Analyzing Effects of Large and Rare Events with an Augmented Synthetic Control Method

Ribal Abi Raad

A thesis

In the Department

of

Economics

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Economics) at

Concordia University

Montreal, Quebec, Canada

December 2022

© Ribal Abi Raad, 2022

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Ribal Abi Raad

Entitled: Analyzing Effects of Large and Rare Events with an Augmented Synthetic Control Method

and submitted in partial fulfillment of the requirements for the degree of

Doctor Of Philosophy (Economics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

________________________________ Chair
Dr. Sara Kennedy

________________________________ External Examiner
Dr. David Benatia

________________________________ Examiner
Dr. Paul Gomme

________________________________ Examiner
Dr. Axel Watanabe

________________________________ Examiner
Dr. Melina Mailhot

________________________________ Thesis Supervisor
Dr. Tatyana Koreshkova

Approved by ________________________________
Dr. Christian Sigouin, Graduate Program Director

10/14/2022 ________________________________
Dr. Pascale Sicotte, Dean
Faculty of Arts and Science

Abstract

Analyzing Effects of Large and Rare Events with an Augmented Synthetic Control Method

Ribal Abi Raad, Ph.D.

Concordia University, 2023

This dissertation consists of four chapters on applying the Synthetic Control Method to rare events with a significant impact. Initially pioneered by Abadie, Diamond, and Hainmueller (2010a), the Synthetic Control Method is a policy analysis tool developed to tackle the weaknesses of traditional policy analysis models such as the Difference-in-Difference. In the first chapter of this dissertation, this model is used to reconcile a recurring issue in the disaster literature: why some countries recover better than others from disasters and the role that political institutions play in this recovery. The results show that regulatory power is the most significant institutional quality variable that determines post-disaster recovery. Ranking in the top 30% of countries regarding regulatory power is linked to GDP recovery rates from disasters that are higher than predicted GDP. The variable with the most negligible impact was corruption, as proxied by the corruption perception index. A 1-point increase in this index was linked to a 0.05% increase in the recovery rates compared to predicted GDP. On the other hand, the degree of democratization or level of democracy is insignificant in determining the size or level of recovery. Finally, over five years after the occurrence of a disaster, countries that experienced negative recovery rates of GDP per capita had this value shrink by about 13%. In contrast, countries with positive recovery rates of GDP per capita ended up with a GDP per capita ahead of its predicted value by 8% over the same five-year period.

One of the most devastating disasters of the last 50 years is the COVID-19 pandemic that put the entire world at a standstill. The second chapter of this dissertation summarizes the literature surrounding anti-contagion policies and highlights a gap in the literature in untangling the impact of individual anti-contagion policies. This gap is tackled in the third chapter, which investigates the relative importance and impact of individual anti-contagion policies in reducing death rates in the United States. Restrictions on gatherings proved to be the most significant policy in reducing

death rates, lowering them on average by four out 100,000 COVID deaths per day 60 days after the implementation of such a policy. School closings and public transportation closings were the least effective policies reducing death rates by 0.2 and 0.5 per 100,000 over the same period.

In the fourth chapter, the traditional Synthetic Control Method is modified to account for cumulative and interrupted events through the Multi Synthetic Control Method. This method is tested on previous examples used in the literature and is shown to be robust to uninterrupted events. When applied to anti-contagion policies in the United States, the Multi Synthetic Control Method finds that the standard Synthetic Control Method can underestimate the true impact of a policy by up to 150%. The values obtained from the Multi Synthetic Control Method for the same event as compared to the base Synthetic Control Method were significantly different, ranging between 20% to 150% different in absolute value.

Significant improvements have been made to the original Synthetic Control Method since its inception. In this thesis, additional improvements are proposed to improve this method's accuracy. In the first chapter, a new method of selecting the vector of relative importance (known as the V vector) is discussed. This method improves the accuracy of obtaining this vector for regressions where the variance of the treated variable and the number of co-factor variables are high.

The results of this dissertation show the ability of the Synthetic Control Method to tackle all kinds of policies. Policy-makers aiming to take on upcoming waves or different mutations of the COVID-19 virus should consider the effectiveness of different policies and the implication of their stringency in affecting death rates and economic variables, and the trade-off between them.

Acknowledgements

My deep gratitude goes first to my inspirational supervisor, Dr. Tatyana Koreshkova who wisely guided me through my PhD and whose energy was crucial in successfully completing this milestone. Her drive and motivation kept me engaged in my research and helped push my inquisition into various fields of Economics.

I would like to express my sincerest appreciation to my committee members for their efforts in reviewing my work and attending the defence: Dr. Paul Gomme, Dr. Axel Watanabe, Dr. Melina Mailhot and Dr. David Benatia. I would also like to thank my Graduate Program Director, Dr. Christian Sigouin who supported me on many occasions, and always had my back any time I had any issues throughout my journey.

I am also very thankful to my colleagues and classmates, who are now my life-long friends: Noemie Cabau, Dilek Sayed Ahmed, Faten Siblini, Shadi-el- Ramli, Rouzbeh Ghouchani, Fadi Harb, and Olha Hnatyshyn. I am thankful to all of them for the times they stood by me throughout my time at Concordia and for all the good and bad moments we shared together. I am also thankful to the wonderful friends that pushed me towards a better life: John Tarakidjian, Samer Matta, Elie Constantine, Ayca Koseoglu, Christian Boukhater, Carl Kreidy, Mickey Hamade, Samkoum, Karim Kaddoura, Rayan Chatila, Nicholas Ayoub, Benjamin Bradley, Carly Roberts, Selin Demir, Leo Williams, and many more.

I am indebted to my mom who has always been an inspiration to me. I am also indebted to my ”best” friend Dimitri Ayoub. You never ceased to push me to be better in your own ways. I appreciate the wisdom and friendship that you provided in my academic and personal life and I am looking forward to sustaining and deepening our friendship.

Finally, I would like to acknowledge Sara Kheireddine. You have always been there for me, and you have always supported me when I doubted myself the most. None of what I have achieved here would have been possible without your unwavering faith in me, it is more than what anyone can hope to have. Your tenacity, meticulousness, and joy are things to admire and aspire to have.

Contents

List of Figures

List of Tables

Chapter 1

Disasters and institutions: recovering from large-scale events the right way

Abstract

Countries affected by large and rare shocks experience different effects on their economies over time, even if they are relatively similar regarding factors such as GDP per capita, unemployment, or other macroeconomic variables. While rare, these disasters have a significant impact on long-run growth. The synthetic control model is used to find the effects of these disasters for countries that have experienced somewhat comparable disasters. The results show that the degree of democratization is insignificant, but the regulatory quality is vital in determining post-disaster recovery. On average, countries with poorer regulatory authorities experience a 10% decrease in GDP 5 years after a disaster, while countries in the top 30% of regulatory power experience an 8% increase in GDP over its predicted value over the same period.

1.1 Introduction

Disasters are low-frequency large-scale events that affect a country or a region and result in significant destruction of capital or loss of life. These could be any number of events such as natural (earthquakes, floods, hurricanes), political (revolutions, instability, war), or financial. The effect of these events on the economy has garnered much attention over the years. Traditionally, the approach to understanding the impact of rare destructive events has focused on the damage and loss of physical and human capital, usually up to three years after their occurrence. Focusing on the short term is a myopic view of the impact of disasters as countries tend to overspend to counteract the disaster's negative impact leading to contradicting results amongst researchers. For example, Jaramillo (2009) finds that for specific countries, the impact of natural disasters lasts beyond the first five years in which recovery happens, with some disasters impacting GDP up to 10 years after their occurrence.

Historically, a negative correlation between disasters and economic growth has been established. Researchers have identified various channels through which disasters can limit growth. This includes lowering interest rates and returns on assets (Barro, 2001), (Ikefuji and Horii, 2012), human capital destruction (Baez, Fuente, and Santos, 2010), or simply permanently contracting GDP through significant loss of capital in smaller countries (Heger, Julca, and Paddison, 2008). When looking at regions that have experienced (or are more at risk of experiencing) such rare, devastating events, researchers have observed lower long-run growth rates than their less risky counterparts across and within countries. Disasters lead to loss of labor, whether direct (death) or indirect (relocation), and with less labor, an influx of capital may not be very effective. It is better to look at various disasters rather than ones that only impact capital.

Some evidence also shows that disasters do not affect long-term growth, as any change is only temporary (Skidmore and Toya, 2002). These conflicting results are often due to model specifications, data selection, or modeling techniques. For example, post-disaster recovery usually implies that capital stock could increase due to over-investment, even more than before the disaster. Five years later, that might not be the case. While GDP is temporarily boosted, the scope of the analysis could lead to contradicting results. This variation in experiences is partially attributed to the quality of political institutions across countries. Old and antiquated capital is replaced by newer and more efficient capital, the 'creative destruction ' of capital (Ikefuji and Horii, 2012). The channels through which these disasters affect growth are many, and in this chapter, the focus is on several institutional quality indices as accelerators or dampeners to post-disaster recovery. These include but are not limited to the degree of democratization, corruption levels, and ease of doing business.

The first motivation for tackling this topic is the increased frequency of disasters, even when controlling for the increased reporting of disasters and their expanding impact over the years. As societies interact more and as we start exploiting more of the limited amount of land we have, large-scale events will have a bigger and more significant impact, something all the more evident with the recent COVID-19 pandemic. Accelerating climate change means that certain disasters will become more frequent and severe, and it is crucial to focus on them moving forward. Growing inequality also leads to turbulent political upheavals, which leads to a spiraling sequence of even more inequality. The other motivation behind this chapter is reconciling and explaining the differences found in the literature.

Using the synthetic control method for a panel of selected countries and GDP data from 1975 to 2018, I model counterfactuals to the occurrence of a disaster and identify the impact of 96 significant and sizeable events. The percentage deviation difference between the counterfactual model and the real data is calculated and used to assess the impact of such events. Results indicate that disasters played a statistically significant role in altering GDP per capita for the countries in the data set. This chapter identifies the impact that a disaster has on GDP, and it finds that on average, in countries where recovery was slower than predicted due to the disaster, GDP was 10% lower than expected up to ten years after the disaster, while in countries where recovery was faster, GDP was on average 8% higher ten years after the occurrence of the disaster. I then look at those deviations and find that specific institutional quality indices explain the size and movement of those deviations. Data limitations do not allow a proper analysis of the trade-off between investing in infrastructure before a disaster and the recovery post-disaster.

While the degree of democratization was insignificant in determining the recovery from a disaster, regulatory power and corruption perception are the most significant factors in the recovery phase, contributing to around 20% of the recovery size. Looking at disasters as events with long-term repercussions, where the impact can be positive or negative, depending on specific co-factors, can help provide policy recommendations for post-disaster relief. This chapter finds that the size of the disaster, identified by the value of capital lost as a percentage of GDP calculated by the Center for Research on the Epidemiology of Disasters (EMDAT, 2020), does not affect the direction of the change in GDP. Large disasters do not necessarily imply negative GDP growth rates. The results support the fundamental theory that a country with more solid political institutions such as an independent judiciary, will see a boost in its long-run growth after a disaster, despite experiencing a significant loss of capital. The path to achieving these strong institutions is not discussed in this chapter.

This chapter contributes to the literature by quantifying the impact of the quality of institutions on post-disaster recovery and identifying critical institutional quality elements through which the channels impacted by disasters affect growth. This empirical analysis explains the different results in the literature around the impact of disasters through institutional quality. The rest of the paper is organized as follows: in section 1.2, I present an up-to-date review of the literature surrounding disasters. Section 1.3 presents the model used in the analysis, while section 1.4 provides a primary historical empirical analysis of disasters and an explanation of the synthetic regression model, section 1.6 highlights the results from the regressions, and finally concluding remarks are presented in section 1.7.

1.2 Literature review

Authors have found that when controlling for reporting, large-scale events are increasing year-on-year (Ruiter et al., 2020). They find that this increase is in both frequency and scale of disasters over time. Table 1.1 highlights this fact for Canada and shows the largest disasters in terms of economic size for Canada over the last 90 years. Data older than this time period is unreliable due to lack of reporting or record keeping in terms of the occurrence of disasters, and their impact on the economy. Economic data prior to 1920 for a large portion of countries is often unavailable, and does not allow for comparison across countries. These data points are therefore excluded from this analysis. Two of the top three disasters that occured in Canada happened in the last 10 years, something not unique to just Canada, but seen all over the world. The heat wave that impacted eastern Canada in July 2020, as well as record temperatures hitting Western Canada in July 2021, are additional examples that extreme events are getting more frequent and more devastating.

Type	Date	Total damage ('000 US$)
Flood	2013	5,700,000
Wildfire	1989	4,200,000
Wildfire	2016	4,000,000
Drought	1977	3,000,000
Extreme temperature	1992	2,000,000

Table 1.1: Estimated disaster damage (top 5 in Canada)

There is much research concerning the short-term impact of disasters on macroeconomic variables, where the focus is usually on the immediate destruction of physical capital (Kajitani and Tatano, 2018), (Benson and Clay, 2000). This view is slowly changing. The impact on long-run growth has been emphasized more in recent years, and this subsection summarizes some of the significant work done in this field. Many mechanisms dictate how disasters affect GDP, and political institutions are crucial in increasing the efficiency of those mechanisms (Drury and Olson, 1998).

One of the main channels through which GDP can be affected post-disaster is the interest rate. Research from Barro (2006) and Barro (2009) tries to explain the equity premium puzzle through disaster risk. He concludes that the allowance of low-probability disasters explains several puzzles about asset returns, specifically the equity-premium and low real rate of return on government bills, or why the expected real interest rates were low in the U.S. during major wars. Gourio (2012) extends his model by adding variable disaster risk. This extension implies that changing disaster risk induces business cycles, mainly through precautionary savings, highlighting how disasters can impact long-run growth by affecting interest rates for a long period. These models expand the definition of disasters to include man-made disasters, such as financial or technological crises.

Additionally, some work has tried to understand the impact disasters have on human capital, such as Baez, Fuente, and Santos (2010), especially since the importance of human capital to long-run growth has been highlighted often (Barro, 2001) (Erosa, Koreshkova, and Restuccia, 2010). Loss of human capital through a disaster can then drive a country to a lower balanced growth path, increasing the gap in income between rich and developing countries.

