ideally, additional variables would have been included in the regression to account for all the vari-

Table 8

Results for Regressions on the State Tax Credit Binary Variable

Variables	(1) Only 2020	(2 & 3) Instrument	(4.1) Between Effects	(4.2) Between Effects
State Tax Credit	-462.2	60,237	-1,082	-77.56***
	(4,690)	(57,566)	(3,527)	(7.039)
Total Population	0.0105***	0.00943***	0.00881***	0.00514***
	(0.00194)	(0.00319)	(0.00138)	(0.000189)
Percent People of Color	-70.55	-158.7	-60.73	-0.131***
	(145.2)	(173.1)	(109.3)	(0.0146)
Own a Home	-0.0371***	-0.0217***	-0.0337***	0.00706***
	(0.00612)	(0.00820)	(0.00463)	(0.000812)
Median Income	-0.0329	1.609	-0.0706	0.000348***
	(0.628)	(1.349)	(0.491)	(3.00e-05)
Graduate Degree	-0.00902	-0.0380	-0.00566	-0.0161***
	(0.0154)	(0.0503)	(0.0116)	(0.00113)
Governor's Political Party	1,740	4,046	-1,290	-81.16***
	(3,553)	(5,204)	(4,172)	(6.984)
Electricity Price	501.7	-1,200	379.4	0.782***
	(510.5)	(1,342)	(409.1)	(0.262)
Statewide Net Metering Policy	5,648	21,709	8,307*	-0.481
	(5,630)	(18,390)	(4,649)	(2.529)
Constant	-6,389	-65,851	-4,610	58.16***
	(20,938)	(54,843)	(14,490)	(8.129)
Level	State	State	State	Tract
Observations	43	472	472	32,959
R-squared	0.825	-0.663	0.835	0.290
Adjusted R-squared	0.777	-0.696	0.790	0.289
F-Stat	17.30	20.26	18.56	283.5

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Std. Errors in parentheses.

Table 9

Results for Regressions on the State Tax Credit Binary Variable

Variables	(5.1) Two-Way Fixed Effects	(5.2) Inverse Hyperbolic Sine
Dollar Credit	-1.380	-813.6
	(0.913)	-597.8
Total Population	0.00491**	0.00492**
	(0.00198)	(0.00198)
Percent People of Color	902.6***	892.3***
	(203.1)	(203.9)
Own a Home	-0.0543***	-0.0541***
	(0.00834)	(0.00835)
Median Income	-0.758	-0.744
	(0.522)	(0.522)
Graduate Degree	0.0988***	0.0987***
	(0.00730)	(0.00730)
Governor's Political Party	1,027	1,014
	(866.6)	(867.1)
Electricity Prices	1,830***	1,826***
	(371.5)	(371.6)
Statewide Net Metering Policy	-5,394**	-5,382**
	(2,317)	(2,318)
Constant	8,801	13,211
	(17,807)	(18,535)
Observations	472	472
R-squared	0.904	0.904
Adjusted R-squared	0.889	0.889
F-Stat	61.91	61.84

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Std. Errors in parentheses.

ation in policies. This would not only include tax credit policies that are structured by percentages rather than dollar amounts, but also variations in how much time a household is given to submit their request for a tax credit and how difficult it is to access the tax credit. Concerns about measurement errors within the data can be addressed using an instrumental variable regression. Unfortunately, the selected instrument of the percentage of Democrats in a state's House of Representatives was a very weak predictor of my variable of interest. A more creative excluded instrument is necessary to solve the endogeneity problem and other concerns discussed in this section.

On another note, a lack of complete and adequate data did not allow for the use of an ideal empirical strategy. Among existing state tax credit policies, most were implemented in the late 1990s or early 2000s. Available data, however, begins in 2010, preventing a comparison of solar panel adoption rates pre and post state tax credit "treatment." For this reason, this paper's analysis of its first variable of interest, the presence of a state tax credit, was constrained to a between effects model. Using a two-way fixed effects model (a continuous difference-in-differences model) with data starting in the 1990s would have been far better.

Conclusion

Although the insignificant results of this paper failed to meaningfully add to existing literature around the effect of solar adoption policies, its findings are still important. They expose the gap in existing data, and this paper serves as a good example for why better data in the residential solar industry need to be collected. Despite the first state solar tax credit being adopted over 40 years ago in 1979, the existing data on solar panel use prior to 2010 are inadequate (DSIRE, 2022). As the world begins to face the effects of a looming climate crisis, it is imperative that lawmakers know which existing climate-based policies are effective. The days of taking a backseat on climate action have passed, and the policies being implemented must resoundingly contribute to building a climate-resilient country. In recent years, as the popularity of solar energy has grown in the United States, groups like the Lawrence Berkeley National Laboratory have begun conducting thorough research and analyses on solar energy, collecting the necessary data along the way. The recent commitment within the United States to study solar energy at the residential level, and create better policies to incentivize solar use, is promising. With the passing of the Inflation Reduction Act by Congress just this year, there have been changes in solar tax credit policies at the federal level. Beyond this most recent change, federal solar tax credit policies have fluctuated since 2016, providing a perfect opportunity to further study the effects of solar tax credits on residential solar adoption in the United States as the data become available. The age of solar energy has just begun.

References

Bauner, C. & Crago, C. (2015). Adoption of Residential Solar Power Under Uncertainty: Implications for Renewable Energy Incentives. *Energy Policy*, 86, 27–35. `https://doi.org/https://doi.org/10.1016/j.enpol.2015.06.009`

Best, R. & Nepal, R. (2022). Saving and Subsidies for Solar Panel Adoption in Nepal. *Applied Economics*, 54(59), 6768–6783.

Bollinger, B., Gillingham, K., Kirkpatrick, J., & Sexton, S. (2020). Visibility and Peer Influence in Durable Good Adoption. *Marketing Science*, 41(3), 453–476. `https://doi.org/https://doi.org/10.2139/ssrn.3409420`

Caldas, M. (2020). An Analysis of the Effectiveness of Tax Incentives for Solar Energy. *University of Texas at Austin Honors Program*, 1. `https://repositories.lib.utexas.edu/bitstream/handle/2152/84129/caldasmariana_4131701_57618453_Caldas_Thesis_2020_Redacted.pdf?sequence=2&isAllowed=y`

Chrobak, U. (2021). Solar Power Got Cheap: So Why Aren't We Using It More? *Popular Science*. `https://www.popsci.com/story/environment/cheap-renewable-energy-vs-fossil-fuels/`

Coffman, M., Wee, S., Bonham, C., & Salim, G. (2016). A Policy Analysis of Hawaii's Solar Tax Credit. *Renewable Energy*, 85, 1036–1043.

Crago, C. & Chernyakhovskiy, I. (2017). Are Policy Incentives for Solar Power Effective? Evidence from Residential Installations in the Northeast? *Journal of Environmental Economics and Management*, 82, 132–151. `https://doi.org/https://doi.org/10.1016/j.jeem.2016.09.008`

Dastrup, S. R., Graff-Zivin, J., Costa, D. L., & Kahn, M. E. (2012). Understanding the Solar Home Price Premium: Electricity Generation and "Green" Social Status. *European Economic Review*, 56(5), 961–973. `https://doi.org/https://doi.org/10.1016/j.euroecorev.2012.02.006`

De Groote, O. & Verboven, F. (2019). Subsidies and Time Discounting in New Technology Adoption: Evidence from Solar Photovoltaic Systems. *The American Economic Review*, 109(6), 2137–2172. `https://doi.org/https://doi.org/10.1257/aer.20161343`

DSIRE (2022). *Programs.* `https://programs.dsireusa.org/system/program/mn/solar`

Durham, C., Colby, B., & Longstreth, M. (1988). The Impact of State Tax Credits and Energy Prices on Adoption of Solar Energy Systems. *Land Economics*, 64(4), 347–355. https://doi.org/https://doi.org/10.2307/3146307

Khurana, R., Elbakidze, L., & Hall, J. (2021). The Political Economy of Solar Initiatives in the Sunshine State. *Applied Economics Letters*, 28(9), 717–720. https://doi.org/https://doi.org/10.1080/13504851.2020.1771269

NASA (2010). *How is Today's Warming Different from the Past?* https://earthobservatory.nasa.gov/features/GlobalWarming/page3.php

Parkman, K. (2022). *How Much Do Solar Panels Cost?* https://www.consumeraffairs.com/solar-energy/how-much-do-solar-panels-cost.html

Parsad, C., Mittal, S., & Krishnankutty, R. (2020). A Study on the Factors Affecting Household Solar Adoption in Kerala, India. *International Journal of Productivity and Performance Management,*, 69(8), 1695–1720. https://doi.org/https://doi.org/10.1108/IJPPM-11-2019-0544

Pickerel, L. (2022). Solar was Nation's Fastest Growing Source of Electricity in 2021. *Solar Power World*. https://www.solarpowerworldonline.com/2022/02/solar-was-nations-fastest-growing-source-of-electricity-in-2021/

Procter, R. J. & Tyner, W. E. (1984). Assessing the Impact of Peak-Load Electricity Pricing and the Solar Tax Credits on the Adoption of Solar Energy. *Land Economics*, 60(1), 49–55. https://doi.org/https://doi.org/10.2307/3146092

Sarzynski, A., Larrieu, J., & Shrimali, G. (2012). The Impact of State Financial Incentives on Market Deployment of Solar Technology. *Energy Policy*, 46, 550–557. https://doi.org/https://doi.org/10.1016/j.enpol.2012.04.032

Sigrin, B., Pless, J., & Drury, E. (2015). Diffusion into New Markets: Evolving Customer Segments in the Solar Photovoltaics Market. *Environmental Research Letters*, 10.

Tibebu, T. B., Hittinger, E., Miao, Q., & Williams, E. (2021). What is the Optimal Subsidy for Residential Solar? *Energy Policy*, 155(1), 1–11. https://doi.org/https://doi.org/10.1016/j.enpol.2021.112326

United Nations (2019). *Only 11 Years Left to Prevent Irreversible Damage from Climate Change, Speakers Warn during General Assembly High-Level Meeting*. https://press.un.org/en/2019/ga12131.doc.htm

University of California Museum of Paleontology (2023). *Burning of Fossil Fuels.* https://ugc.berkeley.edu/background-content/burning-of-fossil-fuels/

Walters, J. P., Kaminsky, J., & Huepe, C. (2018). Factors Influencing Household Solar Adoption in Santiago, Chile. *Journal of Construction Engineering and Management*, 144(6). https://doi.org/10.1061/(ASCE)CO.1943-7862.0001483

Yokell, M. D. (1979). The Role of the Government in the Development of Solar Energy. *The American Economic Review*, 69(2), 357–361. https://doi.org/https://doi.org/10.4324/9780429051920-6

Zagame, K. (2022). *Will a 6 kW Solar Panel System Work for Your Home?* https://www.ecowatch.com/solar/6-kw-solar-systems

Table 0.10

Robustness Check for State Tax Credit State-Level Between Effects Regressions (3)

Variables	(R.C.1)	(R.C.2)	(R.C.3)	(R.C.4)	(R.C.5)	(R.C.6)	(R.C.7)
State Tax Credit	2,044	-2,181	-2,156	-2,253	-2,018	-2,758	-1,082
	(4,800)	(3,053)	(3,093)	(3,231)	(3,374)	(3,508)	(3,527)
Total Population	0.00154***	0.00828***	0.00826***	0.00814***	0.00817***	0.00799***	0.00881***
	(0.000260)	(0.000874)	(0.000888)	(0.00130)	(0.00132)	(0.00135)	(0.00138)
Percent People of Color	87.18	-20.96	-26.64	-26.76	-29.81	-64.84	-60.73
	(153.7)	(97.23)	(100.6)	(102.0)	(103.8)	(112.7)	(109.3)
Own a Home		-0.0335***	-0.0334***	-0.0333***	-0.0330***	-0.0323***	-0.0337***
		(0.00427)	(0.00435)	(0.00443)	(0.00461)	(0.00471)	(0.00463)
Median Income			0.102	0.0672	0.0529	-0.0849	-0.0706
			(0.370)	(0.467)	(0.476)	(0.507)	(0.491)
Graduate Degree				0.00129	0.000134	0.000824	-0.00566
				(0.0104)	(0.0113)	(0.0114)	(0.0116)
Governor's Pol. Party					1,174	459.7	-1,290
					(4,073)	(4,184)	(4,172)
Electricity Price						345.3	379.4
						(421.6)	(409.1)
Statewide Net Metering							8,307*
							(4,649)
Observations	472	472	472	472	472	472	472
Adjusted R-Squared	0.477	0.795	0.790	0.784	0.779	0.777	0.790
F-Stat	13.78	41.75	32.60	26.45	22.10	19.24	18.56

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Std. Errors in parentheses.

Table 0.11

Robustness Check for State Tax Credit Tract-Level Between Effects Regressions (4.2)

Variables	(R.C.1)	(R.C.2)	(R.C.3)	(R.C.4)	(R.C.5)	(R.C.6)	(R.C.7)
State Tax Credit	-1.007	-1.610*	-1.358	-1.264	-78.73***	-77.29***	-77.56***
	(0.829)	(0.825)	(0.834)	(0.823)	(6.883)	(6.896)	(7.039)
Total Population	0.00539***	0.00421***	0.00437***	0.00499***	0.00514***	0.00514***	0.00514***
	(0.000123)	(0.000171)	(0.000187)	-0.000191	(0.000189)	(0.000189)	(0.000189)
Percent People of Color	-0.194***	-0.138***	-0.138***	-0.134***	-0.128***	-0.131***	-0.131***
	(0.0138)	(0.0149)	(0.0149)	(0.0147)	(0.0145)	(0.0145)	(0.0146)
Own a Home		0.00689***	0.00603***	0.00789***	0.00679***	0.00706***	0.00706***
		(0.000698)	(0.000807)	(0.000809)	(0.000807)	(0.0008ll)	(0.000812)
Median Income			4.41e-05**	0.000335***	0.000346***	0.000348***	0.000348***
			(3.03e-05)	(3.03e-05)	(3.00e-05)	(3.00e-05)	(3.00e-05)
Graduate Degree				-0.0148***	-0.0159***	-0.0161***	-0.0161***
				(0.00113)	(0.00112)	(0.00l12)	(0.00113)
Governor's Pol.q1 Party					-78.53***	-81.19***	-81.16***
					(6.930)	(6.982)	(6.984)
Electricity Price						0.783***	0.182***
						(0.261)	(0.262)
Statewide Metering							-0.481
							(2.529)
Constant	3.610***	0.384	-1.069	-10.48	67.54***	57.69***	58.16***
	(0.799)	(0.858)	(1.098)	(1.301)	(7.004)	(7.733)	(8.129)
Adjusted R-squared	0.242	0.253	0.254	0.273	0.288	0.289	0.289
F-Stat	666.9	532.2	426.9	393.8	362.7	318.9	283.5

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Std. Errors in parentheses.

Table 0.12

Robustness Check for State-Level Dollar Credit Two-Way Fixed Effects Regressions (5.1)

Variables	(R.C.1)	(R.C.2)	(R.C.3)	(R.C.4)	(R.C.5)	(R.C.6)	(R.C.7)
Dollar Credit	-1.108	-1.191	-0.918	-1.077	-1.165	-1.384	-1.380
	(1.083)	(1.113)	(1.120)	(0.933)	(0.938)	(0.918)	(0.913)
Total Population	0.0128***	0.0134***	0.0135***	0.00431***	0.00434***	0.00494***	0.00491***
	(0.00141)	(0.00230)	(0.00230)	(0.00203)	(0.00203)	(0.00199)	(0.00198)
Percent People of Color	947,9***	985.9***	946.6***	901.6***	900.8***	911.6***	902.6***
	(222.5)	(250.3)	(250.6)	(208.8)	(208.8)	(204.1)	(203.1)
Own a Home		-0.00307	-0.00330	-0.0551***	-0.0551***	-0.0534***	-0.0543***
		(0.00921)	(0.00919)	(0.00856)	(0.00856)	(0.00838)	(0.00834)
Median Income			1.133*	-0.451	-0.412	-0.740	-0.758
			(0.620)	(0.530)	(0.532)	(0.525)	(0.522)
Graduate Degree				0.101***	0.100***	0.0976***	0.0988***
				(0.00745)	(0.00746)	(0.00732)	(0.00730)
Governor's Pol. Party					840.6	974.2	1,027
					(890.6)	(870.9)	(866.6)
Electricity Price						1,637***	1,830***
						(364.1)	(371.5)
Statewide Net Metering							-5,394**
							(2,317)
Constant	-96,310***	-96,200***	122,942***	20,130	19,106	5,824	8,801
	(9,508)	(9,524)	(17,454)	(17,983)	(18,018)	(17,856)	(17,807)
Observations	472	472	472	472	472	472	472
R-squared	0.851	0.851	0.852	0.897	0.898	0.902	0.904
Adjusted R-squared	0.830	0.830	0.831	0.883	0.883	0.888	0.889
F-Stat	42.19	41.37	40.94	61.08	60.06	62.17	61.91

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Std. Errors in parentheses.

Are There Significant Differences in the Number of Crimes Committed During Days of Extreme Heat?

Grace Generous

ECON 381: Introduction to Econometrics (Advisor: Amy Damon)

Climate change is an increasingly severe problem with extensive and deeply consequential impacts for society and the economy. As this issue progresses, the frequency and magnitude of extreme weather events and natural disasters such as heat waves, forest fires, and hurricanes will only continue to increase. The logical question that follows, then, is how do we expect people to react to these impacts that are fundamentally changing our environment and our lives?

Extreme weather events are a concern because their societal impacts have serious implications for social and political well-being. The relationship between extreme weather events and social unrest have been analyzed through the lens of different historical conflicts. A litany of historical conflicts have been, at least in part, attributed to climate shocks. For example, Yeeles (2015) observed a significant relationship between extreme heat and urban social disturbances in their research on Africa and Asia. Bai & Kung (2011) observed similar patterns when studying Sino-Nomadic conflict, finding decreased rainfall and severe drought were associated with a greater frequency in aggressive encroaches. Extreme drought has also been linked to food insecurity and increased social unrest in England (De Juan & Wegenast, 2020).

Concerning high temperatures are just one of many extreme weather events that will continue to become more frequent in the near future. The United States' Environmental Protection Agency (EPA) predicts that climate change will increase global temperatures and the incidence of heat waves. This increase in heat wave frequency is already happening, for example, the top ten warmest years on record have occurred in the last twenty years (EPA, 2022).

Hotter temperatures have macro-level effects on global economies with serious implications for wealth accumulation. A negative relationship is observed between weighted average temperatures in a country and GDP per capita (Heal & Park, 2013). In a study looking at temperature deviations away from biological optimums and country-level wealth,

Grace Generous is a senior with majors in Economics and Environmental Studies. Correspondence concerning this article should be addressed to ggenerou@macalester.edu.

Heal & Park (2013) found that excessive heat negatively impacts labor productivity. They argued that this relationship can partially explain historical and contemporary disparities in countries' economic outcomes. In fact, Heal & Park (2013)) argue that different temperatures may help predict economic development potential in a country. Better understanding the effects of temperature on labor productivity would improve the accuracy of cost-benefit analyses performed when policymakers weigh whether to invest in climate change mitigation strategies. Overall, these findings suggest that increased frequency in heat waves, as a result of climate change, will cause a decline in labor productivity (Heal & Park, 2013).

In addition to its effect on labor productivity, increases in heat wave frequency are linked to increases in aggression and impulsivity. Using Los Angeles as a case study, researchers have observed a statistically significant seasonal pattern in aggressive and impulsive behavior that correlates to hotter times of the year Simister & Cooper (2005). An increase in violent behavior is observed not only during hotter times of the year, but also in hotter geographical regions (Anderson, 2001). These researchers presented two different theories to explain this relationship between heat and aggression, otherwise known as the "heat hypothesis" or "thermal stress." The first theory is that adrenaline is released as part of the body's stress response to extreme temperatures, increasing aggression and the second is that experiencing extreme heat can increase hostility and thus aggressive thoughts and behavior (Simister & Cooper, 2005; Anderson, 2001). Another observational study researching the relationship between weather and aggressive crime in Cleveland, Ohio also found evidence of the "heat hypothesis." Researchers found that the highest rates of aggressive crime occurred during the summer, while the lowest rates of aggressive crime occurred during the winter (Butke & Sheridan, 2010).

Increased aggression as a result of heat exposure has also been observed in controlled experiments. In an experiment run at the University of California, Berkeley and the University of Nairobi, researchers found that heat had a statistically significant effect on destructive behavior (Almås et al., 2019). Researchers asked participants to play a series of games where they could destroy their opponents' assets. They observed that players were more likely to engage in destructive behavior when they played in hotter rooms (Almås et al., 2019).

Finally, there is a positive relationship between extreme heat and crime. A study conducted in India found that high daily temperatures were associated with an increase in violent crimes, with a slightly higher increase in property crimes as well as "unnatural deaths" (Blakeslee et al., 2021). Identity-based crimes, crimes against women, and inter-group conflict also increased in frequency during hotter days (Blakeslee et al., 2021). Another study on the heat hypothesis in LA found that, on average, crime rates increased by 2.2% and violent crime rates increased by 5.7% on days when temperatures were above

85°F (Heilmann & Kahn, 2019). Of course, crime is not driven solely by temperature, and can be influenced by personal motivations such as pride, revenge, and anger. Research pushing back on the heat hypothesis ran a similar set of quasi-experiments while controlling for thermal climate and found that intentional homicide is driven more by income disparities than temperature (Coccia, 2017)]. These results suggest that further analysis is required to move beyond broad understandings of seasonal temperature changes and aggression and more towards understanding how daily, unexpected, and extreme temperatures affect crime rates to get a clearer picture of predicted behavior as climate change worsens.

This paper expands on previous studies by analyzing temperature deviations from monthly averages, rather than seasonal or geographic temperature trends. I focus on temperature deviations from monthly averages to see how extreme, unexpected, and uncommon temperatures affect people's behavior regarding crime. I use Seattle as a case study for two reasons. First, Seattle experienced the worst heat wave in its history in June, 2021. I believe these extreme cases of heat are important to research, to better understand how populations can prepare and fortify themselves against extreme heat from climate change. Second, Seattle is an interesting case study because most homes do not have air conditioning since the city has had a historically temperate climate. This lack of air conditioning might contribute to a clearer relationship between extreme heat and violence compared to a city like LA, where people anticipate and are prepared for extreme heat. Thus, using Seattle as a case study, this paper attempts to answer the question: are there significant differences in the number of crimes committed during heat waves?

Economic Theory

Economic crime theory can help explain why this question is important to study for crime cost calculations. McCollister et al. (2010) offers a comprehensive framework for calculating crime costs that includes both tangible and intangible costs. Examples of tangible costs include the loss of productive capacity and income for the victim, as well as hospital bills if the victim is injured. Intangible costs include mental health and well-being costs to the victim. Societal costs include criminal justice system costs, from dealing with the police, courts, and incarceration system, as well as the intangible costs of lost productivity and fear. These costs are used to calculate the desirability of passing new laws that target criminal behavior. If extreme heat affects people's decision to commit a crime, this finding is important to understand so economists can adjust their cost-benefit analysis models to better advise politicians on what criminal justice policies most effectively reduce people's incentive to commit crime.

Equally as important, there are costs associated with climate change that are an essential part of policymaker's decisions to invest in climate mitigation policy. If a clear link between hotter temperatures and crime is established, the costs associated with this increase in crime would have to be incorporated into the cost benefit analysis associated with a myriad of policies that would help address climate change.

Data Description

I use crime data from the Seattle Police Department (SPD) which contains every crime ever handled by the SPD since 2008. These data were obtained from the City of Seattle Open Data page. I combined these data with 2011-2021 daily temperature data for Seattle, collected from Weather Underground, a commercial weather service website. Because I was focusing specifically on the ten-year period between 2011-2021, I dropped crime data outside of this time range. I calculated daily temperature deviations by finding the average temperature for each month, then subtracting daily temperatures from its respective monthly average.

The main outcome variable in this study is the number of daily crimes. The main independent variable of interest is daily temperature deviation.

Table 1

Descriptive Statistics

Variable	Obs.	Mean	S.D.	Min	Max
Avg. Temp. Dev.	754,882	0	5.266	-29.73	29.701
Daily Crimes	754,882	194.83	49.262	65	764

Notes: : Descriptive statistics of temperature deviation from a calculated monthly average temperature and the number of crimes recorded by the Seattle Police department per day.

Table 1 summarizes the descriptive statistics for both average temperature deviation and number of crimes per day. Average temperature deviations range from about 29 degrees below average to 29 degrees above average. The mean of average temperature deviation is close to zero indicating that, on average, daily temperatures deviate very little from monthly temperature averages. Number of daily crimes range from as low as 65 crimes per day to as high as 764 crimes per day. There is an average of about 194 crimes per day over this ten-year period.

Figure 1 and Figure 2 visualize the frequency of average temperature deviations and number of total crimes. Average temperature deviations are mostly normally distributed, while the number of daily crimes skews significantly to the right. This skew indicates that there are many more outliers of days with higher-than-average total crimes.

Figure 3 displays the relationship between daily temperature deviation and daily number of crimes. There is a clear, positive relationship between the two. There is more variation in the number of crimes during days when temperature deviations are close to

Figure 1

Frequency count of daily temperature deviation away from average monthly temperature

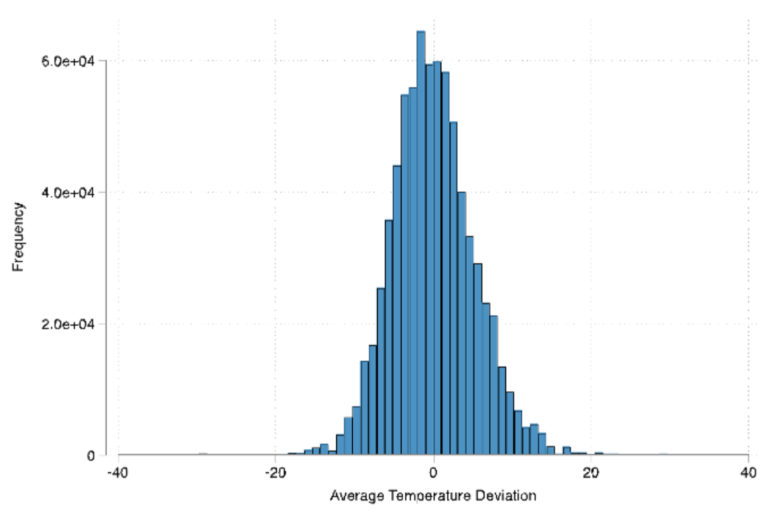


Figure 2

Frequency count of total daily crimes

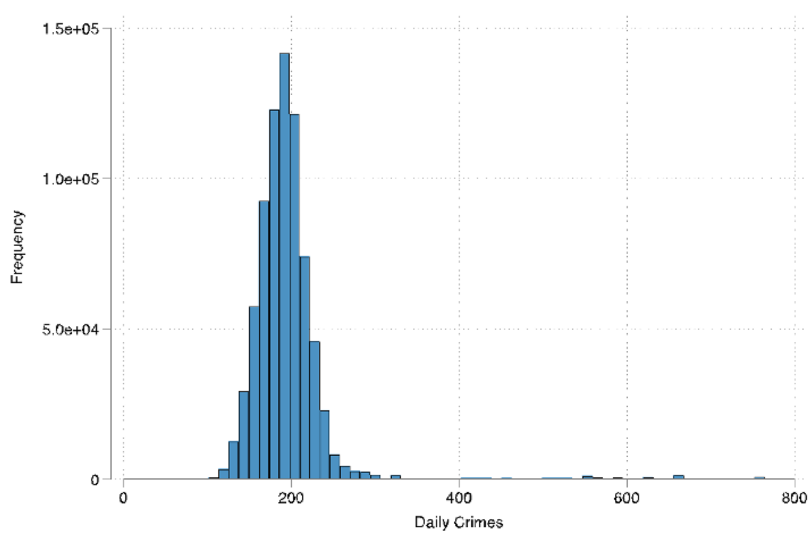
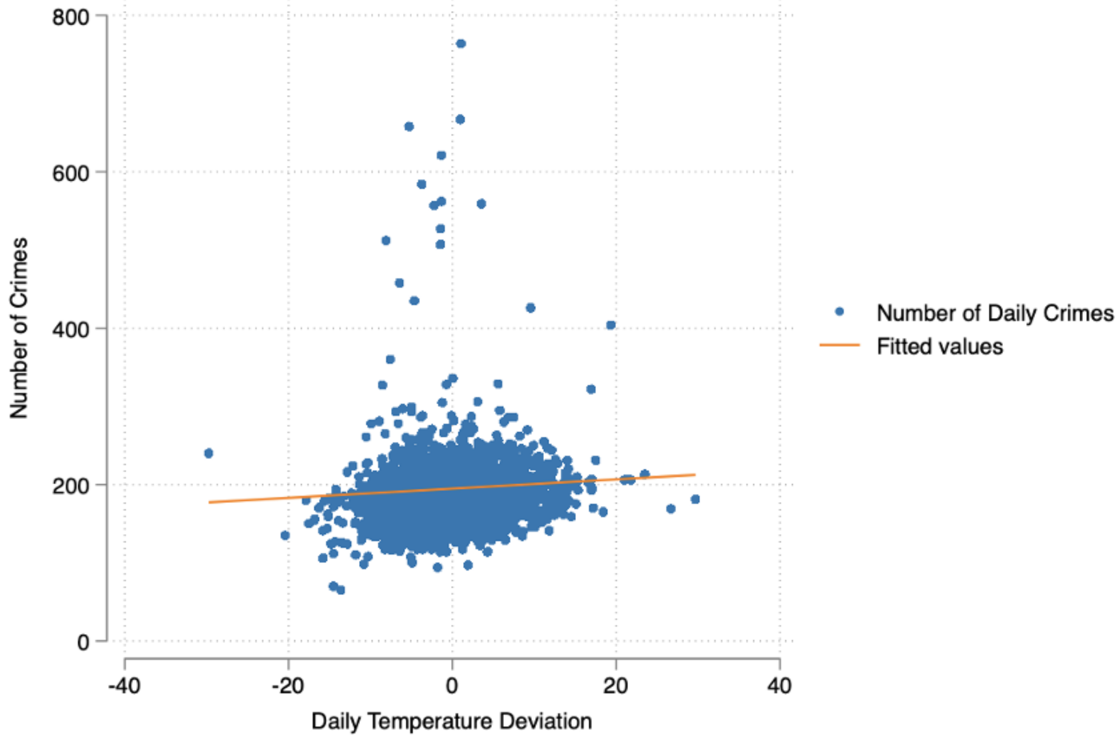


Figure 3

Daily average temperature deviation compared to number of total crimes



zero and, also shown by the previous histograms, the main outliers are days where crime rates are extremely high. This visualization, however, is a simple linear visualization of the relationship between temperature deviation and number of crimes and a more in-depth exploration of this relationship will be conducted in the remainder of this paper.

Empirical Analysis

To assess the effect of extreme heat on crime rates, I use several linear regression models that focus on the effect of daily temperature deviation on the number of crimes per day. The simplest regression, without any controls, is the following:

$$Crime_t = \beta_0 + \beta_1 TempDev_t + \varepsilon_t \tag{1}$$

where $Crime_t$ is the number of crimes committed on day t while $TempDev_t$ is the temperature deviation in degrees Fahrenheit. β_0 describes the predicted number of crimes committed if the temperature deviation is zero. β_1 describes the additional number of crimes the model predicts will happen on top of β_0 with every additional degree that temperature deviates from the monthly average. Finally, ε_t describes the difference between actual number of crimes associated with a particular temperature deviation and predicted number of crimes.

Previous research suggests some controls should be added to the regression. Almås et al. (2019) identify gender and location as two important control variables in their experimental research on heat and aggressive behavior. Coccia (2017) research indicates income should be included. Crime data collected from the SPD does not include identifiable characteristics, so there was no data on the gender or income of the individual committing the crime. The data does, however, include precinct boundaries, identifying within which designated police precinct the crime occurred. The Seattle Police Department has five precincts, North (N), East (E), South (S), West (W), and Southwest (SW). Additionally, the label $UNKNOWN$ is given to crimes that are committed in unknown precincts.[1] I added precincts to my regression to get the following:

$$DailyAssaults_t = \beta_0 + \beta_1 TempDev_t + \beta_2 Precinct_t + \varepsilon_t \tag{2}$$

In this regression, $Precinct_t$ describes the precinct where the crime occurred on day t. β_2 describes the predicted number of additional crimes committed in precincts compared to the Eastern precinct, E, holding all other variables constant.

Simister & Cooper (2005) found a seasonal pattern in aggressive behavior, where aggressive behavior increased in hotter months. This observed relationship led me to include month as a control variable to get the following regression:

$$Crime_t = \beta_0 + \beta_1 TempDev_t + \beta_2 Precinct_t + \beta_3 Month_t + \varepsilon_t \tag{3}$$

where $Month_t$ specifies the month in which the crime on day t occurred and β_3 is the predicted additional number of crimes committed compared to January, holding all other variables constant. Adding month as a fixed effect allowed me to look at how increases in temperature deviations predicted crimes per day, accounting for seasonal differences.

Moreover, Butke & Sheridan (2010) research suggests that weekdays should be added as a control because of the way people behave differently on weekends (more free time, drinking, drug use, etc.). Adding this control accounts for this "partying" effect that occurs on the weekend, where inebriated people are more likely to act impulsively and thus commit a crime. Butke & Sheridan (2010) include Friday as a "weekend," so I did the same, creating a binary variable where days were categorized into weekdays and weekends. This additional

[1] Precinct specification also included the categories $Null$ and OOJ (standing for "out of jurisdiction"). There was only one observation for each of these categories, so I dropped both.

control variable produced the following regression:

$$Crime_t = \beta_0 + \beta_1 TempDev_t + \beta_2 Precinct_t + \beta_3 Month_t + \beta_4 Weekday_t + \varepsilon_t \tag{4}$$

where $Weekday_t$ is a dummy variable that equals "1" if the crime committed on day "t" occurred on a weekday and "0" if the crime occured on a weekend. β_4 is the predicted additional number of daily crimes committed on a weekday, compared to a weekend, holding all other variables constant.

Finally, because of the seasonal pattern observed in Simister & Cooper (2005) research and Heilmann & Kahn (2019)'s research, I ran one final linear regression that included all previous controls, but restricted months to June, July, and August within the ten year period of interest. This choice was also influenced by Heal & Park (2013)'s research that found people's productivity peaks within a certain bandwidth of temperatures and drops off on either extreme of this temperature range. In other words, temperatures that are twenty degrees higher than average in January are not the same as in July, because they are unlikely to be out of this "optimal temperature range" referenced in Heal & Park (2013)'s research. Restricting the regression to the summer allowed for a more specific understanding of crime and heat waves. This regression looks identical to the previous regression, but months only include June, July, and August:

$$Crime_t = \beta_0 + \beta_1 TempDev_t + \beta_2 Precinct_t + \beta_3 Month_t + \beta_4 Weekday_t + \varepsilon_t \tag{5}$$

I also ran a regression with a quadratic term. This choice was influenced both by Simister & Cooper (2005)'s inclusion of a quadratic term in their research and the observation from Figure 3 that the number of daily crimes begins to drop off at the very hottest temperature deviations. Heal & Park (2013) found that these extreme temperatures can result in irritable and sluggish behavior. To account for this phenomenon, I included a quadratic term to my model. This change produced the following regression:

$$\begin{aligned} Crime_t = \beta_0 + \beta_1 TempDev_t + \beta_1 TempDev_t^2 + \beta_3 Precinct_t \\ + \beta_4 Month_t + \beta_5 Weekday_t + \varepsilon_t \end{aligned} \tag{6}$$

where $TempDev_t^2$ is the quadratic term in this regression, and β_2 describes the rate of change of β_1. If the relationship between increased temperature deviations and daily crimes eventually drops off, this coefficient is presumably negative. A negative β_2 would imply that the rate of change described by β_1 occurs at a decreasing rate. For example, an additional degree of temperature deviation when temperature deviation is already low will see a faster increase in daily crimes than when temperature deviation is already high.

In addition to testing how temperature deviations affect all rates of crime, Blakeslee et al. (2021) research on violent crime and high temperatures suggests that a more narrowed regression should be run on specific categories of crime. Blakeslee et al. (2021)'s research looked specifically at identity-based crimes and inter-group conflict whereas I chose to focus on the largest categories of crime. I included crime categories that comprise at least 2% of all crimes over the ten year period of interest (2011-2021). These crime categories include assault, burglary and breaking & entering, destruction of property, drug & narcotic offenses, larceny & theft, motor vehicle theft, robbery, and trespassing. I ran a similar linear regression as before, with all controls outlined previously, for example, the regression for assault crimes looked like the following:

$$Assaults_t = \beta_0 + \beta_1 TempDev_t + \beta_2 Precinct_t + \beta_3 Month_t + \beta_4 Weekday_t + \varepsilon_t \quad (7)$$

Results

The first set of regressions that focused on adding controls, restricting the regression to summer months, and accounting for the drop off of crimes at higher temperatures produced the following results:

Table 2

	(1)	(2)	(3)	(4)	(5)	(6)
Temp. Dev.	.591***	.592***	.592***	.59***	.53***	.637***
	(.01)	(.01)	(.009)	(.009)	(.01)	(.01)
sqdev						-.031***
						(.001)
Constant	195***	195***	188***	191***	191***	192***
	(.057)	(.139)	(.181)	(.19)	(.175)	(.191)
Observations	754,882	754,876	754,876	754,876	191,333	754,876
R-squared	.004	.005	.074	.077	.024	.078

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Std. Errors in parentheses. Linear regression models with (1) no controls (2) adding precinct as a control (3) adding precinct and month as controls (4) adding precinct, month, and weekday as controls, (5) adding all controls, but restricting time period to June, July, and August.

The primary regression of interest is regression (4), a linear regression of daily crimes against temperature deviations with precinct, month, and weekday as a control. The p-values across all temperature deviation coefficients indicate they are all statistically significant. For the first four models that did not include a quadratic term, the predicted increase in crime ranges from 0.53 to 0.59. These results support the hypothesis that hotter than normal temperatures are associated with more crimes per day.

The final regression (6), with a quadratic term, also displays a statistically significant positive relationship between crimes per day and temperature deviation. There is a slightly higher temperature deviation coefficient of 0.637 which is expected since this model accounts for the dampening effect that occurs as temperature deviations continue to rise. The negative sign associated with the quadratic term, -0.031, captures this effect, indicating that the number of daily crimes increases at a decreasing rate. This result is consistent with previous findings, that humans have trouble functioning outside of a certain bandwidth of temperatures Heal & Park (2013). This coefficient, however, is extremely small, indicating that the number of crimes per day will increase for quite a while as temperature deviation increases, before eventually starting to decrease. Simister & Cooper (2005)'s research also used a quadratic term. The authors found 0.001 for their temperature coefficient and -0.000003 for their temperature squared coefficient. Because this research studied temperatures rather than temperature deviations and used degrees Celsius rather than degrees Fahrenheit, it is challenging to compare the results. These results are similar to the ones in this study, however, in that the magnitude of the primary coefficient far outweighs the magnitude of the quadratic coefficient.

The consistency across all six regressions, both in terms of statistical significance and magnitude of the temperature deviation coefficient, expresses how robust these results are. These results support the hypothesis that higher temperature deviations will result in more crimes per day.

The full regression output with all controls is included in Table A2 of the appendix. As expected, warmer months predict higher crime rates compared to January, holding all else equal, and weekdays predict lower crime rates than weekends, holding all else equal. Interestingly, May had an extremely high, statistically significant coefficient of about 45, implying that, all else held equal, an equivalent temperature deviation in May compared to January predicts around forty five more crimes per day. This result may, in part, be influenced by the George Floyd riots that occurred at the end of May, 2020, when cities around the country, including Seattle, were arresting a staggeringly high number of people Kornfield et al. (2020). Without further knowledge of Seattle's culture, specific to certain areas, not much can be discerned from the precinct coefficients.

The crime-category regression results are displayed below. A table with all outputs, including controls, is included in Table A3 of the appendix.

For every category of crime except drug & narcotic offenses, there is a statistically significant increase in crime associated with an increase in temperature deviation. This increase is largest for assault crimes and larceny & theft crimes, while robbery, motor vehicle theft, trespassing, burglary, and destruction of property had smaller increases for every degree increase in temperature deviation. The lack of statistical significance for drug & narcotic offenses makes sense because drug use is largely influenced by addiction which would

Table 3

	Assaults	Burglary	Destruction	Drug and Narcotic Offenses	Larceny and Theft	Motor Vehicle Theft	Robbery	Trespassing
Avg. Temp. Dev.	.172***	.047***	.073***	.001	.275***	.029***	.019***	.052***
	(.002)	(.002)	(.001)	(.001)	(.003)	(.001)	(0)	(.001)
Constant	27.733***	21.874***	17.382***	4.636***	71.427***	11.554***	4.347***	6.95***
	(.037)	(.032)	(.025)	(.016)	(.063)	(.021)	(.01)	(.024)
Observations	754,876	754,876	754,876	754,876	754,876	754,876	754,876	754,876
R-squared	.109	.032	.071	.038	.065	.053	.021	.012

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Std. Errors in parentheses. Regression output for crime categories, including (1) assault crimes (2) burglary and breaking & entering (3) destruction of property (4) drug and narcotic offenses (5) larceny & theft (6) motor vehicle theft (7) robbery and (8) trespassing.

likely dominate any external environmental factors, like temperature. The consistent statistical significance of the temperature deviation coefficient across crime categories further emphasizes how robust this link between increased temperature deviations and increased crime is.

These results align with findings in Heilmann & Kahn (2019)'s research. Heilmann & Kahn (2019) found that on average, crime rates increased by 2.2% and violent crime increased by 5.7% on days when temperatures exceeded 85°F. Similarly, violent crime like assaults and destruction of property increased more than any other category of crime except larceny & theft, in response to increases in temperature deviations.

Figure 4

Residual plot of daily crime rates versus temperature deviation with precinct, month, and weekday as controls.

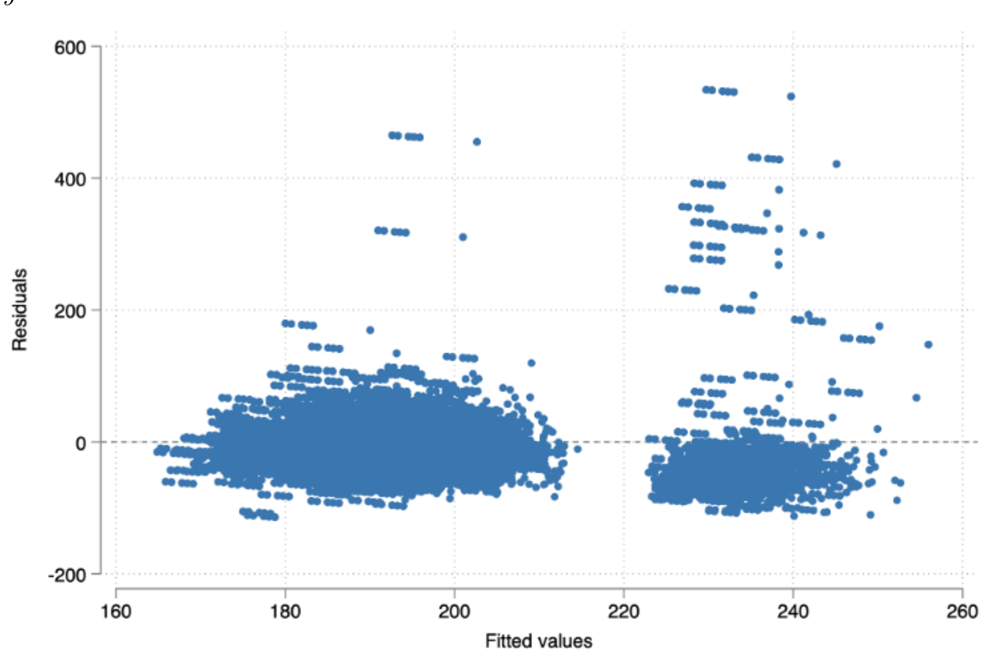


Figure 4 shows there is heteroskedasticity present, which was expected given the outliers observed in Figure 3. To account for this heteroskedasticity, I used robust standard errors.

In terms of economic significance, the first set of results indicate there is about a "half crime" increase in crimes per day for every degree increase in temperature deviation in degrees Fahrenheit. For the quadratic model, this increase in crimes per day is slightly higher at first, but decreases as temperature deviations get more severe. While "half a crime" may seem small, temperature deviations in Seattle have ranged as high as thirty degrees and heat waves are predicted to only get worse in magnitude over time. At a temperature deviation of thirty degrees, the linear regression model predicts about fifteen additional crimes per day, holding all else equal. This concerning increase in the number of daily crimes committed during heat waves highlights another reason why investing in climate change mitigation policies is imperative.

Limitations

This research has established a statistically significant and economically significant relationship between extreme heat and increased rates of crime, however there are several limitations to consider when evaluating the validity of this research.

The first concern is omitted variable bias. There are many unobservable reasons why one might commit a crime such as mental health, support from friends and family, and more. Observable reasons might include gender, race, education level, and income, which were not reported in SPD data. Income and mental health are particularly important to consider when looking at Seattle, which has a notoriously high population of people using illicit drugs and one of the highest homeless populations in the United States. Parsing out results between different categories of crime helps clear through some of this murkiness. For example, larceny & theft crimes increase more than assault crimes. Larceny & theft crimes might be driven more by financial necessity than assault crimes. However, assault crimes increase more than motor vehicle theft or robbery, which are also likely to have financial incentives. A potential area for further research might include these observable characteristics and some proxy for unobservable characteristics.

Additionally, as Butke & Sheridan (2010) point out in their research, some types of crime, especially crimes against persons, are driven by incentives outside of temperature. Revenge, financial opportunity, pride, are just some of the countless alternative reasons why someone might commit a crime. Considering excessive heat and thermal stress is a subconscious driver of crime, these alternative crime incentives make it challenging to identify the mechanism driving this relationship. However, psychology research cited by other researchers looking at crime rates and heat indicate thermal stress is likely the mechanism.

Simister & Cooper (2005) also indicate that humidity should be included in this analysis, because humid heat may cause a stronger thermal stress response than dry heat. I chose not to study humidity, rather temperature deviations because of the significance of heat waves in the Pacific Northwest. A valid area for further research might include humidity as a control variable or even a main variable of interest.

Additionally, there are data limitations to this study since only crimes within the past ten years were studied. Seattle has undergone significant changes to its economy and society in the last ten years as the introduction of companies like Microsoft and Amazon have changed the landscape of Seattle communities. Other factors like increased housing costs, homelessness, gentrification, and more have also changed dramatically in the past ten years and could influence these results. Further research might include more years of data, stretching back decades, to increase result validity and help account for these changes over time.

Finally, there are limitations to the external validity of this study. The data in this study comes exclusively from Seattle which has unique characteristics, particularly regarding homelessness, drug addiction, and gentrification, which means the results cannot be generalized to other contexts. Rather, this analysis contributes to the wide range of literature that has found similar results in controlled and quasi-experimental studies around the world.

Conclusion

If the world continues to emit fossil fuels like normal, the Earth's surface temperature is expected to increase by 5-10.2°F compared to 1901-1960 averages by the end of the century Lindsey et al. (2023). Climate scientists have uncovered a litany of negative impacts associated with this warming on ecosystems and societies around the world, one of which is the increased incidence of heat waves. This paper found that increases in extreme heat in Seattle predicts higher rates of crime per day. These results build on research from previous studies that found links between increased heat and increased violence, impulsive behavior, and crime, as well as decreases in productivity. In addition to overall crime, this relationship was observed within specific categories of crime, particularly assault crimes and larceny & theft crimes.

The robustness of these results indicate that extreme heat is a primary driver of daily crime rates, at least in Seattle, and that is important to account for in climate models. I have studied how other factors that could explain differences in crime rates might affect this model and the inclusion of all of these factors has still produced similar results.

This paper builds on previous research of thermal stress, underscoring the importance of investing in comprehensive, immediate climate change mitigation strategies. Further research should be conducted to get a better estimate of the costs associated with this increase in crime, so it may be incorporated into cost-benefit models that policymakers use to determine whether to invest in green policies. Ultimately, this observed increase in daily crimes in response to hotter temperatures is just one of the many impacts that climate change will have on our lives. We must demand stronger policy action against climate change to protect our communities, our environment, and our world.

References

Almås, I., Auffhammer, M., Bold, T., Bolliger, I., Dembo, A., Hsiang, S. M., Kitamura, S., Miguel, E., & Pickmans, R. (2019). Destructive Behavior, Judgment, and Economic Decision-making Under Thermal Stress. *NBER Working Paper*, 25785.

Anderson, C. A. (2001). Heat and Violence. *Current Directions in Psychological Science*, 10(1), 33–38.

Bai, Y. & Kung, J. K.-s. (2011). Climate Shocks and the Sino-nomadic Conflict. *The Review of Economics and Statistics*, 93(3), 970–981.

Blakeslee, D., Chaurey, R., Fishman, R., Malghan, D., & Malik, S. (2021). In the heat of the moment: Economic and non-economic drivers of the weather-crime relationship. *Journal of Economic Behavior & Organization*, 192, 832–856.

Butke, P. & Sheridan, S. (2010). An Analysis of the Relationship between Weather and Aggressive Crime in Cleveland, Ohio. *Weather, Climate, and Society*, 2(2), 127–139.

Coccia, M. (2017). A Theory of general causes of violent crime: Homicides, income inequality and deficiencies of the heat hypothesis and of the model of CLASH. *Aggression and Violent Behavior*, 37, 190–200.

De Juan, A. & Wegenast, T. (2020). Temperatures, food riots, and adaptation: A long-term historical analysis of England. *Journal of Peace Research*, 57(2), 265–280.

EPA (2022). *Climate Change Indicators: Weather and Climate.* `https://www.epa.gov/climate-indicators/weather-climate`

Heal, G. & Park, J. (2013). Feeling the Heat: Temperature, Physiology & the Wealth of Nations. *NBER Working Paper*, 19725.

Heilmann, K. & Kahn, M. E. (2019). The Urban Crime and Heat Gradient in High and Low Poverty Areas. *NBER Working Paper*, 25961.

Kornfield, M., Ramsey, A. R., Wallace, J., Casey, C., & Valle, V. D. (2020). Data show George Floyd protesters mostly local, faced minor charges. *The Seattle Times.* `https://www.seattletimes.com/nation-world/data-show-george-floyd-protesters-mostly-local-faced-minor-charges/`

Lindsey, R., Dahlman, L., & Blunden, J. (2023). *Climate Change: Global Temperature.* `https://www.climate.gov/news-features/understanding-climate/climate-change-global-temperature`

McCollister, K. E., French, M. T., & Fang, H. (2010). The cost of crime to society: new crime-specific estimates for policy and program evaluation. *Drug and alcohol dependence*, 108(1/2), 98–109. `https://doi.org/https://doi.org/10.1016/j.drugalcdep.2009.12.002`

Simister, J. & Cooper, C. (2005). Thermal stress in the U.S.A.: effects on violence and on employee behaviour. *Thermal stress in the U.S.A.: effects on violence and on employee behaviour*, 16(1), 3–15.

Yeeles, A. (2015). Weathering unrest: The ecology of urban social disturbances in Africa and Asia. *Journal of Peace Research*, 52(2), 158–170.

Appendix
Additional Tables

Table 0.1

Tabulation of "Offense Parent Group"

Offense Parent Group	Freq.	Percent	Cum.
ANIMAL CRUELTY	81	0.01	0.01
ARSON	1,273	0.17	0.18
ASSAULT OFFENSES	108,340	14.35	14.53
BAD CHECKS	4,148	0.55	15.08
BRIBERY	11	0.00	15.08
BURGLARY/BREAKING&ENTERING	86,624	11.48	26.56
COUNTERFEITING/FORGERY	4,030	0.53	27.09
CURFEW/LOITERING/VAGRANCY VIOLATIONS	545	0.07	27.16
DESTRUCTION/DAMAGE/VANDALISM OF PROPERTY	68,176	9.03	36.19
DRIVING UNDER THE INFLUENCE	13,257	1.76	37.95
DRUG/NARCOTIC OFFENSES	17,986	2.38	40.33
DRUNKENNESS	12	0.00	40.34
EMBEZZLEMENT	1,345	0.18	40.51
EXTORTION/BLACKMAIL	670	0.09	40.60
FAMILY OFFENSES, NONVIOLENT	7,473	0.99	41.59
FRAUD OFFENSES	45,936	6.09	47.68
GAMBLING OFFENSES	13	0.00	47.68
HOMICIDE OFFENSES	348	0.05	47.73
HUMAN TRAFFICKING	34	0.00	47.73
KIDNAPPING/ABDUCTION	898	0.12	47.85
LARCENY-THEFT	282,395	37.41	85.26
LIQUOR LAW VIOLATIONS	1,147	0.15	85.41
MOTOR VEHICLE THEFT	46,989	6.22	91.63
PEEPING TOM	148	0.02	91.65
PORNOGRAPHY/OBSCENE MATERIAL	276	0.04	91.69
PROSTITUTION OFFENSES	2,817	0.37	92.06
ROBBERY	17,181	2.28	94.34
SEX OFFENSES	4,757	0.63	94.97
SEX OFFENSES, CONSENSUAL	112	0.01	94.98
STOLEN PROPERTY OFFENSES	4,703	0.62	95.61
TRESPASS OF REAL PROPERTY	26,188	3.47	99.08
WEAPON LAW VIOLATIONS	6,969	0.92	100.00

Notes: : Tabulation of "Offense Parent Group" of all crimes in the dataset from 2011-2021 with frequency, percentage, and cumulative percentages.

Table 0.2

All Crime

	(1)	(2)	(3)	(4)	(5)	(6)
Temp. Dev.	.591***	.592***	.592***	.59***	.53***	.637***
	(.01)	(.01)	(.009)	(.009)	(.01)	(.01)
N		1.332***	1.249***	1.318***	.34**	1.328***
		(.176)	(.169)	(.169)	(.169)	(.169)
S		-2.032***	-2.018***	-1.927***	-.402**	-1.922***
		(.196)	(.19)	(.19)	(.203)	(.19)
SW		.468**	.514**	.616***	-.41*	.62***
		(.238)	(.228)	(.228)	(.223)	(.228)
UNKNOWN		9.259***	8.014***	8.079***	2.204***	8.123***
		(.98)	(.922)	(.922)	(.706)	(.921)
W		-1.211***	-1.224***	-1.249***	.482***	-1.242***
		(.171)	(.165)	(.165)	(.174)	(.165)
Feb			-8.308***	-8.363***		-8.338***
			(.172)	(.172)		(.171)
Mar			6.854***	6.763***		6.574***
			(.291)	(.289)		(.29)
Apr			.975***	.937***		.839***
			(.177)	(.176)		(.176)
May			45.55***	45.531***		45.58***
			(.468)	(.469)		(.469)
Jun			1.821***	1.798***		2.085***
			(.167)	(.167)		(.168)
Jul			5.922***	5.892***	4.089***	5.753***
			(.162)	(.162)	(.139)	(.162)
Aug			5.684***	5.692***	3.869***	5.521***
			(.159)	(.16)	(.137)	(.16)
Sep			6.588***	6.525***		6.446***
			(.164)	(.163)		(.165)
Oct			8.987***	8.972***		8.553***
			(.167)	(.167)		(.167)
Nov			3.618***	3.586***		3.597***
			(.172)	(.172)		(.172)
Dec			.602***	.576***		.611***
			(.168)	(.168)		(.168)

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Std. Errors in parentheses. Full output for (1) no controls (2) precinct controls (3) precinct and month controls (4) precinct, month, and weekday controls (5) all controls, but restricted to summer months only (6) all controls with a quadratic term.

Table 0.2 (cont.)

	(1)	(2)	(3)	(4)	(5)	(6)
Weekday				-5.432***	-3.188***	-5.437***
				(.109)	(.113)	(.109)
Sq. Dev.						-.031***
						(.001)
Constant	195***	195***	188***	191***	191***	192***
	(.057)	(.139)	(.181)	(.19)	(.175)	(.191)
Observations	754,882	754,876	754,876	754,876	191,333	754,876
R-squared	.004	.005	.074	.077	.024	.078

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Std. Errors in parentheses. Full output for (1) no controls (2) precinct controls (3) precinct and month controls (4) precinct, month, and weekday controls (5) all controls, but restricted to summer months only (6) all controls with a quadratic term.

Table 0.3

	Assaults	Burglary	Destruction	Drug and Narcotic Offenses	Larceny and Theft	Motor Vehicle Theft	Robbery	Trespassing
Avg. Temp. Dev.	.172***	.047***	.073***	.001	.275***	.029***	.019***	.052***
	(.002)	(.002)	(.001)	(.001)	(.003)	(.001)	(0)	(.001)
N	-.103***	.106***	-.021	-.015**	-.047	.016	-.012	.06***
	(.023)	(.023)	(.018)	(.007)	(.046)	(.016)	(.008)	(.016)
S	-.129***	-.11***	-.05**	.008	-.329***	.007	-.001	-.166***
	(.028)	(.028)	(.022)	(.009)	(.055)	(.019)	(.009)	(.019)
SW	-.086***	-.003	-.024	-.001	-.341***	-.043**	-.011	-.002
	(.03)	(.031)	(.024)	(.009)	(.061)	(.02)	(.01)	(.021)
UNKNOWN	.47***	.205**	.489***	.061**	.015	-.099	.102***	.3***
	(.099)	(.097)	(.078)	(.03)	(.199)	(.066)	(.032)	(.066)
W	-.015	-.232***	-.075***	.023***	.109**	-.125***	.006	.108***
	(.024)	(.024)	(.019)	(.007)	(.048)	(.016)	(.008)	(.016)
Feb	-.787***	-1.902***	-.1***	.136***	-2.993***	-.49***	-.422***	-.098***
	(.042)	(.037)	(.029)	(.012)	(.073)	(.024)	(.012)	(.029)
Mar	-.349***	-1.136***	.145***	-.214***	-2.89***	-.639***	-.245***	-.325***
	(.04)	(.035)	(.028)	(.012)	(.071)	(.022)	(.012)	(.027)
Apr	-.141***	-.348***	.345***	-.137***	-.867***	-.131***	-.209***	.124***
	(.04)	(.038)	(.027)	(.011)	(.072)	(.023)	(.012)	(.028)
May	2.569***	.81***	1.723***	-.25***	-.235***	-.516***	-.129***	-.371***
	(.04)	(.036)	(.028)	(.012)	(.071)	(.023)	(.012)	(.026)
Jun	2.323***	-.997***	1.783***	-.34***	.296***	1.048***	.136***	-.845***
	(.041)	(.034)	(.027)	(.012)	(.075)	(.026)	(.012)	(.026)
Jul	4.103***	-1.349***	2.779***	-.509***	1.897***	1.07***	.086***	-.651***
	(.043)	(.034)	(.029)	(.012)	(.07)	(.023)	(.012)	(.026)

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Std. Errors in parentheses. Regression output for crime categories, including (1) assault crimes (2) burglary and breaking & entering (3) destruction of property (4) drug and narcotic offenses (5) larceny & theft (6) motor vehicle theft (7) robbery and (8) trespassing.

Table 0.3 (cont.)

	Assaults	Burglary	Destruction	Drug and Narcotic Offenses	Larceny and Theft	Motor Vehicle Theft	Robbery	Trespassing
Aug	3.448***	-.351***	2.199***	-.473***	3.106***	1.146***	.468***	-.95***
	(.041)	(.035)	(.03)	(.011)	(.068)	(.023)	(.012)	(.025)
Sep	2.478***	.457***	1.018***	-.684***	5.124***	2.233***	.251***	-.878***
	(.04)	(.037)	(.03)	(.011)	(.07)	(.026)	(.012)	(.025)
Oct	.991***	1.494***	1.785***	-.484***	6.002***	2.28***	.497***	-.476***
	(.041)	(.038)	(.03)	(.011)	(.073)	(.026)	(.012)	(.026)
Nov	.322***	1.619***	.816***	-.451***	4.486***	1.365***	.42***	-.224***
	(.041)	(.037)	(.029)	(.012)	(.077)	(.024)	(.012)	(.027)
Dec	-.919***	1.602***	.381***	.552***	3.548***	.96***	-.06***	-.395***
	(.04)	(.037)	(.028)	(.013)	(.073)	(.025)	(.012)	(.027)
weekday	-2.562***	.284***	-2.019***	-1.85***	-2.346***	-.601***	-.156***	.459***
	(.015)	(.016)	(.012)	(.005)	(.03)	(.01)	(.005)	(.01)
	(.041)	(.035)	(.03)	(.011)	(.068)	(.023)	(.012)	(.025)
Constant	27.733***	21.874***	17.382***	4.636***	71.427***	11.554***	4.347***	6.95***
	(.037)	(.032)	(.025)	(.016)	(.063)	(.021)	(.01)	(.024)
Observations	754,876	754,876	754,876	754,876	754,876	754,876	754,876	754,876
R-squared	.109	.032	.071	.038	.065	.053	.021	.012

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Std. Errors in parentheses. Regression output for crime categories, including (1) assault crimes (2) burglary and breaking & entering (3) destruction of property (4) drug and narcotic offenses (5) larceny & theft (6) motor vehicle theft (7) robbery and (8) trespassing.

Climate and Cowboys: The Effects of Drought and Climate Change on the Cattle Cycle and Ranch Decision Making Patterns in the Southwestern United States

Mathilda Barr
ECON 239: Global Food Problems (Advisor: Amy Damon)

Does climate change come at the expense of cowboys? According to reports from Summer 2022, it might. In the September publication of the Beige Book, the Dallas Fed reported that "severe drought and higher costs have prompted significant culling of cattle herds," across the region (Federal Reserve Bank of Dallas, 2022). Lloyd Masayumptewa, a rancher on Arizona's Hopi reservation noticed the changing environment this summer and recalled the tribal government asking ranchers to reduce or eliminate their herds to accommodate the conditions (Nowell, 2022). He had previously downsized his own herd from 100 to 70 head to account for drought conditions (Nowell, 2022). Masayumptewa is running what is called a cow-calf operation. These operations breed and raise herds with the goal of selling the steers (male) to slaughter and maintaining the best heifers (female calves) to continue the breeding operation. Climate conditions are a critical determinant of cow-calf ranching success. Rangeland and forage conditions need to be suitable enough for cattle to graze, while ponds or streams need to have sufficient water levels to provide readily available drinking water to a herd. Droughts have also caused an increase in input costs of supplemental feed and grains, such as corn, oats, and barley (Nowell, 2022), which are used as additional sources of energy to fatten the physical condition (and subsequently, prices) of adult stock. This burden is passed to ranchers like Masayumptewa who are faced with making difficult financial decisions. Many ranchers look to the "cattle cycle," an inventory of the total cattle in the nation (USDA, 2022b), to determine a series of decisions that will have the least financial stress on their ranch. However, this current problem they are facing, a combination of drought and high input prices, is frightening. And its effects might be here to stay.

This article investigates existing literature on the significant impact of droughts on micro-ranching decisions and evaluates the risk that climate change poses to the profitability

Mathilda Barr is a sophomore with a major in Economics. Correspondence concerning this article should be addressed to mbarr@macalester.edu.

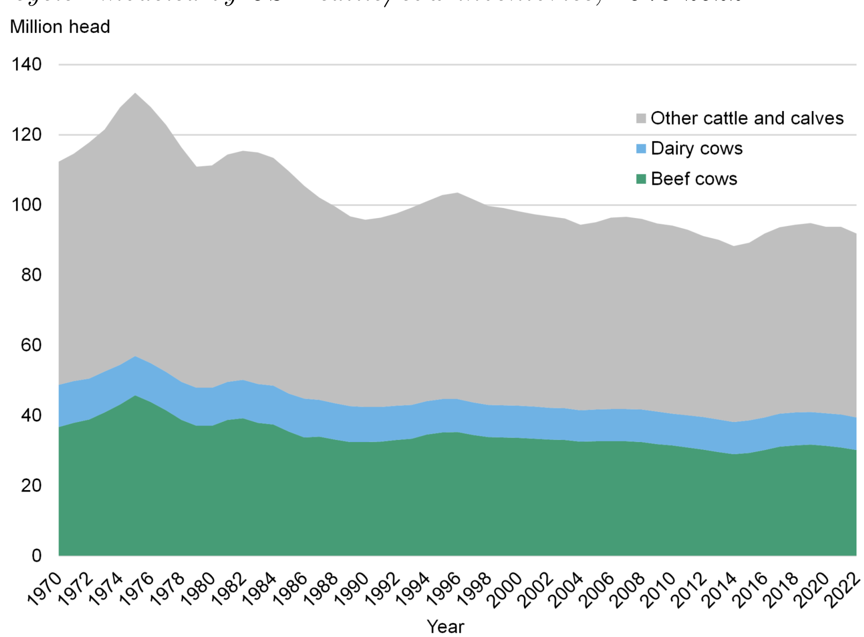
of these decisions. I will explore ways trends in the cattle cycle might be affected or altered if droughts such as the 2022 drought continue to shock the market more frequently. I will evaluate the effectiveness of drought strategies during climate change conditions on producer and consumer welfare. I focus heavily on adverse effects for ranchers because of the volatility in the cow-calf stage of the supply chain. However, the ripples from this beef production problem represent a broader trend in the ways climate change shocks have prompted conversations about sustainability in global food production.

An introduction to the cattle industry: the role of decision-making

Beef is a risky business. Fluctuations in the beef cattle market/industry in response to external shocks are modeled by a periodic time series called the "cattle cycle." This cycle, modeled in Figure 1, represents the total number of cows (size of the total herd) in the USA at one time, and its changes are measured from low-point to low-point over periods of around 8-12 years (USDA, 2022a). The current stage in the cattle cycle is affected by aggregate cost factors including market price, gestation period, and climate, which affect the revenue of ranchers (USDA, 2022a). In order to maximize profit, ranchers make decisions about when to increase head, depopulate their herd, or send stock to feedlots (USDA, 2022a). In order to succeed in this volatile industry, ranchers must be efficient and timely about manipulating the size of their herd.

Figure 1

The "cattle cycle" modeled by USA cattle/cow inventories, 1970-2022



Source: USDA (2022b)

Understanding the fundamentals of cow-calf farming is essential to understanding the decisions made by ranchers. Calves are weaned from their mothers after 6-10 months and enter the “stocking” or “backgrounding” process where they graze and fatten with the assistance of supplemental feed (Countryman et al., 2016). After this process, they are sold in auction markets and sent to finishing feedlots for slaughter (PA Beef n.d.). One of the most important decisions ranchers make is deciding which calves to keep for breeding, and which to send to slaughter. Retaining a cow for breeding depends on two factors: There must be “relative profitability” that she will produce sufficient calves and the expected return from her retainment must exceed the immediate value of selling the cow for slaughter (Rosen, 1987). When rates of heifers sent for slaughter increase, ranchers decide that the immediate cash value of the cow exceeds the potential revenue from breeding. These choices potentially lead to adverse short and long-term effects on both the rancher’s operation and the cow-calf sector as a whole. Cow-calf producers are dealing in the most volatile part of the market (Countryman et al., 2016). During the time it takes to raise, wean, and feed a calf, the cattle cycle can change. Ranchers are expected to estimate their returns based on past patterns, but the supply is much more inelastic at the rancher level, causing them to absorb a greater proportion of unexpected shocks such as drought and input prices (Countryman et al., 2016).