The type of disaster is also an essential factor in the recovery process. While considering large-scale events without filtering for types, Raddatz (2007) found that natural disasters lowered real GDP by 2 to 4% for a panel of low-income countries. This is not the case when focusing on specific types of disasters. Hsiang and Jina (2014) find that cyclones have small but significant negative implications on long-term growth, leading to a GDP loss of about 6% for a panel of U.S. states. Barone and Mocetti (2014) find that earthquakes in Italy reduced GDP by about 10% for specific regions compared to what they should have been. Using a two-sector endogenous growth model, Ikefuji and Horii (2012) argue that theoretically, the negative effects of disasters could be mitigated, and economic growth could be improved if a per unit tax on polluting inputs is imposed. Institutional quality is crucial in ensuring tax policy is implemented properly and applied, as the revenue from these taxes should be invested in replenishing capital lost from the disaster. This also proves important before a disaster occurs, as tax revenue can be used to insure against disasters.

The negative impact of disasters is not always consistent, and there has been more evidence that disasters can lead to higher GDP in the long run. 'Creative destruction' of capital is the most prominent theory supporting positive effects from disasters. Several authors have researched this theory and put forth the idea that while the short-term effects of disasters are adverse, there are observable positive spillovers on long-run growth rates (Ikefuji and Horii, 2012) and (Skidmore and Toya, 2002). When such an event occurs, inefficient capital is destroyed, and an increase in precautionary savings with investment directed towards newer, more productive capital is observed. The replacement of capital would not have occurred at such a pace had the country not been affected by a disaster. This could be due to many reasons such as corruption, bureaucracy, or insufficient funding (Matta, Bleaney, and Appleton, 2022). Disasters can then act as catalysts for change and provide opportunities to grow faster in the long run. Jaramillo (2009) shows that these events have permanent adverse effects for a few geographically small countries typically affected by consistently large disasters. In contrast, larger countries typically experience better recovery rates since they can replace capital better. They attribute this to the relative size of the country shielding certain areas from the disaster and the ability to reallocate more resources (physical and human) towards recovery.

More recently, Akao and Sakamoto (2018) attempted to reconcile this contradiction in the disaster literature. They used an endogenous growth model with both aggregate and idiosyncratic shocks to find that if resources are allocated efficiently, disasters will not hurt long-run growth. A caveat of their research is that they do not mention the source of these inefficiencies. A potential reason those other authors put forth is political institutions. Barone and Mocetti (2014) found that the same disaster had a different impact on growth for different regions in Italy. They focused on the quality of the political institutions of these regions and found that the region with the better quality of political institutions experienced higher than expected growth rates.

Cavallo et al. (2013) argued that disasters have no long-run effects except under particular circumstances, where a disaster is followed by significant political change, highlighting the importance of political institutions in post-disaster recovery. This idea had been supported in other papers. Drury and Olson (1998) find that increased development and regime responsiveness dampen or increase post-disaster political unrest. These varying channels could also explain why we observe different impacts of disasters across different countries. Existing conditions such as corruption levels, capital availability, and institutional quality might play significant roles in determining post-disaster outcomes. Jong-A-Pin (2009) analyzes how some political instability measures affect growth. He finds that instability of the political regime hurts growth, while instability within the political regime works the other way around. This body of work is crucial in identifying the variables used for the model, while the focus remains on political institutions as critical drivers of growth post disasters.

The initial section of this work follows that of Cavallo et al. (2013) and Barone and Mocetti (2014). The former conducted a cross-country study on the impact of disasters on GDP growth rates, while the latter researched the impact of the same disaster on two regions in Italy. Both works use the synthetic control method to understand a disaster's impact on GDP, which is a well-suited tool for studying such events. Differences between both papers include model specification (such as co-factor selection) and inference methods. Cavallo et al. (2013) find that for a disaster to impact long-run GDP, it needs to be followed by drastic political change. Barone and Mocetti (2014) find that the Italian region with better institutions recovered faster and stronger, and GDP was even higher than predicted had they not been exposed to the disaster.

There is excellent support for the idea that institutions play an essential role in the process of recovery post-disasters. Attention needs to be paid to these variables (Matta, Bleaney, and Appleton, 2022). While Cavallo et al. (2013) focus on post-disaster political turmoil, this chapter expands the scope of their research. In this chapter, the period and disasters considered are more considerable. Variables related to political institutions and institutional quality are selected to provide a more accurate analysis of the recovery, and the model is improved to be more accurate. The Synthetic Control Method allows then for both in-sample and out-of-sample forecast.To the best of my knowledge, the model presented in this chapter is unique in its approach to the joint analysis of disasters, growth, and institutions.

1.3 Synthetic Control Method

This section introduces the synthetic control method (SCM) initially proposed by Abadie, Diamond, and Hainmueller (2010b). This model allows the user to generate counterfactual data series and compare them to the actual development of that series. This chapter tackles a given disaster's impact on GDP growth rates per capita. A disaster is defined as a treatment in the context of the Synthetic Control Method. The accuracy of this methodology is improved for data sets where the variance of the treated variable is significant. This is done by grouping the regions selected into different tiers depending on where they are in the initial distribution of the treated object. For this chapter, this treated object is GDP per capita. This is expanded on in the following subsections.

1.3.1 Why Synthetic Control Method

There are several benefits of using the Synthetic Control Method in analyzing regional treatments compared to the more typical regression methods used in the literature, such as Difference-in-Difference regressions. The most important of those is circumventing the problem of credible untreated observations by allowing the use of weighted averages of other units. The only requirement when conducting the analysis is that an appropriate amount of pre-treatment observations exists. For the Synthetic Control Method, this number is relatively small and could be as low as ten to fifteen observations per treated unit (region or country). This advantage is evident when studying yearly GDP observations as long-time series data on GDP for many countries is not always available. Synthetic Control Method also allows country-by-country analysis compared to more general regression methods. Including appropriate covariate variables eliminates biases that control units could potentially have. For example, if certain control units experienced a regional disaster similar to the one experienced by the main region being studied, or if there is a global event (such as a pandemic) that affects many regions, at the same time, the selection of donors can be adjusted to suit the needs of the researcher. The synthetic control method is extremely well suited for regional policy analysis (Barone and Mocetti, 2014).

1.3.2 Synthetic Control Method model

A brief overview of the mathematical implications of the mode is presented, and the improvements proposed in this chapter are discussed. A partial mathematical intuition behind the optimization sequence is discussed in the appendix. I start with a set $I = \{1, \ldots, N\}$ of so-called "regions" (these can be countries, states, counties). One of the regions is exposed to the "treatment", such as a disaster, where $N - 1$ regions are not treated and a region $i = tr$ is the region exposed. The model includes an outcome variable (GDP) referred to as the treated outcome y_i, and a set of predictors. I assume $y_{i,t}$ is the outcome (or treated) of region i at time t, is the GDP per capita of region i. The outcome variable is observed over T periods. At a point $t = T_0 < T$, the treatment occurs (disaster happens), but only for the affected region $i = tr$, leaving $T - T_0$ of treated periods moving forward, meaning the treatment is uninterrupted. In our case here, the treatment only occurs at $t = T_0$, but Synthetic Control Method also works for treatments that occur after that period as explained by (Abadie, Diamond, and Hainmueller, 2010b). We assume that:

$$y_{i,t} = \hat{y}_{i,t} + \alpha_{i,t} D_{i,t}$$

where

$$D_{i,t} = \begin{cases} 1 & \text{if } i = tr \text{ and } \quad t > T_0. \\ 0 & \text{otherwise} \end{cases}$$

where $y_{i,t}$ is the observed value and $\hat{y}_{i,t}$ is the predicted variable. In other words $y_{i,t} = \hat{y}_{i,t}$ for $t < T_0$ and after the treatment

$$\alpha_{i.t} = y_{i,t} - \hat{y}_{i,t} \qquad \text{and} \quad t > T_0.$$

Our goal is then to estimate $\hat{y}_{i,t}$ to be able to estimate $\alpha_{i,t}$. (Abadie, Diamond, and Hainmueller, 2010b) make the assumption that $\hat{y}_{i,t}$ can be estimated through the following factor model

$$\hat{y}_{i,t} = \beta_0 + \theta_i \boldsymbol{z}_{i,t} + \lambda_i \boldsymbol{x}_{i,t} + \epsilon_{i,t}$$

where $\boldsymbol{z}_i$ is a vector of observed covariates (not affected by the intervention), $\boldsymbol{x}_i$ is a vector of unknown factor loadings, with θ_i unknown parameters, and λ_i a vector of unobserved common factors. To solve this model, we build a set of positive weights w_n where $n = 1, \ldots, N-1$ and $i \neq tr$, such that $\sum_{i=1}^{N} w_i = 1$. There are ideal weights w_i^* such that

$$\sum_{i=1}^{N-1} w_i^* y_{i,t} = y_{tr,t} \quad \forall t \ \in T \quad \text{and} \quad \sum_{i=1}^{N-1} w_i^* z_i = z_{tr} \quad \forall t \ \in T$$

$y_{i,t}$ is defined as any linear combination of the outcome variable at time t for region i using the outcome variable of the other regions. We can then use

$$\hat{\alpha_{i,t}} = y_{i,t} - \sum_{i=1}^{N-1} w_i^* y_{i,t}$$

as a way to estimate $\alpha_{i,t}$ where $t \in [T_0 + 1, \dots, T]$.

The vector $\boldsymbol{z}_i$ for region i is built such that

$$\boldsymbol{z}_i = (\boldsymbol{x}_i; \boldsymbol{y}_i^L)$$

where $\boldsymbol{y}_i^L$ is a vector of pre-treatment outcomes for the treated region. Y_i^L could include any combination of $\boldsymbol{y}_i$ up until the treatment, in other words $\boldsymbol{y}_i^L = \{y_i^0, \dots, y_i^{tr}\}$. In building the vector $\boldsymbol{z}_i$, Cavallo et al. (2013) and Abadie (2021) use the first half of the pre-treatment period outcome observations, and reserve the other half for out-of-sample validation or

$$\boldsymbol{z}_i = (\boldsymbol{x}_i, y_i^0, \dots, y_i^{tr/2})$$

This is the most common out-of-sample validation method, and has been used in literature from Bouttell et al. (2018), Donohue, Aneja, and Weber (2019), Mills and Rüttenauer (2022), Alfano, Ercolano, and Cicatiello (2021), and Li and Shankar (2020a). The training periods chosen for this thesis are not always chosen as the first half of the pre-treatment observations. The training set is determined depending on the amount of available pre-treatment observations and the value of the Mean Square Predicted Error. Clearly there is a trade-off between the two as a short training period is more likely to result in a higher Mean Square Predicted Error, and the lowest Mean Square Predicted Error is obtained by using the full pre-treatment period for training. In his paper Abadie (2021) does not recommend a specific way to set the training data period.

Taking all of this into consideration, the training data points are chosen randomly between 1/4 at the least and 3/4 at most of the pre-treatment observations, as long as at least 8 observations are possible for training. For example, if the number of data points available for the pre-treatment observations (y) is less than 15, in other words, the treatment starts in 1985, the minimum training period can be is half of the length of the values pre-treatment, guaranteeing at least 8 out of 15 observations will be used in training, otherwise Mean Square Predicted Error is large. The training period with the lowest mean-square predicted error for pre-treatment training values is chosen, and then out of sample-validation is conducted on the rest of the pre-treatment data.

If the distance between 0 and tr is odd, I round up for more accuracy. X_i is the set of predictor variables, as described earlier, with the explanation already provided for their use. The vector Z_i is then defined as the vector of covariates. To get around the lack of data for certain disasters, especially since a lot of the data is missing with regards to the corruption perception index (starts from 1995), certain covariates are included only if data is available. The vector V is key in finding the optimal weights, and improvements to the selection of this vector is explained in the following subsection.

1.3.3 Improvements to selection of vector of relative importance

The weights for the donor pool are chosen in such as a way as to minimize a penalty function

$$\underset{W^*}{argmin} ||\boldsymbol{z}_1 - W\boldsymbol{z}_0|| = \sqrt{(\boldsymbol{z}_1 - \boldsymbol{z}_0 W)' V (\boldsymbol{z}_1 - \boldsymbol{z}_0 W)}$$

where $\boldsymbol{z}_1$ is a vector of pre-treatment variables relevant for the treated region and $\boldsymbol{z}_0$ is the same vector of variables for the non-treated regions, and V is a positive semi-definite matrix that highlights the relative importance of every co-factor variable in determining the treatment variable. The choice of V is done in such a way as to replicate the path of the outcome variable of the treated "region", by minimizing the distance between the variables of concern, which means it cannot be arbitrary.

A general case for the initial guess for V is for it to be data driven, based on the treated region, i.e. including the data values as a guess for V. The values in this vector are all positive, with the first element of the vector always having a value of 1, and the remaining values reflect the relative importance of the other variables in determining my outcome observations. For this chapter, this includes the pre-treatment outcome variable or the GDP per capita prior to the occurrence of the disaster, as well as the covariates selected for this regression. The selection of the non-treated variables is also crucial in ensuring that the values of the weights sums up to 1. The values obtained from the matrix V reflect how important each variable in $\boldsymbol{z}$ is in determining the synthetic treated variable. These V values we obtain help us define the relative importance of the covariates. Once V^* is obtained, I then find the vector of weights W^* that minimizes the following distance:

$$\underset{W^*}{argmin}(\boldsymbol{y}_1 - \boldsymbol{y}_0 W^*(V^*))'(\boldsymbol{y}_1 - \boldsymbol{y}_0 W^*(V^*)).$$

An improper selection of V may lead to an improper solution to the minimizing problem and a choice of V that does not minimize the mean squared prediction error (Mean Square Predicted Errors) of the outcome variable. I augment the standard Synthetic Control Method technique of selecting V^* by using a two-step selection method. Countries are first grouped by the similarity of their dependent variable, in this case, GDP per capita PPP, prior to the treatment. For the selection of countries available three groups are created low income ($<$ \$5000), middle income($<$ \$15000 and $>$ \$5000), and high income($>$ \$15000). This selection is made because the importance of covariates is likely to vary between these groups. For example, the importance of government expenditure or secondary school enrolment in determining GDP could vary significantly between a low-income and a high-income country. Once countries have been sorted into these groups and for a given treated period, I run a synthetic regression for all these countries given the specified treated period and for an initial guess of V being data-driven. The average Mean Square Predicted Errors (Mean square predictive error) of the pre-treated periods is then calculated for every country in every group prior to the treatment period T_0. The optimal V^* of the country with the lowest Mean Square Predicted Errors in their respective region is then chosen, and this vector is used as a guess for any future regressions for that particular group. When conducting an Synthetic Control Method regression for the desired period and country, the guess is then selected from a pool of already calculated vectors. This thesis provides a new way of obtaining the optimal vector of relative importance, and highlights the benefits of using this two-step procedure. This method leads to faster and more accurate convergence to an optimal V^* than the data-driven guess when running multiple regressions and placebos while either lowering or not changing Mean Square Predicted Errors for pre-treatment outcome variables. This comes at the cost of increasing the time needed to conduct successful regressions for smaller data sets. As the number of regressions conducted increases, the time required to conduct the regressions decreases as the initial guess of the vector V is more accurate.