Drought economy & decision-making

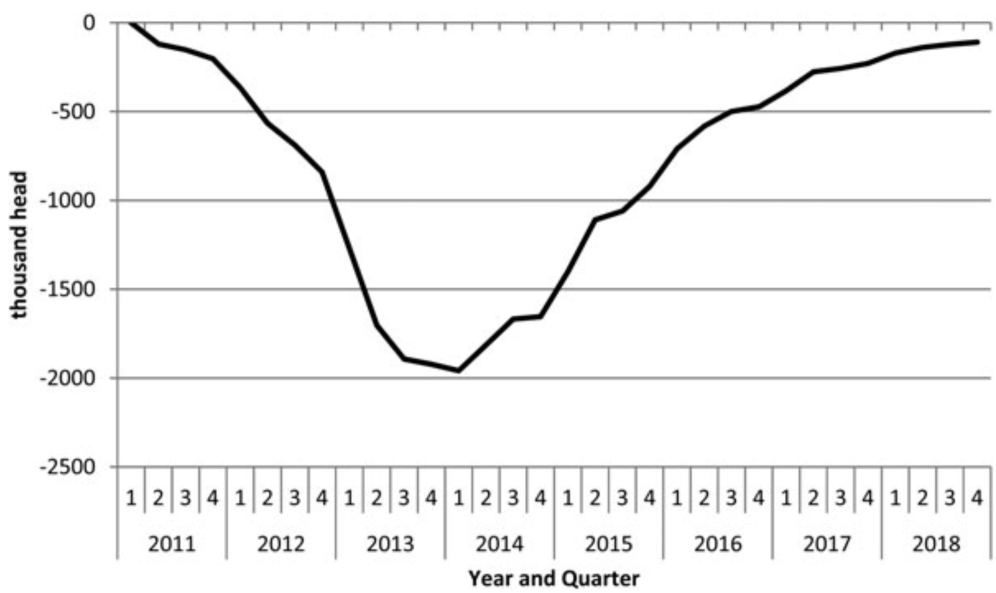
Ranchers are highly attuned to climate patterns because of the relationship between pasture conditions and profits. Cattle ranchers are price-takers and when they sell their cattle to slaughterhouse feedlots, ranch profit margins heavily depend on the size and condition of the product (Bastian, 2018). The backgrounding stage of the cow-calf process is critical because suitable forage conditions means high energy intake for a herd, which ensures that their stock will reach adequate weight for sale or breeding. Ranchers will choose an optimal amount of stock for the size of their pasture and can preserve their grazing resources by feeding supplementary grains and roughage to cattle during drier periods (Rusche, 2021).

Drought years pose a significant threat to the process of backgrounding. Deficits in regional soil moisture, streamflow, and precipitation levels cause rangeland conditions to deteriorate and reduce water availability (Bates, 2021). Ranchers are often left without land suitable enough for forage. When this condition arises, ranchers are forced to make difficult financial decisions. Typically, ranchers respond by “depopulating” their herd or selling a high volume of calves for slaughter, even those who haven’t reached substantial weight (Hughes, 2012). In the short-term, this “fire-sale” depopulation might result in immediate rancher cash income, and a glut of cheap beef on supermarket shelves Horsley (2022), however, these decisions have adverse long-term financial consequences.

A dynamic analysis of severe regional droughts in 2012 serves as a model for the long-term micro-effects of depopulation. The 2012 droughts predominantly affected the arid Central and Western regions of the USA which faced an increased lack of precipitation (Wallander et al., 2017). In response, economist Harlan Hughes published a series of articles on Drought Strategies and Drought Economics in BEEF magazine, analyzing the economic costs of droughts, and advising ranchers on optimal drought strategies. He referenced a period of earlier droughts in 2002 and 2006 which had resulted in low points in the cattle cycle (Hughes, 2012). Hughes identified three substantial costs of depopulation: selling heifers or cows at a reduced value, needing to raise or buy back breeding replacements at high prices, and the opportunity cost of having "fewer calves to sell in the years following depopulation" (Hughes, 2012). Costs from drought liquidation or depopulation, have the potential to affect ranch revenue for the next decade. Figure 2 shows the magnitudes of the changes in breeding herd inventory (measured in thousand head of cattle) during an 8-year period before and after the 2012 drought (Countryman et al., 2016). Many ranchers chose to sell their breeding stock, leading to a substantial fall in national inventory. In doing so, they accepted the costs of having to either buy back breeding stock or spend 3-5 years raising new heifers. It took the full 8 years, nearly an entire cattle cycle, before breeding herd inventories returned to near-baseline levels.

Figure 2

Changes in U.S breeding herd inventory, 2011-2018



Source: Countryman et al. (2016)

During the first quarter of a drought period, high depopulation rates flood auction markets with steer, heifer, and calf supplies. This means that ranchers will receive lower prices for their stock, and consumers will pay lower market prices due to the temporarily increased supply. Ranchers in less affected regions who did not depopulate are aware of these trends and might invest in increasing head because they expect that cattle prices will rise after an initial market flood due to drought herd liquidations (Tronstad & Feuz, 2002). In the next few quarters, as drought effects lessen, ranchers who choose to depopulate will increase the demand for breeding stock and drive prices high. Consequently, the cattle cycle turns upward and those ranches with sufficient stock will sell calves at high prices. These trends confirm Hughes' claim that ranchers who are affected by droughts will face the economic burdens of that drought years beyond initial depopulation (Tronstad & Feuz, 2002).

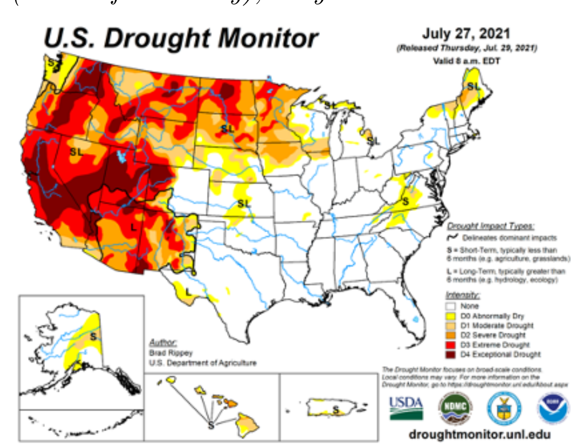
The other choice that ranchers can make in response to drought is "destocking" their herd or moving cattle from drought areas to pastures in less-affected regions with better forage (Hughes, 2012). Ranchers who can afford to destock can sell their cattle during the post-drought periods when the supply of steers is low and demand for beef and breeding stock is high. Ranchers must look to past cattle cycle trends to consider whether the profitability of selling cattle at post-drought prices is expected to exceed the cost of relocation. This is why Hughes argues that destocking is most effective during times of simultaneous drought and high calf prices when ranchers can more clearly estimate high profitability (Hughes, 2012). It is more common that cattle prices are low during drought and farmers must depopulate and accept the immediate cash value of the cow in order to make ends meet (Hughes, 2012).

In response to the 2012 droughts, the Livestock Forage Disaster Program (LFP) and the Livestock Indemnity Program (LIP) were introduced in the 2014 Agricultural Act (Countryman et al., 2016). The LFP provided compensation for ranchers who suffered grazing land losses from drought. The LIP provided compensation for mortality due to adverse weather conditions such as drought and fire. While these federal assistance programs are accessible, there is debate about whether they are effective. The LIP and LFP were leveraged again in 2021, and the LFP offered payment of up to $35 per head for a steer or heifer over 500 lbs (USDA, 2019). Based on statistics from the USDA Economic Research Service, the average price for 500-pound stock in 2021 was $669.5 (or $133 per cwt averaged over January- December 2021). This means that the program only subsidized up to 5.2 percent of the steer or heifer's value. However, in 2022, rising costs of fuel, grain, and other inputs due to drought and inflation are expected to increase the cost of producing a calf by $75 (Drovers 2022). While these safety nets provided by the government might cushion the immediate effects of a drought, they do not appear to be providing adequate assistance to support the rising climate and production costs of this industry.

The 2022 Droughts and Climate Change

Figure 3

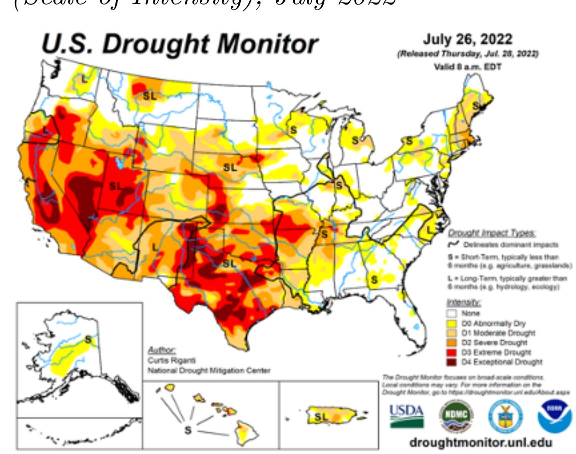
U.S. Drought Monitor (Scale of Intensity), July 2021



Source: National Drought Mitigation Center, University of Nebraska Lincoln

Figure 4

U.S. Drought Monitor (Scale of Intensity), July 2022



Source: National Drought Mitigation Center, University of Nebraska Lincoln

Long-term climate change shocks to the Southwest, America's hottest and driest region, pose a significant threat to the region's livestock economy and the welfare of ranchers. Figures 3 and 4 show drought intensity across the U.S, comparing July 2021, to July of 2022. In July 2022, drought conditions were most severe in California, Texas, Oregon, Nevada, Utah, and New Mexico, and "32 percent of land in Western states was classified as experiencing extreme or exceptional drought" (USDA, 2022a). This problem is only getting worse. Given the A2 emissions scenario, which makes predictions based on current trends in global emissions, Southwestern annual average temperatures are projected to rise by up to 5.5 °F in the next 50 years (U.S. Global Change Research Program, 2009). Climate change will continue to have a significant impact on animal agriculture in the southwest, particularly by tightening the feed grain supply chain, and reducing forage crop production (U.S. Environmental Protection Agency, 2014). I predict that costs associated with drought depopulation and destocking decisions will increase and there is a serious threat of long-term firm exit in the cow-calf industry.

This risk of firm exit is because aggregate effects of adverse climate conditions result in a very likely decrease in profitability for the cow/calf sector. Climate change effects on feed grain crop production increase variable costs for ranchers and fixed pasture lands are devalued as they yield less forage (Lawrence & Strohbehn, 1999). Ranchers will increase their outsourcing of supplemental feed grains such as corn, oats, and barley. The demand for these crops is increasing, however, climate change is reducing their production yield. Ranches will have to pay higher prices for grains, which means these variable costs will become a larger portion of the unit costs for raising cattle for slaughter (Lawrence & Strohbehn, 1999).

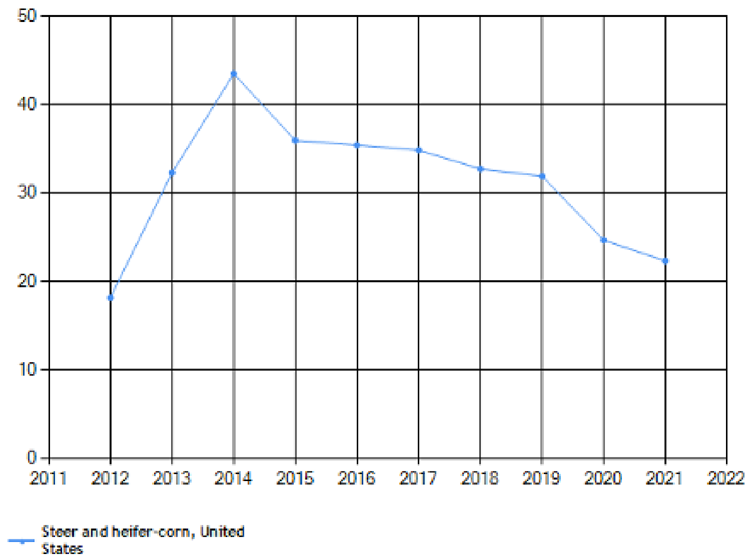
These trends are exhibited by the price ratios of steers and heifers to corn during the period of 2012-2022. USDA calculates these ratios by dividing the price of a steer per hundredweight (cwt) by the price of corn per bushel. Since 2014 these ratios have declined due to rising corn prices. Feed grain accounts for a larger portion of the costs of production of a single calf. There was a 9.4 percent decrease in price ratio from 2020 to 2021 and a 48.7 percent decrease in price ratio from 2014 to 2021 (USDA, 2021). These trends demonstrate that the unit cost of production for calves risks climbing closer to or above the anticipated profit.

Strategic drought management becomes difficult when expected returns to land and breeding stock productivity are affected by longer drought periods which can have longer-term and more dynamic outcomes on production. Destocking strategies are less feasible because even if ranchers can afford the costs of relocation, 32 percent of drought in the Western USA is experiencing extreme or exceptional drought, and land suitable for grazing is very limited. Ranchers have resorted to depopulation strategies to cope with the 2022 drought effects. In the first quarter of 2022, cows were slaughtered in "the greatest numbers since the 1980s" (McFall-Johnsen, 2022). When profitability is rapidly decreasing, choosing depopulation doesn't guarantee short-term security. Ranchers who depopulate will likely need a second source of income to cover their losses. (McFall-Johnsen, 2022).

The cattle cycle is already starting to reflect the industry's response to these climate conditions. In January 2022, the total number of cattle in the nation had fallen by 2 percent (Bates, 2021). New Mexico, one of the most severely affected drought regions,

Figure 5

Price Ratio of Steers/Heifers to Corn



Source: USDA Economic Research Service Feed Grains Database, (USDA, 2021)

has experienced a steady significant decline in state herd inventory, with the average herd size shrinking by 43 percent in 2022 (Nowell, 2022). This is reflective of an overall market contraction. Economist David Anderson predicts that climate change has "the potential to make some areas uneconomic for cattle production" (McFall-Johnsen, 2022). As areas like New Mexico become economically unsustainable for production, ranchers will be forced to liquidate their herds and exit the market. The USDA already expects beef production to drop by 7 percent in 2023 (McFall-Johnsen, 2022). I expect there to be an overall decline in beef cattle supply in the USA over the next 10 years. The cattle cycle might become less periodic because farmers might not fully repopulate their herd to pre-drought size because of the volatile climate and the large economic risks of carrying a full herd.

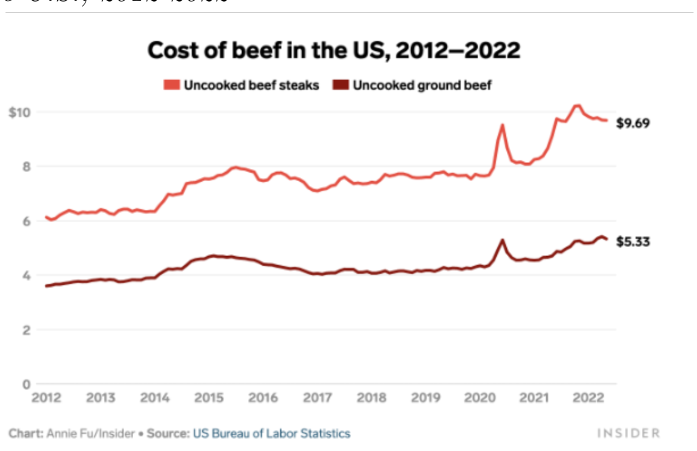
Price Effects

The welfare losses from adverse climate conditions are eventually transferred to consumers. Price trends for beef in 2022 are strong indicators of the correlation between market prices and cattle cycle patterns. Prices for uncooked beef steak dipped to $9.15 per lb in April 2022 after droughts caused ranchers to send high volumes of calves to slaughter (U.S Bureau of Labor Statistics, 2022). This influx of calves slaughtered resulted in temporarily lower market prices. Despite the dip, the April 2022 prices were still 11.72 percent greater than beef prices in April of the previous year (McFall-Johnsen, 2022). Looking forward, the decreased number of calves in the USA in the months post-depopulation means that consumers will be paying high prices for beef for years to come (McFall-Johnsen, 2022).

A long-term economic analysis of the effects of the 2012 droughts showed that over an 8-year period, consumers faced $8.8 billion in surplus losses from higher beef prices (Countryman et al., 2016). Customers were most affected in the first four years after the 2012 drought and beef prices reached their peak for this period in mid-2015 (Countryman et al., 2016). However, rather than prices falling back to pre-2012 baselines, beef prices have continued to increase significantly in the last two years. While spikes in 2020 prices were affected by Covid-19 shocks, Figure 6 indicates that beef prices in 2021 and 2022 have continued to rise (McFall-Johnsen, 2022). Based on the literature, I expect that due to shrinking regional cow-calf production markets, feed grain supply tightening, and subsequent increasing production costs, consumers will continue to pay higher prices to account for the declining supply of beef cattle in the nation.

Figure 6

Cost of Beef in the U.S., 2012-2022



Source: U.S Bureau of Labor Statistics (2022)

Trends also indicate changes in consumer demand for beef. Over the past 20 years, there has been an increased consumer demand for meats such as poultry and pork, while the demand curve for beef has remained flat (Cowley, 2021). Due to environmental externalities of beef production, such as emissions of methane gasses, plant-based meat alternatives have grown as a popular substitute. These shifting consumer demands away from beef due to price, preference, and climate have the potential to further reduce long-term production and cause millions of losses in cattle producer welfare (Cowley, 2021). However, some economists suggest that further expansion into international markets could help mitigate this declining domestic demand for beef (Cowley, 2021).

Conclusion

The financial strains of climate change on cow-calf ranchers reflect that current systems of decision-making are unsustainable. Ranchers face business and welfare uncertainties while consumers face higher prices. Market tightening in the beef industry represents a broader pattern of financial and social consequences which climate change is having on agricultural systems. Climate change affects beef production and livestock agriculture globally. The Southwestern USA is a domestic hotspot for droughts, and other heavily climate-impacted regions of the world, such as central Asia and East Africa are seeing similar effects of long-term drought on their pastoral livelihoods (Orisbayev, 2021).

It is evident that failure to adapt strategies to climate change will result in financial losses and welfare stress for cow-calf ranchers. Thus, there is a strong incentive for ranchers to look toward more sustainable methods of production. Certain green initiatives focus on grassland and prairie conservation. Land stewardship is integral to rancher identity and necessary to maintain a successful business. Practices such as regenerative grazing through pasture rotation allow a herd to sustain itself on forage while also maintaining the health of the rancher's lands (Ryan, 2021).

It is clear that climate change is a global shock, particularly to agricultural industries which rely on the land to support the production of their goods which yield profit and income. Strategies that have helped mitigate financial losses during past drought periods will become less effective as climate change becomes more pervasive and frequent. Frequent drought, intense heat waves, and unpredictable weather create volatile conditions which force many ranchers to lower their stocking numbers, driving the cattle cycle downwards and tightening the national beef supply. Cowboys will need to turn to climate-aware production methods, re-evaluate herd numbers, and land management, and adapt to high input prices in order to maintain their business.

References

Bastian, C. (2018). *Ranch Profitability Given Increased Precipitation Variability and Flexible Stocking.* https://grassland.unl.edu/images/seminar/BastianSlides10082018.pdf. Place: Center for Grassland Studies: University of Nebraska Lincoln

Bates, S. (2021). *Drought Makes its Home on the Range.* http://www.nasa.gov/feature/goddard/esnt/2021/drought-makes-its-home-on-the-range. Publication Title: NASA Type: Text

Countryman, A. M., Paarlberg, P. L., & Lee, J. G. (2016). Dynamic Effects of Drought on the U.S. Beef Supply Chain. *Agricultural and Resource Economics Review*, 45(3), 459–484. https://doi.org/10.1017/age.2016.4

Cowley, C. (2021). Long-Term Pressures and Prospects for the U.S. Cattle Industry. *The Federal Reserve Bank of Kansas City Economic Review*, First Quarter 2022. https://doi.org/10.18651/ER/v107n1Cowley

Federal Reserve Bank of Dallas (2022). *The Fed - Beige Book - September 7, 2022.* https://www.federalreserve.gov/monetarypolicy/beigebook202209.htm

Horsley, S. (2022). Droughts are hitting cattle ranchers hard – and that could make beef more expensive. *NPR.* https://www.npr.org/2022/09/01/1120080654/drought-climate-change-ranchers-cattle-beef-steaks-inflation-food-prices

Hughes, H. (2012). The Economics Of Various Drought Strategies. *Beef.* https://www.beefmagazine.com/blog/economics-various-drought-strategies. Publication Title: Beef Magazine

Lawrence, J. D. & Strohbehn, D. R. (1999). Understanding and Managing Costs in Beef Cow-Calf Herds. *Iowa Beef Center, Iowa State University.*

McFall-Johnsen, M. (2022). Steak and burgers cost more these days. Drought may keep prices high for years. *Insider.* https://www.businessinsider.com/beef-prices-higher-drought-fewer-cattle-less-meat-2022-8. Publication Title: Business Insider

Nowell, C. (2022). The end of a way of life? Ranchers struggle to survive the south-west's megadrought. *The Guardian.* https://www.theguardian.com/environment/2022/oct/02/ranchers-struggle-survive-megadrought-us-south-west

Orisbayev, Z. (2021). Kazakh social media rings alarm about drought, livestock die-off \textbar Eurasianet. *Eurasianet*. https://eurasianet.org/kazakh-social-media-rings-alarm-about-drought-livestock-die-off

Rosen, S. (1987). Dynamic Animal Economics. *American Journal of Agricultural Economics*, 69(3), 547–557. https://doi.org/10.2307/1241690

Rusche, W. (2021). *Supplementing Cows on Pasture to Stretch Forage Supplies*. https://extension.sdstate.edu/supplementing-cows-pasture-stretch-forage-supplies

Ryan, B. (2021). Keeping Cattle on the Move and Carbon in the Soil. *The New York Times*. https://www.nytimes.com/2021/10/31/climate/cows-grassland-carbon.html. Section: Climate

Tronstad, R. & Feuz, D. (2002). Impacts of the 2002 Drought on Western Ranches and Public Land Policies. *Western Agricultural Economics Association*, 1(2).

U.S Bureau of Labor Statistics (2022). *Databases, Tables & Calculators by Subject*. https://data.bls.gov/timeseries/APU0200FC3101?amp%253bdata_tool=XGtable&output_view=data&include_graphs=true

U.S. Environmental Protection Agency (2014). *Climate Impacts on Agriculture and Food Supply*. https://doi.org/10.7930/J0RF5RZW

U.S. Global Change Research Program (2009). Agriculture. *Global Climate Change Impacts in the United States*, 188. Cambridge University Press. https://nca2009.globalchange.gov/agriculture/index.html

USDA (2019). *Livestock Indemnity Program*. https://www.fsa.usda.gov/programs-and-services/disaster-assistance-program/livestock-indemnity/index

USDA (2021). *Feed Grains Custom Query*. https://data.ers.usda.gov/FEED-GRAINS-custom-query.aspx#ChartPanel

USDA (2022a). *Drought in the Western United States*. https://www.ers.usda.gov/newsroom/trending-topics/drought-in-the-western-united-states/

USDA (2022b). *Sector at a Glance*. https://www.ers.usda.gov/topics/animal-products/cattle-beef/sector-at-a-glance/

Wallander, S., Marshall, E., & Aillery, M. (2017). Farmers Employ Strategies To Reduce Risk of Drought Damages. *USDA Economic Research Service*. https://www.ers.usda.gov/amber-waves/2017/june/farmers-employ-strategies-to-reduce-risk-of-drought-damages/

Does Campus Culture Affect Student Behavior in Public Goods Games?

Abbie Natkin and Siri Hoff
ECON 490: Behavioral and Experimental Economics (Advisor: Pete Ferderer)

Over the past few decades, economists have started to explore the influence of culture on economic development. Researchers have identified factors such as cooperation, trust, and other social preferences as contributors to economic growth (Alesina & Giuliano, 2015; Koyama & Rubin, 2022). Cooperation is particularly important because many relevant collective action problems, such as climate change and the alleviation of poverty, require coordination across large groups of unrelated people. Thus, understanding how culture impacts willingness to cooperate can help researchers design better policies.

In economic literature, economists use public goods games to simulate cooperation dilemmas. Public goods embody non-excludability, meaning that people cannot be barred from accessing the good even if they don't contribute to its production, and the good is nonrivalrous because consumption of it by someone does not reduce its accessibility to others. Though game arrangements vary, the basic version of a public goods game examines how contributions to the public good change over time. The game is structured so that the welfare-maximizing outcome occurs when everyone contributes all of their endowment, but self-interest causes individuals to free-ride on the contributions of others (Fehr & Fischbacher, 2003).

Evidence from Gächter et al. (2010)), suggests that cultural differences account for variation in public goods game behavior across the globe. The authors define culture using classifications developed by Inglehart & Baker (2000) that identify cultural differences using World Value Survey responses. The survey includes questions about general and interpersonal trust, a necessary precursor to cooperation. Gächter et al. (2010) find that the behavioral variation across cultural groups is greater than within cultural groups. They conclude, based on this result, that culture influences these differences in public goods game behavior. Our online experiment is motivated by this important finding and explores the impact of campus culture on cooperation in public goods games. The Gächter et al. (2010)

Abbie Natkin is a senior with majors in Economics and Mathematics (Applied Mathematics and Statistics). Siri Hoff is a senior with a major in Economics and a minor in Political Science. Correspondence concerning this article should be addressed to lnatkin@macalester.edu and shoff@macalester.edu.

study compares cultures at a city and supranational level, which we attempt to replicate on a micro-level.

There is a large concentration of private colleges around Saint Paul, MN, with perceived differences in campus cultures. We investigate how and if this manifests in a public goods game by testing whether contributions vary across colleges, with each institution serving as a proxy for a distinct cultural environment.

Another driver of cooperation is social identity. This aspect is explored experimentally by Lankau et al. (2012) who find that people contribute more to the public good when paired with people from their own identity group than with others from a different identity group. While Lankau et al. (2012) create arbitrary identity groups and use team-building tasks as priming, our experiment tests the saliency of campus affiliation as the relevant identity group. In particular, we examine whether cooperation declines when students play public good games with students from other colleges.

This experiment combines the findings of Gächter et al. (2010) and Lankau et al. (2012) to test the null hypothesis that behavior in a public goods game is identical (a) between college campuses and (b) for same-college and mixed-colleges groups. We test these hypotheses using experimental and survey data from nine groups made up of students from three Minnesota private colleges.

This paper starts by reviewing the literature related to altruism, cooperation, trust and public goods games that motivate our hypotheses. The next section covers the experimental design, data and methods. Lastly, the paper concludes with the results, limitations and suggestions for future research.

Literature Review

Fehr & Fischbacher (2003) explore the biological and evolutionary roots of altruism. Altruistic phenomena, wherein individuals willingly incur a cost to achieve a better outcome for others, challenge traditional economic models that describe humans as inherently selfish. Fehr & Fischbacher (2003) argue that gene-based evolutionary theories fail to explain why humans practice altruism and identify cultural evolution and gene-culture coevolution as possible explanations. The willingness of certain groups to practice altruism offers important insights into cooperation dilemmas because cooperation requires a degree of compromise.

Koyama & Rubin (2022) investigate whether cooperation impacts economic outcomes. They find that countries with high levels of trust, a necessary precursor to cooperation, tend to have higher GDP per capita[1]. The authors argue that cultural beliefs about trust and altruism contribute to these outcomes because they facilitate the cooperation

[1] GDP per capita is a crude measure of economic outcomes because it fails to account for inequality. However, it is widely used as an indicator in economic literature and is therefore referenced here.

needed to achieve large-scale economic outcomes.

Alesina & Giuliano (2015) identify surveys, experimental evidence and second-generation immigrant behavior as the three standard methods to measure culture. The authors assert that although surveys allow researchers to ask about specific values, results are subject to administration bias. Their paper goes on to examine the two-way causal relationship between culture and institutions and concludes that culture and institutions co-evolve over time, with "mutual feedback effects". Thus, while institutions can serve as a proxy for culture, the causal direction is difficult to determine.

Gächter et al. (2010) expand on Herrmann et al. (2008) which shows statistically significant differences in public goods game outcomes in different cities across the world. In the Herrmann et al. (2008) study, participants played an anonymous 10-round public goods game. At the start of each round, participants received 20 tokens and were given the choice to keep or contribute them to a common pool. When the round finished, communal contributions were multiplied by 1.6 and those earnings were evenly distributed back to participants. The first 5 rounds were played without punishment and then the option to anonymously punish free-riders was introduced for the last 5 rounds. Herrmann et al. (2008) find that punishment behavior differs significantly among the 16 cities. The authors find that cities such as Muscat and Athens tend to engage in high levels of antisocial punishment, while cities such as Boston and Melbourne tend to engage in altruistic punishment.

Gächter et al. (2010) hypothesize that cultural differences account for the variation in observed cooperative behavior. To investigate this hypothesis, the authors classified the 16 cities used in Herrmann et al. (2008)'s study into six different cultural categories developed by Inglehart & Baker (2000) using the questions from the World Values Survey. Using ANOVA models, Gächter et al. (2010) then tested whether there were empirically distinct contribution outcomes both within and across these cultural groups. Their analysis finds that the variation in contributions across cultures was greater than the variation within cultures, and thus concludes that cultural background has a significant influence on cooperation in public goods games.

Lankau et al. (2012) expand on previous findings that culture impacts public goods game behavior and examine the impact of social identity on cooperation preferences. They hypothesize that participants exhibit higher levels of positive reciprocity when matched with in-group participants as compared to out-group participants. As predicted, Lankau et al. (2012) find that participants display greater conditional cooperation when paired with in-group members. This provides evidence that shared social identity impacts conditional cooperation in mixed-group settings. We explore this finding in our college student sample by investigating if behavior differs when students are placed with others from their same school or others.

Experimental Design

Our primary research question explores whether campus culture affects student behavior in a public goods game. We structure our game like Herrmann et al. (2008), but use schools instead of cities as proxies for different cultural environments. Unlike Gächter et al. (2010), we do not classify the schools into different cultural groups prior to experimentation. Instead, we use behavior and survey data to test our predictions. Our primary null hypothesis, H_0^1 is that cooperative behavior is identical across schools, with the alternative hypothesis, H_A^1 being they differ. Additionally, we expect to find evidence that cooperation levels vary between punishment and non-punishment treatments. We predict based on the literature that cooperation levels will be lower in the non-punishment rounds and higher in the punishment rounds (Fehr & Fischbacher, 2003).

Our secondary research question tests the Lankau et al. (2012) findings that cooperation, measured in public good contributions, increases when participants are matched with in-group members. Thus, our secondary null hypothesis, H_0^2, is that behavior is identical in mixed groups and homogeneous groups, with the alternative hypothesis, H_A^2 being that they differ. Based on the Lankau et al. (2012) results, we expect to find evidence that cooperation increases when students play the game with students from their home institution relative to mixed groups.

To recruit participants, we contacted behavioral economics professors and/or economics department chairs at six Minnesota private colleges. Students then received a Google Form from their professors to collect demographic-qualifying information and availability. To participate, students needed to be taking an economics course, attend one of the selected schools and be between the ages of 18-25. These criteria were established to make the sample relatively homogenous. For the three schools with more than 20 responses, we sorted students into similarly sized groups. Some were placed with students from the same school, while others were placed with students from different schools based on their availability. Students were then invited to join anonymously via Zoom on a pre-established date and time to participate in the experiment.

The game itself was a voluntary contribution version of a public goods game conducted through VeconLab, an online experimental platform. Following Gächter et al. (2010), participants Anonymously played a total of 10 rounds. The first 5 rounds had no punishment, while the second 5 rounds introduced anonymous punishment. Groups received standardized instructions that mentioned explicitly the name(s) of the school(s) participants attended to prime them to consider group identity in their decision-making. At the start of each round, participants were endowed with 25 tokens valued at \$1 each and an investment return of \$0.50. In the latter 5 rounds, participants received 10 punishment points that cost the sender \$2 for each point sent and decreased the earnings of the receiver by 10%.

When the game ended, students completed an exit survey consisting of 11 demographic and World Values Survey questions, asking participants various questions related to their own personal level of trust and ideas of individualism.

Data and Methods

In total, we had 10 experiment groups and 86 participants. Eight of the ten groups were with students from the same school and the other two were mixed. The average group size was nine people and we cut one session because only three participants attended. Participants who completed the study were entered into a lottery to win 1 of 4 $100 prizes. Additionally, 1 person from each of the 10 groups were selected to win $10 from their in-game earnings. The data on contributions and punishment decisions were collected from the game and merged with the exit survey data.

The exploratory graphs shown in Figures 1 - 3 below show the variation in the exit survey responses. We notice in Figure 2, that most of the students who have taken only 1 economics course are concentrated in School A, while Schools B and C both have a relatively consistent range from 1 to 6+. In Figure 1, we can see that Schools A and C have a similar proportion of trusting to non-trusting students, with a greater number being untrusting, while School B seems to have a nearly equal split. Finally, evidence from Figure 3 shows that most students have no prior Behavioral Economics or Game Theory experience. This suggests that most participants are not following a learned optimization strategy, but making decisions based on their personal beliefs.

Figure 4 illustrates how the three schools deviate from the mean average contribution. While Schools A and B closely follow the average contribution line, School C has much lower average contributions.

Figures 5 - 8 show the variation across groups in their contribution and punishment levels. Group average contributions per round are shown with circular points following the scale on the left y-axis, while group average punishment points are shown with triangular points following the scale on the right y-axis. School B, has more across-group variation in the shape and level of contributions between groups, while Schools A and C have more similar contribution trends across groups.

To minimize sample variation, we included demographic questions to use as controls in our exit survey. Control variables included gender, citizenship status, number of economics classes taken, prior experience with game theory and school. Our basic conceptual model is as follows:

$$\widehat{\text{Contributions}} = f(\text{Round}, \text{Lagged Punishment}, \text{Controls})$$

We expect contributions to be a function of two main variables: current round and the amount of punishment received in the previous round, measured using lagged punishment. Additionally, because individuals have different propensities to give and unique utility functions, we included a variety of control variables.

Through preliminary regression analyses, we find that many of our exit survey controls are insignificant. To simplify our model, our end specification is as follows:

$$\begin{aligned}\widehat{\text{Contributions}} = \beta_0 = \beta_1\text{Round} + \beta_2\text{School} + \beta_3\text{Mixed Group} + \beta_4\text{Number of Econ Classes}\\ + \beta_5\text{Trust Level} + \beta_6\text{Game Theory Knowledge} + \beta_7\text{Lagged Punishment},\end{aligned}$$

where β_2 and β_4 are the vectors of coefficients associated with each level of the related variable.

Results

Table 1 below shows the results of our regressions using our full sample of 81 students. We create 2 separate models conditioning on the treatment in order to best illustrate the differences we find, though the full model is included in column 1. In column 2, we see that during treatment one, (no punishment) round is not a statistically significant predictor of contribution, with a p-value of 0.959. However, in line with our predictions, we see that school (specifically for School C at the 5% level and School B at the 10% level) is a statistically significant predictor of contributions. Relative to School A, for any given round, keeping the other covariates fixed, we see that on average students from School B contribute 2 tokens fewer and students from School C contribute 4 tokens fewer. Additionally, students sorted into a mixed group, for fixed covariates, will contribute on average 3 tokens more. This contradicts the Lankau et al. (2012) prediction that in-group contributions are higher than out-group contributions.