1.3.4 Inference

In terms of statistical significance, the synthetic control method does not rely on traditional inference tests but rather on "placebo" tests. As per Abadie, Diamond, and Hainmueller (2010b) and Cavallo et al. (2013), the p-value of the level of significance of a disaster is obtained using the following:

$$p\text{-value} = \frac{\sum_{np=1}^{N_{pl}} I(\bar{\alpha}_l^{pl(np)} < \bar{\alpha}_i)}{N_{pl}}$$

where N_{pl} is the number of placebo tests conducted. In the context of synthetic regression, placebo refers to the counterfactuals obtained from running an Synthetic Control Method regression on a region that was not treated for the selected treatment period. The goal of such a process is to see whether the deviation of the synthetic country's GDP from the actual country's GDP is larger than that of a potential placebo region. In other words, in a random country not affected by this disaster. To conduct this inference Cavallo et al., 2013 propose the following method:

1. For every disaster, compute the place effect using the available controls for the corresponding disaster

2. At every point in time following the occurrence of the disaster (called leads) compute all the placebos, and then take the average across all placebos

3. The actual lead average is ranked in the distribution of placebo averages

4. The lead specific p value is given by the following formula

$$p\text{-value} = \frac{\sum_{np=1}^{N_{pl}} I(\bar{\alpha}_l^{pl(np)} < \bar{\alpha}_i)}{N_{pl}}$$

where α_l is the effect of the disaster on the country in question and α_i is the placebo effect

There are caveats to using this particular method. First, the full data set cannot be considered as potential placebos. Some regions in the pool of donors could have also experienced disasters or significant events that impacted their GDP during the same period as the treated region. These countries need to be excluded from the list of placebos since a significant event in two regions needs to be compared to insignificant events. Second, the placebo test cannot include countries that are considered outliers as the model cannot find enough donors to build the counterfactual of the treated variable (GDP per capita PPP) accurately, given the constraint of $0 \leqslant w_i \leqslant 1$, since it would be impossible to build a counterfactual for an outlier such that $\sum_{i=2}^{N} w_i y_{i,t} = y_{tr,t}$ This constraint can be relaxed and placebos can be formed using those outlying regions, and Li and Shankar (2020b) show that this does not alter the results. The only noticeable drawback of relaxing this assumption is significantly higher computational times. Given the large enough data set, and the small amount of outliers, inference can be successfully conducted without having to include the outlying regions.

From the possible set of countries used in most inferences tests done in this chapter, the following countries are excluded: Burundi, Qatar, Luxembourg, Sierra Leone, Botswana, and Switzerland. These countries display the most extreme outcome variable values for a large portion of the time series considered. This means that it is almost impossible for the model to build an exact synthetic counterpart for each one without a noticeable increase in computational time. Adding to this, any country that experienced a significant disaster (as mentioned previously) three years prior to the treatment period (not just those that experienced an event during the same year) is excluded from being in the placebo pool. The reasoning is simple since a placebo is supposed not to have experienced treatment at T_0. Given the nature of disasters and their prolonged effect, it is expected that if a country experienced a disaster, this effect might still be noticeable several years after its occurrence, impacting the accuracy of the placebo.

Finally, a limit to the inference period is imposed and is one of two criteria: 10 years post-disaster or the closest significant disaster, depending on the shortest period. For example, Mexico experienced massive earthquakes in 1985 and 1995, and the closest selected disaster to the 1995 earthquake was in 2010. Therefore, the inference period for the 1985 earthquake is nine years, the inference period for the 1995 earthquake is ten years, and the inference period for the 2010 disaster is eight years as it is data-limited. This is done for every inference test. A treatment is considered significant at p-values less than 0.1 or 90% significance. Changing this 0.05 results in the loss of only three significant events.

To sum up, the contribution to the existing literature is two-fold. In terms of the model, the selection of the V-vector improves accuracy of the simulations but increases computational time. This vector reflects the relative importance of every co-factor in the model, so an accurate selection of this vector results in more accurate regressions. Several propositions have been made to estimate this vector, such as out-of-sample validation (Abadie, Diamond, and Hainmueller, 2015). This is improved by grouping all the units (regions) depending on the values of their treated observations at a specified point in time (the occurrence of the treatment). The number of these groups is chosen to be three because it allows for he best trade-off between accuracy and speed, meaning the lowest 1/3 of the outcome observations in terms of values are put in one group, the second 1/3 in another, and the final 1/3 in another. A representative vector for each group is found, which would then be the initial guess for each simulation for a member of that group. This technique is robust to out-of-sample validation, as all the simulations conducted in this chapter relied on an out-of-sample selection for the vector V with no loss of accuracy in predicting the pre-treatment outcome variable.

1.4 Empirics and data

1.4.1 Data sources

To conduct this analysis, data were obtained from several sources. The GDP data selected is available from 1975 to 2018. Co-factor variables are selected if they are statistically significant in determining GDP per capita. The literature on their importance is well established, so this subsection only summarizes the variables and their sources. The list is as follows:

- *Share of value-added (Agriculture and Industry)*: I use the World Bank Database for a panel series dataset from 1970 to 2018, that describes the share of two different sectors in the economy in the value added.

- *Secondary enrolment rate:* I use the World Bank Database for a panel series dataset from 1980 to 2018.

- *PolityIV:*[1] I use the Polity5 project that codes the degree of democratization of a country. The values range from -10 (full autocracy) to +10 (full democracy), and data is available from 1960 to 2013.

- *Capital stock at current PPP:* I use the capital stock data obtained from FRED that uses the perpetual inventory method of calculation.

- *Corruption Perception Index:* I use the corruption perception index values calculated by Transparency international.[2] This data set is based on several sources that have to qualify to certain criteria. I also use this data set in my regression. This is an additional control variable in the synthetic regression model.

- *Population, Labor Force, Enrolment rates, and Trade openness:* I use the World Bank Database to obtain these data series from 1975 to 2018. Not all countries have sufficient data points.

- *Disaster Data:* I use the EM-DAT database from the Catholic University of Louvain in Belgium. The data covers disasters from 1970 until 2018. This data is not used in the synthetic regression calculations, but is used to generate the disaster list, which includes the country and the time of the disaster of interest.

Both *polity* and corruption perception (cpi) are used as proxy measures of institutional quality in the regression analysis done in section 1.5.2. Several other variables are added to the analysis of the results that are not included in the Synthetic Control Method, as there is not enough data for the entire sample of countries. Despite having data for 98 countries, and those countries being the pool of donors, this chapter only focuses on the impact of disasters on a list of 30 countries. This list is separated into two regions, South East Asia (SEA) and South and Central America (SCA). Initial work had been done on the full set of 98 countries prior to narrowing down the list. Simulations were inaccurate for pre-treatment periods for some countries that either lacked sufficient covariate data or were on the edge of the convex set of countries concerning their

[1] https://www.systemicpeace.org/polityproject.html
[2] https://www.transparency.org/en

GDP. Not enough suitable donor countries were found to build counterfactuals due to the extreme values of the dependent variable of these countries. The final countries selected do not exhibit such qualities, making the estimations much more reliable. Countries in one of the regions mentioned before tend to have similar economies, especially regarding GDP per capita in 2017 international U.S. dollars using purchasing power parity rates.

Regressions on richer countries found no impact of disasters on GDP or insignificant results. The final list of countries and the relevant and significant disasters can be found in Appendix A. These regions were chosen for two reasons. First, the disasters affecting each region are very similar in size and nature, mainly due to their proximity. Second, the quality of institutions between these regions is significantly different, with outliers in each region. So while SA countries are on the lower end of the institutional quality spectrum, except Peru, Ecuador, and Chile, most of the SEA countries chosen are on the higher end. Other minor reasons for selecting these regions are the availability of complementary data (GDP per capita, consumption per capita, trade immigration/migration data, accuracy of the data reported, and the frequency/impact of large-scale disasters). Improvements in this chapter could include more regions for robustness.

1.4.2 Disasters

In this subsection, the selection process of disasters is described in detail. As mentioned previously, the impact of disasters is more than just on physical capital, which is why in this chapter, the criteria for disaster are expanded to consider the impact on human capital by considering the number of affected individuals and capital destroyed. Data for disasters is obtained from the Emergency Database (EM-DAT), collected by the Catholic University of Louvain, which defines a disaster as an event that satisfies at least one of the following criteria:

- 10 or more people are reported killed.
- 100 people are reported affected (displaced); a state of emergency is declared.
- A call for international assistance is issued.

This is a broad definition of disasters and is biased towards smaller countries. The two main criteria I care about are affected individuals and economic impact. The latter refers to the direct value of the destruction caused by the disaster and does not consider any long-term effects, as it is hard to track economic loss months or years after the disaster. The database records direct estimated damage in terms of losses to GDP. The database includes various types of disasters. For this research, I include natural disasters as a general classification group, which includes geophysical disasters such as earthquakes, hydrological such as floods or landslides, biological such as an epidemic, climatological such as droughts and wildfires, and meteorological such as storms. I exclude technological or financial disasters from this list.

It does not make sense to single out unique disasters for two reasons. First, disasters tend to be followed by an influx of financial aid, meaning losses in potential GDP may not be adequately measured. Second, richer countries have been historically more prepared over time for disasters (and become even more prepared the more disasters to happen), which contributes to the decreasing impact of these disasters on GDP over time. This could be due to over-investing in disaster-prone areas, another trade-off that countries must consider. Richer countries will be able to provide this level of investment and often do not rely on external aid to alleviate the impact of the disaster. This means that policy implications are different between rich and developing countries. There is a sizeable gap in the maturity of the insurance markets between the rich and developing countries. This chapter also analyzes areas where these markets are underdeveloped to discuss alternative policy implications for these countries. I argue that the well-preparedness of these richer countries, which includes more mature financial and insurance markets, has contributed to the income gap between them over time, despite significant moral hazard problems. This preparedness is partly due to higher institutional quality that takes advantage of the increase in spending after a disaster and the influx of new, more advanced capital.

In this chapter, disasters are selected depending on the percentage impact on population or GDP. A sample selection criteria for a disaster is the event with the largest impact as a percentage of GDP, without considering the impact on the population. The selection process is similar to that designed in Cavallo et al. (2013) to select a minimum of three disasters in the pool of countries for each country. Using the EM-DAT database, a list of all the disasters for the 98 countries occurring between 1975 and 2018 is obtained. I select disasters specific to the countries in one of the two regions mentioned in this list. The economic and human impact of every disaster recorded is calculated as a percentage of the country's nominal GDP during the disaster and percentage of the total population, respectively. All the disasters for the 98 countries in the pool are obtained to ensure that when conducting inference, no country is also experiencing an event in the same time frame as the treated country. For every country, two lists are made, the economic list, which ranks the disasters from highest to lowest depending on their impact on GDP, and the human list, which ranks the disasters from highest to lowest similarly. The top three disasters from each list are then selected. This means that three to six disasters are selected per country, depending on whether the economic disaster list selection overlaps with the human disaster one. The disasters obtained are from 1975 to 2018, and for the 98 countries, the result is a total of 457 disasters selected. For the specific regions, the total is 106 disasters or an average of 3.5 disasters per country. For the panel of countries considered and out of the 96 disasters considered, 72 caused significant deviations from GDP (positive or negative) up to 10 years after the occurrence of the disaster.

A key addition presented by this chapter and that has been overlooked by Cavallo et al. (2013) and Barone and Mocetti (2014) is that countries have had to deal with a series of disasters with no predictable interval for their occurrence. Therefore, any analysis of the impact on long-term growth needs to take that into account, which means that any subsequent disasters limit the scope of the impact of a particular disaster. By ranking these events in terms of their effect on GDP and population and choosing the top disasters such that there is a minimum of five years between each disaster considered for the Synthetic Control Method, I guarantee that the events chosen are infrequent and significant enough to be classified as "disasters", but also providing an adequate time frame to analyze the long term effect of these events.

It is important to note that disaster impact is biased towards smaller or poorer countries, meaning the sample of disasters studied disproportionately includes poorer countries. Richer countries with better institutions have likely been able to mitigate the effect of these disasters over time through proper planning. This bias is tackled by not including minimum thresholds for disasters to be considered. In other words, in previous research, a disaster needed to have a certain minimum level of impact on GDP to be included in the sample selection. Given that the sample of countries in the two regions selected includes richer countries, not setting a minimum threshold allows for a broader sample. The results from the Synthetic Control Method indicate that their inclusion is significant, even if the impact is not as pronounced compared to poorer countries.

While in this chapter only natural disasters are included, a possible extension would be the addition of financial, technological, or political disasters. This would also expand the possible sample of countries studied. However, it would require a change to the covariates and the model.

1.4.3 Region selection

Two sets of countries were chosen for analysis in this chapter. The properties set for this selection were done in such a way to ensure that events were significant and comparable. The first group is South and Central America (SCA), and the second is South East Asia (SEA). The choice of groups was decided by the similarity of events that affected each country in this group and the high variance in institutional quality and GDP per capita. When comparing institutional quality, a wide range is observed both between and within groups. Most SCA countries are significantly lower ranked than their SEA counterparts in terms of corruption, regulatory power, and stability rankings. This, however, does not translate similarly to GDP per capita values. Even within each group, countries exhibit radically different institutional quality values. These values are also not static and varied over time.