We find no statistical significance in the contributions of students who have taken 2-3 or 4-6 classes relative to 1 class, but students who have taken 6+ classes contribute on average 4 more tokens than students who have taken 1 class, which contradicts our prediction. We expected that students exposed to more economics courses would be more welfare-maximizing for themselves, but instead, we find that students with more economics experience tend to prioritize group welfare. Additionally, we notice that relative to students with no prior behavioral economics or game theory courses, students who have taken these classes contribute on average 3 tokens less in a given round with fixed other covariates at a statistically significant level. Finally, our most surprising result is that relative to students who think most people can be trusted, those who feel they need to be very careful with people contribute on average 3 tokens more, at a statistically significant level. This contradicts evidence from the literature that higher trust levels predict more cooperation Alesina & Giuliano (2015); Koyama & Rubin (2022).

Among rounds in treatment 2 which include punishment, the round number now becomes statistically significant at the 5% level which is shown in column 3. When fixing all other covariates, on average under the punishment treatment, subjects contribute almost 1 token more for each subsequent round. This finding aligns with prior research that shows punishment tends to incentive group collaboration. Moreover, in the punishment treatment, both School B and School C are statistically significantly different from School A at the 1% level. Both contribute on average 3-4 fewer tokens than students from School A. This supports our first hypothesis that contribution behavior differs across schools.

Furthermore, in the punishment condition, the mixed group indicator is no longer statistically significant, and neither is trust level or exposure to game theory. However, the number of economics classes taken is statistically significant at the 5% level for 4-6 classes and 6+ classes, relative to 1 class. On average, for fixed covariates, students who have taken 4-6 economics classes contribute 6 tokens more than those who have taken 1, and students who have taken 6+ economics classes contribute on average 4 tokens more than those who have taken 1.

Finally, we find a counterintuitive result for the lagged punishment coefficient. At a statistically significant 5% level, on average, for fixed covariates, every 4 punishment points received in the previous round results in a contribution of 1 less token. This suggests participants are engaging in antisocial punishment wherein higher contributors are punished instead of low contributors. Though they contradict our predictions, these results align with Herrmann et al. (2008), who show that antisocial punishment occurs in some places.

We notice generally similar significance levels when we stratify by homogeneous vs. mixed groups. The results of those regressions are included in the appendix, but whose results will not be discussed here.

Limitations

While we took steps to maintain anonymity, randomize our sample, and control for external confounding factors, there remain numerous limitations to our study. The first problem arises from the sampling technique. Some professors we contacted had connections with our school, which may have introduced bias. Another sample limitation arises from the student participants at our school. Since we attend a small school, we know most of the students personally, which may have introduced administration bias.

A second limitation stems from our group assignment method. While we attempted to randomize the process, our sample size and participant availability restricted this ability. Students who only indicated one available time tended to receive that time assignment.

Additionally, based on various class times or extracurriculars, we may have funneled similar people into the same time slot. However, we expect this to be a small limitation because our analysis is less dependent on group assignment time than it is on the school.

The experiment itself raises another set of limitations. Though all participants were kept anonymous and muted in the Zoom room, we had no control over their physical location. Thus, we had no way to check if they collaborated with other students or consulted outside resources. Additionally, we allowed students to ask questions, and in certain groups, the questions that were asked may have indicated some of the goals of the game to other participants. No students were given answers that indicated the experiment question, but the non-responses may have cued them in.

The biggest limitation is in the challenge of isolating and quantifying culture in order to make a causal claim. Though we attempted to minimize this shortcoming by following past experiments, it is impossible to fully avoid it given the small-scale nature of our experiment. For example, to keep the experiment brief and maximize the possibility that people would finish the survey, we only used four questions from the World Values Survey to try to understand campus culture. Additionally, we omitted questions that might have helped us control for differences in students' geographic origins and socioeconomic status.

Conclusion

Our paper has an underlying reverse causality problem, but this is inherent in any study of this nature. In our experiment, it is impossible to ascertain whether students select their school based on its culture or whether the students are the ones bringing the culture to the school. Research from Alesina & Giuliano (2015) shows that culture and institutions co-evolve so, it is difficult to design a study that isolates which of the two has the greater effect. Gächter et al. (2010) partially solve this problem by using cities that are larger cultural environments, but we cannot replicate this on such a small scale.

Our primary research question asks whether campus culture affects student behavior in public goods games. Due to the complexity of measuring campus culture, it is impossible to make a direct causal claim based on our results. However, our analysis suggests that there is statistically significant evidence that campus affiliation can predict student behavior in a public goods game. In both treatments, average contributions differed statistically significantly by school, with students from Schools B or C giving less than students from School A.

The findings for our secondary research question about mixed versus homogeneous groups, are more inconclusive. We find statistically significant evidence that in the no-punishment treatment, students in the mixed group contribute a higher number of tokens than students in the homogeneous group. However, we do not find the same level of significance in the punishment treatment. Additionally, relative to the entire sample, the proportion of students who were in a mixed group was small at about 17%, so it is harder to make any conclusions to this question.

Further research is needed to determine what specific factors create campus culture so that it can be measured more accurately. Furthermore, a broader experiment could include students from other majors and control for other demographic factors like socioeconomic status.

References

Alesina, A. & Giuliano, P. (2015). Culture and Institutions. *Journal of Economic Literature*, 53(4), 898–944. https://doi.org/10.1257/jel.53.4.898

Fehr, E. & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960), 785–791. https://doi.org/10.1038/nature02043

Gächter, S., Herrmann, B., & Thöni, C. (2010). Culture and cooperation. *Philosophical Transactions: Biological Sciences*, 365(1553), 2651–2661.

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial Punishment across Societies. *Science*, 319(5868), 1362–1367. https://www.jstor.org/stable/20053530

Inglehart, R. & Baker, W. E. (2000). Modernization, Cultural Change, and the Persistence of Traditional Values. *American Sociological Review*, 65(1), 19–51. https://doi.org/10.2307/2657288

Koyama, M. & Rubin, J. (2022). *How the World Became Rich: The Historical Origins of Economics Growth* (1 ed.).

Lankau, M., Bicskei, M., & Bizer, K. (2012). Cooperation Preferences in the Provision of Public Goods: An Experimental Study on the Effects of Social Identity. *University of Göttingen Working Papers in Economics*, 148. https://doi.org/10.2139/ssrn.2191469

Figures and Tables

Table 1

Full Model Regression Outputs

Variables	(1) Both Treatments	(2) Treat. 1 (No Punishment)	(3) Treat. 2 (Punishment)
Round Number	0.688**	0.0135	0.688**
	(2.039)	(0.0514)	(2.039)
College or University = 2, B	-3.332***	-1.920*	-3.332**
	(-2.860)	(-1.694)	(-2.860)
College or University = 3, C	-3.769***	-4.025**	-3.769***
	(-3.768)	(-4.113)	(-3.768)
Mixed-Group Indicator (1/0)	0.922	3.472***	0.922
	(0.924)	(3.550)	(0.924)
Number of Econ Classes Taken = 2, 2-3	0.588	0.927	0.588
	(0.583)	(0.937)	(0.583)
Number of Econ Classes Taken = 3, 4-6	6.191***	1.090	6.191***
	(4.065)	(0.730)	(4.065)
Number of Econ Classes Taken = 4, 6+	4.502***	4.209**	4.502***
	(3.523)	(3.361)	(3.523)
Trust Towards People = 2, Need to be very careful	0.116	2.819**	0.116
	(0.144)	(3.548)	(0.144)
Lagged Punishment	-0.236***		-0.236***
	(-2.707)		(-2.707)
Taken Behavioral Econ/Game Theory (Yes/No) = 2, Yes	0.636	-2.749**	0.636
	(0.479)	(-2.150)	(0.479)
Constant	7.071**	6.150***	7.071**
	(2.356)	(5.383)	(2.356)
Observations	328	409	328

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. T-statistics in parentheses.

Figure 1

Trust Towards People

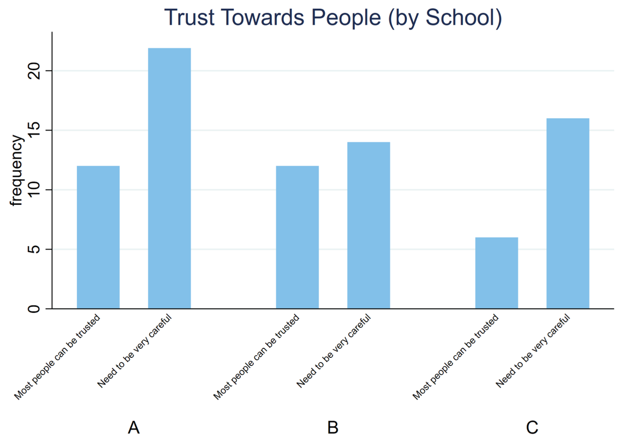


Figure 2

Number of Economics Classes Taken

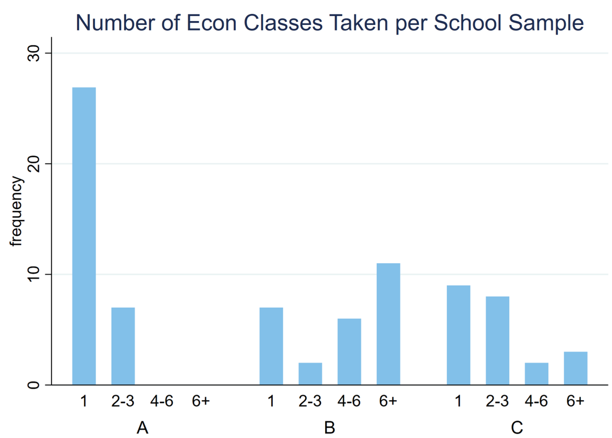


Figure 3

Behavioral Economics/Game Theory Taken?

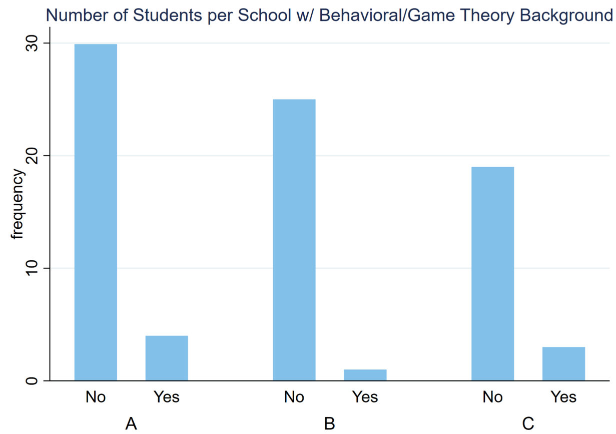


Figure 4

Average Contributions by School

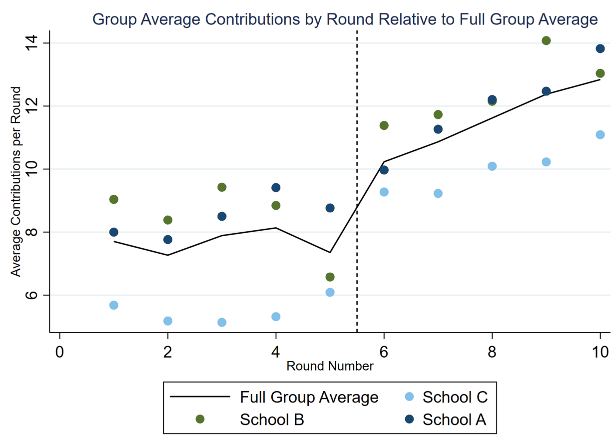


Figure 5

School A Average Contribution & Punishment

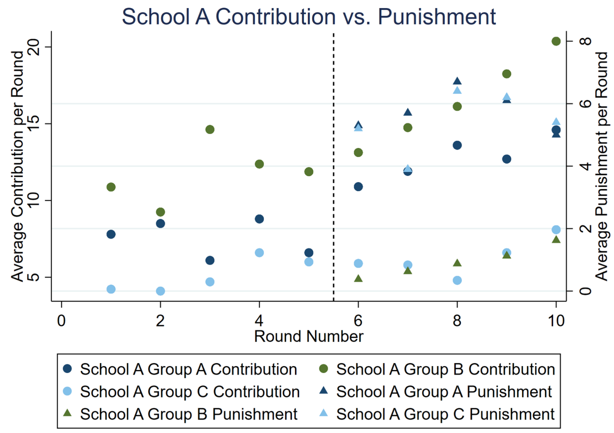


Figure 6

School B Average Contribution & Punishment

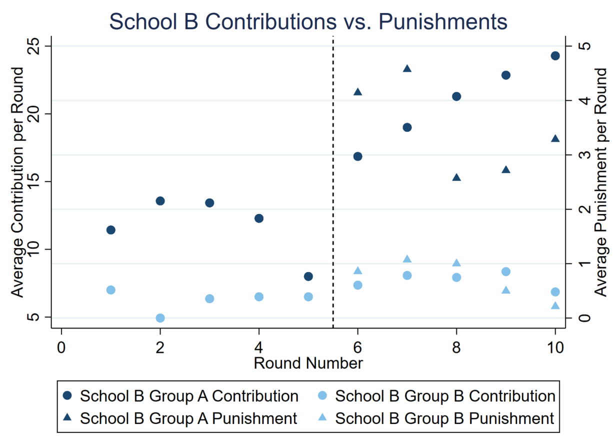


Figure 7

School C Average Contribution & Punishment

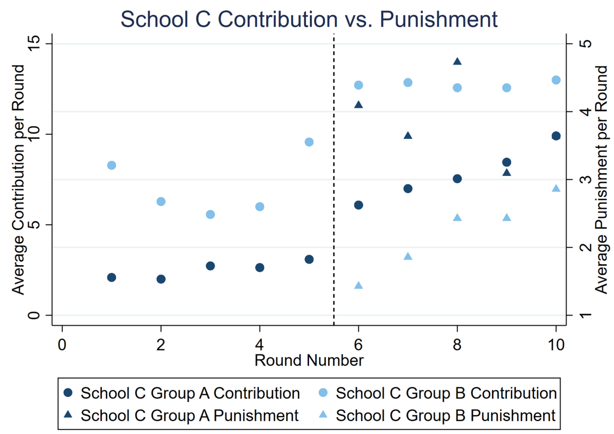
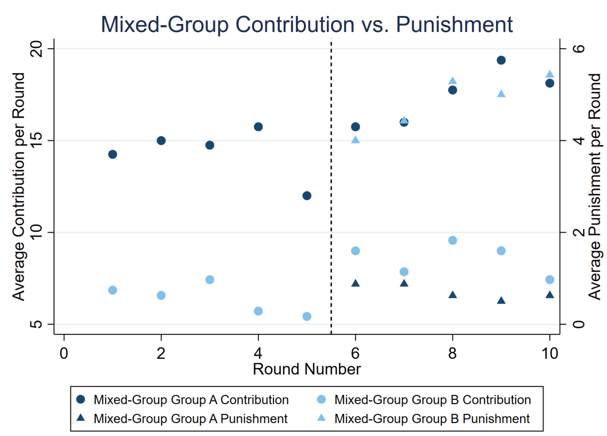


Figure 8

Mixed-Groups Average Contribution & Punishment



Appendix
Additional Regressions

Table 0.1

School A Homogeneous Group Regressions

Variables	(1) Treat. 1	(2) Treat. 2
Round Number	0.640**	1.258**
	(3.810)	(2.258)
Number of Econ Classes Taken = 2, 2-3	-1.912*	-1.565
	(-1.682)	(-0.999)
Trust Towards People = 2, Need to be very careful	0.862	-1.906
	(0.854)	(-1.392)
Lagged Punishment		-0.379***
		(-3.274)
Taken Behavioral Econ/Game Theory (Yes/No) = 2, Yes	3.463*	6.145***
	(1.761)	(2.717)
Constant	5.489**	3.833
	(4.599)	(0.789)
Observations	229	112

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. T-statistics in parentheses.

Table 0.2

School B Homogeneous Group Regressions

Variables	(1) Treat. 1	(2) Treat. 2
Round Number	-0.0797	0.523
	(-0.411)	(0.715)
Number of Econ Classes Taken = 2, 2-3	4.367**	5.381*
	(2.278)	(1.679)
Number of Econ Classes Taken = 3, 4-6	2.034	9.177***
	(1.284)	(3.840)
Number of Econ Classes Taken = 4, 6+	3.455**	6.533***
	(2.395)	(2.898)
Trust Towards People = 2, Need to be very careful	2.056*	0.237
	(1.694)	(0.128)
Lagged Punishment		0.548*
		(1.837)
Taken Behavioral Econ/Game Theory (Yes/No) = 2, Yes	-4.211*	-3.394
	(-1.701)	(-0.817)
Constant	5.193***	1.927
	(3.063)	(0.287)
Observations	175	84

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. T-statistics in parentheses.

Table 0.3

School C Homogeneous Group Regressions

Variables	(1) Treat. 1	(2) Treat. 2
Round Number	1.300***	0.607
	(6.421)	(1.220)
Number of Econ Classes Taken = 2, 2-3	3.225**	1.086
	(2.313)	(0.801)
Number of Econ Classes Taken = 4, 6+	11.81***	6.667***
	(6.653)	(3.612)
Trust Towards People = 2, Need to be very careful	-1.84	0.559
	((-1.449)	(0.431)
Lagged Punishment		-0.0645
		(-0.445)
Taken Behavioral Econ/Game Theory (Yes/No) = 2, Yes	-6.899***	-4.401**
	(-4.264)	(-2.640)
Constant	0.183	3.859
	(0.136)	(0.882)
Observations	125	72

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. T-statistics in parentheses.

Table 0.4

Mixed-Group Regressions

Variables	(1) Treat. 1	(2) Treat. 2
Round Number	-0.0879	0.410
	(-0.361)	(0.690)
College or University = 2, B	5.373**	4.889*
	(2.549)	(1.942)
College or University = 3, C	-3.404	-8.840***
	(-1.627)	(-3.783)
Number of Econ Classes Taken = 2, 2-3	10.47***	9.452***
	(4.332)	(3.884)
Number of Econ Classes Taken = 3, 4-6	-4.672**	-0.102
	(-2.199)	(-0.0422)
Number of Econ Classes Taken = 4, 6+	-4.943**	-4.530
	(-2.106)	(-1.640)
Trust Towards People = 2, Need to be very careful	8.416***	8.799***
	(4.956)	(4.955)
Lagged Punishment		-0.660***
		(-2.688)
Taken Behavioral Econ/Game Theory (Yes/No) = 2, Yes	-8.342**	-1.493
	(-2.529)	(-0.487)
Constant	6.189***	6.870
	(3.321)	(1.298)
Observations	110	60

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. T-statistics in parentheses.

The Price of Russian Gas: An Event Study of the Russian Gas Pipeline Shutdowns Following the Russian Invasion of Ukraine

Dino Weinstock and Patrick Fuchs

ECON 421: International Trade and Multinational Corporations (Advisor: Felix Friedt)

In the wake of Russia's invasion of Ukraine in early 2022, natural gas flowing through pipelines from Russia to Europe have come to a complete halt. This has been Russia's response to the punitive sanctions much of the Western world have placed on the country since their attack on Ukraine started on February 24, 2022. Since then, Electricity prices across Europe have exploded, and in some countries more than others. In this paper, we use an Event Study model to explain the variances in electricity prices across European countries and pin much of the blame on the level of reliance certain European states have on natural gas flowing from Russia through pipelines like that of Nord Stream 1 which was shutoff entirely on June 15, 2022.

The shutoff of the Russian pipelines is important because for many European countries, natural gas is a key source of electricity generation. Italy, Hungary, and the Netherlands, for example, get more than 40% of their electricity from natural gas. Because of this, it is easy to see how the gutting of the European natural gas supply, as was achieved by the shutdown of Russia's pipelines, could affect electricity markets and cause prices to spike. This in turn presents a serious threat to not only natural gas reliant states, but more specifically those most reliant on Russia as an exporter of that natural gas who at this very moment are grappling with the widespread effects of high electricity prices, not just economically, but also politically.

There are many historical parallels between the energy crisis we see today caused by Russia's war in Ukraine and previous energy crises, including the 1973, 1979, and 1980 oil crises which we can use to analyze the current crisis. Each of these crises were the result of a war or revolution which global oil markets predicted would have a serious impact on the output of the countries involved, thus dramatically increasing the price of crude oil as well as the likelihood of oil shortages around the world. In the wake of Russia's invasion of

Dino Weinstock is a senior with majors in Economics and Computer Science. Patrick Fucks is a senior with a major in Economics. Correspondence concerning this article should be addressed to dweinsto@macalester.edu and pfuchs@macalester.edu.

Ukraine and the resulting sanctions enacted by the Western world on Russia as punishment, natural gas markets have reacted similarly: with prices rising to their highest ever in recent history. In fact, European natural gas markets nearly immediately priced in the fact that the two Nord Stream pipelines, one of which was under construction and the other of which provided much of Russia's gas to Europe, would have zero throughput for at least the short to medium term. That market sentiment has been proven by the fact that natural gas prices in Europe hardly even registered the mysterious sabotage of both Nord Stream pipelines in the Baltic Sea in late September.

The question we seek to answer, "how has reliance on Russian pipeline natural gas affected electricity prices in Europe this year?", is an essential one. The answer has far-reaching policy implications, not only when it comes to economic and more specifically trade policy, but also foreign policy. The shutting down of Russian gas flows into Europe has been used as a political tool in conjunction with Russia's invasion of Ukraine to economically punish Europe for standing up to Russia and assisting Ukraine. Examining the extent of the effect Russia's shutdown of its pipelines has had on Europe will hopefully illustrate the importance of choosing your trading partners wisely and diversifying your energy mix properly. Take Germany for example, one of the most infamous offenders when it comes to turning away from a diverse energy mix. At the start of the conflict in Ukraine, Germany was wrapping up an almost decade long plan to decommission nearly all nuclear power plants and entirely phase out nuclear power by the end of 2022 under the auspices of going "all-in" on renewable energy sources. Since the start of that plan, passed on July 1, 2011, nuclear has been almost entirely phased out, but the renewables never filled that vacuum - Russian natural gas did, and it was flowing through the brand new Nord Stream 1 right to Germany's front door. Germany is of course an extreme case, it's not even the most Russia-reliant European country, but it and other reliant European countries who for years were enticed by promises of cheap Russian energy are quickly learning that Russia cannot be counted on to be a reliable trading partner, much less a friend. The significant increases in electricity prices experienced by these countries and the effect they will have not just economically, but also politically, should absolutely be seen as a risk associated with purely following an economically beneficial trade liberalization policy without regard for having a reliable, diverse mix of trading partners.

Literature Review

Precautionary Demand

When it comes to energy crises, the knee jerk reaction is typically to automatically categorize the shock as a supply shock. Kilian (2009) disputes this and suggests the dramatic jump in price of a commodity such as oil, as seen throughout the energy crises of the 1970's and 1980's, is due to an immediate spike in demand that is precautionary and speculative in nature. The jump in price is notably not due to a delayed and transitory drop in the supply of oil.

This model of energy supply and demand, however, is not a perfect match for the current situation in Europe. With the near total throttling of Nord Stream 1 from 100% output before the invasion of Ukraine down to 20% after, and now finally zero since the unexplained explosion of the underwater pipeline, there is a very tangible supply shortage of natural gas in Europe that has naturally caused prices to rise. Kilian (2009) addresses this type of situation using the example of the 1979 Iranian Revolution as well as the 1990 Gulf War where crude oil output was meaningfully affected. In these instances, Kilian (2009) argues that the drop in supply does play a role in the price increase, but it is exacerbated and overshadowed by uncertainty-fed precautionary demand.

Effects of shock on household demand

When it comes to supply shocks such as the shutdown of Russian pipelines to Europe, followed by the destruction of Nord Stream 1, there is more to consider than just a decrease in supply. Reiss & White (2008) found that when looking at California's energy crisis during the year 2000, public officials were able to effectively lean on public appeals for energy conservation and keep electricity prices stable without any painful shortage or rationing.

Alcott (2011), investigating similar strategies for conservation that were used in California, also found that there was some efficacy and usefulness of appeals to consumers. He found that when letters were sent to customers which compared the recipient's electricity consumption with that of their neighbors', consumption on average fell by 2% which Alcott (2011) estimated was equivalent to the drop in consumption that would be seen had electricity prices increased by 11% to 20%.

Public appeals for conservation are something we are very likely to see not just in Germany, but throughout Europe as the continent hunkers down for a cold winter and governments try to avoid the most painful effects of this crisis. As to how effective these conservation "nudges" will be in terms of avoiding shortages and nightmare scenarios like rolling blackouts, has yet to be seen. In Germany, a 1°C reduction in thermostat levels is expected to result in 10 billion cubic meters (bcm) less natural gas consumption per year. However, with a harsh winter approaching Europe in the latter half of 2022, behavioral changes are not expected to contribute much to the required reduction in gas consumption.

Previous Expectations of Gas Crises

The annexation of the Crimean peninsula in 2014 sparked similar fears over both Russia's encroachment and the EU's reliance on Russia for resources like oil and gas. Richter

& Holz (2015) use the Global Gas Model (GGM) which incorporates trade and distinct elasticities for gas demand by the type of consumer to solve for equilibrium levels within countries with varying characteristics and endowments. Though certain key pipelines including the Nord Stream 1 and 2 which are the focus of the current Russia conflict were not yet completed at the time of this study, the study examined a scenario where Russia enters a prolonged state of decreased or entirely ceases exports of gas to Europe. Assuming these policies are taken through the mechanism of reducing gas flows through existing interconnecting pipelines, aggregate European gas demand would only partially be offset through imports from other countries and ramping up domestic production, and equilibrium gas prices would increase significantly.

Outline

In the paper that follows, we will review the Event Study model we use to investigate the relationship between a state's level of reliance on Russian natural gas and the electricity prices they have experienced since the invasion of Ukraine on February 24, 2022 and the shut-off of Nord Stream 1 and the elimination of all Russian pipeline natural gas exports on June to Europe on June 15, 2022. Next, we will review the data we will be using for our regression analysis, pulled from a number of sources to allow us to relate daily wholesale electricity prices in a number of European countries with their respective rate of reliance on Russian gas. Following the data section we will then explain the results of our regression, highlighting our preferred specification and then explaining the implications and policy relevance of our findings.

Conceptual Framework

In order to isolate and analyze the effect of Russia turning off Nord Stream 1 on European electricity prices, we will be using a theory of energy supply shocks borrowed from Kilian (2009). In this paper, the author makes a distinction between energy supply shocks where supply actually falls (as in the case of Nord Stream 1 being shutdown on June 15, 2022 and then promptly blown up) and precautionary demand shocks which are the result of the anticipation of limited supply (as in the case of energy prices spiking following Russia's invasion of Ukraine on February 24, 2022). We are interested in the effect of the actual supply shock, the shuttering of Nord Stream 1, on European electricity prices and how that effect varies between countries most and least reliant on Russian natural gas.

In order to isolate this effect, we must control for some endogeneity, primarily that first precautionary demand shock that resulted from the invasion of Ukraine on February 24, 2022. The invasion caused widespread uncertainty about the future of Europe's natural gas imports and caused electricity prices throughout Europe to rise well before the natural gas supply had actually suffered any shock. In fact, Russia would not touch gas exports to Germany until June 15th, 2022, nearly 4 months later. In order to identify and compare these two effects, we will use an Event Study wherein the invasion of Ukraine on February 24, 2022 serves as our first event and the actual shut-off of Nord Stream 1 and other Russian gas-carrying pipelines serves as our second event.

$$p_{ct} = \sum_{k=-55, k\neq 0}^{K} \gamma_t f_{t+k} + \alpha_c + \varepsilon_{ct} \tag{1}$$

This yields our model for an event study predicting price at a given time for a specific country with pre ($k = -55$) and post-treatment ($k = 55$) effects for the 55 days leading up to and following each of the invasion and shutdown events. We include country fixed effects, denoted as α_c, controlling for differing electricity price levels across the 17 countries. We incorporate reliance on Russian gas through regressing on subsets of our dataset consisting of two groups of countries. The event study is run with different scenarios for the group of countries with above and below median reliance on Russian piped natural gas. We acknowledge that other factors impacting electricity price may be unaccounted for, but the limited time frame of just under 2 months before and after each event reduces the magnitude of the resulting omitted-variable bias.

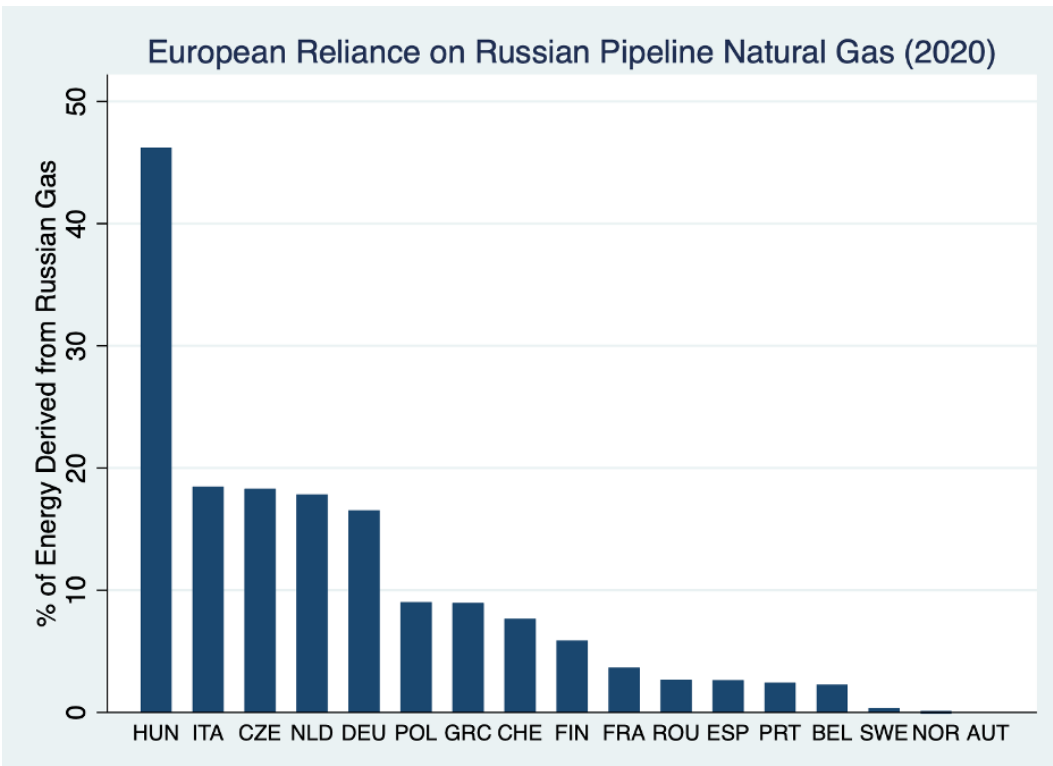
Data Description

Data Summary

To study the effect of the shut-off of Nord Stream 1 on European electricity prices and the effect of a state's level of reliance on Russian natural gas on the magnitude of that price change, we create a panel of data using daily wholesale electricity prices received by generators on the spot market dating back to January 1, 2020 and up through October 31, 2022 for each EU member state. This data was generously aggregated and cleaned by EMBER, a global energy think tank "dedicated to shifting the world from coal to clean energy". The data originates from the ENTSO-E or European Network of Transmission System Operators for Electricity, an EU Commission entity responsible for overseeing electricity transmission systems across the EU. It is important to note that the wholesale price of electricity, measured in Euros per megawatt hour, is different from the price faced by households as it does not include taxes, network charges, subsidies, supplier profits, etc.

In order to measure the level of reliance each European country's energy mix has on Russian gas, we pulled together energy mix data from two sources: the 2020 BP Statistical Review of World Energy and Eurostat. From the BP Statistical Review, we were able to get the total energy consumption for 21 European countries for the year 2020 as well as how much of that energy came from natural gas. From Eurostat, we pulled data for 29 European countries on how much pipeline natural gas they imported from Russia in 2020. With these three pieces of data, we are able to calculate the level of reliance on Russian natural gas for 17 European countries that are shared between the two datasets, which is essentially the percentage of energy used in 2020 that originated from Russian pipelines like Nord Stream 1.

Figure 1

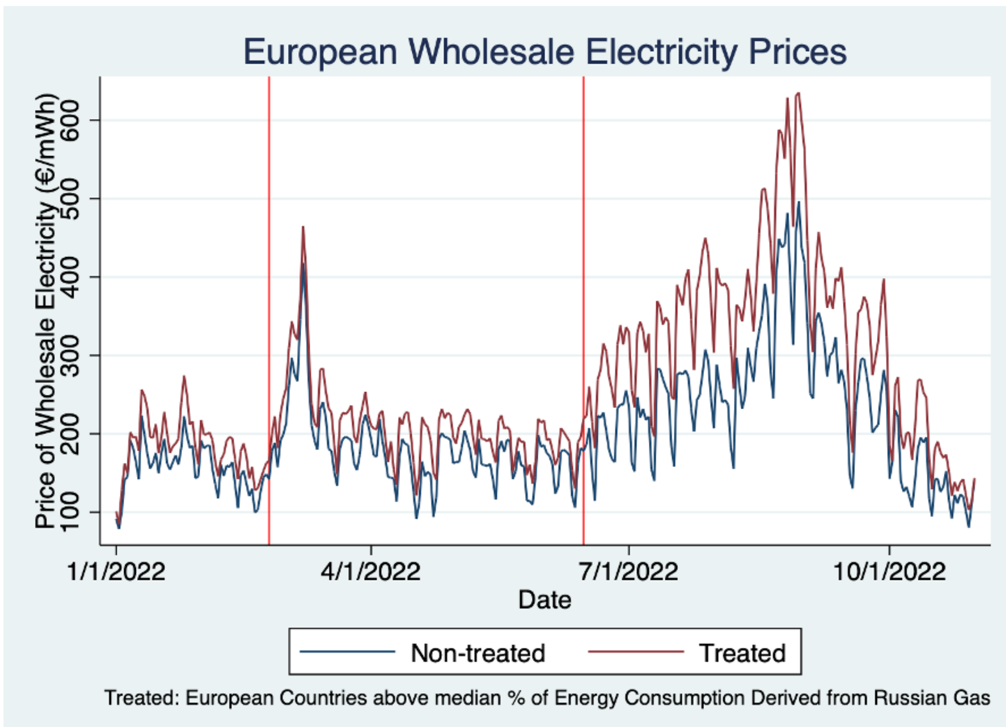


Notes: Authors' Calculations from BP's Statistical Review of World Energy BP (2021).

When talking about Russian natural gas, it is important to differentiate between pipeline natural gas (PNG) and liquified natural gas (LNG). This is due to the fact that since the invasion of Ukraine and despite the shut-off of their pipelines, Russia continues to sell copious amounts of natural gas to customers around the world, including those in Europe, just not through their pipelines. They do this using LNG which is typically shipped via ship and received at purpose-built facilities called LNG terminals. LNG exports to Europe from Russia have largely been untouched by either side when it comes to retaliation (unlike Russian PNG) which is why we choose to focus solely on Russian PNG in this paper.