Country	*Polity*	*Corruption Perception Index*	*year*
Mexico	8	3.5	2005
Argentina	-8	n/a	1988
Chile	10	7.2	2010
Peru	9	3.5	2007
Ecuador	-9	n/a	1992

Table 1.2: Polity and corruption index for select SCA countries

Country	*Polity*	*Corruption Perception Index*	*year*
Indonesia	-7	n/a	1994
Philippines	-8	3.6	2013
Australia	10	8.8	1996
Malaysia	4	n/a	1985
Singapore	-2	9.1	2000

Table 1.3: Polity and CPI index for SEA countries

1.5 Results

The limitations of the data used and the selection criteria for the disaster list allow the analysis of no more than ten years of post-disaster impact across all disasters. This is because the interval of the occurrence of a disaster for some countries is sometimes less than ten years. However, in the selection process of the disasters, a minimum of five years between disasters was imposed. The post-disaster average deviation for all countries is calculated for a maximum of 10 years. The final selection is the three-year and five-year averages. The three-year average is considered by the literature as usually the long-run effect, and the five-year is where most disasters end up peaking in terms of their effect, as seen in figure 1.2. This subsection presents the results of both the simulations from the Synthetic Control Method and the regressions on the deviations (or α) from the Synthetic Control Method. The first subsection discusses the average impact of a disaster on GDP, while in the second subsection, and using an OLS regression; different institutional quality variables are regressed on these deviations to quantify their impact and importance.

1.5.1 Results from changes to V-selection

This subsection highlights the impact of the changes to the vector of relative importance through stylized examples. The goal is to understand how these changes affect the Mean Square Predicted Errors (MSPE) for pre-treatment outcome variables. A smaller value of this number leads to more robust conclusions about the impact of the treatment. The selection of regions and treatment periods impacts the results of this robustness test, however, the conclusion is that Mean Square Predicted Errors are on either reduced or the same as the base Synthetic Control

Model. Computational time depends on the number of regressions conducted. All regressions were conducted and recorded using a 2021 Apple Macbook Pro (M1 Max chip), with 32 GBs of RAM, running Python 3.

Three tests were conducted to compare the Mean Square Predicted Errors from the base Synthetic Control Method to the augmented version of the model (grouped version). The tests were done using $N = \{20, 30, 40\}$ regions as donor pools. For each pool, three groups were selected for the tests. To stay consistent with the method proposed in this chapter, the groups were chosen based on the ranking of the outcome variable for each country, where the first group included the highest one-third values, the second, the middle one-third, and the third group was composed of the bottom one-third values. This grouping can be altered to include more or fewer groups or by changing the grouping criteria. This grouping is ex-post because obtaining the data used in obtaining the optimal weights for the Synthetic Control Simulations is only up to the treatment i.e. T_0. Using ex-ante data (or post-treatment) data may lead to higher Mean Square Predicted Errors and bad fits. The dataset of regions and treatment periods was selected randomly from the available data used in this chapter.

Tests were conducted twice on a sample of three, five, seven, and ten random countries from these pools, with the rest of the countries used in the simulation of the Synthetic Control Model. Significance does not matter in this case since the purpose of the test is to simply find the Mean Square Predicted Errors pre-treatment. Since the number of pre-treatment periods impacts the convergence of the Synthetic Control Method, the results displayed in Table 1.4 capture the average Mean Square Predicted Errors difference between the grouped version of the Synthetic Control Method and the base model. A negative value indicates that the base Synthetic Control Method has a higher Mean Square Predicted Errors than the grouped version.

Pool size \ Sample Size	3	5	7	10
$N = 30$	-10.2**	-23.4***	-1.5**	-4.3**
$N = 40$	0.8	-30.2***	-20.3**	-0.05
$N = 50$	0.01	-5.8**	-3.4**	-0.85*
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table 1.4: Mean Square Predicted Errors differences

In every case, the augmented Synthetic Control Method resulted in lower or similar pre-treatment Mean Square Predicted Errors for the pre-treatment variables. This implies that the weights calculated are a better fit for the model and would provide more accurate results for forecasted variable. This adds confidence to the validity of the results from the Synthetic Control Simulation, leading to more robust conclusions. The trade-off, in this case, comes at the cost of higher computational time. The average computational time to complete regression is highlighted in Figure 1.1. The augmented model is significantly slower, but the average time per regression decreases as the number of regressions increases. This augmented version is therefore helpful for large data sets.

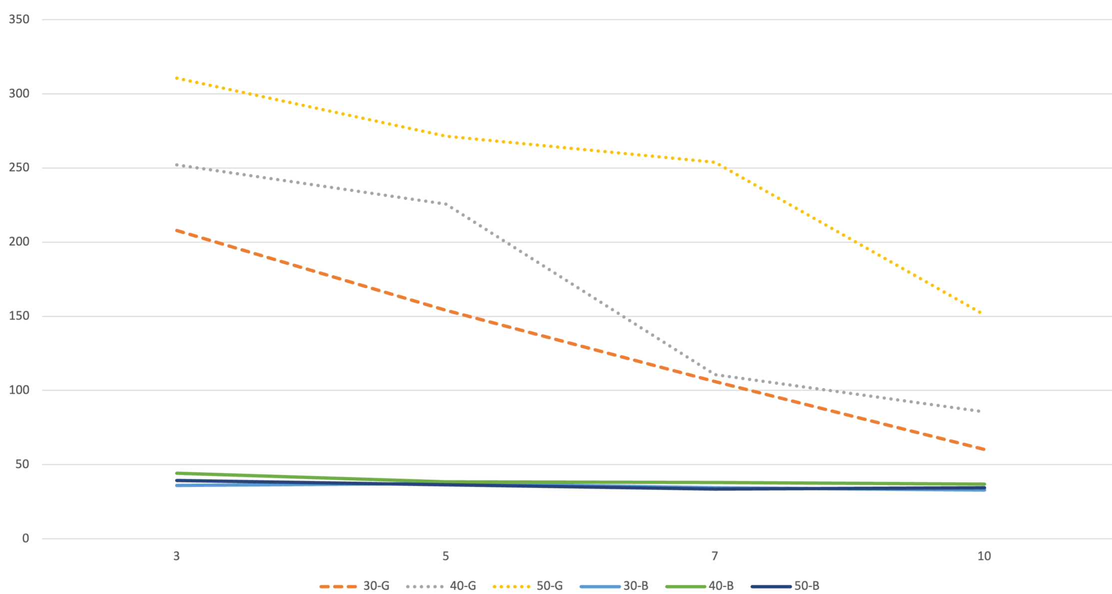
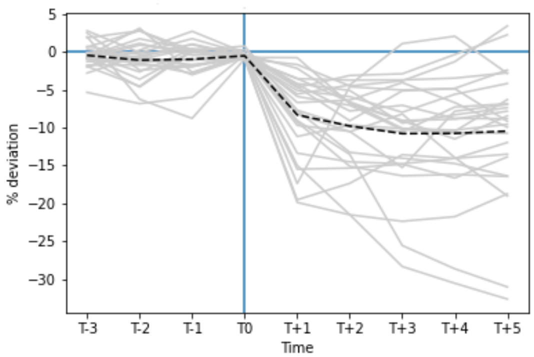
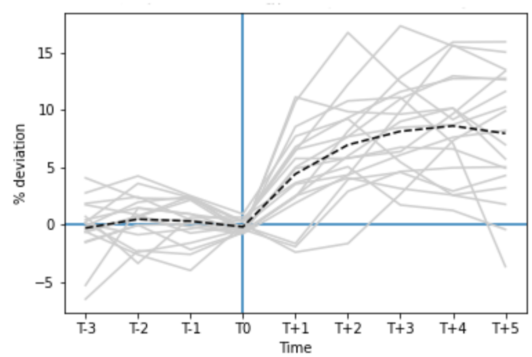


Figure 1.1: Average time per regression (in seconds)



(a) Disasters with negative recovery



(b) Disasters with positive recovery

Figure 1.2: Percentage deviations for significant events

1.5.2 Impact on GDP from Synthetic Control Method

Running the synthetic simulations described above for all possible countries provides 45 statistically significant disasters from the set of 106 disasters chosen. The deviations of the synthetic model from the actual data are saved for all periods after the occurrence of a significant event. Out of the 45 observed events, 26 resulted in negative post-recovery differences (three-year and five-year average), indicating that the synthetic regression predicted a higher value of GDP per capita than its data counterpart. The split in the results matches the differences in results in previous research (Ikefuji and Horii, 2012) (Akao and Sakamoto, 2018), and the results include disasters that had both a positive and a negative impact. Figures 1.2b and 1.2a show the various percentage deviation between the predicted minus the treated variable five years after the occurrence of an event. Each grey line represents one statistically significant event, and the dotted line represents the average of all events. This dotted line is also presented in table 1.5. While this difference becomes minimal for certain countries after five years, the majority show the long-term implication of these events. These values are essentially the dotted line in figures 1.2a and 1.2b.

Several conclusions can be made from these results. First, while the most significant positive recovery was around 15% higher than the actual GDP, this amount was more than half the negative recovery values. It is a result that is to be expected, as it is harder to rebuild. While rebuilding can be more beneficial than staying on the post-disaster path under certain conditions, disasters will often just exacerbate an already underperforming system, heavily stifle growth, and impact

GDP for extended periods.

To understand what these conditions are and what are the underlying factors behind this split, several OLS regressions are done on the data obtained from the Synthetic Control Method using both three and five-year average differences. These regressions include various proxies of institutional quality. Table 1.5 shows the average disaster value for the data set, for both the negative and positive values.

	Negative		Positive	
Time period	Mean	SD	Mean	SD
$T-3$	-0.46	1.69	-0.29	2.36
$T-2$	-1.09	2.57	0.47	2.02
$T-1$	-0.97	2.38	0.30	1.84
T_0	-0.52	0.42	-0.19	0.508
$T+1$	-8.32	5.46	3.41	1.78
$T+2$	-9.79	5.27	6.95	2.90
$T+3$	-10.79	5.96	8.13	3.96
$T+4$	-10.76	5.76	7.79	2.44
$T+5$	-9.96	4.47	6.25	5.35

Table 1.5: Average impact of disasters

1.5.3 Importance of political institutions

In analyzing the impact of political institutions on recovery, simple OLS regressions are conducted, once for the three-year post-disaster GDP averages and another for the five-year average. The goal is to understand the impact of *polity*, corruption perception index, and other institutional quality variables on this deviation. The model is as follows :

$$D_{i,t} = \alpha + \beta_i X_i + \gamma_i Z_i + e_i$$

where D_i is the average deviation from actual GDP for the countries in the sample for the inferred post-treatment dates for country i and periods t, depending on the model, this can be the three year average or five year average. X is a matrix of dependent variables and Z is a matrix of controls. I include the following variables in X:

- **Regulatory power:** I use data calculated by (Kaufmann, Kraay, and Mastruzzi, 2010), to estimate the regulator power in a country. It captures perceptions of the ability of the government to formulate and implement policies and regulations that develop the private sector. The data is normally distributed, and ranges between -2.5 to 2.5.

- **Government effectiveness:** I use data calculated by Kaufmann, Kraay, and Mastruzzi (2010). The data captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures. It is normally distributed, and ranges between -2.5 to 2.5.

- **Violence:** I use data calculated by Kaufmann, Kraay, and Mastruzzi (2010). The data captures frequency of acts of violence and civil disobedience. It is normally distributed, and ranges between -2.5 to 2.5.

- **CPI and Polity:** Data for the Corruption perception index is limited, and only extends to 1994. If a disaster happened during a time where there is no corruption perception index data, the average of total CPI is used as a proxy. Dropping this and lowering the sample does not lead to any change in significance in the model. *polity* is an index that refers to the degree of democratization of a country. It is between -10 and 10, where 10 indicates a full democracy, and -10 refers to a full authoritarian regime.

- Z **is a vector of controls:** This vector includes population, labor force, land size, and interest rates, as explained in earlier section, as they are significant variables when determining GDP.

The motivation behind including *polity* in addition to the other institutional quality is to highlight the importance of democratization in disaster recovery. *Polity* measures patterns of authority, and is essentially an index of the level of authority in a country. A score of 10 indicates full democracy, but is not fully reflective of the quality of institutions, so while Denmark and Cyprus have the same polity score, the quality of the regulatory institutions are not the same. The results of the regression can be seen in tables 1.6 and 1.7. The conclusions that can be drawn from these table is that political government regulatory power and corruption play a significant role in determining how well a country recovers from a disaster.

This is an opposite conclusion to that of Cavallo et al. (2013) who finds that large enough disasters induce political revolutions, linking the impact of disasters on GDP to these revolutions. The results in this chapter challenge this idea to a certain extent. The *polity* index for the affected countries seems insignificant, and regime changes are not necessarily the direct reason why disasters impact GDP, but rather the deterioration of political-institutional quality. This is based on the fact that a revolution leads to a significant decrease in the polity index of a country.

Insurance markets play a crucial role in hedging against disasters for both households. There is a lack of comparable insurance data for the countries selected in this chapter, which does not allow for a direct comparison of the insurance markets between countries. Historical values for variables such as regulatory power, government effectiveness, and the corruption perception index serve as valuable proxies for the ability of a government to direct capital to where it is needed most and the confidence of citizens as well as foreign governments in funneling aid to governmental agencies. This analysis does not extend, however, to the trade-off between preparing for a disaster prior to its occurrence and the impact of the disaster.

The results from the regressions shown in Tables 1.6 and 1.7 provide evidence that, while controlling for various factors, cpi and regulatory power are the only statistically significant variables in determining post-disaster recovery, with regulatory power being the strongest of the two and the more significant. A one standard deviation off the distribution, in other words, being in the top 30% of countries regarding regulatory power, leads to a 13% increase in recovery rate. This can be observed in the row titled *reg* in the regression tables. A country is expected on average to recover 13% above their path prior to the disaster, five years after the disaster occurs if they are in the top 30% of countries in regulatory power. Similarly, being on the opposite end will lead to lagged recovery and a lower GDP 5 years after the disaster than had the disaster not occurred.

For example, Panama experienced two disasters of relatively similar impact. One was in 1989 and the other in 2013. Rocked by political turmoil prior to the first disaster, Panama's recovery for the following five years proved to be difficult, almost 16% less than the synthetic model. Without

	Reg. 1	Reg. 2	Reg. 3	Reg. 4	Reg. 5
cpi	1.358*	1.311*	1.281*	1.295*	1.271*
	(0.774)	(0.816)	(0.774)	(0.831)	(0.824)
pol	0.2281	0.2240	0.2281	0.2273	0.2191
	(0.3147)	(0.3320)	(0.4157)	(0.3523)	(0.3451)
gov	-7.8913	-7.9961	-7.8521	-7.1261	-8.1961
	(5.4229)	(5.7210)	(5.3216)	(5.4253)	(6.3147)
reg	13.0713**	13.3504**	13.2310**	13.4154**	13.2152**
	(6.2258)	(6.5680)	(6.4312)	(6.3156)	(6.1283)
vio	3.7847	4.0835	4.52261*	4.0835	4.1422*
	(2.8122)	(2.9668)	(2.0431)	(2.512)	(2.1318)
labor force			1.2041	1.4512	1.041
			(1.4511)	(1.410)	(1.325)
land size				0.002351*	0.002281*
				(0.001161)	(0.001025)
interest rate					0.5281***
					(0.01147)
R-squared	0.2305	0.2220	0.2013	0.2512	0.4211
R-squared Adj.	0.1343	0.1248	0.1281	0.1247	0.2255
N	45	45	45	45	45
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 1.6: OLS results for three-year averages

	Reg. 1	Reg. 2	Reg. 3	Reg. 4	Reg. 5
cpi	1.632	1.512	1.415	1.421	1.364
	(0.912)	(0.879)	(0.945)	(0.913)	(0.925)
pol	0.142	0.135	0.131	0.131	0.148
	(0.121)	(0.113)	(0.107)	(0.101)	(0.098)
gov	-7.315*	-7.321	-7.451	-7.422	-7.962
	(4.381)	(4.765)	(4.915)	(4.453)	(4.221)
reg	10.522**	10.415**	10.387*	11.327**	13.2152**
	(4.314)	(4.217)	(4.597)	(4.512)	(5.323)
vio	6.821	6.587	6.891	6.098	6.922*
	(5.210)	(6.124)	(5.997)	(6.014)	(5,326)
labor force			0.665	0.621	0.258
			(1.331)	(1.521)	(1.425)
land size				0.0032**	0.0025**
				(0.0041)	(0.0056)
interest rate					0.4219**
					(0.1427)
R-squared	0.2014	0.2025	0.2142	0.2314	0.4211
R-squared Adj.	0.1285	0.1275	0.1264	0.1235	0.3187
N	45	45	45	45	45
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 1.7: OLS results for five-year averages

going into the political details of the disaster, the drop in politicization, regulatory power, and governmental effectiveness prior to the 1989 disaster contributed to the slow recovery. The period from 1990 to the mid-2000s marked an increase in government effectiveness and a reduction in corruption as the country became more democratic. Compared to 1989, Panama's recovery from the disaster in 2013 is better than expected, averaging 12% over its synthetic counterpart over the following five years.

1.5.4 Robustness check

In order to test the validity of the results, several robustness checks are conducted. First, the data is split into two different groups. The first group includes the countries where the post-disaster recovery leads to a positive difference, i.e., the predicted GDP is higher than the actual GDP path, and the other group is the inverse. In addition, the Quality of Government variable obtained from the quality of government project run by the Department of Political Science at Gteborg University is added to the regression. It is a variable that considers the core features that determine the QoG such as impartiality, bureaucratic quality, and corruption, as well as measures that are broader such as the rule of law and transparency. This is done to account for omitted variable bias. The results in the appendix show that splitting the data or adding variables do not add to the significance of the model or affect the significance of the previously mentioned variables. Regulatory quality and cpi are still the only statistically significant values. The tables can be found in Appendix B.

Reverse causality of the quality of institutions and democratization is another important factor to consider here. Several authors, such as Jong-A-Pin (2009) found evidence of reverse causality between political instability and growth. To test for this, a two-stage regression with instrumental variables is applied. In the first stage, the endogenous explanatory variables are treated as functions of their instruments. The instruments considered are: education levels (primary and secondary), distance to equator (Hall and Jones, 1999), and mortality rates (Acemoglu, Johnson, and Robinson, 2001). The predicted value is used as an explanatory variable in the second stage in the equation. In this case, the original function's endogenous variable is the regressor. In testing for endogeneity, a few instrumental variables are selected to control for the endogeneity between the exogenous variables and endogenous variables.

	polity
cpi	-0.24
polity	1.0
gov	-0.18
reg	-0.25
vio	-0.03

Table 1.8: Correlation matrix for polity and other explanatory variables

Another potential issue is the link between any explanatory variables and *polity*. I test for multicollinearity and show the correlation matrix for *polity* and the other variables. No significant evidence of collinearity exists in this case. To test for endogeneity a simple Hausman test is conducted and for both models listed above the p-valeues are respectively $p_3 = 0.24$ and $p_5 = 0.31$. The null cannot be rejected and no evidence of endogeneity is found for the quality of institutions and growth rates. While things like *polity* and interest rates are crucial in GDP determination, the variables selected in the model do not display any statistically significant sign of endogeneity. While the literature does mention political instability, *polity* itself is not necessarily a measure of instability, but rather democratization, which explains the lack of endogeneity in the model.

1.6 Simulations from Synthetic Control Method

This section presents the results of the Synthetic Control Method regressions conducted on both regions. A sample of these results is provided in figures 1.3 and 1.4, and the rest are shown in the appendix. The dotted lines show the predicted GDP per capita values, while the solid red lines show the actual values of the treated variable. Some notable observations for each region. are also provided.

1.6.1 South and Central America

In the SCA region, the biggest disaster impacting the following countries are considered: Argentina, Brazil, Bolivia, Mexico, Chile, Ecuador, Dominican Republic, Honduras, Colombia, Paraguay, Panama, Uruguay, and Peru. Table 1.2 presents an exciting range of the *polity* and corruption levels of these countries during highlighted periods. For example, in the case of Mexico, three major natural disasters are chosen. These occurred in 1985 (earthquake), in 2005 (earthquake), and finally in 2010 (flood), but only two of these were statistically significant. The earthquake in 1985 led to an increase in public scrutiny and increased expenditure on infrastructure. However, prior to this, the quality of institutions was relatively poor, which is why the predicted values for Mexico are higher than the actual values, meaning this disaster had adverse effects, even ten years after its occurrence. There was also a severe restructuring of public safety codes, infrastructure investment, and improvements in relief and response efforts which would be necessary in dealing with upcoming disasters (UNDRR, 2017). Despite a worldwide financial crisis in 2008, the disaster that hit Mexico City, one of Mexico's biggest cities in 2010, seemed not to have as big of an impact on the GDP per capita in Mexico. However, Mexico's infrastructure investment, regulatory power, and cpi index prior to this disaster were relatively poor, even with higher *polity* values (in 1985, polity was -3, in 2007, it was 9). The recovery is statistically significantly lower than expected, similar to what happened in 1985.

For the countries in this group, the disasters that seem to have the most significant impact are earthquakes, which is a benefit for comparing across countries and regions. In contrast, the most devastating type of disasters in SEA are hurricanes and floods.[3] When conducting the OLS regression on the deviations and controlling for the various types of disasters, there is no statistically significant difference between institutional quality and the impact of the disaster on average.

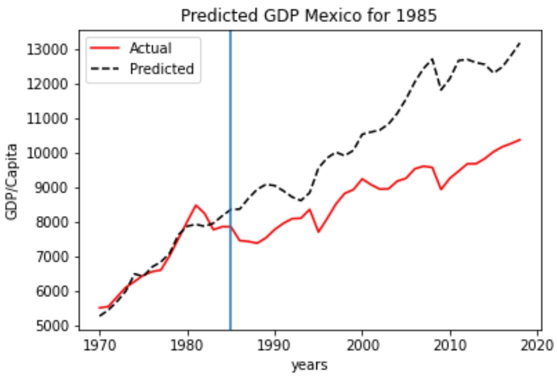
Figure 1.3 shows the synthetic analysis for a selection of countries from this subset. As mentioned in section 1.3, only the first half of the pre-treatment outcome variable was used to estimate the weights, thereby testing the model's validity over the second half of that period as an added

[3] I cannot state whether the type of disasters can make a difference in the recovery process, but this is work that can be developed in future research.

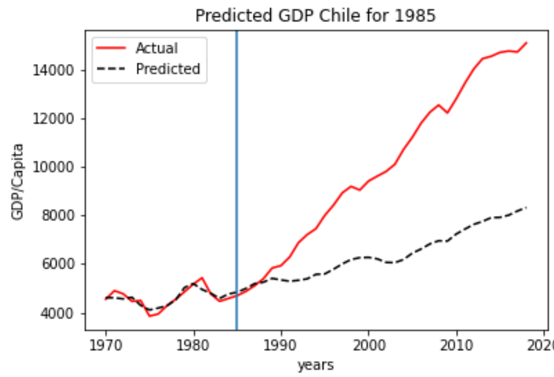
measure to check the accuracy of the simulation. When looking at figure 1.3, the results of the Synthetic Control Method display a fairly good pre-treatment fit to the data. This lends even more credibility and robustness to the inference results. Except for Argentina, most regressions were good with very low pre-treatment Mean Square Predicted Errors values. Inference tests are conducted for every treated period regardless of pre-treatment Mean Square Predicted Errors values, and the p-values for the forecasted periods are obtained for each simulation. These are displayed below the graphs for the events chosen. Since the objective is to calculate the impact of these disasters, the percentage deviation of GDP for a given period post-treatment is calculated depending on the maximum post-treatment periods possible. The results of these deviations are shown in Table 1.5, which provides the average impact of disasters up to five periods after one has happened.

In terms of politics, the SCA countries display many fluctuations in different institutional quality variables, as seen when looking at *polity.* For example, Ecuador, one of the poorer countries in Latin America, was subjected to a financial crisis in 1998 which led to a political crisis soon after, where the president was ousted and replaced by his vice president a year later. The result was a drop in polity over the years that improved over time. However, the quality of the political institutions was rather low, which meant that Ecuador performed relatively poorly in recovering from three disasters in 1982, 1987, and 1993.

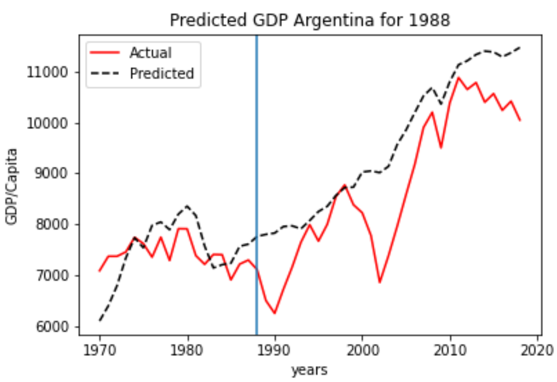
While Mexico and Chile were affected by devastating earthquakes in 1985, Chile's post-disaster recovery is better than Mexico's. At the same time, Mexico underperformed relative to its synthetic counterpart, and Chile over-performed. As seen in table 1.2, Chile's polity index is -1. A few years later, Chile's polity index jumped from -6 to 8. In context, Chile's military regime was ousted in 1989. The 1985 earthquake caused more than 1 Billion dollars worth of damages and killed 177 people. Due to the military's inability to deal with this earthquake, the Finance minister at the time implemented reforms in the form of privatization of the construction sector, loosening the military's grip on a crucial sector of the economy (Vogler, 2010). In 2010 Chile experienced another devastating earthquake, and the predicted GDP 8 years after was 6% lower than the actual GDP, potentially owing to these reforms that occurred almost two decades before.



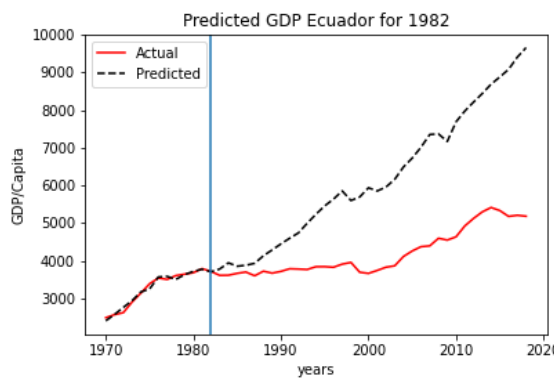
(a) Mexico 1985, p-value = 0.08



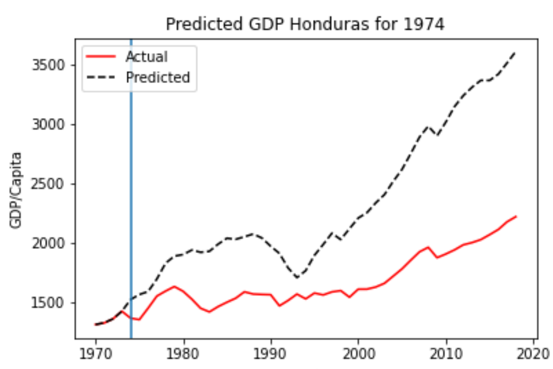
(b) Chile 1985, p-value = 0.04



(c) Argentina 1988, p-value = 0.45



(d) Ecuador 1982 , p-value = 0.07



(e) Honduras 1974, p-value = 0.21

Figure 1.3: Synthetic Control regressions (SCA)

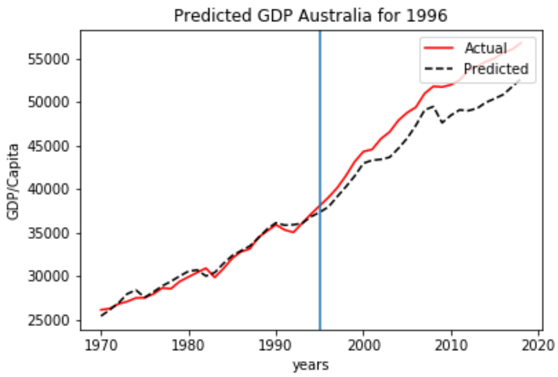
1.6.2 South and South-East Asia

In the SEA region, similar to the SCA region, a mixture of countries is chosen on the spectrum of the *polity* and corruption indices. These are Australia, Philippines, Malaysia, Thailand, Taiwan, Indonesia, and Singapore. Table 1.3 shows the different *polity* levels at the time of their most significant disasters and their corruption perception index for those periods. These values show the disparity within these countries and highlight an important fact. A higher polity value does not correlate with a higher CPI. This can be seen when looking at Singapore and Australia. Singapore has a higher CPI value but a lower *polity* score than Australia. A -2 score indicates an authoritarian regime that is also perceived to be less corrupt than the full democracy of Australia. This will be crucial in understanding how each of these countries will be impacted by disasters.