There is the issue of course that the data with which we calculate each state's level of reliance comes from 2020, which is the most recent year of the BP Statistical Review. This could cause some complications and bias in our estimates due to the fact that the percentage of total energy coming from Russian pipelines may have changed for many of these countries between the time of the BP survey and the time of Russia's invasion. It is our opinion that because we eventually break these countries into treated and untreated groups based on above or below-median levels of reliance, most changes in the level of reliance are not necessarily relevant as any change in the level would a) be relatively small and b) be due to broader causes affecting the entirety of Europe fairly uniformly, meaning that we can assume most likely that none of our 17 countries jump from treated to untreated or vice versa in a single year.

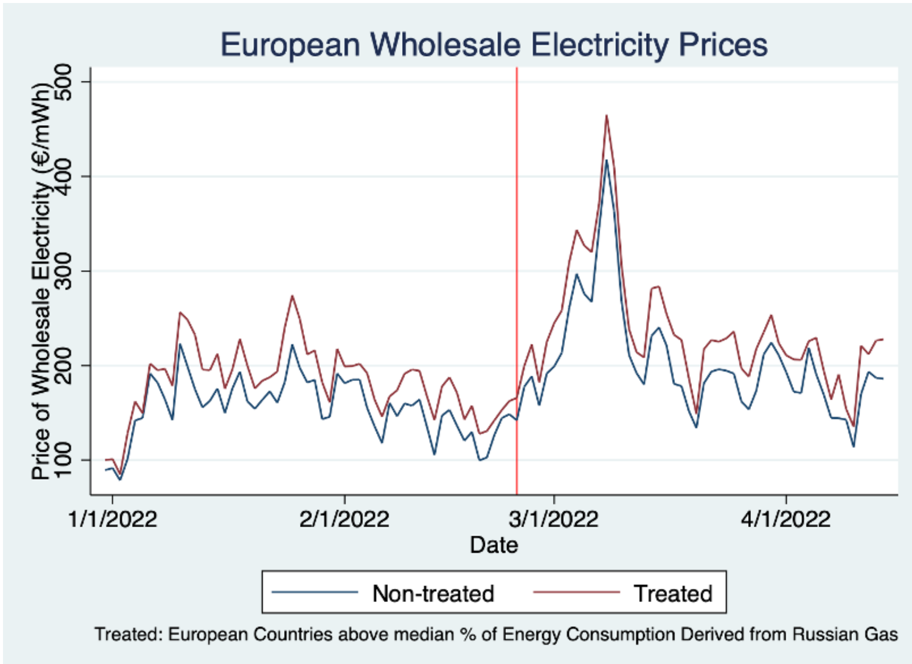
Figure 2



Notes: Average European wholesale electricity price among the treated and untreated groups of European countries in 2022 through October. The invasion and shutdown event dates are indicated by the red vertical lines. Treated countries include: Hungary, Italy, Czechia, the Netherlands, Germany, Poland, Greece, and Switzerland. Untreated countries include: Finland, France, Romania, Spain, Portugal, Belgium, Sweden, Norway, and Austria.

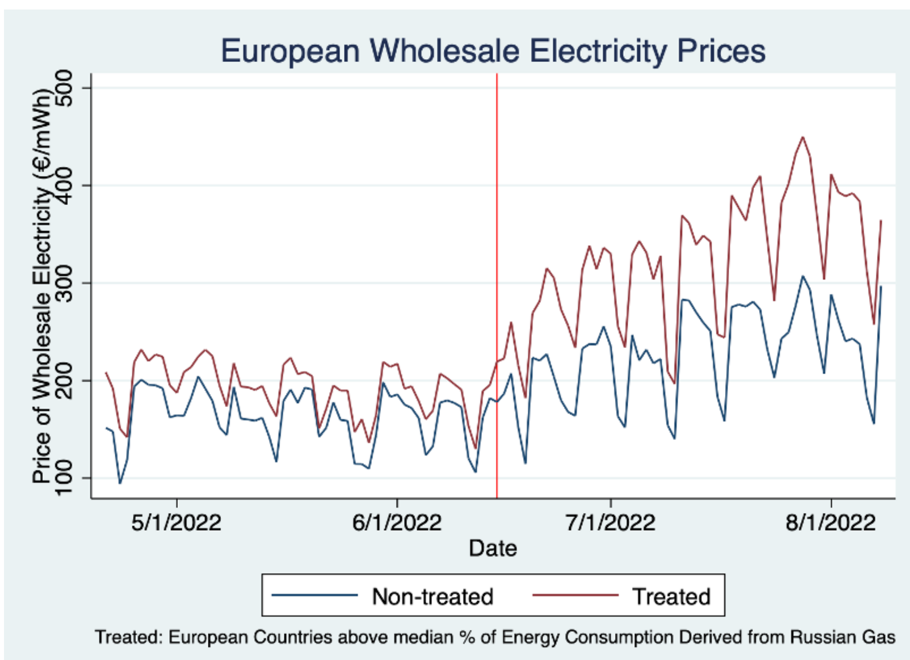
After merging the reliance data from the BP Statistical Review and Eurostat with the wholesale electricity price data from EMBER, we are left with 17 European countries for which we have daily price data from 12/31/2021-10/31/2022 and reliance data for the year 2020. In Figure 1 you can see these 17 countries and their respective rates of reliance on Russian PNG, defined as the percent of total energy consumption in the year 2020 that comes from Russian pipelines. Hungary is the obvious frontrunner and a bit of an outlier with about 46% of their energy consumption coming from Russian PNG, more than twice

Figure 3



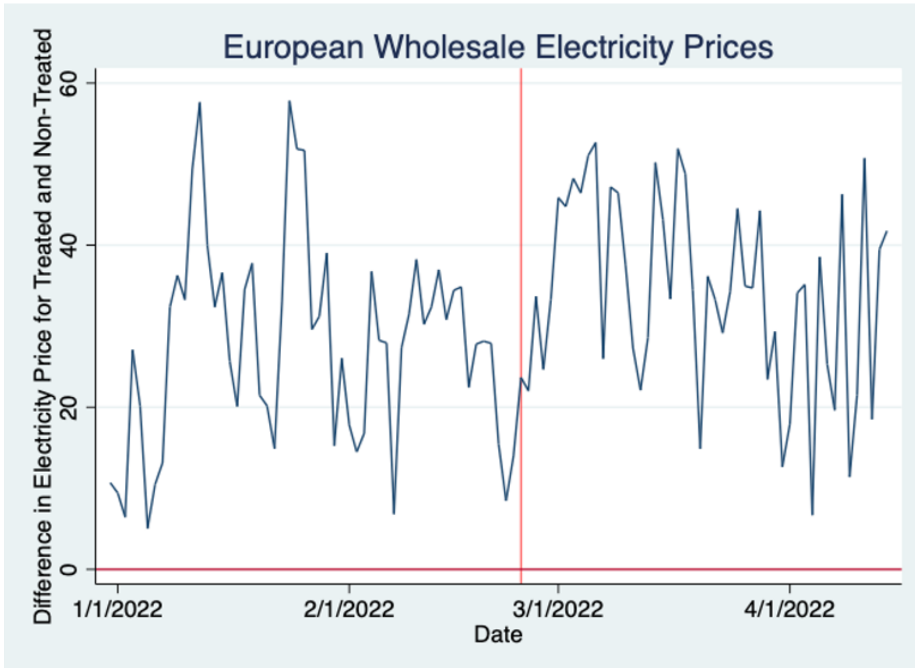
Notes: Average European wholesale electricity price among the treated and untreated groups of European countries 55 days before and after the outbreak of the Russian invasion of Ukraine on February 24, 2022.

Figure 4



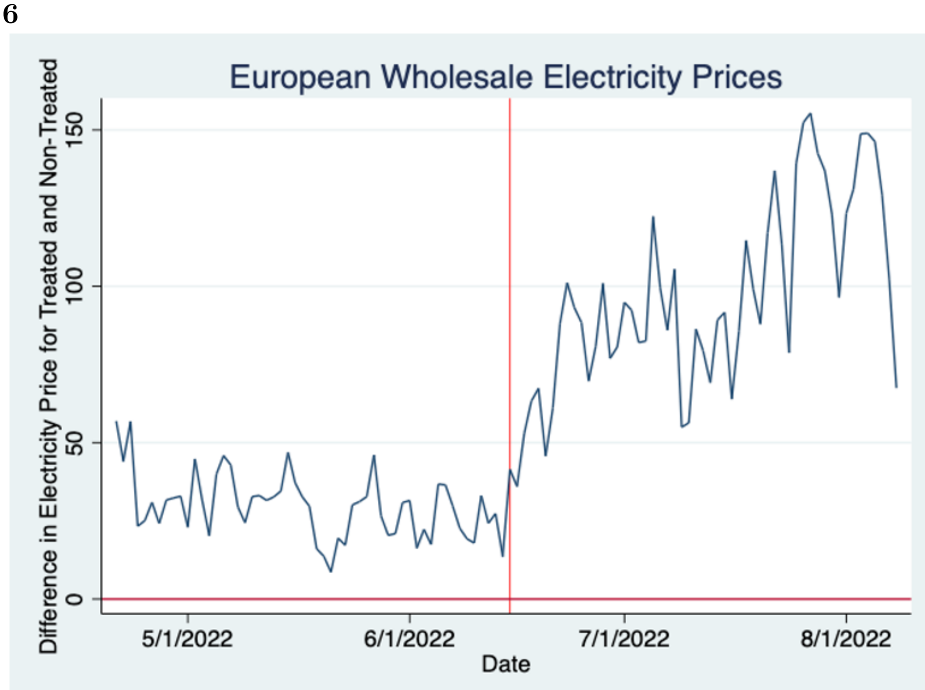
Notes: Average wholesale electricity price among the treated and untreated groups of European countries 55 days before and after the outbreak of the shutdown of the Nord Stream 1 Pipeline on June 15, 2022.

Figure 5



Notes: Difference in electricity prices between treated and non-treated countries before and after the February 24th invasion of Ukraine.

Figure 6



Notes: Difference in electricity prices between treated and non-treated countries before and after the June 15th pipeline shutdowns.

that of Italy, the second most reliant country in Europe for which we have data.

These figures illustrate that after the invasion, the two groups seem to react very similarly as precautionary demand spikes across the board, but when it comes to the pipeline shut-off and supply is actually tangibly affected, the treated group reacts much more harshly with much higher electricity prices than those of the untreated group. This is possible evidence of the differentiation of energy crises described by Kilian (2009) and previously discussed in this paper and warrants further exploration. Despite these two groups' differences though, near the end of our data things seem to calm down with electricity prices returning to around €100/mWh which is around where they started in the beginning of 2022.

Results

To identify the impact of the invasion and shutdown events, we perform our event study with four scenarios. We investigate the relationship between these countries' rates of reliance on Russian PNG and their electricity prices following the Russian invasion and pipeline shutdowns and find results that are consistent with what the theory might suggest and are highly statistically significant. We structure our analysis of the effects of reliance on Russian PNG around two groups of countries that are above or below the median reliance on Russian PNG relative to total energy consumption. The event studies are centered on each of these groups in each of the two event periods, yielding results suggesting broadly significant effects among the above-median group and largely inconclusive evidence of an impact across the below-median group.

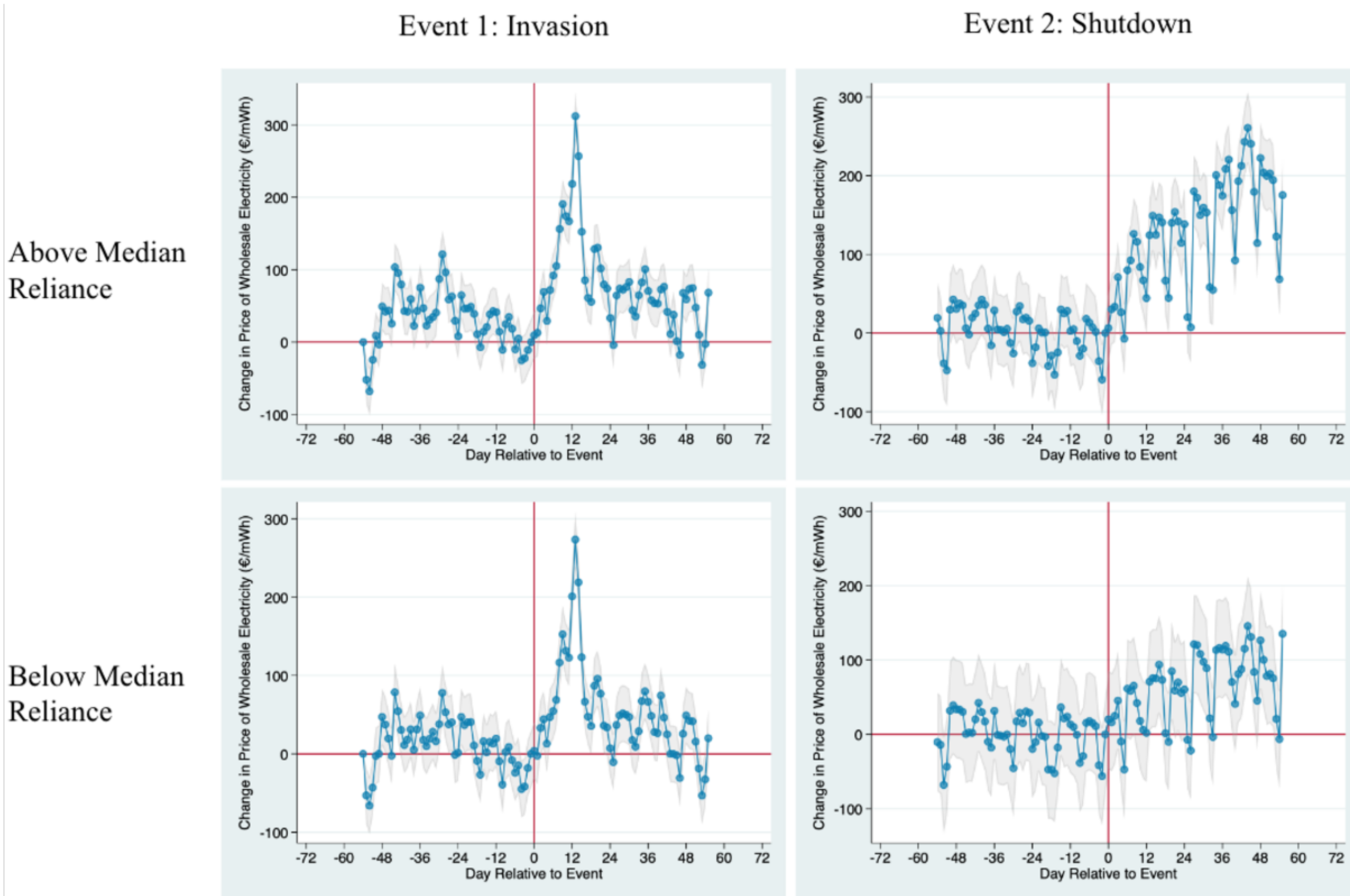
The results of our regression and our estimated pre and post fixed effect coefficients are plotted in Figure 7. Our pre-fixed effects tend to not be significant while the vast majority of our post fixed effects are highly significant. Each point on the graph corresponds to each day's change in wholesale electricity price relative to day 0, the day of the event.

Comparing the graphs in Figure 7 between events and between groups, one can see the differences in price effects. In the left column, prices initially react to the invasion much in the way they would be expected to in a precautionary demand shock with prices spiking and relatively quickly reverting back to their pre-event trend, as we see here. There is also little difference in the first column between groups and this is most likely due to the speculative nature of the shock which doesn't necessarily take into account the levels of reliance these countries have because the gas supply has yet to actually be affected at this point in time.

The right column shows the effect of an actual supply shock – supply has been meaningfully affected by the shutdowns and electricity prices are slowly but surely reflecting that with a steady rise. The difference in effect between country groups in this right

Figure 7

Event Study Regression Results



Precautionary demand shock - uncertainty fed, reverts back to pre-invasion levels.

Supply shock - tangible drop in supply, no reversal to prior trend.

column is also more apparent as energy prices, and by extension, electricity prices are more accurately pricing in the situation on the ground in these countries and more accurately pricing in these countries' levels of reliance on russian gas. In the wake of each event, the price shock is larger in magnitude and more persistent for the more Russian-reliant group of countries, however following the shutdown of pipelines, the price shock is particularly more extreme for the above median reliance group than the below median reliance group, with price increases being at times more than €100/mWh greater than those experienced by the less reliant group. This leads us to believe that our results are not only highly statistically significant, they are also highly economically significant. The difference in price effect following the shutdowns between above and below median reliance countries is large and means that, at times, the treatment group is paying nearly 50% more for electricity than the below median reliance countries on average.

This is significant because higher electricity prices, aside from raising the cost of living for everyday people which has its own political repercussions, lead to higher input prices in an economy and can contribute significantly to inflation as well as cause industry to leave countries and seek lower costs abroad. It should also be noted that countries that were previously so enticed by cheap Russian natural gas and who are now struggling to fill the gap have been forced to compromise their energy transition goals and, in many cases, have turned back on their previously phased out coal fired power plants.

It can of course be seen in Figure 2 that near the end of our available electricity price data, around October, 2022, that electricity prices have come down significantly from their post-invasion and post-shutdown highs and have reverted back to their pre-invasion levels for the most part. This is definitely puzzling as the shutdowns of Russian are very much still in effect as Europe has stood defiant in the face of Russia's extortion attempts. Is natural gas from elsewhere filling the gap left by Russia? Have government campaigns, pleading the public to conserve energy, been successful? Are other energy sources picking up the slack, like coal and nuclear? The answer is likely yes to some extent for all of the above. Liquified natural gas (LNG) originating from the US has begun pouring into Europe at record levels since around the time of the shutdowns. Around that time, many European countries embarked on ambitious plans to stock up on gas ahead of what many thought would be a cold, as well as expensive, winter and a defining moment for Europe's tough stance on Russia. As of the end of November, 2022 however, EU countries were largely able to successfully fill gas storage for the winter heating season, despite the shutdowns. Europe also had luck on their side it seems. Not only were temperatures particularly mild during the fall and early winter, they also benefited from significantly depressed demand for energy from China which is, and has for some time, been in the throes of severe COVID-19 lockdowns. All of these reasons combined with the sincere commitment and determination to save energy and make tough choices by European governments and their citizens have allowed Europe to avoid the worst case scenarios predicted by many following Russia's invasion of Ukraine.

Limitations

We believe the results of our analysis suggest significant evidence for the hypothesis that the events impacted electricity prices at varying magnitudes depending on the degree of reliance on Russian gas. However, there are several limitations particularly concerning external validity and relevance to energy crises more broadly. First, the analysis is based on two events in a short time period, the Russian invasion of Ukraine and subsequent halt of natural gas flowing through pipelines to Europe. While this period provides a useful case study, it may not be representative of other situations or broader trends in the energy market. Additionally, the study is based on data as of October 2022, and the rapidly evolving situation may have changed since then.

Furthermore, our data set examining the electricity price impacts contains just 17 countries across Europe. While we argue the countries are representative of the larger developed nations within the greater European energy market, an ideal set of data would include all European countries as well as unaffected nations beyond the 'treated' vicinity of Russia. Data availability limited the inclusion of certain countries such as Ireland and Ukraine as only countries with complete observations for both Russian gas import and electricity prices were retained in the dataset. The analysis and its interpretation should therefore be constrained to the highly specific event period and regions represented through our data.

A second factor restricting the applicability of our results for other events is the uniqueness of both the Russian invasion and shutdown events. Previous literature studying energy shocks focuses on energy crises that line up loosely with the precautionary demand shock of the initial February invasion. However, the tightly woven energy trade relationships between affected countries and Russia that was actually shut off in the case of the pipeline closures was abrupt and unprecedented in scale. Other widely publicized energy market shocks can often be described more as precautionary demand shocks, with disproportionately large shifts in energy market pricing that revert close to pre-event levels soon after the event. Thus, our results surrounding the actual supply shock of the shutdown event should be cautiously extrapolated to other crises that lack a significant actual supply shock. Future energy crises driven by halted trade flows might also have dampened immediate impact due to increased preparation of alternative sources of energy in response to the Russian invasion.

Finally, the analysis is based on an event study model, which is a useful tool for analyzing the effects of specific events on market prices, but it has its own limitations. The event study does not directly examine the difference in price impact for higher and lower reliant countries and cannot provide statistically significant evidence for or against the different outcomes for the two groups of countries. We acknowledge this conclusion could have been robustly explored through a DiD setup, but we were unable to generate bins of countries that satisfied the parallel trends assumption. Additionally, the event study model assumes that the effects of an event are immediate and measurable, which is likely not the case as either of our events likely took place within a single day.

Conclusion

The Russian invasion of Ukraine and subsequent halt of natural gas flowing through pipelines to Europe has resulted in significant increases in electricity prices across the continent. We used an Event Study model to analyze the variances in electricity prices in European countries and found that those with a higher reliance on Russian natural gas were particularly affected, experiencing price increases that were at times more than €100/mWh higher than those experienced by less reliant countries. The shutdown of Russian pipelines has had far-reaching economic and political implications, highlighting the importance of diversifying energy sources and choosing trading partners wisely.

References

Alcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9/10), 1082–1095. https://www.sciencedirect.com/science/article/abs/pii/S0047272711000478

BP (2021). *Statistical Review of World Energy 2020*. https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html

Kilian, L. (2009). Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *The American Economic Review*, 99(3), 1053–1069. http://www.jstor.org/stable/25592494

Reiss, P. C. & White, M. W. (2008). What Changes Energy Consumption? Prices and Public Pressures. *The RAND Journal of Economics*, 39(3), 636–663. http://www.jstor.org/stable/25474390

Richter, P. M. & Holz, F. (2015). All Quiet on the Eastern Front? Disruption Scenarios of Russian Natural Gas Supply to Europe. *Energy Policy*, 80, 177–189. https://doi.org/https://doi.org/10.1016/j.enpol.2015.01.024

FDI, Exports, and GDP in East and Southeast Asia: Evidence from Time-Series and Causality Analyses

Zefan Qian

ECON 421: International Trade and Multinational Corporations (Advisor: Felix Friedt)

As the world FDI inflows have increased steadily in recent decades, there was a wide discussion on the effect of FDI on the economy of the host country. Literature indicates that there is a mixed result of causal relationships between GDP and FDI. In the endogenous growth theory, foreign direct investment (FDI) has been considered as one factor to promote economic growth in the host country. De Mello Jr (1997) found two ways for FDI to stimulate the economy. First, through capital spillovers, new technology was encouraged to be used in the production process. Second, knowledge was easier to be transferred through labor training and better management (De Mello Jr, 1997).

However, some literature showed that this positive effect was not significant, and some even found a negative impact due to the crowding out effect (Carkovic & Levine, 2002). This motivated Hansen and Rand to adopt a bivariate vector autoregressive model to re-evaluate the model specification Carkovic and Levine changed in their project using a sample of 31 developing countries from 1970 to 2000. They found a strong causal link from FDI to GDP in both the short-run and long-run in developing countries, but not in the opposite direction. Also, using the standard Solow model as a benchmark, Hansen and Rand observed that FDI had similar growth enhancement as domestic investments (Hansen & Rand, 2006). These all indicated that there is a mixed relationship between FDI and GDP, varying from country to country.

Hsiao & Hsiao (2006) examined the Granger causality relations between GDP, exports, and FDI among eight economies with similar backgrounds in East and Southeast Asia using time series and panel data between 1986 and 2004. The inclusion of exports in the analyses further demonstrated the FDI's reinforcing effects on GDP through exports. This motivated me to further examine the causal relations between GDP, exports, and FDI in East and Southeast Asia both at individual and aggregate levels Hsiao & Hsiao (2006). Specifically, this article will contribute to Hsiao & Hsiao (2006) in two ways. First, since

Zefan Qian is a senior with majors in Economics and Mathematics and has a minor in Data Science. Correspondence concerning this article should be addressed to zqian@macalester.edu.

previous literature indicated that causal relations vary over time, this article extends the time period to 2018. The changes in the causal relations between these three variables of interest will explain how these economies change over time and the underlying reasons behind the development of various policies such as China's export-led-growth policy. Second, Hsiao & Hsiao (2006) added random effects to the VAR model to conduct panel data analyses. Instead, this article adopts the Dumitrescu–Hurlin test by Dumitrescu & Hurlin (2012), which allows panel data Granger causality tests for a relatively short period of time.

In this article, Section 2 develops the models used to examine the causal relations between exports, FDI, and GDP. Section 3 summarizes the data and elaborates on sample selection. Section 4 demonstrates the empirical results and contextualizes the causal relations. Finally, Section 5 discusses the limitations and Section 6 concludes this article.

Model

Analytical Framework

Hsiao & Hsiao (2006) examined the national income model in their article. By assuming that the money sector and government sector are at their equilibria, the equilibrium condition of the Keynesian model of aggregate demand and aggregate supply is

$$Y = C(Y) + I(Y,r) + F + X - M(Y,e) \tag{1}$$

where Y, C, I, F, X, M, r, and e are real GDP, real consumption, real domestic investment, real FDI inflows, real exports, real imports, nominal interest rate, and nominal exchange rate of foreign currency in terms of the domestic currency, respectively. Thus, $X - M(Y,e)$ is the current account surplus of the domestic country. To explore the real aspect of the economy, Hsiao & Hsiao (2006) indicated that it is needed to ignore the financial variables. Here, we should expand the non-linear functions $C(Y)$, $I(Y,r)$, and $M(Y,e)$ logarithmically around the origin by the Taylor expansion. Thus, (1) can be written in a more general implicit function form

$$f(Y,X,F) = 0 \tag{2}$$

where GDP, Exports, and FDI are the variables of interest. Here, we can regress each of the three variables on the other two variables, and take the lags of each variable. This is the prototype of the vector autoregression (VAR) form for the Granger causality tests (Hsiao & Hsiao, 2006).

VAR Granger Causality Tests

This article adopts VAR(p) model to test the Granger causality relations between variables of interest (Hsiao & Hsiao, 2001).

The VAR(p) model is specified as

$$\mathbf{y}_t = \mu + \Gamma_1 \mathbf{y}_{t-1} + \Gamma_2 \mathbf{y}_{t-2} + \ldots + \Gamma_p \mathbf{y}_{t-p} + \varepsilon_t. \tag{3}$$

where $\mathbf{y}_t$ is a (3×1) vector of endogenous variables Exports, FDI and GDP, μ is a (3×1) constant vector, p is the order of lags, each of $\Gamma_1, \Gamma_2, \ldots, \Gamma_p$ is a (3×3) coefficient matrix, each of $\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p}$ is (3×1) vector of the lagged endogenous variables, and ε_t is a (3×1) vector of the random error terms in the equation system. Therefore, the hypotheses of Granger non-causality between each pair of variables of interest are

$$H_0(\mathbf{y}_i \nrightarrow \mathbf{y}_j) : \gamma_{ji(k)} = 0, \quad k = 1, 2, \ldots, p$$

where $\mathbf{y}_1$ is exports, $\mathbf{y}_2$ is FDI, $\mathbf{y}_3$ is GDP, and $\gamma_{ji(k)}$ are the (j, i)-elements in the Γ_k matrix. For instance, the null hypothesis of Granger non-causality from FDI to GDP is given by

$$H_0(\text{GDP} \nrightarrow \text{FDI}) : \gamma_{32(k)} = 0, \quad k = 1, 2, \ldots, p$$

Panel Data VAR and Granger Causality Tests

Panel data analysis uses information regarding cross-section and time-series analyses. Different from VAR model for individual countries, we need to make assumptions about the intercept, the slope coefficients, and the error term when we estimate panel data regression models. This article adopts the Dumitrescu–Hurlin test by Dumitrescu & Hurlin (2012). The underlying regression is

$$y_{i,t} = \alpha_i + \sum_{k=1}^{K} \gamma_{ik} y_{i,t-k} + \sum_{k=1}^{K} \beta_{ik} x_{i,t-k} + \varepsilon_{i,t} \quad \text{with } i = 1, \ldots, N \text{ and } t = 1, \ldots, T \tag{4}$$

Specifically, the DH test detects causality at the panel level such that causalities don't necessarily exist for all individuals. Thus, the alternative hypothesis is given by

$$\begin{aligned} H_1: \quad & \beta_{i1} = \cdots = \beta_{iK} = 0 && \forall i = 1, \ldots, N_1 \\ & \beta_{i1} \neq 0 \text{ or } \ldots \text{ or } \beta_{iK} \neq 0 && \forall i = N_1 + 1, \ldots, N \end{aligned}$$

To test this hypothesis, Dumitrescu and Hurlin propose the average Wald statistic $\bar{W}$ calculated from the individual Wald statistic W_i. They also introduce standardized statistics $\bar{Z}$ and approximated standardized statistic $\tilde{Z}$ for long periods of time and relatively shorter periods. Since this article looks at time periods shorter than 30 years, I will look at the approximated standardized statistics to test the hypothesis.

Data

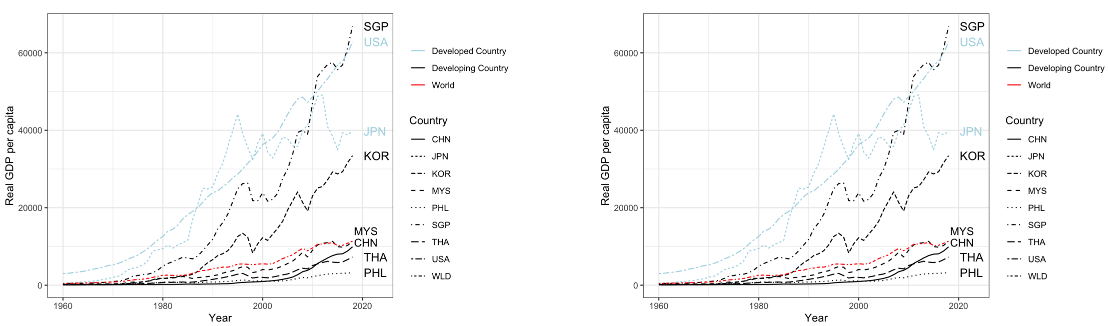
East and Southeast Asia in the World Economy

This paper mainly looks at six Asian economies: mainland China, South Korea, Malaysia, the Philippines, Singapore, and Thailand, to avoid considering countries with different backgrounds and stages of development in cross-section analysis. The main variables of interest include real GDP, real GDP per capita, real FDI inflow, and real Export, in 2022 US dollars which are annually reported by the World Bank at the national level.[1] These six economies have a unique development position in the world economy due to two reasons.

Figure 1

Real GDP per capita of Eight Economies and World

a *Eight Economies and World* b *Four Economies and World*



First, the real GDP per capita of these six developing countries grew rapidly, even at a greater pace compared to developed countries as shown in Figure 1. Figure 1a presents the real GDP per capita of these six economies between 1960 to 2018. The levels of real GDP per capita of the world, the United States, and Japan are also presented for comparison. From the figure, it is obvious that in 1960, the levels of real GDP per capita of South Korea and Singapore were close to that of the world. However, they quickly exceeded the world average with thirty years of development, and gradually caught up with developed countries. After 2010, the level of real GDP per capita of Singapore even went beyond that of Japan and kept pace with that of the United States.

[1] The current values of variables of interest are deflated by the GDP deflator of each country; and denoted as real values.

While the other four countries didn't experience such large rapid growth, the growths in their levels of real GDP per capita were still conspicuous. As shown in 1b, while the levels of real GDP per capita of these four economies were far below the world average in 1960, Malaysia and mainland China quickly converged to the world average in the twenty-first century. To avoid the heterogeneity problem of the early rapid growth of GDP, as indicated by Nair-Reichert & Weinhold (2001), and COVID-19, this article will choose the data from 1986 to 2018. It might not be convincing that the growths in Thailand and the Philippines were great enough to be considered as the ones with a unique development position. However, they are included mainly because they are large recipients of FDI inflows.

Table 1

Summary Statistics of FDI Inflow (2022 US$, Billion) of Six Economies

Year	CHN	KOR	MYS	PHL	SGP	THA	Total	Dev.
[1986, 1988]	2.46	0.94	0.54	0.46	2.73	0.57	7.71	11.63
[1989, 1991]	3.75	1.30	2.67	0.55	4.45	2.08	14.78	13.27
[1992, 1994]	24.15	0.99	4.84	1.02	5.15	1.76	37.91	24.27
[1995, 1997]	40.49	2.86	4.80	1.41	13.03	2.77	65.34	49.43
[1998, 2000]	42.92	9.41	3.28	1.87	13.44	5.59	76.51	88.71
[2001, 2003]	52.68	6.34	2.32	1.01	13.41	4.55	80.29	75.42
[2004, 2006]	98.77	12.03	5.33	1.65	27.61	7.66	153.1	175.0
[2007, 2009]	152.9	9.68	5.59	2.11	28.12	7.87	206.3	280.9
[2010, 2012]	255.0	9.59	11.63	2.10	53.26	10.04	341.6	336.4
[2013, 2015]	267.2	8.71	10.59	5.04	67.62	9.95	369.1	389.3
[2016, 2018]	192.1	14.07	10.38	9.49	81.92	8.32	361.2	392.7
All	102.9	6.90	5.63	2.43	28.25	5.56	139.1	153.1

Notes: The table reports the annual average of FDI inflow in the given time period. The developing countries include low- & middle- income countries defined by the World Bank and graduated developed economies such as Singapore and South Korea.
Source: World Bank.

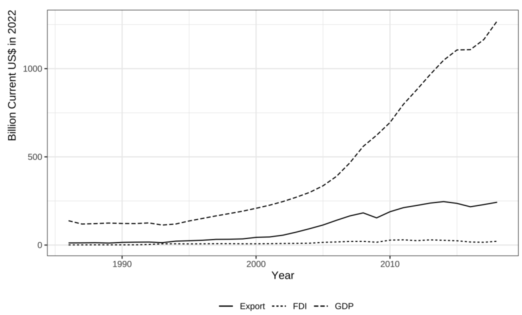
Second, these six economies were the largest recipients of FDI inflow. Table 1 presents the summary statistics of the real FDI inflow of six economies between 1986 to 2018. Since this table aims to show the scale of FDI inflow in these six economies, instead of exhibiting detailed variation over years, the year is disaggregated in 3-year intervals for simplicity. By comparing the last two columns, the total real FDI inflow of these six economies is close to, or even more than, the real FDI inflow of developing countries.