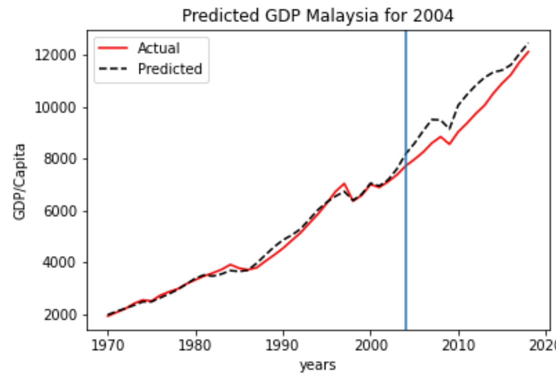
The following graphs are selected results of conducting synthetic regressions on the aforementioned countries. Australia and Singapore also present a contrasting image to the majority of the other countries discussed so far in this chapter. They are the only countries in the SEA region where the synthetic control country under-performs the actual country in all disasters selected. These two countries rank lowest in terms of corruption but vary differently when it comes to *polity*, their degree of democratization. For SEA countries, the majority of disasters are hurricanes or typhoons. The result of the inference tests are available in the appendix, showing the distribution and type of significant disasters.

1.7 Concluding remarks

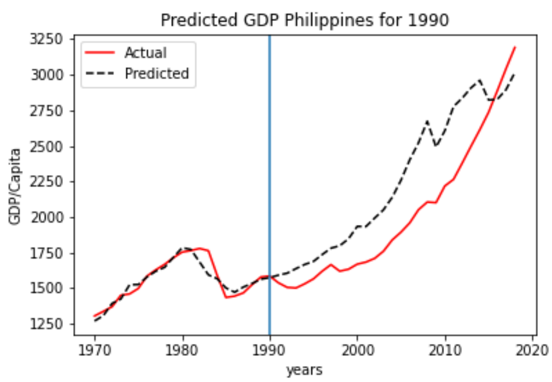
In this chapter, I find that institutional quality plays a significant role in the recovery of GDP per capita after the occurrence of a disaster for countries where GDP per capita is less than $35,000. The two most important measures of institutional quality are regulatory power and level of corruption. These two variables, combined contribute to around 20% to the recovery from disasters. The scope of the analysis is limited to 5 years for most of the countries chosen. Therefore, a claim cannot be made about the optimal post-disaster recovery path with confidence, especially since there is no criteria discussed about the optimal path for growth in this analysis.



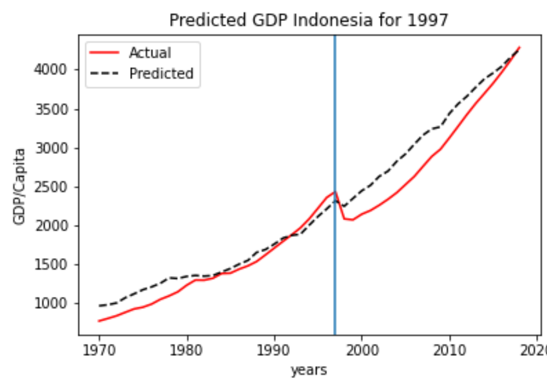
(a) Australia 1996, p-value = 0.04



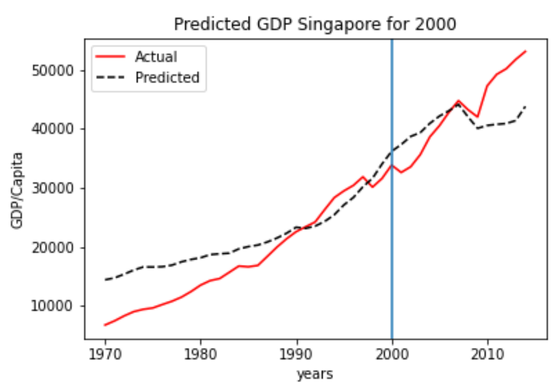
(b) Malaysia 2004, p-value = 0.08



(c) Philippines 1990, p-value = 0.03



(d) Indonesia 1997, p-value = 0.07



(e) Singapore 2000, p-value = 0.33

Figure 1.4: Synthetic Control regressions (SEA)

While the importance of political institutions prior to the occurrence of the disaster is highlighted in this research, there is no claim made about the cost required to establish these institutions and the impact on GDP prior to the disaster. Further research could be directed towards answering this question.

This chapter attempts to tackle an issue raised by various studies conducted on the long-term impact of disasters. Since there is no clear consensus on the long-run effect of disasters on GDP, this chapter contributes to the literature by studying the impact of disasters on a large panel of countries using the Synthetic Control Method and quantifying the significance of various factors on this impact. The findings in this chapter suggest that politicization levels are insignificant in affecting GDP during the recovery phase. There is no evidence that democracies perform better than autocracies recovering from a disaster. Even when considering institutional quality variables, this research finds that only regulatory power seems to be the most significant factor. Regulatory power and corruption are vital components in achieving a positive recovery from disasters and even outgrowing its pre-disaster outcome. In practical terms, countries in the top 30% of the distribution of regulatory authority will recover from a disaster in such a way that they will overtake the path of their pre disaster GDP.

On average, and over five years after their occurrence, disasters will have a more significant negative impact than positive impact, with an absolute difference of around 3% between the average positive and negative values. While the negative impact can be pretty significant, there is a smaller limit regarding how much more a country can recover after a disaster. The results from the Synthetic Control Method regressions on richer countries show that disasters do not affect long-run GDP. For countries where GDP per capita is greater than 35,000$ U.S. per year, the effect of the disaster is negligible one year after its occurrence. Finally, I extended the methodology used in performing regressions from the Synthetic Control Method. Grouping regions according to their outcome variable leads to an efficient selection of the vector V, the vector of relative importance. This results in faster and more accurate regressions for research that includes many events and regions.

Future work can extend the model by looking at the consecutive impact of disasters or adding different types of disasters. In terms of the former, ignoring the accumulated effect of disasters over time and their interaction with institutional quality seems unwise. Repeated disasters could lead to a significant overhaul in institutional quality. The framework presented by this chapter only studies the impact of a particular disaster until another occurs. In practical terms, this means that the impact of some disasters can only be studied for less than five years. Adding multiple steps to the synthetic control regression will provide a deeper insight into the true impact of these disasters.

Appendix A

Table 1.9 summarizes the list of countries and dates selected for the final regressions.

Country	Date	Country	Date
Bolivia	1992	Indonesia	2004
Bolivia	1997	Honduras	1974
Brazil	1984	Honduras	1993
Colombia	1985	Mexico	1985
Colombia	1999	Mexico	1995
Colombia	2011	Mexico	2005
Costa Rica	1988	Malaysia	2004
Costa Rica	1996	Malaysia	2007
Chile	1985	Panama	1988
Chile	2010	Panama	2013
Dominican Republic	1979	Paraguay	1983
Dominican Republic	1998	Paraguay	1998
Ecuador	1982	Paraguay	2007
Ecuador	1987	Philippines	1990
Ecuador	1993	Philippines	1995
Indonesia	1997	Philippines	2013
Thailand	2011	Thailand	2013
Thailand	1984	Singapore	2000
Singapore	2003	Uruguay	1999
Singapore	2007	Uruguay	2002

Table 1.9: List of countries and dates selected where a disaster was significant

Appendix B

Tables 1.10 and 1.11 are summaries of the regressions conducted on the data that was split between positive and negative values using the five-year averages

VARIABLES	Reg 1	Reg 2	Reg 3	Reg 4	Reg 5
corruption perception index	1.314*	1.221	1.314*	1.289	1.264*
	(0.665)	(0.542)	(0.684)	(0.712)	(0.613)
polity	0.458	0.551	0.528	0.671	0.591
	(0.853)	(0.546)	(0.485)	(0.543)	(0.613)
government authority	-6.241	-7.112	-6.895	-7.1261	-8.1961
	(4.223)	(4.387)	(4.335)	(5.425)	(6.314)
regulatory power	12.891***	13.251**	13.449**	13.124**	13.667**
	(4.321)	(6.531)	(6.351)	(5.951)	(6.220)
violence		4.025	4.985	4.835	4.231
		(4.221)	(4.877)	(4.996)	(4.212)
labor force			1.315	1.312	1.381
			(1.332)	(1.322)	(1.455)
land size				0.0021*	0.0019*
				(0.0011)	(0.0011)
interest rate					0.455***
					(0.032)
R-squared	0.182	0.195	0.201	0.205	0.253
R-squared Adj.	0.180	0.173	0.169	0.168	0.155
N	26	26	26	26	26
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 1.10: OLS results for five-year averages for negative events

VARIABLES	R 1	R 2	R 3	R 4	R 5
corruption perception index	1.271	1.277	1.753*	1.741	1.887
	(0.852)	(0.899)	(0.923)	(1.132)	(1.453)
polity	0.422	0.452	0.531	0.557	0.581
	(1.214)	(1.227)	(1.315)	(1.087)	(1.023)
government authority	-4.321	-5.221	-5.432	-5.257	-7.885
	(3.891)	(4.047)	(3.987)	(3.996)	(7.243)
regulatory power	13.921***	13.857***	13.449***	12.985***	12.775***
	(2.483)	(2.335)	(2.875)	(2.645)	(2.578)
violence		2.520	2.557	2.894	3.024
		(3.8511)	(3.781)	(3.857)	(4.002)
labor force			1.224	1.386	1.251
			(1.459)	(1.231)	(1.557)
land size				0.0020**	0.0021***
				(0.0012)	(0.0009)
interest rate					0.583***
					(0.0342)
R-squared	0.210	0.235	0.241	0.252	0.267
R-squared Adj.	0.208	0.204	0.203	0.198	0.205
N	19	19	19	19	19
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 1.11: OLS results for five-year averages for positive events

Tables 1.12 shows the results of the regression with the added QoG value.

	R 1	R 2	R 3	R 4	R 5
corruption perception index	1.632	1.512*	1.415*	1.421*	1.271*
	(0.774)	(0.712)	(0.7740)	(0.8165)	(0.8165)
polity	0.2281	0.2240	0.2281	0.2273	0.2191
	(0.3147)	(0.3320)	(0.4157)	(0.3523)	(0.3451)
government authority	-7.8913	-7.9961	-7.8521	-7.1261	-8.1961
	(5.4229)	(5.7210)	(5.3216)	(5.4253)	(6.3147)
regulatory power	13.0713**	13.3504**	13.2310**	13.4154**	13.2152**
	(6.2258)	(6.5680)	(6.4312)	(6.3156)	(6.1283)
violence	3.7847	4.0835	4.52261*	4.0835	4.1422*
	(2.8122)	(2.9668)	(2.0431)	(2.512)	(2.1318)
labor force			1.2041	1.4512	1.041
			(1.4511)	(1.410)	(1.325)
land size				0.002351*	0.002281*
				(0.001161)	(0.001025)
interest rate					0.5281***
					(0.01147)
R-squared	0.2305	0.2220	0.2013	0.2512	0.4211
R-squared Adj.	0.1343	0.1248	0.1281	0.1247	0.2255
N	19	19	19	19	19
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 1.12: OLS results for five-year averages for all events with added variables

Appendix C

The Synthetic Control Method works by minimizing a loss function defined as $f(V)$ using quadratic optimization with equality constraints defined as follows

$$\frac{1}{2}V'PV + q'V$$

subject to

$$GV \leqslant h$$

$$AV = b$$

where $P \in R_{n\times n}$ is a symmetric matrix. D is defined as the diagonal of the V vector. We then have

$$P = X_0' D X_0.$$

and

$$q = X_1' D X_0.$$

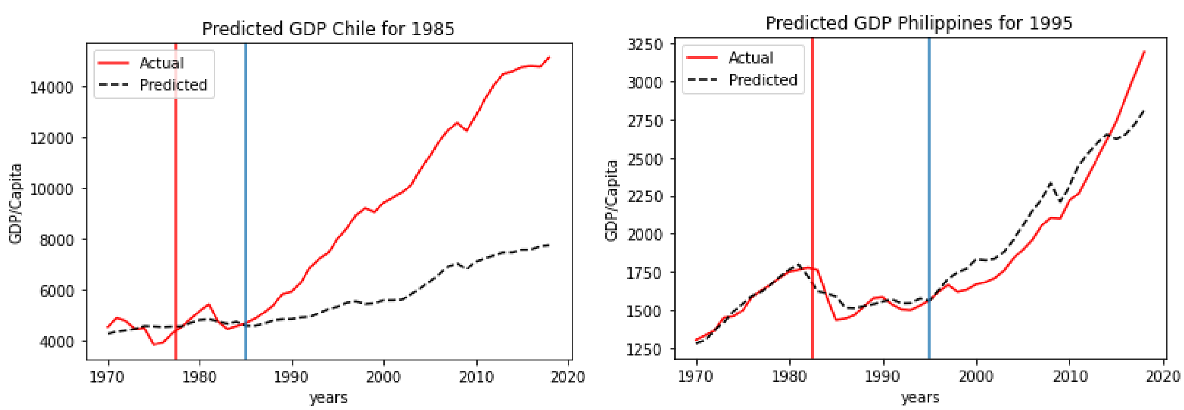
with l being the column size of Z_0. In this case $h = 1$ and so G is bounded by 0 and 1. A is a matrix of ones with a size $1 \times l$. The Lagrange function of this problem is

$$L(V, u) = \frac{1}{2} V' P V + q' V + v'(GV - h) + \mu'(AV - b).$$

In order to minimize this loss function an initial guess of V is assumed called V_{guess}. The closer the guess is to the actual optimizing value of V the faster and more accurate the convergence is. This function is minimized using a Sequential Least Squares Programming method. This iterative process requires that both the objective function and constraints are twice continuously differentiable. Fu, Liu, and Guo (2019) provide more details on the SLSQP method and its application.