This doesn't mean that other developing countries have no FDI inflow. FDI inflow can be negative if the value of disinvestment by foreign investors was more than the value of capital newly invested in the reporting economy. However, this informs that 1) real FDI inflow in these six economies grew rapidly from 1986 to 2018, and 2) they received most of the FDI inflow among developing nations.

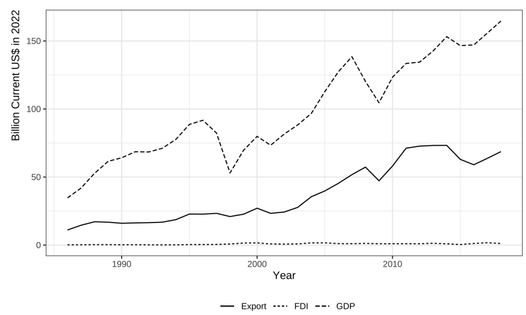
Figure 2

Real GDP, Real FDI, and Real Exports of Six Economies between 1986 to 2018

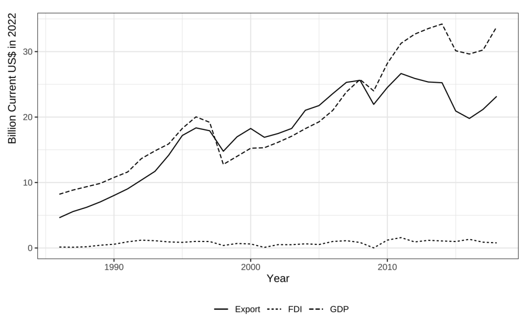
a *China*



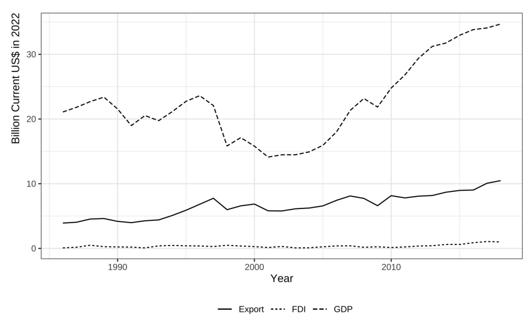
b *Korea*



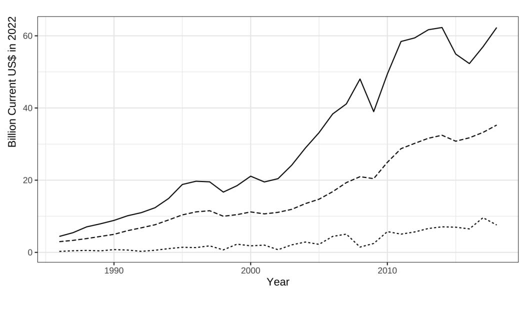
c *Malaysia*



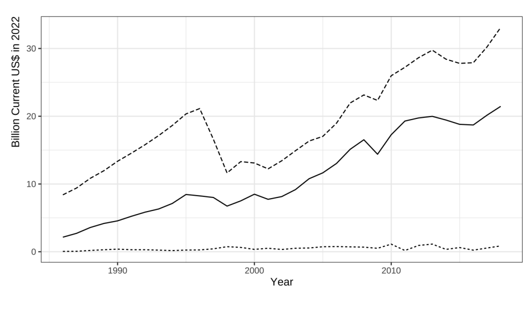
d *the Philippines*



e *Singapore*



f *Thailand*



For Thailand and the Philippines, whose growths in real GDP per capita were not rapid enough, we can see that they experienced a similar degree of growth in real FDI inflow across this time period, even compared to South Korea which experienced one of the most rapid growth in real GDP per capita. Thus, this article aims to examine the causal relations of GDP, FDI, and exports in these six countries between 1986 and 2018.

Characteristics of the Country Data

Figure 2 visualizes the real GDP (rGDP), real FDI (rFDI), and real Exports (rEX) in 2022 US$ for each of these six economies between 1986 to 2018. Since the figures are not used to compare the values across countries, the scales of the vertical axes are not controlled. From the figure, we can see that there is continuous growth in GDP and exports. Except for China, the levels of GDP were negatively affected by the 1997 Asian financial crisis and the 2008 Great Recession. They also negatively affected the exports, while this effect was negligible to Singapore. More interestingly, different from other countries, Singapore's export is constantly higher than its GDP, and Malaysia's export went beyond its GDP between the two events previously mentioned. In general, we can see that the patterns of rGDP and rEX for each country are similar, respectively, such that they are strongly correlated.

In terms of scale, it seems like FDI is much less significant to account for economic growth. However, Hsiao & Hsiao (2006) indicated that there are two channels from FDI to GDP growth. First, as mentioned, mainland China was not affected by the Asian financial crisis and the Great Recession as badly as other nations. This might happen because the events redirected FDI to mainland China and it reduced the influence of FDI on GDP. This, as one of many reasons, explained why other nations are more vulnerable to the financial crisis. Second, FDI mainly goes to key industries such as high-tech manufacturing sectors. Thus, it is crucial in promoting productivity growth and exports in these industries (Hsiao & Hsiao, 2006). Thus, it is also important to consider the causal relationships between FDI and GDP, and FDI and exports.

Results

Hsiao & Hsiao (2006) examine that for a sample of eight economies across a period of twenty years, they can choose between VAR(1) and VAR(2) models. Thus, in this article, for a sample of six economies across thirty years, I assume I can also choose between VAR(1) and VAR(2) models. In this article, the lag order is 2. In terms of lag selection, the AIC of lag 1 is higher for three countries and lower for the other three countries compared to that of lag 2. The only difference is that the unidirectional causality from FDI to GDP is

Table 2

ADF Unit Root Test: Six Individual Economies, 1986-2018

	Level Series			First-difference			Level Series			First-difference	
	Lag	TS		Lag	TS		Lag	TS		Lag	TS
China						**Singapore**					
ex	0	-4.202***	dex	0	-4.010***	ex	3	-3.316**	dex	0	-3.516***
		(0.00)			(0.00)			(0.06)			(0.01)
fdi	0	-1.315	dfdi	0	-6.096***	fdi	1	-2.370	dfdi	1	-4.664***
		(0.88)			(0.00)			(0.40)			(0.00)
gdp	2	-2.226	dgdp	1	-4.155***	gdp	0	-5.760***	dgdp	0	-8.038***
		(0.48)			(0.00)			(0.00)			(0.00)
Korea						**Malaysia**					
ex	1	-3.816***	dex	1	-5.103***	ex	3	-2.662	dex	0	-4.678***
		(0.02)			(0.00)			(0.25)			(0.00)
fdi	1	-2.538	dfdi	0	-5.010***	fdi	2	-1.853	dfdi	0	-3.974***
		(0.31)			(0.00)			(0.68)			(0.00)
gdp	1	-2.609	dgdp	1	-6.061***	gdp	0	-4.776***	dgdp	0	-8.365***
		(0.28)			(0.00)			(0.00)			(0.00)
Philippines						**Thailand**					
ex	3	-1.760	dex	0	-4.534***	ex	3	-2.656	dex	0	-3.876***
		(0.72)			(0.00)			(0.25)			(0.00)
fdi	3	-2.956	dfdi	0	-5.902***	fdi	0	-3.214*	dfdi	0	-4.292***
		(0.14)			(0.00)			(0.08)			(0.00)
gdp	0	-3.254*	dgdp	0	-6.995***	gdp	0	-4.699***	dgdp	0	-8.623***
		(0.07)			(0.00)			(0.00)			(0.00)

In level series, the test equation includes the constant and the linear trend. In the first-difference series, the test equation includes the constant. Lag is selected by the minimum AIC with maximum Lag = 3. The p-values are in the parenthesis. *** (**, *) denotes rejection of null hypothesis at the 1% (5%, 10%) level of significance, respectively.

unobserved when the lag order is 1. Thus, both models should provide robust results and this article chooses the lag order to be 2 to observe variations from more lags.

Moreover, I can't perfectly replicate the work of Hsiao & Hsiao (2006). When I compare their data summary with mine, data from World Bank and ICSEAD are not exactly the same. However, since there is no significant change in the results and due to the page limit, replication results are not presented and I follow the results in Hsiao & Hsiao (2006).

Individual Economy's Granger Causality Tests

As explained in Section 3.1, the values of the real variables should be transformed into log values, which are denoted by ex, fdi, and gdp. Thus, the growth rates of these variables are denoted by dex, dfdi, and dgdp.

Unit Root Tests

Table 2 presents the Augmented Dickey-Fuller (ADF) unit root test results to examine the stationarity of each time series. Specifically, it presents the results for the level series and first-difference series. The test results for the level series are mixed for all six economies. China's and Korea's logs of exports are stationary series; Malaysia's and Thailand's log of GDP are stationary series; the Philippines's log of GDP is stationary at a 10% level of significance; and Thailand's log of GDP is stationary and log of FDI is stationary at 10% level of significance.

Table 3

Summary Statistics of the Growth Rates of Real Export, FDI, and GDP, 1987-2018

	dex			dfdi			dgdp		
Country	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
China	0.094	0.486	-0.223	0.101	0.859	-0.342	0.069	0.183	-0.147
Korea	0.057	0.274	-0.192	0.055	1.062	-0.847	0.049	0.273	-0.440
Philippines	0.031	0.210	-0.258	0.079	1.626	-1.311	0.016	0.169	-0.333
Singapore	0.083	0.259	-0.209	0.104	1.188	-1.233	0.077	0.200	-0.142
Malaysia	0.050	0.193	-0.192	0.053	4.483	-4.128	0.044	0.162	-0.408
Thailand	0.072	0.275	-0.173	0.088	1.633	-1.822	0.043	0.151	-0.356
Panel Data	0.064	0.486	-0.026	0.080	4.483	-4.128	0.050	0.273	-0.440

Notes: The growth rates are calculated by the first-differences of the logarithmic values of variables.

Therefore, I need to continue to apply the ADF unit root test on the first-difference series. I find that all first-difference series are stationary at the 1% level of significance. This also implies that there is no need to apply the Johansen cointegration test because no country has non-stationary level series of logs of exports, FDI, and GDP at the same time. Thus, the tests suggest the use of stationary first-difference series in the VAR model for causality tests, and they are summarized in Table 3. Consistent results are also obtained from the sample with the time period between 2005 to 2018 such that the same methodology can be used as well.

VAR Granger Causality Tests

Table 4 presents the results for VAR(2) Granger Causality Tests. The Granger causality relations are examined using the Wald test of coefficients (F-test) and each null hypothesis is stated in the footnote of the table. Specifically, the table only reports the causality directions corresponding to statistically significant wald tests of coefficients at the 10% level.

For China, there is bidirectional causality between FDI and exports and unidirectional causality from GDP to FDI. This indicates that in these decades, the large volumes of GDP and exports in China attracted FDI inflow. It in turn led to a greater level of GDP. This can be explained by China's early export-led-growth policy which promoted its economy to a scale that was large enough to attract significant vertically-oriented FDI.

Furthermore, it is interesting to find that while there were no causal relations between these three variables of interest between 1986 to 2004 (Hsiao & Hsiao, 2006), I do observe some causal relations when this time period is extended to 2018. For Korea, there is unidirectional causality from GDP to exports. This shows that the level of Korea's GDP has already been great enough to attract exports. For the Philippines, I have found a unidirectional causality from FDI to export growth, from FDI to GDP, and from exports to GDP. Recall what was shown by Figure 2: the Philippines experienced great economic growth after 2004. The result reveals that FDI was a main contributor to such growth and exports could be one channel to lead to this relationship. For Malaysia, there is unidirectional causality from GDP to exports. Malaysia's GDP grew at a faster rate compared to its exports. Thus, we can imply that the large scale of GDP in Malaysia promoted export growth.

For the other two economies, new causal relationships, which were not found by VECM(2) model examining the time period between 1986 to 2004, are also observed. For Singapore, there exists a unidirectional causality from exports to GDP and from FDI to GDP. Between 1986 to 2004, there was a bidirectional causality between FDI and exports and a unidirectional causal relation from GDP to FDI. Thus, we can tell that in the early decades, the scale of Singapore's economy and exports attracted FDI, and this increase in FDI inflow promoted export growth. With years of accumulation of FDI and exports, they reached a level large enough to promote economic growth in Singapore. For Thailand,

Table 4

Granger Causality Test: Six Southeast Asian Economies, 1986-2018

	Vector Autoregression: VAR(2)							Wald Test of Coefficients		Causality Direction	Wald Test of Coefficients		Causality Direction
	Dependent	dex(-1) c1	dex(-2) c2	dfdi(-1) c3	dfdi(-2) c4	dgdp(-1) c5	dgdp(-2) c6	H_0:	F-stat		H_0:	F-stat	
China	dex	0.222	0.177	-0.041	0.176	-1.762	1.347	C	9.139	gdp → ex ***	B	3.579	
		(0.26)	(0.33)	(0.66)	(0.06)	(0.01)	(0.00)		(0.01)			(0.17)	
	dfdi	0.366	0.734	0.284	-0.365	-2.099	0.103	C	5.20	gdp → fdi **	A	5.366	ex → fdi **
		(0.30)	(0.02)	(0.09)	(0.03)	(0.08)	(0.90)		(0.07)			(0.07)	
	dgdp	0.117	-0.012	-0.025	0.055	0.412	0.217	B	3.557		A	4.166	
		(0.07)	(0.83)	(0.39)	(0.07)	(0.05)	(0.15)		(0.17)			(0.13)	
Korea	dex	0.135	0.214	-0.020	-0.064	0.142	-0.541	C	11.06	gdp → ex ***	B	2.241	
		(0.38)	(0.00)	(0.67)	(0.16)	(0.53)	(0.27)		(0.00)			(0.33)	
	dfdi	0.143	0.510	0.004	-0.484	-0.806	-0.723	C	2.885		A	0.529	
		(0.86)	(0.49)	(0.98)	(0.00)	(0.19)	(0.26)		(0.24)			(0.77)	
	dgdp	-0.225	0.354	0.034	-0.028	0.180	-0.407	B	0.494		A	2.537	
		(0.42)	(0.16)	(0.59)	(0.64)	(0.39)	(0.07)		(0.78)			(0.28)	
Philippines	dex	-0.234	-0.048	0.087	0.010	0.215	-0.234	C	1.088		B	11.732	fdi → ex ***
		(0.31)	(0.84)	(0.00)	(0.72)	(0.39)	(0.37)		(0.58)			(0.00)	
	dfdi	-0.376	2.856	-0.320	-0.253	2.764	-2.369	C	3.836		A	3.844	
		(0.79)	(0.05)	(0.04)	(0.15)	(0.07)	(0.14)		(0.15)			(0.15)	
	dgdp	-0.512	-0.195	0.068	0.037	0.441	0.102	B	8.974	fdi → gdp **	A	7.558	ex → gdp **
		(0.02)	(0.37)	(0.00)	(0.16)	(0.06)	(0.67)		(0.01)			(0.02)	
Singapore	dex	-0.341	0.314	0.023	-0.050	1.023	-0.704	C	4.295		B	3.878	
		(0.27)	(0.30)	(0.56)	(0.20)	(0.05)	(0.18)		(0.12)			(0.14)	
	dfdi	1.517	0.775	-0.660	-0.531	-0.665	-1.645	C	0.730		A	1.276	
		(0.33)	(0.61)	(0.00)	(0.01)	(0.80)	(0.53)		(0.69)			(0.53)	
	dgdp	-0.163	0.198	0.004	-0.058	0.765	-0.226	B	8.417	fdi → gdp **	A	1.646	ex → gdp *
		(0.40)	(0.30)	(0.87)	(0.02)	(0.02)	(0.49)		(0.02)			(0.44)	
Malaysia	dex	1.015	-0.137	-0.025	-0.017	-0.607	0.044	C	5.950	gdp → ex **	B	2.537	
		(0.00)	(0.67)	(0.15)	(0.25)	(0.02)	(0.88)		(0.05)			(0.28)	
	dfdi	-0.083	0.715	-0.556	-0.343	-0.143	-1.204	C	0.138		A	0.041	
		(0.99)	(0.86)	(0.01)	(0.07)	(0.97)	(0.74)		(0.93)			(0.98)	
	dgdp	0.662	-0.204	-0.031	-0.013	-0.170	-0.000	B	2.380		A	3.291	
		(0.09)	(0.59)	(0.13)	(0.45)	(0.57)	(1.00)		(0.30)			(0.19)	
Thailand	dex	-0.179	0.328	0.004	0.009	0.542	-0.498	C	11.88	gdp → ex ***	B	0.134	
		(0.45)	(0.14)	(0.88)	(0.72)	(0.01)	(0.03)		(0.00)			(0.94)	
	dfdi	1.022	3.089	-0.600	-0.247	-2.005	-0.494	C	2.502		A	5.029	ex → fdi *
		(0.50)	(0.03)	(0.00)	(0.12)	(0.12)	(0.73)		(0.29)			(0.08)	
	dgdp	-0.132	0.557	-0.022	-0.026	0.480	-0.574	B	0.743		A	4.243	
		(0.66)	(0.05)	(0.49)	(0.42)	(0.06)	(0.04)		(0.69)			(0.12)	

Notes:
1. The p-value is in the parenthesis.
2. *** (**, *) denotes rejection of null hypothesis at the 1% (5%, 10%) level of significance, respectively.
3. In Wald test of coefficients, the null hypothesis A is c1 = c2 = 0, B is c3 = c4 = 0, and C is c5 = c6 = 0, respectively.

I have found a unidirectional causality from GDP to exports and from exports to FDI. However, the 1986 to 2004 analyses found a bidirectional causality between exports and GDP. My observation tells us that exports no longer play a role as crucial as they used to be to promote the economy. Instead, the scale of the economy in Thailand started to attract export growth, and it in turn promoted FDI.

Combining the above results and results in Hsiao & Hsiao (2006)], there are roughly five stages of development for the economies. At the beginning of the development, there is no significant association between these three variables in Korea and Malaysia in Hsiao & Hsiao (2006). With years of development, when the GDP reaches a certain level, it will start to promote exports as Thailand in Hsiao & Hsiao (2006) and Korea in my results. Then, the FDI inflow will start to promote economic growth through the channel of export growth when the economies have a certain level of exports as in China in Hsiao & Hsiao (2006) and the Philippines in my results. When the GDP is accumulated to a large scale, it will start to attract FDI inflow as Singapore in Hsiao & Hsiao (2006) and China in my results, and FDI inflow will in turn promote economic growth and export growth as Singapore in my results.

Table 5 presents the results for the structural break tests. There was no significant change in causal relationships in Korea, the Philippines, Singapore, and Malaysia. However, it is surprising to observe a significant change in China and Thailand. In China, while both Hsiao & Hsiao (2006) and my results reveal that FDI, GDP, and exports are interdependent in a complex way, there is no causality among these three variables between 2005 to 2018. In recent decades, China no longer replies its economic growth on exports and FDI inflow and has already shifted its policy to stimulate the economy.

In Thailand, the causal relationships among these three variables become much more complicated between 2005 to 2018. There is bidirectional causality between GDP and exports and between FDI and exports. Also, there is unidirectional causality from FDI to GDP. This is consistent with the observations from Section 3 that the patterns of GDP and exports were nearly the same after 2005.

Panel Data Granger Causality Tests

As explained in Section 4, these six countries in East and Southeast Asia have similar backgrounds and are at similar stages of development. Thus, I pool their six cross-sectional data over 33 years (1986 to 2018) into a panel dataset and examine the causal relations using panel data regressions.

Table 5

Granger Causality Test: Six Southeast Asian Economies, 2005-2018

	Vector Autoregression: VAR(2)							Wald Test of Coefficients		Causality Direction	Wald Test of Coefficients		Causality Direction
	Dependent	dex(-1) c1	dex(-2) c2	dfdi(-1) c3	dfdi(-2) c4	dgdp(-1) c5	dgdp(-2) c6	H_0:	F-stat		H_0:	F-stat	
China	dex	0.059	0.498	0.250	-0.046	-1.776	0.743	C	2.357		B	1.574	
		(0.91)	(0.15)	(0.26)	(0.79)	(0.17)	(0.51)		(0.31)			(0.46)	
	dfdi	0.844	0.764	-0.322	-0.617	-1.172	1.696	C	0.412		A	2.24	
		(0.51)	(0.36)	(0.55)	(0.14)	(0.71)	(0.54)		(0.81)			(0.33)	
	dgdp	0.183	0.012	-0.004	-0.028	0.494	0.093	B	0.275		A	1.645	
		(0.27)	(0.91)	(0.95)	(0.60)	(0.22)	(0.80)		(0.87)			(0.44)	
Korea	dex	-0.030	0.571	0.013	-0.040	0.866	-1.183	C	17.90	gdp → ex ***	B	0.757	
		(0.91)	(0.02)	(0.80)	(0.42)	(0.01)	(0.00)		(0.00)			(0.69)	
	dfdi	-0.482	-0.011	-0.182	-0.575	-0.389	0.301	C	0.115		A	0.209	
		(0.65)	(0.99)	(0.41)	(0.01)	(0.77)	(0.83)		(0.94)			(0.90)	
	dgdp	-0.311	0.413	0.048	0.036	0.642	-0.717	B	1.856		A	5.766	fdi → gdp *
		(0.14)	(0.04)	(0.27)	(0.38)	(0.01)	(0.01)		(0.40)			(0.06)	
Philippines	dex	-0.452	-0.217	0.128	0.010	-0.130	-0.546	C	3.537		B	35.73	fdi → ex ***
		(0.14)	(0.22)	(0.00)	(0.71)	(0.67)	(0.17)		(0.58)			(0.00)	
	dfdi	-1.811	2.525	-0.272	-0.067	2.107	-7.373	C	2.932		A	1.723	
		(0.59)	(0.20)	(0.27)	(0.82)	(0.52)	(0.09)		(0.23)			(0.42)	
	dgdp	-0.120	-0.275	0.094	-0.027	0.715	-0.157	B	15.14	fdi → gdp ***	A	2.695	
		(0.71)	(0.15)	(0.00)	(0.33)	(0.03)	(0.71)		(0.00)			(0.26)	
Singapore	dex	-0.694	0.561	0.087	-0.024	1.709	-0.986	C	4.719	gdp → ex *	B	2.917	
		(0.10)	(0.18)	(0.16)	(0.66)	(0.03)	(0.26)		(0.09)			(0.23)	
	dfdi	2.891	2.854	-0.525	-0.872	-1.095	-5.816	C	3.248		A	4.643	ex → fdi *
		(0.11)	(0.11)	(0.05)	(0.00)	(0.75)	(0.13)		(0.20)			(0.10)	
	dgdp	-0.192	0.376	0.033	-0.071	0.963	-0.580	B	9.491	fdi → gdp ***	A	3.730	
		(0.40)	(0.10)	(0.33)	(0.02)	(0.03)	(0.22)		(0.01)			(0.16)	
Malaysia	dex	0.217	0.506	-0.034	0.010	0.406	-1.284	C	5.950		B	3.219	
		(0.67)	(0.35)	(0.80)	(0.55)	(0.58)	(0.09)		(0.23)			(0.20)	
	dfdi	-1.544	3.336	-0.655	-0.290	2.210	-8.675	C	0.386		A	0.111	
		(0.87)	(0.74)	(0.08)	(0.34)	(0.87)	(0.55)		(0.83)			(0.95)	
	dgdp	0.209	0.257	-0.031	-0.013	0.386	-0.561	B	3.466		A	0.658	
		(0.65)	(0.61)	(0.08)	(0.38)	(0.56)	(0.43)		(0.18)			(0.72)	
Thailand	dex	-0.875	1.362	-0.054	0.015	2.156	-2.382	C	20.08	gdp → ex ***	B	8.197	fdi → ex **
		(0.04)	(0.00)	(0.05)	(0.56)	(0.00)	(0.00)		(0.00)			(0.02)	
	dfdi	-3.174	9.417	-0.982	-0.449	5.549	-8.524	C	2.743		A	7.444	ex → fdi **
		(0.37)	(0.02)	(0.00)	(0.03)	(0.30)	(0.18)		(0.25)			(0.02)	
	dgdp	-0.643	1.235	-0.051	-0.006	1.568	-1.902	B	9.913	fdi → gdp ***	A	23.62	ex → gdp ***
		(0.03)	(0.00)	(0.01)	(0.73)	(0.00)	(0.00)		(0.01)			(0.00)	

Notes:
1. The p-value is in the parenthesis.
2. *** (**, *) denotes rejection of null hypothesis at the 1% (5%, 10%) level of significance, respectively.
3. In Wald test of coefficients, the null hypothesis A is c1 = c2 = 0, B is c3 = c4 = 0, and C is c5 = c6 = 0, respectively.

Panel Data Unit Root Tests

Similarly, I first test the stationarity of the three level series. Specifically, this article adopts IPS W-test by Im et al. (2003) and the ADF-Fisher Chi-square test by Maddala & Wu (1999).

Table 6

Panel Data Unit Root Tests, 1986-2018

	Panel Level Series			**Panel First-difference Series**	
	IPS W-stat	ADF-Fisher Chi-square		IPS W-stat	ADF-Fisher Chi-square
ex	-0.291	(0.39)	dex	-9.576***	130.702***
	(0.39)	(0.68)		(0.00)	(0.00)
fdi	-4.594***	43.42***	dfdi	-14.754***	250.725***
	(0.00)	(0.00)		(0.00)	(0.00)
gdp	-1.47*	8.755	dgdp	-7.615***	90.155***
	(0.07)	(0.72)		(0.00)	(0.00)

Notes:
1. In level series, the test equation includes individual effects and individual linear trends.
2. In the first-difference series, the test equation includes individual effects.
3. Lag is selected by the minimum AIC with maximum Lag = 3. The p-value is in the parenthesis.
4. *** (**, **) denotes rejection of the null hypothesis that panel series has a unit root at the 1% (5%, 10%) level of significance, respectively.

Table 6 presents the results for the panel data unit tests. The results demonstrate that the panel-level series exports are non-stationary, the panel-level series FDI is stationary, and there is a mixed result for GDP. However, both tests indicate that all three panel first-difference series are stationary. Thus, we should conduct panel data VAR Granger causality tests using panel first-difference series.

The unit root test results for the period between 1984 to 2004, from 2005 to 2018, and from 2008 to 2018 are consistent such that we can conduct panel data VAR Granger causality tests for each time period as well. The results are not presented due to the page limit.

Panel Data VAR Granger Causality Tests

I first look at the time period between 1984 to 2004 and observe a bidirectional causality between GDP and FDI, which is different from the results in Hsiao & Hsiao (2006). Besides that my data is not exactly the same as that of Hsiao & Hsiao (2006), I was not able to replicate their panel data tests for two other reasons. First, my panel doesn't include Hong Kong and Taiwan. Thus, it is natural that when the panel is different, the results from different panels are different. Second, the Dumitrescu-Hurlin test works differently compared to the random effects. It is not surprising that different approaches generate different results.

Table 7 presents the results for panel data Granger causality test. There is only one unidirectional causality from exports to FDI. This is a surprising finding because Hsiao & Hsiao (2006) found out that there was a bidirectional causality between GDP and exports, and there was a unidirectional causality from FDI to both GDP and exports when they looked at the data from 1986 to 2004 and my replication also observes a bidirectional causality between GDP and FDI.

If all results are valid, our result implies that with 15-year development, these six economies have already developed to a certain stage where FDI is no longer crucial to promote the economy and exports. At this stage of development, the existing large scale of exports attracts FDI inflows to these six economies.

Nonetheless, this result indicates that one possible hypothesis raised by Hsiao & Hsiao (2006): "FDI's reinforcing effects on GDP through exports" might no longer hold in the most recent decade. Since it is not clear whether FDI promotes export growth in most countries in the world or not, we do not have enough evidence to say that export growth is the channel from FDI to GDP in the 2010s.

However, when I take a closer look at the time period between 2005 and 2018, I noticed that causal relationships appear. There are bidirectional causal relationships between GDP and FDI and between export growth and GDP. In addition, there is unidirectional causality from FDI to export growth. Due to the page limit, I did not report the structural breaks of the 2008 Great Recession. However, between 2008 to 2018, there is no longer a unidirectional causality from FDI to GDP. Thus, this implies that the bidirectional relationships between GDP and exports, the unidirectional causality from GDP to FDI, and the unidirectional causality from FDI to exports are robust.

Here, we can learn that compared to Hsiao & Hsiao (2006)], the only difference is that the unidirectional causality from FDI to GDP is reversed in this time period. In recent decades, the large scale of GDP in East and Southeast Asia started to attract FDI inflow and exports. Also, FDI could promote export growth and in turn stimulate economic growth. Thus, while FDI no longer plays as crucial of a role as exports to promote economic growth, it still exists as an implicit channel to stimulate the economy.

Limitations

One possible limitation of this article is that the 1997 Asian financial crisis and the 2008 Great Recession are not considered in the models. This is only valid when all six economies recover quickly from these two events. Hsiao & Hsiao (2006) have revealed that the 1997 Asian financial crisis has negligible effects on the causality analyses and Section 4

Table 7

Panel Data Granger Causality Test

		Wald Test of Coefficient		Causality Direction	Wald Test of Coefficient		Causality Direction
		H0:	F-stat		H0:	F-stat	
1986-2018	dex	C	2.073 (0.92)		B	1.529 (0.49)	
	dfdi	C	2.474 (0.03)	gdp → fdi**	B	1.476 (0.55)	
	dgdp	B	1.695 (0.35)		A	1.933 (0.80)	
1986-2004	dex	C	0.161 (0.87)		B	-0.524 (0.60)	
	dfdi	C	6.089 (0.00)	gdp → fdi***	B	-0.991 (0.32)	
	dgdp	B	1.736 (0.08)	fdi → gdp*	A	0.063 (0.95)	
2005-2018	dex	C	3.573 (0.00)	gdp → ex***	B	4.701 (0.00)	fdi → ex***
	dfdi	C	-1.708 (0.09)	gdp → fdi**	B	-1.212 (0.23)	
	dgdp	B	2.404 (0.02)	fdi → gdp**	A	1.951 (0.05)	ex → gdp*

Notes:
1. Lag is selected by the minimum AIC with maximum Lag = 2. The p-value is in the parenthesis.
2. *** (**, *) denotes rejection of the null hypothesis that panel series has a unit root at the 1% (5%, 10%) level of significance, respectively.
3. In Wald test of coefficients, the null hypothesis follows denotations in Table 4 that A is $c1 = c2 = 0$, B is $c3 = c4 = 0$, and C is $c5 = c6 = 0$, respectively.

has explained the 2008 Great Recession doesn't affect the causality analyses significantly as well through structural break tests. However, a wiser way to incorporate these two events in the causality tests would further improve the validity of the results.

Moreover, regarding the panel data Granger causality tests, the sample size is relatively small in two ways. First, based on Hsiao & Hsiao (2006), this sample size only allows us to choose between VAR(1) and VAR(2) models. Thus, we are unable to capture the variations provided by a greater lag order. Second, the number of six economies is too small as a panel. Dumitrescu-Hurlin test allows for two types of panels: micro-panels, with large N and small T, and macro-panels, with large N and large T, and doesn't leave space for panels with small N, which is the panel with six economies in this article.

This is one way to explain the next limitation: there is only one unidirectional causality from exports to FDI at the 5% level of significance. In addition, as explained in Section 4.1.2, the large scale of the economy in China has started to attract FDI, but countries like the Philippines, are still at the stage when FDI promotes economic growth. Thus, it is hard to observe a shared Granger causality across the panel given the small number of countries included in the panel. However, there are complicated causal relationships among these three variables when we only look at the time period between 2005 and 2018. Thus, another reason that there is only one causality is that the causality relationships change over time. Thus, it is more difficult to find a shared causal relationship over a longer period of time.

Finally, it will be helpful to discuss the impulse response function. It will explain the evolution of other variables in reaction to the shock of one variable. However, due to the page limit, it is not presented in this article.

Conclusion

In conclusion, this article examines the Granger causality between FDI, exports, and GDP among six economies in East and Southeast Asia: China, Korea, Philippines, Singapore, and Malaysia between 1986 to 2018. In addition, this article compares its results to the results provided by Hsiao & Hsiao (2006) which examines the same countries between 1986 to 2004. The article finds that the causal relations between these three variables of interest do shift at both national and aggregate levels.

For China, due to its export-led growth policy, there used to be a bidirectional causality between FDI and GDP and a unidirectional causal relation from exports to GDP. However, with years of accumulation of the economy, China's large-scale of economy and exports start to attract FDI inflows such that there is a bidirectional causality between FDI and exports and unidirectional causality from GDP to FDI.

In addition, in some countries such as Korea and the Philippines, no causal relations were found. However, this article found that with the inclusion of 15 more years in the sample, the causality from FDI to exports and GDP starts to exist. Thus, the FDI inflows started to play an important role to promote economies in many countries in recent decades, while their economies were not influenced by FDI significantly back in the twentieth century.

The article finds that for most economies, there are five stages of development. There were no causal relationships among these three variables at the initial stage. After years of development, when they have a certain level of the economy, GDP will start to promote export growth. Then, FDI will start to promote economic growth through the channel of export growth. Afterward, when the scale of GDP is large enough, it will start to attract FDI inflow and promote export growth. Finally, when FDI reaches a large scale, it will promote both economic growth and export growth.

At the aggregate level, the causal relations used to imply the "FDI's reinforcing effects on GDP through exports." However, the related causalities are no longer observed, and there is unidirectional causality from exports to FDI. Thus, with the development of the economy, the causal relations between FDI, exports, and GDP also change over time, and when we look at a long period of time, it is difficult for us to find a shared causality.

References

Carkovic, M. & Levine, R. (2002). Does foreign investment accelerate economic growth. *University of Minnesota Department of Finance Working Paper*.

De Mello Jr, L. R. (1997). Foreign direct investment in developing countries and growth: A selective survey. *The Journal of Development Studies*, 34(1), 1–34. Publisher: Taylor & Francis.

Dumitrescu, E.-I. & Hurlin, C. (2012). Testing for Granger non-causality in heterogeneous panels. *Economic Modelling*, 29(4), 1450–1460. Publisher: Elsevier.

Hansen, H. & Rand, J. (2006). On the causal links between FDI and growth in developing countries. *World Economy*, 29(1), 21–41. Publisher: Wiley Online Library.