Appendix D

The results from selected different training periods



(a) 2/3 training period (b) 1/3 training period

Chapter 2

A critical review of the literature on COVID-19 anti-contagion policies

Abstract

This chapter presents an extensive overview of the literature surrounding the pandemic and the different models used in analyzing the impact of anti-contagion policies on reducing death rates and their impact on the economy. It also highlights the essential results of these policies and shows that the overall impact on the economy was not as large as anticipated in terms of reducing GDP. Understanding the impact of individual policies on reducing the death rates from COVID-19 is raised by various researchers, and is highlighted throughout this chapter. This chapter sets the stage for the rest of the thesis and highlights the importance in using the Synthetic Control Method to tackle such a topic.

2.1 Introduction

To better understand how anti-contagion policies developed in the pandemic and their effects, it is crucial to provide a critical review of the literature surrounding COVID-19. This paper discusses the different models used to predict the impact of these policies and the research that has discussed their effects. I highlight some key areas where research is still lacking. More specifically, I focus on discussing research surrounding individual policies. I show that despite the vast wealth of literature, there are still gaps in the knowledge around the impact of individual policies. Properly preparing for future waves of any future pandemic requires a thorough review of the data and results of further research. Showcasing gaps in the literature highlights the contributions of the rest of this thesis to the scientific literature.

Early in the pandemic, the literature surrounding COVID exploded and covered many topics. Initially, the focus was on predicting the potential impact of the pandemic and what are some possible responses to the virus. Eventually, the research shifted towards assessing the impact of these responses. Given the significant uncertainty surrounding this event, its impact on people, its ability to spread, and concerns over potential mutations, researchers have attempted to tackle all the different facets of the contagion. Over two years since the start of the pandemic and concerns about the impact of variants of the COVID-19 virus are taking center stage, more specifically with regards to containment measures and their effectiveness.

Vaccine mandates, mask mandates, social distancing, and restrictions on gatherings are just some of the policies used throughout the pandemic by policy-makers, which are still in effect today. While the public consensus initially was that containment measures were needed and essential in “flattening the curve” of fatalities, over the last year, there has been severe push back against a lot of these anti-contagion measures, citing lack of effectiveness as shown by Montreal Gazette (2021) and Public Broadcast Service (2021). Economists have attempted to tackle the pandemic’s economic and human impact using various models and quantitative methods. With much research being done in a short time, it is vital to take a step back and understand what has been found so far and what is still missing. This chapter aims to show and discuss the other research that has been done about the pandemic, current gaps in the literature, why they are there, and how these gaps could be tackled.

Initial attention was focused on the SIR, or SIRD (Susceptible, Infected, Removed, Dead) model of infectious diseases. The model was developed by Kermack and McKendrick (1927) to mathematically characterize and track the spread of infectious diseases and their impact on a population. In terms of modeling the spread of the COVID-19 virus using this model, two seminal papers by Atkeson (2020) and Stock (2020) introduced the basics of the SIR model to economists and explored potential applications of this model. This was then followed by research from Atkeson, Kopecky, and Zha (2020), Acemoglu et al. (2021), Fernández-Villaverde and Jones (2020), among many. The focus was on the spread of the virus and the impact that containment and social distancing measures had on its propagation, but not necessarily tackling the economic impact of the pandemic. Abel and Panageas (2021) improve the model by adding waning immunity and vaccinations.

In contrast, Pindyck (2020) augments how the model reacts to deaths and shows how reducing the reproduction number (or infection rate) reduces the number of deaths. This addition helps quantify the effects of lockdowns on the death rates. Eichenbaum, Rebelo, and Trabandt (2021) extend this by incorporating household decisions during the pandemic to account for self-selection into anti-contagion policies and quantify the impact on consumption. Finally, Acemoglu et al. (2021) expands the model to account for targeted policies that focus on particular groups of individuals (young versus elderly) and find that these policies are more effective than general lockdown procedures. These models, however, do not track all the policies implemented. So much of the research has consistently included only a few policies without capturing the effect of the full scope of anti-contagion policies.

Much empirical work has been done on analyzing the performance of these anti-contagion policies on “flattening the curve” as well. Researchers such as Deb et al., 2020, George et al., 2020, or Alvarez, Argente, and Lippi, 2020 use traditional econometric models or policy models such as difference-in-difference, OLS, or maximum-likelihood to study these policy effects. This research was often not country-specific and provided a basis to assess the performance of different political entities worldwide. Because of this, it was a non-trivial exercise to isolate specific policies, and so the bulk of the work done either focused on using an aggregate stringency index developed as a way of measuring the strictness of aggregated policies across countries or research was specific to understanding the impact of specific isolated policies. Often times the issues that limited methodologies such as difference-in-difference in tackling the impact of anti-contagion policies were the same for OLS. In this chapter, I will highlight some of the main issues hindering methodologies such as difference-in-difference and Vector Auto Regression, and these issues will also be the reason why OLS is also not suitable for such analysis.

Several gaps exist in the current body of work surrounding the impact of these policies. The vast majority of the literature concludes that containment measures effectively reduce the spread of diseases. The degree of effectiveness has varied between research. The first gap is the lack of research surrounding the impact of each policy on its own in lowering death rates. This is due to the difficulty in isolating policies, especially because of variations across and within countries regarding policy implementation. Most countries imposed their anti-contagion policies around the same time, but individual policies varied wildly (Hale et al., 2021a). Researchers have then had to rely on comparing aggregate responses, namely the stringency index. On the other hand, studying the removal of policies proved to be a more manageable task. Researchers such as Cho (2020) could quantify the impact of the removal of lockdown measures simply because it was done asymmetrically. In their research, they study the impact of removing lockdowns in Sweden relative to other European countries.

This chapter aims to provide a detailed and somewhat updated review of the current literature surrounding the pandemic. As the virus devastated healthcare systems all across the globe and with hospitals reaching total capacity at alarming rates, much research was done on the impact that policies have had in curbing infection rates and death rates. Crucially, how policies were implemented, and the public pressure put on policymakers resulted in a myriad of research about the efficacy of these policies, with sometimes contradicting results. A more considerable concern relevant to economists was the trade-off between lowering death rates and affecting the economy. Economists such as Acemoglu et al. (2021) and Atkeson, Kopecky, and Zha (2020) show that this trade-off does exist, but that for the U.S., it is minimal. Others, such as Davis, Liu, and Sheng

(2021) show that the economic cost of market activity was not as severe as expected in countries that implemented a complete lockdown, such as South Korea.

The paper is structured as follows. In section 2.2 a review of all the models that have been used in researching the link between anti-contagion policies and lowering death rates from the virus is presented. Each model's strengths and weaknesses are discussed, and an explanation for the gap in the literature that fails to address the impact of specific anti-contagion policies is proposed. In subsection 2.3.1 I showcase several major papers that researched the direct impact on macroeconomic variables, such as interest rates, economic activity, and government spending. The goal of this subsection is to show how widespread this impact of the pandemic was. In subsection 2.3.2, I summarize the literature on the impact of the pandemic on the labor market. With work-from-home mandates and wage compensation policies put in place for substantial amounts of time, many authors aimed to quantify the impact of these policies. One of the larger bodies of work around the impact of the virus revolves around its impact on financial markets and assets. I explore this in subsection 2.3.3. I end the chapter in section 2.4 with concluding thoughts and provide a brief motivation about the work in the rest of the chapters in this thesis.

Before proceeding, an important note about the scope of the chapter is needed. As mentioned, the literature around the pandemic is vast, and in this chapter, the focus is mainly on understanding the impact of policies on reducing death rates. In 2021 alone, more than 350 working papers were presented to NBER with the COVID-19 virus as the central focus. Tackling the full range of literature around the pandemic would be a monumental task. I choose to focus on anti-contagion policies to show that despite the extensiveness of this research on this topic, gaps still exist. In this chapter, there is no attempt to address the efficacy of these policies. In other words, there is no argument being presented about the efficacy of the policies but rather the importance of studying the impact on the loss of life, as the impact on GDP was not as bad as initially assumed or predicted.

2.2 Containment measures research

At the onset of the pandemic, researchers scrambled to understand the impact of the pandemic on GDP and the impact of anti-contagion measures on the economy as a whole. Given the different containment measures imposed early in the pandemic, there was serious concern about an unparalleled depression of major economies, in the vein of the Great Depression of 1929. The earliest work to tackle this was from Eichenbaum, Rebelo, and Trabandt (2021) who use the standard SIR framework, which will be expanded on further in this chapter, to study how household consumption decisions would be affected during the pandemic as a response to lockdown procedures. They find that by implementing the best simple containment policy in the U.S., lockdowns increase the severity of the recession, with the trade-off reducing the death toll by around 500,000 for the year 2021. The focus on the impact of anti-contagion policies is important for economists since the research highlights a certain trade-off between death rates and GDP when imposing anti-contagion policies, at least in the short run (Acemoglu et al., 2021) (Casares, Gomme, and Khan, 2022). Narrowing down the impact of individual policies on death rates provides the necessary data to conduct this trade-off analysis for individual policies rather than general lockdown or stringency policies.

It would not be harsh to say that many policymakers implemented anti-contagion measures without considering their effects. With a lack of data, tools, and conflicting recommendations, government responses to the pandemic varied widely. To that extent, George et al. (2020) provide a guide in measuring the response of political entities to this pandemic, which includes the impact these policies had on the environment and the total deaths caused by COVID-19. More specifically, they cite a mix of the Oxford Stringency Index and the OECD policy tracker as critical tools for comparing government responses and their outcomes. It is essential to highlight the stringency index, as it is the foremost measure used by most researchers when conducting work on policies. Despite being a solid indicator of the combined effectiveness of anti-contagion policies, this index is insufficient in comparing policy effectiveness due to how it is constructed, but it does provide a foundation to be able to compare these policies.

This index was developed by Hale et al. (2021a), using a set of 9 different anti-contagion policies to build a composite numerical value. These policies are assigned values individually depending on their stringency (from no policy being implemented to the policy being imposed with fines applicable for not abiding by the policy). These numerical values range from 0 to 4, although not all policies can go up to 4. The final index is built as the average of these values.

$$S = \frac{1}{k}\sum_{j=1}^{k} I_j$$

where I is the indicator value of policy j with k possible policies. For containment measures, $k = 9$ comprehensive policies are considered, but these can be containment, economic, or health policies. Since indicators have different maximum values defined as N_j and some indicators were regionally specific (applied to a capital city or only specific cities), they are assigned flags f_j to represent the impact of these indicators (targeted versus un-targeted). The authors rework the index I_j to normalize it across policies.

$$I_{j,t} = 100\frac{v_{j,t} - 0.5(F_j - f_{j,t})}{N_j}$$

where $v_{j,t}$ is is the recorded policy value at time t, F_j is an indicator if the policy F_j has a flag or not, and $f_{j,t}$ is the actual flag value, with N_j being the maximum value for the policy in question. The myriad of research that has used this indicator finds a strong correlation between higher stringency and lower death rates.

Maximum values N_j for each policy are set according to the policy's different implementation methods. Restrictions on gatherings can go from 0 (no restrictions) to 4 (restrictions on gatherings of 10 people or less), with 3 being restrictions on gatherings between 11-100 people, for example. These values are set through observations in the data and are not country-specific. While policies are normalized, there is no inherent distinction or inclusion of the efficacy of each policy. All scales are ordinal; therefore, the final stringency index is not enough to identify the impact of individual policies on death rates. The generalization of the scale is one of the main reasons why much research has focused on the mix of restrictions rather than individual policies.

When using this index, it is impossible to disentangle specific policies' impact. The same stringency score value could be due to wildly different policy mixes. Controlling for these values in regressions is impossible because of the differences between policies. This is key as to why using the stringency index on its own is insufficient and is a possible explanation of different efficiency values.

The following two subsections provide a relatively thorough examination of the literature that covers both the SIR model and the traditional econometric regressions and provides detailed results from this research.

2.2.1 SIR Model

This subsection describes the SIR model, the potential insights that can be obtained from it, and a summary of the most prominent research results from this model. The SIR model is a basic infectious disease model that is widely used in the literature. What follows is a description of its basic version. A population N at time t is split into four groups. $S(t)$ refers to the group of individuals that are susceptible to the disease (individuals that can be infected), $I(t)$ for those that are infected, $R(t)$ for those that are resistant (or recovered from the disease), and $D(t)$ individuals that die due to the virus. $t = 0$ is the start of a pandemic. At this time, several assumptions are made about the resistant individuals, such as the fact that they can never be infected again, or for example, in the case of COVID that $S(0) \approx 1$ and $D(0) = 0$, but these can be relaxed for robustness purposes. In its most basic form, the total number of people in the economy is then.

$$S(t) + I(t) + R(t) + D(t) = N(t)$$

These segments of the populations evolve over time according to the following equations

$$\begin{aligned}
\partial S(t)/\partial t &= -\beta(t)\frac{S(t)}{1-D(t)}I(t) \\
\partial I(t)/\partial t &= \beta(t)\frac{S(t)}{1-D(t)}I(t) - \gamma I(t) \\
\partial R(t)/\partial t &= (1-\upsilon)\gamma I(t) \\
\partial D(t)/\partial t &= \upsilon\gamma I(t)
\end{aligned}$$

Where β is the infection rate or rate with which disease is spread from sick infected individuals to susceptible individuals, γ is the rate at which agents recover from infections, and υ is the fatality rate. All of these parameters are positive and agents flow only in one direction, from S to D

This model allows the modelling of social distancing by reducing the infection rate β through whatever policy is appropriate. When analyzing this basic model for a single treatment of social distancing, Sadeghi, Greene, and Sontag (2021) find that this mandate enacted early on will significantly reduce the peak of the infection by around 60%. However, this does not tackle the social distancing measure used to maintain social distancing. In other words, social distancing refers to a broad set of rules and their severity, but not the actual rule being implemented, such as a mask or work-from-home mandates. Understanding the impact of different policies is essential to calibrate the parameter β properly. The following subsection shows the empirical research around specific policies.