Hsiao, F. S. & Hsiao, M. (2006). FDI, exports, and growth in East and Southeast Asia: Evidence from time-series and panel data causality analyses. *international conference on Korea and the World Economy V, Seoul, Korea*.

Hsiao, F. S. & Hsiao, M.-C. W. (2001). Capital flows and exchange rates: recent Korean and Taiwanese experience and challenges. *Journal of Asian Economics*, 12(3), 353–381. Publisher: Elsevier.

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1), 53–74. Publisher: Elsevier.

Maddala, G. S. & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and statistics*, 61(S1), 631–652. Publisher: Wiley Online Library.

Nair-Reichert, U. & Weinhold, D. (2001). Causality tests for cross-country panels: a New look at FDI and economic growth in developing countries. *Oxford bulletin of economics and statistics*, 63(2), 153–171. Publisher: Wiley Online Library.

Impact of Refugee Arrival on Norwegian Employment

Mahmoud Majdi
ECON 381: Introduction to Econometrics (Advisor: Amy Damon)

The tragedy that refugees face all over the world is two-faceted. On the one hand, they suffer from the loss of all that is familiar to them when they are displaced. On the other hand, they are stigmatized and dehumanized by their host communities due to rising populist rhetoric. The most recent instance of such anti-immigrant discourse is the rejection of an immigrant ship by the new far-right Italian government in violation of what international law dictates (Beardsley, 2022). Unfortunately, the list of such occurrences is long. The Syrian refugee crisis in the Middle East and Europe, the Rohingya crisis in Myanmar and neighboring countries, and the Venezuelan refugee crisis in the Americas are all examples of phenomena that have fueled the same talking point by every right-leaning politician: refugees and immigrants occupy the jobs that nationals would have gotten, worsening the host community's welfare.

The immigration debate is perhaps as contentious in economic circles as it is in political ones. According to economic models, the effect of a labor supply increase in labor markets corresponds to a decrease in wages, and a more competitive labor market leads to a decrease in employment. This is a short-run scenario when labor is the only mobile factor of production. In the long-run, as capital has enough time to adjust to the additional influx of labor, there is no impact on wages and productivity will increase. So far, empirical studies testing these theories disagree in their findings. My task throughout this paper will be to explore the impact of refugee arrival on natives' employment levels using Norway as a case study.

Migrant influxes to high-income countries became fertile ground for examining the impact of migrants on host labor markets, and the results vary depending on the scope and focus of each conducted study. Peri (2011) as well as Mayda et al. (2017) authored a paper that found no evidence supporting adverse effects to employment and wages caused by refugees and immigrants in the long-term with respect to any skill group in the US. Card (1990), who studied the short-run impact of the mass migration of Cubans into Miami in the

Mahmoud Majdi is a junior with a major in Economics, and minors in Mathematics and Music. Correspondence concerning this article should be addressed to mmajdi@macalester.edu.

1980s, also found that the city was able to absorb the large influx of migrants in a relatively short time and its labor market wages and unemployment rates remained unchanged for low-skilled workers particularly.

Internationally, some of the findings are the same. Venturini & Villosio (2004) posit that in the North of Italy, the place where immigrants are mostly concentrated, native workers overall did not experience any significant negative effects on employment, and in Denmark, it was shown that low-skilled immigration increased occupational mobility for native counterparts as less-educated workers took on more complex jobs with higher wages (Foged & Peri, 2016). Opposingly, after compiling evidence from four historical refugee influxes into Europe, the US, and Israel, Borjas & Monras (2016) show that exogenous labor supply shocks have adverse effects on same-skilled workers in the host economy and positive effects on complementary native workers. In other words, their findings support the theory that immigration harms native workers who are considered substitutes and benefits complementary workers through increasing demand for their labor. This pattern, according to Angrist & Kugler (2003), is intensified when labor laws are more stringent, meaning that if there are high costs for employers with regards to the hiring and firing process, the impact of a larger labor force will affect same-skill-level-employment more negatively than if those laws were more flexible. In general, there remains debate over the real effect of migration on high-income countries' labor markets.

While the literature concerned with high-income countries focuses on workers in the formal sector of the economy, studies conducted in low and middle-income countries deal with the informal sector as well, given its preponderance in the Global South, not to mention that refugees and migrants with no work permits tend to work in the informal labor market. Syrian refugees in Turkey have had a negative impact on the Turkish informal labor market primarily in terms of increased unemployment and decreased labor force participation, especially among women, young workers and low-skilled workers (Ceritoglu et al., 2017). The same refugee crisis in Lebanon led to higher unemployment among low-skilled Lebanese workers and did not affect Lebanese workers in the high-skill group (David et al., 2019). Further evidence on the impact of migrants on native workers' outcomes is available in studies about Cameroon, Ghana, South Africa (Viseth, 2020), and Venezuela (Olivieri et al., 2022), where impact was established and the importance of substitutability and complementarity in determining the direction of the impact is reiterated.

Another major variation in the literature is the methodological approach each study takes in order to avoid selection bias in the samples. This bias comes from the assumption that immigrants move to places where wages are already high, and where employment is available. In order to address this bias, some studies resort to correctional methods using fixed effects regressions. Others move towards "quasi-experimental" techniques by which they observe a forced migration such as mass repatriation or persons seeking asylum. The argument made here is that when immigrants do not get to choose where to settle but are rather randomly assigned to certain areas by a greater force, one can effectively address the selection bias hindering accurate results. Foged & Peri (2016) clearly illustrate this approach by using a refugee dispersal policy between 1986 and 1998 enacted by the Danish Refugee Council, which allocated refugees into specific municipalities based only on their nationality and family size, not accounting for any municipality-specific labor market conditions. Such a policy represents an exogenous shock that corrects for the reverse causality between labor market conditions and refugee share in a municipality and isolates the impact of refugees on the host community such that it can be accurately measured. All the studies listed above fall into one category or the other in terms of their empirical approach.

My paper will use a "quasi-experimental" approach similar to the one in the Foged & Peri (2016) paper, as well as a fixed effects regression. I will rely on Norway's refugee resettlement policy of 2002 to measure the effect refugees have on native Norwegians' employment levels (Hernes et al., 2019). In other words, I will be looking at **how refugee arrival in Norway impacted natives' employment levels**. By doing so, I will add another case study to the high-income country category of papers exploring the same question, thereby providing more evidence for the on-going debate. Furthermore, since a study of the same phenomenon exists in a Danish context, arguably quite similar to the Norwegian one, it will be meaningful to compare the results of the two neighboring case studies.

Economic Theory: Factor Proportions Model

The model used in any discussion about the effect of immigration on the labor market outcomes of the host country is the "Factor Proportions" model. The typical assumptions of this model are: 1) immigrant and native workers are perfect substitutes for each other within the same skill-level, 2) labor supply in the host economy is perfectly inelastic, meaning that people will work at any wage level, 3) capital supply is perfectly elastic, implying that capital can be freely manipulated depending on the needs of the economy (Sarzin, 2021).

One of the most simplified versions of the model, and the one most relevant to this study, illustrates an economy with a single industry where production is achieved through a combination of labor and capital. Keeping our assumption in mind, an increase in labor supply will not affect overall employment and average wages since firms will simply increase capital such that the capital to labor ratio is equal to what it was prior to immigration. However, since the labor force is composed of high-skilled and low-skilled workers, the relative change in wages for each skill group will depend on the distribution of skill levels among the newcomers. In other words, if the new workers are high-skilled, they will lower the wages of native high-skilled workers since they are perfect substitutes to each other.

The scenario is the same for low-skilled workers. Nevertheless, if the skill-distribution of immigrants is the same as that of native workers, there will be no impact on wages because the relative sections of the labor force will proportionally increase (Sarzin, 2021).

Data Description

All the datasets I use for my paper come from Statistics Norway, which is the leading provider of national statistics in the country. The statistics center contains multi-level data ranging from national economic indicators to labor market data. The center reports to the Norwegian Ministry of Finance but has full control and autonomy over the data publishing process (Statistics Norway, 2022b). I specifically use three county-level datasets: 1) number of people at the county level, 2) number of people with a refugee background, and 3) the number of employed persons per county in different industry groups as well as their educational attainment, all of which were collected on an annual basis. Additionally, I include a dataset with value added in current prices for each county in Norway as a control to be used in my regressions.

The statistics center's definition of a person with a refugee background is "people with refugee as reason for immigration, as well as immigrants with family as reason for immigration who are reunited with a person with reason refugee" (Statistics Norway, 2022c). This is important to note since it poses a problem to the empirical strategy of my paper which relies on a "quasi-experimental" approach made possible by a change in the refugee settlement policy of Norway. I will address this in a later section in this paper.

The datasets that I chose contain observations for different time spans. For example, the most expansive dataset I have is the county population dataset which holds records from 1986 to 2021. The statistics center, however, only began recording the number of refugees in a given county in 1998 with a two-year gap between 2011 and 2013. I ran into a similar issue when I examined the rest of the datasets I am interested in. Moreover, the data published by Statistics Norway switched the data source their Labor Force Survey was based on beginning from the year 2015, making comparisons with subsequent years difficult. In addition to that, the country underwent a territorial rearrangement where certain municipalities were assigned to different counties in 2019, once again complicating any attempts to compare the years following this rearrangement to previous observations (Statistics Norway, 2022a). For those reasons, I decided to restrict the data analysis to the period between 2000 and 2014. This decision was also motivated by the need to have observations in years prior to and succeeding 2002, which is the year in which Norway adopted a different refugee settlement policy that I will explain further in the empirical strategy section of this paper.

Finally, it is worth noting that Svalbard, which is included as a county in the dataset

I use in the paper is a special territory in Norway. The island is not part of mainland Norway and is not officially a county. Nevertheless, it is treated as such in official statistics (Statistics Norway, 2022a). Because most of the values recorded for Svalbard are "0" for many of the important variables, I chose to drop the region from the dataset.

Table 1

Descriptive Summary of Dependent Variables

Variables	N	Mean	S.D
% Empl in Agri	133	1.798	1.083
% Empl in Mining	133	0.987	1.138
% Employed in Manu	133	5.061	1.684
% Empl in Utilities	133	0.628	0.175
% Empl in Construction	133	4.201	0.513
% Empl in Retail	133	7.222	0.854
% Empl in Transport	133	2.837	0.467
% Empl in Accomodation	133	1.621	0.291
% Empl in Info	133	1.371	0.932
% Empl in Finance	133	0.812	0.398
% Empl in Real Estate	133	2.685	0.960
% Empl in Admin	133	2.383	0.499
% Empl in Public Admin	133	3.124	0.775
% Empl in Education	133	4.180	0.580
% Empl in Health	133	10.53	0.803
% of Low Skilled Empl	285	34.32	2.954
% of High Skilled Empl	285	14.43	3.438
% of Employed Persons per County	285	50.68	2.143

Notes: The Industry names have been shortened for convenience. The full industry category names are available on the Statistics Norway website. The variables have a different number of observations because data available for industry employment spans a shorter time period than skill-level employment and the total number of employed.

The resulting panel data contains 300 county-level observations with the following dependent variables: 1) Total employment per capita per county, 2) high skilled employment per capita per county measured by the educational attainment of workers [1] , 3) low skilled employment per capita per county, and 4) employment per capita per industry in a given county. The independent variable of interest is the share of new refugees in every county,

[1]The original dataset divides education attainment to: primary and lower secondary, upper secondary, tertiary with 4 years or less after high school, tertiary with more than 4 years after high school, and unknown or no complete education. I aggregated these into high skilled and low skilled by combining the observations for tertiary education into high skilled and the rest into low skilled.

and the control variable is GDP per county (measured as value added in current prices). **Table 1** includes summary statistics for the outcome variables in the dataset [2], and **Table 2** provides a summary of the independent variables. **Figure 1** shows the progression of employment per capita in the counties of Aust-Agder and Akershus over time. These two counties have had the highest average shares of new refugees compared to the rest of the counties in Norway. There isn't much that we can infer from looking at **Figure 1** regarding a correlation between the share of new refugees and employment levels. The line-graph shows typical business cycle fluctuations in employment at the level of the two counties with the unsurprising drop in employment during the years of the Great Recession. I will test if there really is no correlation between employment levels and new refugees in the next section of the paper.

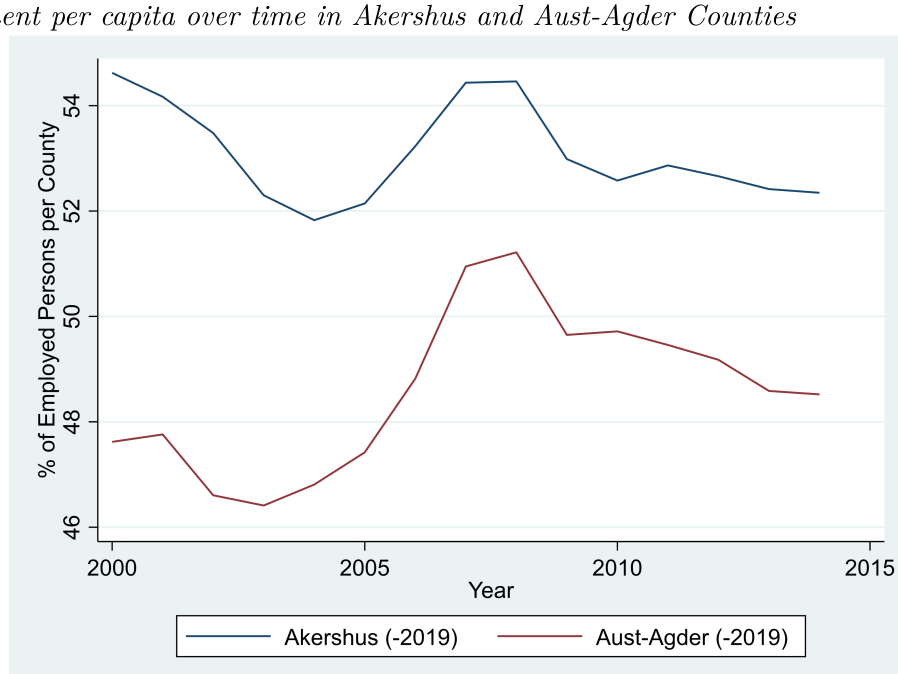
Table 2

Descriptive Summary of Dependent Variables

VARIABLES	N	mean	sd
Value added at basic prices. Current prices (NOK million)	140	97,181	92,810
% of New Refugees per County	228	0.143	0.111

Figure 1

Employment per capita over time in Akershus and Aust-Agder Counties



[2]The reason why employment by industry observations are lower than the skill-level and total employment shares is that I have data that spans a longer period of time for the latter categories.

Empirical Strategy and Concerns

Instrumental Variable Approach

Since there is self-selection in the sample originating from the assumption that refugees generally move to areas where employment is already high, I utilize a Two Stage Least Squares method using an instrumental variable to address this issue. In 2002, Norway adopted a new refugee settlement policy that gave municipalities autonomy in deciding whether to accept refugees or not. After 2002, the decision to settle in an area was no longer that of refugees. Rather, it stemmed from voluntary municipal cooperation with the central authority in charge of settling refugees. Specifically, the central agencies tasked with refugee settlement contact municipalities requesting refugee settlement for the following year. Each municipal council then decides whether they are willing to accept the requested number of refugees (Hernes et al., 2019). I argue that this policy, which is still in place to date, introduced some randomness to the independent variable of interest (share of refugees in a county). I use the share of new refugees in each county in the year 2003 – the year after the policy was enacted – as my instrumental variable to predict the share of refugees in the same counties in the year 2010, with the predicted year chosen arbitrarily. My hypothesis is that the policy introduced in 2002 makes the share of refugees in the following years uncorrelated with employment since the choice is no longer in the hands of the refugees. At the same time, the share of refugees in 2003 is endogenous to the share of refugees in 2010 through the assumption of the existence of networks that refugees have established over time, representing more attractive settlement areas for newcomers as opposed to places with no networks. This makes the share of refugees in 2003 a potential Instrumental Variable to use in my regression. From that, we get as the first stage regression that predicts the share of refugees in 2010:

$$Refugees2010_i = \delta_0 + \delta_1 Refugees2003_i + \delta_2 GDP2010_i + \omega_i \tag{1}$$

The second stage regression uses the predicted share of refugees in a county in the year 2010 and makes it the new independent variable of interest upon which the share of employment in 2010 is regressed. As such, the second stage regression is as follows:

$$Employment2010_i = \beta_0 + \beta_1 \widehat{Refugees}2010_i + \beta_2 GDP2010_i + \epsilon_i \tag{2}$$

Both stages of the Two Stage Least Squares regression are represented in Tables 3 and 4:

The results shown in the tables demonstrate how the instrumental variable chosen for this regression is a weak one. There is an important reason for why this is the case. The

Table 3

First Stage Regression Results

Variables	Percentage of Refugees in 2010
Percentage of Refugees in 2003	0.092
	(0.09)
GDP	-1.13e-07
	(1.80e-07)
Constant	0.166***
	(0.193)
F-value	1.03
Observations	19

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Standard Errors in parentheses.

Table 4

Second Stage Regression Results

Variables	Share of Employment in 2010
Predicted Share of Refugees in 2010	-67.39
	(61.79)
GDP	1.36e-05
	(9.26e-06)
Constant	60.78***
	(10.08)
Observations	19
R-squared	-2.833

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Standard Errors in parentheses.

policy adopted by Norway in 2002 affects the country at the municipal level. The data that I was able to find is at the county level, each county including many municipalities. Therefore, the level of analysis does not match the level at which the decision to settle refugees is made. Additionally, since the regression is a cross sectional, the number of observations is low, which lowers the likelihood of getting results at any level of significance. To address the latter problem particularly, I make use of panel data methods that are explained in the next sub-section.

Fixed Effects Regression

$$E_{it} = \alpha_i + \varphi_t + \beta_1 NewRefugees_{it} + \beta_2 GDP_{it} + \epsilon_{it} \tag{3}$$

where E_{it} is the share of people employed in county i in year t, α_i and φ_t represent the county fixed effects and the year fixed effects respectively. $NewRefugees_{it}$ is the number of new refugees in a given county at a given year, GDP_{it} is the value added in current prices (in millions of Norwegian Kroners) for county i at year t, and ϵ_{it} is the error term at the county level in year t.

The Two-Way Fixed Effects model allows me to account for time and county invariant factors obscuring the real effect of new refugee arrival into counties in Norway. Furthermore, I can use the relative randomness in the allocation of refugees to counties introduced by the settlement policy that the country adopted in 2002 to address some of the reverse causality between the number of new refugees in a county and the change in its employment level. The results from These regressions are presented in the following section.

Results and Limitations

Regression Results and Discussion

Not much can be said about the instrumental variable results for reasons that are discussed in that section. The instrument proved weak and was not able to properly predict my endogenous variable. The most likely reason for this is that when dealing with the data as a cross section, there is a very small number of observations and hence not much room for variation.

As for the Fixed Effects regressions, the results are more interesting. Since my results differ greatly between the regressions that have controls and the same regressions without controls, I will compare the different sets of regressions and discuss the source of this variation. It is important to keep in mind that the story these regressions tell is highly dependent on the assumptions that we make on the nature of the labor market in Norway outlined in the economic theory section.

Table 5 displays employment percentage and skill-based employment as the dependent variables. This regression includes GDP per county as a control and indicates that refugees on all outcomes have a positive and significant impact on employment levels. Specifically, it suggests that the number of new refugees in a county increases overall employment in a county by 2.09%, low-skilled employment by 1.07%, and high-skilled employment by 0.75%. These unconventional results may be due to the imperfect substitutability of refugees and native workers. Refugees may have skills that are different than either high-skilled native workers or low-skilled ones, therefore lowering competition between the groups and increasing overall wages.

The second regression is concerned with the impact of new refugees in a county on that county's employment level by traditionally high-skill industries. **Table 6** shows that results are insignificant, implying that for high-skill industries, refugees do not cause any change in employment levels. This is indeed consistent with theory when we assume that

Table 5

Fixed Effects for Employment per Capita by Skill Level

Variables	% of Employed Persons per County	% of Low Skilled Employees	% of High Skilled Employees
% of New Refugees per County	2.085**	1.067**	0.748**
	(0.793)	(0.480)	(0.286)
Value added at basic prices (Current prices in NOK million)	1.26e-05***	3.43e-06*	1.31e-05***
	(3.67e-06)	(1.88e-06)	(1.34e-06)
Year = 2009	-1.333***	-1.394***	0.291***
	(0.110)	(0.0721)	(0.0296)
Year = 2010	-1.540***	-2.054***	0.365***
	(0.100)	(0.0736)	(0.0415)
Year = 2014	-1.920***	-3.202***	1.814***
	(0.217)	(0.131)	(0.0719)
Constant	51.21***	34.61***	13.48***
	(0.422)	(0.197)	(0.142)
Observations	76	76	76
R-squared	0.862	0.980	0.986
Number of ID	19	19	19
Controls	Yes	Yes	Yes

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Standard Errors in parentheses.

refugees have an overall lower skill level and work in low-skill industries. If the additional labor supply is disproportionately distributed mainly in low-skill industries, it makes sense to find that average employment levels in high-skill industries would not be affected.

Finally, the regression in **Table 7** exploring the relationship between new refugees in a county and employment levels in low-skilled industries shows that there is no significant impact in the construction and transportation industries, but a positive impact on the retail industry of 0.82%. It is difficult to make a conclusion about the overall results of this regression, but the impact within two out of the three industries displayed suggests that even in low-skill employment industries, new refugees have a negligible impact on employment levels in Norway. Once again, this could be attributed to refugees being imperfect substitutes for native workers, but the story is inconclusive in this case.

Running the same regressions without GDP per county as a control reveals certain

Table 6

Fixed Effects for Employment per Capita by Traditionally High Skill Industries

Variables	% Employed in Info	% Employed in Finance	% Employed in Real Estate
% of New Refugees per County	0.0747	-0.0164	-0.0850
	(0.104)	(0.0550)	(0.136)
Value added at basic prices (Current prices in NOK million)	1.26e-06*	-1.79e-06***	1.81e-06**
	(6.82e-07)	(2.63e-07)	(7.51e-07)
Year = 2009	-0.0548***	-0.0259***	-0.0387*
	(0.0173)	(0.00377)	(0.0221)
Year = 2010	-0.0588***	-0.0328***	-0.0641**
	(0.0200)	(0.00597)	(0.0296)
Year = 2014	-0.111***	-0.0702***	0.0187
	(0.0341)	(0.0121)	(0.0466)
Constant	1.296***	1.041***	2.523***
	(0.0735)	(0.0307)	(0.0767)
Observations	76	76	76
R-squared	0.378	0.899	0.500
Number of IDs	19	19	19
Controls	Yes	Yes	Yes

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Standard Errors in parentheses.

inconsistencies relative to the regressions that include it. I include these additional regressions in the appendix. Since each set of regressions is impacted differently by whether or not GDP per county is included as a control, I will not interpret the regressions separately. The overall conclusion from examining these additional regressions is that several stories about the bias that GDP per county causes emerge and can only be addressed hypothetically. If refugees are often settled in high-GDP counties, with the assumption that these counties have higher employment levels, then GDP per county applies an upward bias on our estimate of interest. If, on the other hand, refugees are settled in low-GDP counties, and these counties have lower employment levels, then GDP causes downward bias that underestimates our estimate of the effect of refugees. The most important thing to note about GDP per county, however, is that it represents a significant omitted variable if not included in the regression. Unfortunately, since these additional regressions obfuscate the real impact of refugees on employment overall, I am unable to draw a clear conclusion from

Table 7

Fixed Effects for Employment per Capita by Traditionally Low Skill Industries

Variables	% Empl in Construction	% Empl in Retail	% Empl in Transport
% of New Refugees per County	0.193	0.820***	0.251
	(0.190)	(0.283)	(0.174)
Value added at basic prices. Current prices (NOK million)	9.24e-07	4.17e-07	-4.14e-07
	(7.66e-07)	(1.28e-06)	(9.08e-07)
Year = 2009	-0.147***	-0.340***	-0.140***
	(0.0132)	(0.0325)	(0.0234)
Year = 2010	-0.145***	-0.408***	-0.160***
	(0.0186)	(0.0314)	(0.0243)
Year = 2014	0.104**	-0.775***	-0.261***
	(0.0384)	(0.0654)	(0.0523)
Constant	4.100***	7.525***	3.012***
	(0.0981)	(0.130)	(0.0974)
Observations	76	76	76
R-squared	0.734	0.906	0.777
Number of ID	19	19	19
Controls	Yes	Yes	Yes

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Standard Errors in parentheses.

the data.

Limitations

There are several factors that may threaten internal validity in the paper that I was unfortunately unable to address. The first concern I have is that there may still be some bias in the fixed effects regression that comes from the reasons a municipality in Norway would accept or decline accepting refugees. Although some of the selection into being a refugee in a certain county is mitigated through the resettlement policy of 2002, the municipality's decision could depend on the state of the labor market at the time the refugee is being settled. For example, if a county's labor market is booming in a given year, they might be more inclined to accept refugees as the demand for labor would be high. If this is true, then the new refugee arrivals would be correlated with the error term. Unfortunately, I was unable to find data about how many refugees different municipalities accepted after the introduction of the policy. The existence of this data would have made the results of my regressions stronger since I would have been able to control for the differences in refugee acceptance rates for municipalities.

Another selection into the new refugee variable is that "A Person with a Refugee Background" can include people reunifying with family who are refugees according to Statistics Norway (Statistics Norway, n.d.-c). In the case of reunification, the policy change is useless since the settlement of refugees is certainly not random. People reunified with family who are refugees are part of the sample I used for this research.

Employment data from Statistics Norway going all the way back to 2000 had the number of employed rather than the employment percentage. To obtain this number as a percentage for each county, I would ideally divide the number of employed people by the labor force of the county. I was unable to find the number of people in the labor force for counties, so I was compelled to divide the number of employed people by the county's population, even the non-working age category. This means that the employment percentage in a county would be lower than expected since we are dividing the number of employed people by a larger number than what the labor force numbers would be.

Finally, as briefly discussed in the instrumental variable section of the paper, the data available and analyzed is at the county level, whereas the scope of the settlement policy is at the municipal level. This weakens the random effect that the settlement policy introduces to the regressions since the results are aggregated at the county level.

Conclusion

In this paper, I used several econometric methods to estimate the impact of refugee arrival in Norway on employment levels in different counties in the country. In order to address certain endogeneity issues, I made use of the new refugee settlement policy introduced in 2002 which allowed municipalities to accept or refuse refugees rather than offering refugees themselves that choice. I argued that this policy mitigated some of the selection bias into my sample and made results more accurate. Unfortunately, since results were inconsistent and relied on assumptions that I have little evidence for, one cannot infer a clear impact that refugees have on employment levels in Norway. The main obstacle faced throughout the research is the small sample size. I believe that with more comprehensive data, much more can be achieved with the same research.

References

Angrist, J. D. & Kugler, A. D. (2003). Protective or counter-productive? labour market institutions and the effect of immigration on eu natives*. *The Economic Journal*, 113(488), F302–F331. https://doi.org/10.1111/1468-0297.00136

Beardsley, E. (2022). France accepts a migrant rescue ship as relations sour with Italy. *NPR*. https://www.npr.org/2022/11/11/1136030142/france-italy-migrants-ocean-viking

Borjas, G. J. & Monras, J. (2016). *The Labor Market Consequences of Refugee Supply Shocks*. https://www.nber.org/papers/w22656

Card, D. (1990). The Impact of the Mariel Boatlift on the Miami Labor Market. *INDUSTRIAL AND LABOR RELATIONS REVIEW*, 14.

Ceritoglu, E., Yunculer, H. B. G., Torun, H., & Tumen, S. (2017). The impact of Syrian refugees on natives' labor market outcomes in Turkey: evidence from a quasi-experimental design. *IZA Journal of Labor Policy*, 6(1), 5. https://doi.org/10.1186/s40173-017-0082-4

David, A., Marouani, M. A., Nahas, C., & Nilsson, B. (2019). *The economics of the Syrian refugee crisis in neighboring countries. The case of Lebanon.* https://www.afd.fr/en/ressources/economics-syrian-refugee-crisis-neighboring-countries-case-lebanon

Foged, M. & Peri, G. (2016). Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data. *American Economic Journal: Applied Economics*, 8(2), 1–34. https://www.jstor.org/stable/24739100

Hernes, V., Arendt, J. N., Joona, P. A., & Tronstad, K. R. (2019). *Nordic integration and settlement policies for refugees* (2019:529 ed.). TemaNord. Nordic Council of Ministers. http://urn.kb.se/resolve?urn=urn:nbn:se:norden:org:diva-5576

Mayda, A. M., Parsons, C., Peri, G., & Wagner, M. (2017). The Labor Market Impact of Refugees: Evidence from the U.S. Resettlement Program. *OCE Working Paper*, 2017-04.

Olivieri, S., Ortega, F., Rivadeneira, A., & Carranza, E. (2022). The Labour Market Effects of Venezuelan Migration in Ecuador. *The Journal of Development Studies*, 58(4), 713–729. https://doi.org/10.1080/00220388.2021.1988077

Peri, G. (2011). Rethinking the area approach: Immigrants and the labor market in California. *Journal of International Economics*, 84(1).

Sarzin, Z. (2021). The impact of forced migration on the labor market outcomes and welfare of host communities. *Reference Paper for the 70th Anniversary of the 1951 Refugee Convention*, 1.

Statistics Norway (2022a). *11615: Employeded persons per 4th quarter, by contents, region, level of educa.*

Statistics Norway (2022b). *About Statistics Norway.* https://www.ssb.no/en/omssb/ssbs-virksomhet

Statistics Norway (2022c). *Persons with refugee background.* https://www.ssb.no/en/befolkning/innvandrere/statistikk/personer-med-flyktningbakgrunn

Venturini, A. & Villosio, C. (2004). Labour Market Effects of Immigration: an Empirical Analysis Based on Italian Data. *Centre for Household, Income, Labor and Demographic Economics (CHILD) Working Paper*, 17/2004. https://ideas.repec.org/p/wpc/wplist/wp17_04.html

Viseth, A. (2020). Immigration and Employment: Substitute Versus Complementary Labor in Selected African Countries. *IMF Working Paper*, 2020/149.

Appendix
Additional Regressions

Table 5.1

Fixed Effects for Employment per Capita by Skill Level without GDP per County

Variables	% of Employed Persons per County	% of Low Skilled Employees	% of High Skilled Employees
% of New Refugees per County	0.674	0.683	-0.712*
	(0.921)	(0.536)	(0.392)
Year = 2009	-1.253***	-1.372***	0.374***
	(0.0910)	(0.0640)	(0.0274)
Year = 2010	-1.422***	-2.022***	0.486***
	(0.0967)	(0.0693)	(0.0338)
Year = 2014	-1.512***	-3.091***	2.236***
	(0.173)	(0.0978)	(0.0916)
Constant	52.53***	34.97***	14.85***
	(0.148)	(0.0912)	(0.0497)
Observations	76	76	76
R-squared	0.826	0.979	0.963
Number of ID	19	19	19
Controls	No	No	No

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Standard Errors in parentheses.

Table 6.1

Fixed Effects for Employment per Capita by Traditionally High Skill Industries without GDP per County

Variables	% Empl in Info	% Empl in Finance	% Empl in Real Estate
% of New Refugees per County	-0.0657	0.184**	-0.287***
	(0.0597)	(0.0703)	(0.0935)
Year = 2009	-0.0468***	-0.0373***	-0.0272*
	(0.0139)	(0.00881)	(0.0150)
Year = 2010	-0.0471**	-0.0495***	-0.0473*
	(0.0179)	(0.00684)	(0.0246)
Year = 2014	-0.0702***	-0.128***	0.0770***
	(0.0212)	(0.0156)	(0.0255)
Constant	1.427***	0.852***	2.713***
	(0.0139)	(0.00904)	(0.0190)
Observations	76	76	76
R-squared	0.309	0.781	0.443
Number of ID	19	19	19
Controls	No	No	No

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Standard Errors in parentheses.

Table 7.1

Fixed Effects for Employment per Capita by Traditionally Low Skill Industries

Variables	% Empl in Construction	% Empl in Retail	% Empl in Transport
% of New Refugees per County	0.0899	0.774**	0.297***
	(0.152)	(0.306)	(0.0875)
Year = 2009	-0.141***	-0.337***	-0.142***
	(0.0144)	(0.0308)	(0.0211)
Year = 2010	-0.137***	-0.404***	-0.163***
	(0.0234)	(0.0323)	(0.0181)
Year = 2014	0.133***	-0.761***	-0.275***
	(0.0274)	(0.0526)	(0.0312)
Constant	4.197***	7.569***	2.969***
	(0.0326)	(0.0495)	(0.0188)
Observations	76	76	76
R-squared	0.729	0.905	0.775
Number of ID	19	19	19
Controls	Yes	Yes	Yes

Notes: *, **, and *** indicate significance at the 10, 5, and 1 percent levels. Robust Standard Errors in parentheses.

Land Use Regulations and Local Responses to Labor Demand Shifts

Zak Yudhishthu

ECON 381: Introduction to Econometrics (Advisor: Amy Damon)

In the past 20 years, land use regulations have become an increasingly prominent topic politically, across scholarly research, and particularly within the field of economics. In the 2000s, empirical economists took increasing interest in local land use regulations and their impacts on local housing and labor markets (Glaeser et al., 2003, 2005; Saks, 2008).. At its core, this research focused on the fact that strict land use regulations (such as minimum lot sizes, single-family-only zoning, or building height limits) restrict supply and distort the types of supply available. This literature has since developed extensively, and more recent scholarship has shown how strict land use regulations reduce U.S. GDP (Hsieh & Moretti, 2019; Herkenhoff et al., 2018), increase regional income inequality by slowing regional migration (Ganong & Shoag, 2017), and maintain local racial patterns of segregation (Trounstine, 2020). At the same time, political movements have formed across the country to reduce these regulations at the state and local level, leading to policy changes in places including (but far from limited to) California, Minnesota, North Carolina, and Oregon. The core arguments made by these advocates include explicitly economic ones about increasing housing supply and reducing construction costs related to procedural requirements.

This paper builds on theory from (Moretti, 2011)'s chapter on Local Labor Markets in the Handbook of Labor Economics. Moretti (2011) creates a simple model of spatial equilibrium between localities, where utility is defined as $U = f(\textit{wages} + \textit{amenities} - \textit{rents})$: people aim to maximize the total value of their wages and amenities, while rents are subtracted from that value. If some areas offer more utility than others, people will move to those areas, driving wages down and rents up until utility is in a cross-city equilibrium. In this framework, Moretti (2011) models the incidence of local productivity shocks based on different cross-city supply elasticities of labor and local elasticities of housing supply. In the extreme of a perfectly inelastic housing supply, all of the benefits of a localized productivity shock go to landowners in that city because more employees cannot move in and landowners

Zak Yudhishthu is a junior with a major in Economics and a minor in Music. Correspondence concerning this article should be addressed to zyudhish@macalester.edu.

simply charge higher rents. Hence, the degree to which housing regulation reduces housing supply elasticity is highly important.

Moretti (2011) further extends the spatial equilibrium model to include low and high-skill workers, with the result that a skill-biased rise in demand for high-skill workers can raise housing costs while low-skill workers see no wage gains. Intuitively, I would expect the result here to be displacement: this demand shift will draw in high-earning workers who outcompete low-earning incumbents for the existing housing stock. This is a similar result to Matlack & Vigdor (2008), who model and find empirically that increased incomes at the top end of the earnings distribution will raise housing prices for low-income residents if housing markets are tight (or perhaps, less able to adjust supply in response to demand shocks). A paradigmatic example would be San Francisco, which has seen huge growth in high-skill labor demand, has strict geographic and regulatory housing supply constraints, and has experienced steep rent growth paired with out-migration of low-income residents. There is also a large racial component of this change. In other words, this is an economic model of displacement and gentrification.

In this paper, I seek to measure how the level of land use regulation affects localities' responses to a shift in labor demand. I construct these demand shifts using a shift-share instrumental variable, also known as a Bartik Instrument, which instruments for local employment growth by estimating the expected employment growth if a region's mix of industries had all grown at the national growth rate. Paired with a cross-sectional index of local land-use regulation created by researchers at the Penn Wharton School (Gyourko et al., 2019), I estimate how different levels of regulation affect rent growth, housing permit growth, and — as described above — economic displacement, measured as the change in population with no college degree.

My results are mixed: they only match my theoretical predictions for rent growth among my three outcome variables, but my results all become insignificant when I include controls for geographic development constraints and regional trends.

Literature Review

Economists have tried different strategies to gauge the impact of zoning restrictions on housing affordability. In two widely-cited early studies, the authors sought to calculate a "regulatory tax" by measuring the gap between costs of housing production and sales prices of housing, because in a market without zoning-related barriers to entry, price should equal marginal cost (Glaeser & Gyourko, 2002; Glaeser et al., 2003). However, this model is static, and likely to capture the location-value of land, thus overestimating the regulatory tax (Murray, 2021).

A spatial equilibrium model, as created by Rosen (1979) and Roback (1982) and more accessibly teased out by Moretti (2011), allows us to think through a more dynamic analysis. Because labor will move until utility (as defined by wages, rents, and amenities) is equal across cities, the incidence of a local change in labor demand will depend on elasticities of labor and housing. Of interest to this paper is the model's prediction that lower housing elasticities of supply will lead to slower housing permit change and a sharper price increase in the event of an increase in labor demand.

With spatial equilibrium models, calculate elasticities and their relation to housing market outcomes becomes relevant. Researchers have used different models and identification methods to investigate this topic. Gyourko et al. (2013) models the concept of "superstar" cities, in which some cities have limited supply elasticities but are highly desirable and productive, meaning that national increases in population or income — even without any regionally specific demand shocks — will lead to superstar cities experiencing faster price growth. Howard & Liebersohn (2021) use estimated supply elasticities from Saiz (2010) to model how a shift in housing demand towards lower-elasticity cities can lead to large price increases in average rents nation-wide, even if total demand is the same.

The instrument of Bartik shocks can be especially useful within this spatial equilibrium model. Also known as shift-share, this instrumental variable approach measures local productivity shocks by using pre-existing industry shares to capture localities' exposures to national industry trends (Goldsmith-Pinkham et al., 2020). Researchers have used these Bartik shift-share changes to instrument for local labor demand shocks (Glaeser et al., 2005; Accetturo et al., 2021). Their results confirm many of the spatial equilibrium model's theoretical predictions: more housing-restricted areas see lower population and employment growth, smaller increases in housing supply, and faster housing price growth. Saiz (2010) extends this analysis by incorporating geographic supply constraints such as steep slopes and water features, and by creating a model in which regulations are endogenously determined by land scarcity and high land values. Like the previous authors, he utilizes Bartik shocks, in this case to estimate housing supply elasticities in 95 Metropolitan Statistical Areas with over 500,000 people. Some recent research has further developed this literature at the hyperlocal level, showing the effect of even highly localized demand shocks on neighborhoods within a city, and the intermediating role of elasticity-reducing regulations (Burley, 2017; Baum-Snow & Han, 2019).

There is also some relevant literature that examines the importance of income inequality, helping to flesh out the displacement mechanism that I will be investigating. (Matlack & Vigdor, 2008) model and find empirically that increased incomes at the top end of the earnings distribution will raise housing prices for low-income residents if housing markets are tight (or perhaps, less able to adjust supply in response to demand shocks). Other research shows that the salary-adjusted cost of living is even across cities for high earners, but not for low earners: the cities where low-income workers can earn the most money are so expensive that they are also the cities where cost-adjusted standard of living is lowest (Diamond & Moretti, 2021; Hoxie et al., 2022). This finding seems to align with the aforementioned "superstar" theory in Gyourko et al. (2013).

My paper seeks to add to these literatures by combining the methods from the research on demand shocks and supply elasticities with a focus on urban inequality and city affordability. By using plausibly exogenous Bartik shocks to local demand, I can then measure the degree to which low-income residents of those localities are then economically dislocated, ideally isolating a causal effect of land use regulations.

Data Description

The key dataset of land use regulation comes from Gyourko et al. (2019), and is called the Wharton Residential Land Use Regulation Index (WRLURI). The authors created this dataset from surveys sent to jurisdictions in 2018; it contains 2844 observations. To create the headline index values, the authors measured a wide range of values, attempting to encapsulate the many different ways that regulatory restrictions can reduce housing construction. There are 12 subindices capturing regulations on housing's built form, such as minimum lot sizes and single-family zoning, as well as procedural regulations, such as the prevalence of veto points or direct political involvement for individual housing projects. My variable of interest is the aggregate value, called *WRLURI2018*, which is a weighted aggregate of the subindices. The authors standardize this aggregate value, so it has a mean of 0 and standard deviation of 1; this will allow for easy interpretations of regression coefficients on WRLURI. It is also right-skewed, as a few places have unusually strict regulations.

The American Community Survey provides data for rents and the population with no college degree. When operating at the census place level (cities and towns), which is the spatial level in the WRLURI survey, I encountered challenges with the margins of error in 1-year American Community Survey (ACS) estimates for census places, which are so large as to make the data almost totally unreliable. This leads to tricky tradeoffs — we can use the 5-year pooled ACS to get more reliable estimates, but this precludes us from measuring year-to-year change. An alternative option is to use the core-based statistical area (CBSA)-level estimates from the WRLURI dataset, which contain WRLURI18 values for 44 CBSAs created by averaging the WRLURI18 measures across the CBSAs that had at least 10 jurisdictions in the survey (Census CBSAs include both metropolitan statistical areas [MSAs] and smaller micropolitan statistical areas [μSAs], but as my dataset only includes 44 MSAs I use CBSA and MSA interchangeably).

I also have a suite of outcome variables. Many of these variables come from the IPUMS National Historical GIS website, which can generate variables for a large dataset of localities. These variables come from the annual American Community Survey. Using 5-digit CBSA codes, I paired the NHGIS dataset with the WRLURI dataset at the level of CBSA. I use this to create two outcome variables of interest: gross rent (which includes all renter expenses, such as utilities), and the amount of non-college graduates in the metropolitan area[1].

My other key outcome variable is number of building permits. These come from CSV files held on the U.S. Census website that contain annual housing permits (measured in units added) for places across the U.S., which I also paired to the WRLURI dataset using 5-digit CBSA codes.

For summary statistics at the MSA level, see Table 1. In my empirical analysis, I am ultimately interested not in the *levels* of these outcome variables but the *change* in them over time within each place. Because my regression will utilize a change in demand for a locality (see Study Design Section), I am interested in how my outcome variables change in response to that demand shift irrespective of their initial levels. This means that though my regressions are not panel regressions, they are panel-like, as the values are demeaned and baseline MSA levels are not incorporated into the analysis. In my summary statistics, I include the levels of my key variables in 2018 and their change from 2018 to 2019.

Study Design

As a suggestive preliminary result, I performed a cross-sectional regression of WRLURI2018 values on rent as recorded in the 2018 American Community Survey, which is the same year as the regulatory survey. The results are both significant and meaningful: a one standard deviation increase in WRLURI value is associated with a $145.35 increase in rent; see Figure 1. This matches with our expectation that rents should be higher in more strictly-regulated areas, although in a cross-section the effect likely goes both directions: stricter regulations reduce supply and drive up prices, but local residents may be more likely to implement them when property values or demand for an area are high in order to prevent new construction nearby (Saiz, 2010).

In order to address some of these concerns, this study investigates how our local outcome variables respond to a shock in labor demand for that area based on different

[1] Measuring gentrification or economic displacement is not straightforward, in part due to varied definitions of the term in the literature (Cohen & Pettit, 2019). While my focus in this paper is specific to a reduction in incumbent low-skill or low-income workers, I also face problems with variable endogeneity, because I cannot track individual members of the population. As a result, a variable like average income is of little use because I cannot tell if that means local workers began earning more, new high-earning workers moved in, and/or low-income workers were forced to move out. I focus on the quantity of workers with no college education, not the proportion, to attempt to find if the amount of such people declined in absolute terms, but am aware that this outcome also faces some policy endogeneity as new people with no college degree will surely also move into the cities during my study period.

degrees of land use regulations. Using a response to short-term fluctuations in labor demand can help account for policy endogeneity of regulations.

My relevant regression is

$$Y_{it} = \beta_0 + \beta 1 WRLURI18_i + \beta_2 LaborDemand_{it} + \beta_3 WRLURI18_i * LaborDemand_{it} + u_i \quad (1)$$

where β_1 would be the main coefficient of interest, interpreted as how a 1 standard deviation increase in measured land use regulatory stringency affects price growth, permit growth, or economic displacement for a given change in labor demand. I include an interaction term to account for the asymmetries between urban growth and decline due to housing supply's low downwards elasticity — for an increase in housing demand, supply will increase more than it would decrease for a fall in demand, because the existing housing stock doesn't disappear (Glaeser & Gyourko, 2005). An interaction term should also be useful because land use regulations are more likely to be binding constraints when demand shocks are higher; in low-demand cities they may do little to affect local outcomes because little housing would be built there in any case.

One way to estimate the change in labor demand would be to measure the change in regional employment. However, such a measurement would run into endogeneity problems — indeed, one of the key papers in this literature finds that stricter land use regulations reduce employment growth for a given change in demand (Saks, 2008). If we simply measure employment growth then we are not capturing the true demand shock for the area, but instead the employment change after variables including land regulation have changed the demand shock's effect. And as Saks (2008) and others do, I address this by using the Bartik instrumental variable (IV), which instruments for the demand shock using a region's baseline industry composition and its ensuing exposure to national-level trends. During the time period, if a region has a disproportionately large share of an industry that nationally performs well (poorly), their employment growth as captured by the Bartik instrument will be large (small), and we capture only the variation in employment that is exogenous. Following Goldsmith-Pinkham et al. (2020)], this instrument is measured as

$$Bartik_{it} = \sum_k z_{ik0} g_{kt} \quad (2)$$

where each industry k's employment share in location i at time 0 is multiplied by the growth rate in that industry in all locations except for location i over the time period. Then, the first stage of the IV is

$$\Delta LaborDemand_{it} = C\delta_0 + \delta_1 WRLURI18 + \delta_2 Bartik_{it} + \delta_3 WRLURI18 * Bartik_{it} + \nu_{it} \quad (3)$$

With some matrix of controls C, and an error term that includes location-specific growth or amenity shocks that should be independent of national-level industry trends. *Bartik* is the excluded variable in the first stage. The second stage is

$$\begin{aligned} Y_{it} = \beta_0 + \beta 1 WRLURI18_i + \beta_2 Labor\widehat{Demand}_{it} \\ + \beta_3 WRLURI18_i * Labor\widehat{Demand}_{it} + C\beta_4 + \mu_i \end{aligned} \tag{4}$$

where $\Delta Labor\widehat{Demand}_{it}$ is the predicted change in employment from the Bartik instrument. Again, we are most interested in β_1, which captures how a one standard deviation increase in measured land use regulatory stringency affects a locality's response to a given increase in labor demand. Importantly, such an estimation is only made possible through the utilization of the Bartik instrument, which "controls" for the predicted exogenous change in local labor demand. WRLURI18 has no time subscript because existing data only measures land use regulation once, not in a repeat panel.

Like Saks (2008), I calculate the Bartik shocks at the 3-digit NAICS level using industry employment from the County Business Patterns Survey, which contains 86 industries for each MSA.

Results

Using this Bartik instrumental variable, I then regress each of my three outcome variables (change in rent, change in permitting, change in population with no college degree) on my instrumented variable for change in local employment demand and the WRLURI18 index. For the three reported regression specifications, I report my results for both the first and second stage of the IV regression. I also include the F-statistic for the first stage in the 2SLS regressions.

The second and third regressions include an an interaction term between the instrumented employment shift and WRLURI18, accounting for asymmetry in housing supply curves as described in the Study Design Section. Additionally, I include two control variables in the third specification. One is the percent of undevelopable land, using data from an earlier working paper version of Saiz (2010). Saiz (2010) creates this variable by calculating how much of MSAs' land has a slope too steep to build on or surface water, with both representing a mostly insurmountable barrier to housing development. This control can help address supply-limiting factors in cities' natural geographies. Three of the MSAs in my dataset are not in Saiz (2010)'s dataset and thus dropped from that specification. The other control is a set of categorical controls that indicate the nine census-designated "divisions," which I use in order to control for regional demand trends[2].

[2] These divisions are New England, Middle Atlantic, East North Central, West North Central, South

Using the Pagan-Hall test statistic on my instrumental variable regressions, the regressions for rent change and permit change do not reject the null of homoskedasticity. The regressions for no college change do show heteroskedasticity, however, and to address this I use heteroskedasticity-robust standard errors for the no college regressions.

Across the specifications, the F-statistics are quite high, assuring us that the Bartik instrument is a sufficiently strong predictor of local employment change. Additionally, the excluded Bartik change in employment always has a strongly significant relationship to on total change in employment.

Given that I theoretically predict WRLURI to have a positive coefficient on rent growth and a negative coefficient on permits change and no college population change, my results are mixed.

Rent change (in Table 2) shows a positive coefficient on WRLURI18 in the first two specifications, significant at the 1 percent level. We would interpret these significant coefficients as telling us that a one standard-deviation increase in in regulation leads to a \$30.33 and \$39.24 increase in rent for a given demand shift, respectively. The interaction terms are not significant, and their inclusion does not sizably change the WRLURI18 coefficients and standard errors. However, upon incorporating our controls for geographic constraints and census divisions, the coefficient on WRLURI18 approximately halves and becomes insignificant, while the percent undevelopable control has significance at the 5 percent level.

Permit change (in Table 3) has negative coefficients on WRLURI18, but none of them are significant and we cannot reject the null of no effect. There is, however, a positive coefficient on the interaction term, significant at the 5 percent level for the no-control regression and the 1 percent level for the specification with controls. I interpret this as indicating that the marginal effect of increased WRLURI18 values on housing permit change is more positive for higher levels of demand shocks — this is an unexpected result, as I would expect the marginal effect of regulations on permits to become more negative as demand shocks become greater, because they become more binding on construction.

No college change (in Table 4) has the most unexpected result. Theoretically, I would predict the WRLURI18 coefficient to be negative here: the more tightly that a city is regulated, the more an increase in demand would push out lower-income residents, because more housing cannot be built to accommodate the influx of residents. In the 2SLS regression without an interaction term, WRLURI18 has no effect on the college-educated share. Once I add an interaction term, however, the coefficients on both WRLURI18 and the interaction term are significant. The WRLURI18 coefficient is positive, telling us that

Atlantic, East South Central, West South Central, Mountain, and Pacific, and they each have between four and nine states.

for a given demand shock, the population with no college actually increases with higher WRLURI18 values. However, the negative interaction term shows a more expected result: with higher demand shocks, WRLURI18 has an increasingly negative effect on the no college population. However, WRLURI18 loses its significance upon adding the controls, similarly to the rent growth variable.

Threats to Validity

There are a few reasons for concern about my study's internal validity. One is that land use regulations are likely endogenous to demand. While in many of these cities, the relevant regulations were established in the 1970s and should therefore be largely unrelated to changes during my study period, Gyourko et al. (2019) document that localities have changed their rules somewhat over the 21st century, so they are certainly not static. Saiz (2010) shows that these regulations are endogenous to land scarcity, because property owners of scarce and valuable land push to protect their land values.[3] This risks biasing my results upwards. Furthermore, recent policy efforts to change zoning codes have explicitly focused on the problem of high rents that supply restrictions are believed to exacerbate — if such policy efforts were successful during my time period, they would instead bias my results downwards. One way I attempt to control for this is via the use of a Bartik shock, which is assumed to be random; even if a region's baseline level of demand is positively correlated with regulation, the regulations should not be causally related to the change in demand based on national growth. I also work only within a one-year time period, which is short enough that a jurisdiction is unlikely to adjust their land use policies in response to demand trends. However, it is plausible that locales that saw an economic boom in one year could quickly respond with supply-constraining policies (or, depending on their political leanings and the influence of homeowners, supply-increasing policies) to prevent (allow) a corresponding development boom. Unfortunately, there are no frequent panel measures of land-use regulation that allow us to systematically track these types of changes. Hence, our estimate of land use regulation's effects may have bias in both directions, although the short time period hopefully precludes much of this.

Additionally, research on this topic would ideally consider other features of cities that limit elasticities but are related to the land use regulations; otherwise my results are likely to suffer from omitted variable bias. My study accounts for some of this problem by using the measure of geographic constraints from Saiz (2010), which accounts for surface water or steeply sloped terrain, but this is not sufficient. Recent research from Orlando & Redfearn (2022) argues that housing stock growth is in large part determined by long-term

[3] The theory that homeowners shape local policy in their own interest was most famously advanced by Fischel (2001) in his book *The Homevoter Hypothesis*.

past development patterns — is the area generally "built out" with few remaining greenfield sites, or are there plenty of available lots for development? For example, the Houston, TX MSA (WRLURI18 of -0.04) is more strictly regulated than the Bay Area, CA MSA (WRLURI18 of 1.18), but also has more open land that can be cheaply developed. This effect could bias us in both directions, and also ought to be controlled for.

Although my use of the Bartik instrument is useful for addressing some concerns about endogeneity of the land use regulations and OVB in a region's population or employment growth, one of this paper's biggest challenges is the lack of more detailed spatial equilibrium modelling in my regression analysis. To accurately understand economic activity across cities, we must account for the fact that labor and capital's movements are affected by all other cities' local characteristics. In other words, spillover effects are abundant: when some localities have lower housing supply elasticities, and thereby see higher rent price growth and lower population growth in the face of a demand shock, this will cause demand to spill over to other cities. For example, research by Erdmann (2021) argues that the 2008 financial crisis was in part caused by strictly regulated "closed access" cities such as San Francisco and New York City forcing many migrants to move into "contagion cities" such as Las Vegas and Tampa, driving intense speculative activity. Further demand shocks in strictly regulated areas are similarly likely to create local demand in "substitute" locations. On a larger scale, there will be less labor mobility (Ganong & Shoag, 2017) and lower aggregate growth (Hsieh & Moretti, 2019), trends affecting all cities relative to the less-regulated counterfactual. Zabel (2012), for example, builds a more complex model that allows for spillover effects in his study of local labor market responses to employment shocks.

My simple regression analysis is unable to account for these spillover effects, and so there is most likely demand not captured in my Bartik instrument that affects my outcome variables but is not randomly distributed across locations, hence biasing my results.

There is also the concern that my instrument for local labor demand violates the exclusion restriction, which would be true if my instrument (local exposure to national sectoral trends due to industrial composition) affected my outcome variables via mechanisms other than local labor demand. One of the most likely possibilities would be if a single MSA represented a very large share of one industry, and so their idiosyncratic employment growth characteristics would considerably affect national rates. Furthermore, it is likely true that *some* of the industries that contribute to my Bartik IV estimation are directly related to my outcome variables — for example, construction's relative success will affect building rates, and growth in different industries will have uneven effects on college and non-college residents. These possible causal pathways from the instruments to the outcomes may invalidate my instrumental variable.

Conclusion

Today, there is consensus among economists that land use regulations create large and meaningful distortions. By distorting labor markets, they reduce overall productivity, and by constraining housing development they increase housing's cost at the local levels.

However, there is more work to be done on the systematic distributional impacts of these policies. There is strong evidence that by biasing housing towards more expensive (i.e. larger, single-unit) forms, they keep lower-income people out of certain neighborhoods (Lens & Monkkonen, 2016; Trounstine, 2020). We should also investigate how these regulations affect low-income people in the places that they currently live. Theoretical work and some empirical work suggests that by reducing supply elasticities, these regulations increase low-income peoples' risk of economic displacement as other people move nearby — or, gentrification (Gyourko et al., 2013; Diamond & Moretti, 2021).

In my paper, I sought to quantify the effects of these regulations at the MSA level, using measures from the 2018 Wharton Residential Land Use Regulatory Index. I constructed MSA-level Bartik shocks from 2018 to 2019, which estimates the expected change in local labor demand as a result of a location's exposure to national product demand. I then used this Bartik shock as an instrument for the change in local employment attaining what I hope to be an exogenous measure of local labor demand. I regressed my three outcome variables (change in rent, housing permits, and people with no college degree) on the WRLURI18 measure, using the instrumented labor demand as a control with an interaction term between the two. I then added controls for geographic restrictions on land and the census division.

In my favored specification using the controls, my results were mostly not significant, failing to reject the null hypothesis that land use regulations do not affect rent, housing permits, and the non-college population. Without the controls, the results were mixed compared to my theoretical predictions. Higher levels of regulation worsened rents, had no effect on permit change, and led to increases in college graduates (the final result is the most unexpected).

However, my study had imperfections. Likely the foremost was the lack of spatial equilibrium modeling, which prevents me from accounting for the ways that people in the aggregate make decisions, and how one city's characteristics affect other cities' outcomes. I also faced problems with policy endogeneity and an inability to account for some important city traits that affect housing supply elasticities, both of which surely led to OVB in my results.

Nevertheless, my results indicate a research agenda focused on the cross-city distributional effects of land use regulations that scholars should continue to investigate. There are many facets of housing markets, labor markets, and their interactions with each other that are shaped by housing policy and are worthy of attention. Furthermore, such questions have been rapidly gaining relevance to politics and policymaking across the nation, and empirical work should continue filling in our knowledge of these topics.

References

Accetturo, A., Lamorgese, A. R., Mocetti, S., & Pellegrino, D. (2021). Housing supply elasticity and growth: evidence from Italian cities. *Journal of Economic Geography*, 21(3), 367–396. `https://doi.org/10.1093/jeg/lbaa034`

Baum-Snow, N. & Han, L. (2019). The Microgeography of Housing Supply. *Atlanta Fed Working Paper*, 67.

Burley, J. (2017). Labour Demand, Housing Markets, and Neighbourhood Change. *Job Market Paper*, 60.

Cohen, M. & Pettit, K. L. S. (2019). Guide to Measuring Neighborhood Change to Understand and Prevent Displacement (National Neighborhood Indicators Partnership). *Urban Institute*.

Diamond, R. & Moretti, E. (2021). Where is Standard of Living the Highest? Local Prices and the Geography of Consumption (Working Paper No. 29533). *National Bureau of Economic Research*. `https://doi.org/https://doi.org/10.3386/w29533`

Erdmann, K. (2021). Shut Out: How a housing shortage caused the great recession and crippled our economy. *ROWMAN & LITTLEFIELD*.

Fischel, W. A. (2001). The Homevoter Hypothesis: How home values influence local government taxation, school finance, and land-use policies. *Harvard University Press*.

Ganong, P. & Shoag, D. (2017). Why Has Regional Income Convergence in the U.S. Declined? *Journal of Urban Economics*, 102, 76–90.

Glaeser, E. L. & Gyourko, J. (2002). The Impact of Zoning on Housing Affordability (Working Paper No. 8835). *National Bureau of Economic Research*. `https://doi.org/https://doi.org/10.3386/w8835`

Glaeser, E. L. & Gyourko, J. (2005). Urban Decline and Durable Housing. *The Journal of Political Economy*, 113(2), 345–375.

Glaeser, E. L., Gyourko, J., & Saks, R. (2003). Why is Manhattan So Expensive? Regulation and the Rise in House Prices (Working Paper No. 10124). *National Bureau of Economic Research*. `https://doi.org/https://doi.org/10.3386/w10124`

Glaeser, E. L., Gyourko, J., & Saks, R. E. (2005). Urban Growth and Housing Supply (Working Paper No. 11097). *National Bureau of Economic Research*. `https://doi.org/https://doi.org/10.3386/w11097`

Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8), 2586–2624. https://doi.org/10.1257/aer.20181047

Gyourko, J., Hartley, J., & Krimmel, J. (2019). The Local Residential Land Use Regulatory Environment Across U.S. Housing Markets: Evidence from a New Wharton Index. *NBER Working Paper*, 26573.

Gyourko, J., Mayer, C., & Sinai, T. (2013). Superstar Cities. *American Economic Journal: Economic Policy*, 5(4), 167–99. https://doi.org/10.1257/pol.5.4.167

Herkenhoff, K. F., Ohanian, L. E., & Prescott, E. C. (2018). Tarnishing the golden and empire states: Land-use restrictions and the U.S. economic slowdown. *Carnegie-Rochester-NYU Conference on Public Policy held at the Stern School of Business at New York University*, 93, 89–109. https://doi.org/10.1016/j.jmoneco.2017.11.001

Howard, G. & Liebersohn, J. (2021). Why is the rent so darn high? The role of growing demand to live in housing-supply-inelastic cities. *Journal of Urban Economics*, 124, 103369. https://doi.org/10.1016/j.jue.2021.103369

Hoxie, P., Shoag, D., & Veuger, S. (2022). Moving to Density: Half a Century of Housing Costs and Wage Premia from Queens to King Salmon. *American Economic Institute Economcis Working Papers*, 34.

Hsieh, C.-T. & Moretti, E. (2019). Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics*, 11(2), 1–39. https://doi.org/10.1257/mac.20170388

Lens, M. C. & Monkkonen, P. (2016). Do Strict Land Use Regulations Make Metropolitan Areas More Segregated by Income? *Journal of the American Planning Association*, 82(1), 6–21. https://doi.org/10.1080/01944363.2015.1111163

Matlack, J. L. & Vigdor, J. L. (2008). Do rising tides lift all prices? Income inequality and housing affordability. *Journal of Housing Economics*, 17(3), 212–224. https://doi.org/10.1016/j.jhe.2008.06.004

Moretti, E. (2011). Local Labor Markets. *Handbook of Labor Economics*, volume 4, 1237–1313. Elsevier. https://doi.org/10.1016/S0169-7218(11)02412-9

Murray, C. K. (2021). Marginal and average prices of land lots should not be equal: A critique of Glaeser and Gyourko's method for identifying residential price effects of town planning regulations. *Environment and Planning A: Economy and Space*, 53(1), 191–209. https://doi.org/10.1177/0308518X20942874

Orlando, A. W. & Redfearn, C. L. (2022). Houston, You Have a Problem: How Large Cities Accomodate More Housing. *Working Paper*.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *Journal of Political Economy*, 90(6), 1257–1278.

Rosen, S. (1979). Wage-Based Indexes of Urban Quality of Life. *In Current Issues in Urban Economics*.

Saiz, A. (2010). The Geographic Determinants of Housing Supply*. *The Quarterly Journal of Economics*, 125(3), 1253–1296. https://doi.org/10.1162/qjec.2010.125.3.1253

Saks, R. E. (2008). Job creation and housing construction: Constraints on metropolitan area employment growth. *Journal of Urban Economics*, 64(1), 178–195. https://doi.org/10.1016/j.jue.2007.12.003

Trounstine, J. (2020). The Geography of Inequality: How Land Use Regulation Produces Segregation. *American Political Science Review*, 114(2), 443–455. https://doi.org/10.1017/S0003055419000844

Zabel, J. E. (2012). Migration, housing market, and labor market responses to employment shocks. *Journal of Urban Economics*, 72(2), 267–284. https://doi.org/10.1016/j.jue.2012.05.006

Figures and Tables

Figure 1

Simple regression of regulatory index and rents at the MSA level

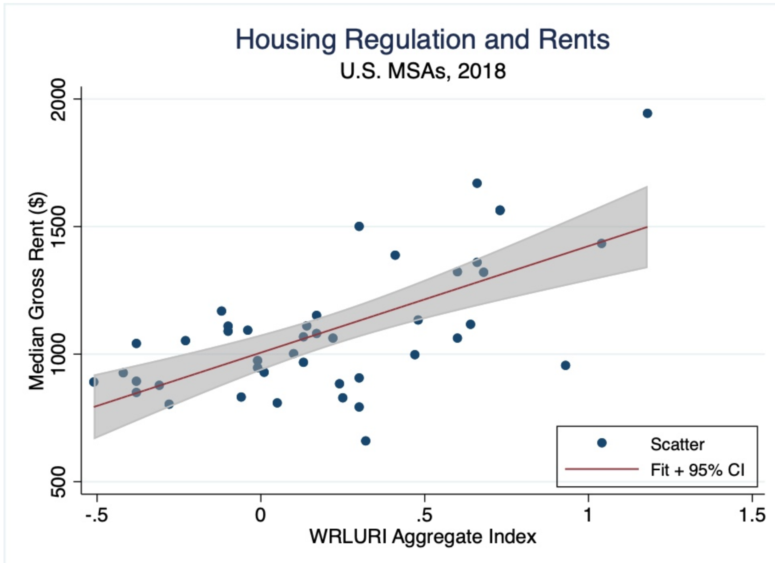


Table 1

MSA cross-sectional summary statistics, 2018

	(1) Mean	SD	Min	Max
WRLURI18	0.21	0.41	-0.51	1.18
Population, Thousands	3529.44	3717.31	535.42	19979.48
Median Rent	1094.30	266.35	660.00	1944.00
Pop. w/ no College, Thousands	860.71	960.76	115.31	5220.20
Housing Permits	14099.89	15443.97	374.00	63893.00
2018-2019 Population Change, Thousands	-7.17	120.12	-763.29	90.24
2018-2019 Median Rent Change	40.59	27.34	-8.00	113.00
2018-19 Pop. w/ no College Change, Thousands	-3.60	32.75	-192.70	30.21
2018-2019 Change in Permits	560.91	2731.99	-6605.00	11476.00

Table 2

Regression Results for Rent Change

	(1) First Stage	(2) Second Stage	(3) First Stage	(4) Second Stage	(5) First Stage	(6) Second Stage
Bartik Change in Emp.	1.092***		0.931***		0.786***	
	(0.107)		(0.175)		(0.212)	
WRLURI18	7976.8	30.33***	1237.7	39.24***	-2425.2	17.71
	(6264.8)	(9.340)	(8536.6)	(10.96)	(10527.2)	(11.30)
Change in Emp.		0.000176		0.000436*		0.000210
		(0.000139)		(0.000247)		(0.000275)
Bartik Change in Emp. × WRLURI18			0.261		0.453*	
			(0.226)		(0.259)	
Change in Emp. × WRLURI18				-0.000387		-0.000151
				(0.000285)		(0.000307)
Pct. Undevelopable Land					-160.4	0.435**
					(178.6)	(0.189)
Constant	-5060.4	30.26***	-2106.6	26.44***	6015.5	21.05*
	(3357.3)	(4.366)	(4207.3)	(5.247)	(9397.7)	(10.81)
Census Division Controls		No		No		Yes
Observations	44	44	44	44	41	41
First Stage F-Statistic		103.6		24.68		19.13

Standard errors in parentheses

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

Table 3

Regression Results for Permit Change

	(1) First Stage	(2) Second Stage	(3) First Stage	(4) Second Stage	(5) First Stage	(6) Second Stage
Bartik Change in Emp.	1.092***		0.931***		0.786***	
	(0.107)		(0.175)		(0.212)	
WRLURI18	7976.8	-228.7	1237.7	-1591.2	-2425.2	-717.2
	(6264.8)	(994.8)	(8536.6)	(1150.0)	(10527.2)	(1427.2)
Change in Emp.		0.0474***		0.00757		0.00139
		(0.0148)		(0.0259)		(0.0347)
Bartik Change in Emp. × WRLURI18			0.261		0.453*	
			(0.226)		(0.259)	
Change in Emp. × WRLURI18				0.0592**		0.0675*
				(0.0299)		(0.0387)
Pct. Undevelopable Land					-160.4	-4.009
					(178.6)	(23.90)
Constant	-5060.4	-444.2	-2106.6	140.7	6015.5	189.0
	(3357.3)	(465.0)	(4207.3)	(550.2)	(9397.7)	(1364.9)
Census Division Controls		No		No		Yes
Observations	44	44	44	44	41	41
First Stage F-Statistic		103.6		24.68		19.13

Standard errors in parentheses

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

Table 4

Regression Results for No College Change

	(1) First Stage	(2) Second Stage	(3) First Stage	(4) Second Stage	(5) First Stage	(6) Second Stage
Bartik Change in Emp.	1.092***		0.931***		0.786***	
	(0.107)		(0.175)		(0.212)	
WRLURI18	7976.8	5542.1	1237.7	28552.5***	-2425.2	12198.0
	(6264.8)	(10662.3)	(8536.6)	(10588.9)	(10527.2)	(13524.7)
Change in Emp.		-0.792***		-0.120		-0.617*
		(0.159)		(0.239)		(0.329)
Bartik Change in Emp. × WRLURI18			0.261		0.453*	
			(0.226)		(0.259)	
Change in Emp. × WRLURI18				-1.000***		-0.508
				(0.275)		(0.367)
Pct. Undevelopable Land					-160.4	281.5
					(178.6)	(226.5)
Constant	-5060.4	12845.6***	-2106.6	2968.5	6015.5	5580.1
	(3357.3)	(4984.6)	(4207.3)	(5066.7)	(9397.7)	(12934.0)
Census Division Controls		No		No		Yes
Observations	44	44	44	44	41	41
First Stage F-Statistic		103.6		24.68		19.13

Standard errors in parentheses

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