In their research, Acemoglu et al. (2021) expand this model to include multiple groups of infected. They split the model population into j groups where the total population of each group is N_j from equation 2.2.1, and the total population in the economy is $\sum_{j=1}^{J} N_j$. In their model, they impose two restrictions, a lockdown, and social distancing measures, to understand whether these effects are different between age groups. They also include the possibility of a vaccine and a cure for the virus at some point. Their results focus less on individual policies since they only consider full lockdowns and social distancing measures as anti-contagion policies but more on how these policies impact different groups. They find that semi-targeted policies perform better with or without the vaccine than general policies by reducing the fatality rates by half. While general policies reduce the fatality rate to 1.83% of the total population, targeted policies lower this number to 1.02%, something that is improved if the groups could be separated, down to 0.71%. As with other SIR models, there is still no clear distinction between policies, meaning they do not distinguish which policy was more effective, but rather how the different policies affected different groups of people. Finally, they find that GDP would be lowered by these policies anywhere between 24% and 10%, depending on the severity of the policies and the assumptions made. They acknowledge that this is a first step in the SIR research and that their model could be improved by understanding how anti-contagion measures could be actually implemented. More recent research from Casares, Gomme, and Khan (2022) provides clearer numbers and indicates that the impact of anti-contagion policies on GDP on the long run is minimal, reducing GDP by 1% if socieconomic restrictions are implemented, and up 4% in the short run.

Alvarez, Argente, and Lippi (2020) test the optimal lockdown policy by focusing on containment measures and using a dynamic planner's control problem within a SIR framework. They find that a severe lockdown that covers 60% of the population, which gradually gets reduced to 20% of the population, is ideal for minimizing death rates from the virus, or what is called flattening the curve. Their work provides a basis for modeling the impact of lockdowns but fails to include other containment measures. Using this model, Atkeson, Kopecky, and Zha (2020) find that for COVID-19, the reduction in death rates is mainly attributed to the reduction of the infection rate rather than the reduction in the susceptible population since this population does not decline quickly enough. With the emergence of the new variants of COVID-19, this work is even more relevant in highlighting the importance of policies that appropriately target the infection rate. Similar to the research mentioned earlier, the gap in their work isolates individual policies or sets of policies.

These models provide a solid theoretical background in modeling the impact of the pandemic on fatality rates and the spread of the disease. They offer the flexibility and adaptability to policies such as vaccine or mask mandates (Alvarez, Argente, and Lippi, 2020), teleworking mandates, or multi-group distinctions (Acemoglu et al., 2021), through various parameters and improvements to the basic SIR model. While these parameters can capture some of the impacts of these policies, a bloated model that considers all the different policies (together or separate) will be difficult to calibrate. It is necessary to complement this work with a quantitative approach that provides a way for economists to properly disentangle the effect of these policies relative to the magnitude of their implementation. When looking at the current state of the empirical research, there is no thorough work done on the individual impact of these policies. The majority of the current quantitative work also uses the stringency index as a proxy for overall policy implementation, which also poses another issue, the policy implication of such research. As mentioned before, this index does not differentiate between the relative efficiency of each of these policies implemented. The same stringency score could be achieved through various policies and outcomes. In order to provide policymakers with the right tools, it is essential then to tackle these individual effects. In the following subsection, I present an overview of the empirical research in terms of both methodologies and results until this chapter is written.

In this subsection, literature relevant to this dissertation's scope are highlighted, by focusing on specific improvements to the SIR model that attempt to model current anti-contagion policies. This model can be extended to many different applications, and this evident in the various research of Abuin et al. (2020), Gevertz et al. (2021), Broek-Altenburg and Atherly (2021), and Albi, Pareschi, and Zanella (2021).

2.2.2 Empirical research

March 11th, 2020, was the day that the World Health Organization declared the COVID-19 outbreak a global pandemic Cucinotta and Vanelli (2020). Containment measures were implemented far and wide across countries, and six months later, empirical literature about the impact of these policies followed. While early on, theoretical models such as the SIR provided predictive analysis about the efficacy of lockdowns, researchers quickly tackled government responses to this pandemic. Most empirical literature relied on either a Difference-in-Difference or a structural Vector Auto Regression (VAR) approach. In this subsection, I present this research, its results, and the shortcomings or issues that researchers faced.

Deb et al. (2020) consider the impact of overall containment measures for a panel of international countries by using the stringency index as an explanatory variable. They use a difference-in-difference method to study these policies' impact. They find that quicker implementation of measures resulted in a 10-20% reduction of deaths compared to later adopters of measures by looking at the timing of the increase in the stringency index. Their research, however, does not quantify the impact of individual policies in reducing these death rates.

Another critical problem in their research is an endogeneity issue that arises when conducting this type of empirical work. Whether the dependent variable was a policy indicator or the stringency index, case rates and death rates can be considered significant determinants for these variables. It would be expected for any government entity to impose these restrictions as a response to high death rates. Expecting the endogeneity to have a similar effect across policies is also a bit naive. However, it is a workaround that some papers used to abstract from this issue purely out of convenience and lack of instruments to be used. Another issue with this and other econometric models is a feedback loop between policy implementation and individual behavior. As governments implement various policies, individuals might see this as a signal to self-impose certain restrictions, which might overstate the actual impact of these policies. Another endogenous issue that cannot be tackled by their Difference-in-Difference model is the response of individuals to the pandemic ahead of government restrictions as highlighted by Casares, Gomme, and Khan (2022). The Synthetic Control Method provides an elegant solution to this problem by a method called backdating that will be discussed further in this dissertation.

Data quality is also varied and incomparable for many countries, especially developing nations. While the stringency index might be more consistent across the board, relevant variables for the Difference-in-Difference regression such as testing data, mobility data, death rates, or air-quality data are questionable. Think testing data for various countries. These values tend to be very politically motivated, with some countries either under-reporting or not providing any testing data (Deb et al., 2020). The same can be said for death rates, where transparency in reporting deaths from COVID-19 is not always assured, as some states under-reported the deaths from the virus. The Difference-in-Difference approach might not be suitable in this case. These measurement errors can result in data not being considered for the regression, an underestimation, or overestimating the impact of these policies.

It is difficult then to assess the true impact of the stringency index's impact, given that it could be composed of different policy mixes. This can lead to biased results when it comes to policy implications. The fact is that the Difference-in-Difference methodology does not allow its users much flexibility when tackling individual policies. However, there are robustness exercises that can be done. Several authors attempt to deal with the shortcomings of the Difference-in-Difference methodology of estimating policy effects.

Some studies have tried to work around these by including more variables such as Jamison et al. (2021). The trade-off is that the model was often too complicated, and in the case of Jamison et al. (2021) was only able to be estimated for a period of two months only. This was due to the lack of data for all the variables considered in this model. Given these issues, and at the time of writing this chapter, no paper has tackled the individual effects of anti-contagion policies within or across countries.

To show the gap in the data, I discuss key results from different empirical research. The consensus is clear that anti-contagion policies have an impact on reducing death rates, arguably only in the short term. In a cross country study spanning 13 countries, Jamison et al. (2021) compare the impact of self-imposed behaviour changes (such as no gatherings, or voluntary isolation), to government imposed restrictions on the death rates from COVID-19. They consider deaths from the virus between 16-20 days after the start of the following form

$$\Delta_{i,t} = \alpha + \beta_1 \text{behavior} + \beta_2 \text{Policy} + \gamma X_i + \theta_u + \mu_1 t + \mu_2 t^2 + e_{i,t}$$

where **behavior** and **policy** are dummy variables that indicate the policies mentioned above. They conduct a simple regression and test for aggregated and disaggregated policies; namely, they focus on the stay-at-home policies. They acknowledge that their approach is simplistic and do not complicate their model for this study because of computational needs. Their main finding is that government-imposed restrictions are more effective than self-imposed restrictions. They also found that only three restrictions were significant. They concluded that workplace closing, event restrictions, and size gathering restrictions are the only significant restrictions by the government. This conclusion is indecisive since their sample size and robustness tests are limited. The key finding here is that policies are not equal in their impact. Endogeneity is once again an issue, as well as omitted variable bias.

In one of the more detailed papers concerning anti-contagion policies, Chernozhukov, Kasahara, and Schrimpf (2021) deviate slightly from the standard empirical literature and use a Wright-type causal path model, a structural model of econometrics, to find the impact of 6 different policies on death rates. These policies are imposing masks for employees, closing K-12 schools, stay-at-home orders, closing restaurants except take-out, and closing movie theatres and businesses. They tackle how policy and behavior impact death rates and the interaction between policy and behavior. Similar to most empirical research, self-selection is present in their model and is acknowledged. While they try to abstract from it, the feedback mechanism mentioned earlier is still likely present. They suggest that had mask mandates been implemented in March 2020, deaths from COVID could have probably been lowered by 19 to 47 thousand over the three months from March to May 2020 in the United States. This research only focuses on the implementation of national policies. As with all the research mentioned earlier, there is no ability to properly analyze the asymptotic implementation and removal of policies. This means that they do not assess the impact of individual policies nor do they compare the impact for different states. This major caveat means that the actual final impact of state policies is ignored.

In a more localized setting, Huber and Langen (2020) study the impact of two anti-contagion policies on death rates in Switzerland and Germany. These are mask mandates and restrictions on gatherings. Once again, they use standard OLS models for estimations for Switzerland and Germany. They find that implementing these policies reduces fatalities by 0.6 cases per 100,000 one month after their implementation, and the effect remains steady for two months (60 days) after their implementation.

Given the already established results from empirical work and the highlighted issues of these models, this paper proposes an alternative model that does not suffer from such issues, that is relatively new, but is also computationally more intensive, and that is the Synthetic Control Method.

2.2.3 Synthetic control

A recent paper by one of the authors of the synthetic control method discusses the relative importance of this model in quantifying the impact of all kinds of policies (Abadie, 2021). In this paper, he provides a literature review on various ways this technique has been used and highlights the topics covered, including economic and non-economic subjects. These subjects range from taxation policies (Kleven, Landais, and Saez, 2013) to right-to-carry laws (Donohue, Aneja, and

Weber, 2019), to election results post large disasters (Heersink, Peterson, and Jenkins, 2017), and even in medical research such as (Wedel and Pieters, 2017) and (Bouttell et al., 2018). I have highlighted this method previously in chapter 1 and outlined the mathematical intuition behind the model. In this subsection, and to be concise, I proceed with showcasing the benefits of using this model in studying the effects of policies throughout the pandemic on reducing death rates from COVID-19. This is contrasted with the previously discussed econometric models and how this method can be used to provide calibration to more theoretical models such as the SIR model.

When considering the traditional econometric models used in the literature, an initial advantage of the Synthetic Control Method over these models is that the issue of omitted variable bias is no longer present, at least not in the same way discussed above. The Synthetic Control Method uses co-factors and the variable of interest prior to the occurrence of an event to model the impact of said event. While issues with this approach exist, omitted variables can be seen by having bad pre-treatment fit the treated variables. In other words, if there is not enough data on the treated variable before the treatment, then the model might over or underfit the data prior to this treatment. This can be easily remedied through an augmented synthetic control model proposed by Ben-Michael, Feller, and Rothstein (2021). Their model shows that when there is a lousy fit prior to treatment in a traditional Synthetic Control Method, it is still possible to achieve accurate results by de-biasing the Synthetic Control Method using ridge regressions for the pre-treatment data.

Since COVID infection and death rates were recorded daily for specific countries or states since the start of the pandemic, the previously mentioned problem is only an issue when trying to predict the impact of policies implemented early on in the pandemic since there is little in-sample data available to build counterfactuals. This is a clear advantage of the Synthetic Control Method over its other counterparts because the need for in-sample data is minimal as compared to something like Difference-in-Difference.

Donor selection is also not an issue when using Synthetic Control Method since . Li and Shankar (2020b) show how, when running through different sample permutations, the pre-treatment fit for the treated variable was significantly more accurate than Difference-in-Difference when applied to retail store sales for the same sample. This is also discussed by Abadie (2021) in his literature review of the synthetic control method, where he shows that the Synthetic Control Method produces a more accurate pre-treatment fit than a Difference-in-Difference approach for any particular sample.

In the context of the pandemic, Cho (2020) first applied the synthetic control method to Sweden to quantify the impact of removing lockdown on death rates and infections. A significant contribution to this paper is that Cho, 2020 shows that using a Difference-in-Difference approach produces similar results to the synthetic control method when applied to Sweden. The advantage, in this case, is that the synthetic control is a more suitable and robust tool for undertaking such a task because of the ability to detect omitted variable bias and correct for endogeneity. In his paper, he highlights donor requirements within the synthetic control method and presents a clear benchmark for applying this method to other pandemic scenarios. However, in his paper, Cho (2020) only studies the impact of completely removing the lockdown in Sweden and its impact on the death rates. Similar to much of the literature, this only compares the impact of a selection of anti-contagion policies to none. In fact, in his paper, the donors are selected based on their stringency index level, regardless of what policies they implemented. He also highlights the importance of undertaking work similar to Chernozhukov, Kasahara, and Schrimpf (2021) using the Synthetic Control Method and calls it a "*worthwhile project to pursue*". In terms of results, he finds that early removal of the lockdown increased death rates by 25%. This number is depressed because people will self-select into these policies even without a requirement imposed on them. An issue not discussed in the paper is that the definition of early removal is vague. There is no expectation that strict lockdowns will last forever, which then begs the question, at what point is it optimal to remove the lockdown? Given that the author only looks at one country and at one particular time, without comparing the effect of this removal to other similar scenarios, it is hard to answer this question without looking at the result of the removal of policies in different locations. In the following chapter of the thesis, this topic is tackled by considering more treatment groups, more treated periods, and more variables, specifically vaccination rates. Alfano, Ercolano, and Cicatiello (2021) use this method to study the impact of school opening on death rates in different regions of Italy. They show that for a particular region, 15 days after the opening of schools, there is an increase in death rates by around 20% compared to what it should have been. This effect persists for around 50 days after the opening of the schools. Using the Synthetic Control Method Mills and Rüttenauer (2022) found that COVID-19 certification (negative test, vaccination certification) led to an increase in vaccination rates in a panel of six countries.

The Synthetic Control Method has gained traction when it comes to studying the impact of the pandemic. Alfano, Ercolano, and Cicatiello (2021) use this method to study the impact of school opening on death rates in different regions of Italy. They show that for a particular region, 15 days after the opening of schools there is an increase in death rates by around 20% compared to what it should have been. This effect persists for around 50 days after the opening of the schools. Using the Synthetic Control Method Mills and Rüttenauer (2022)find that COVID-19 certification (negative test, vaccination certification) led to an increase in vaccination rates in a panel of six countries.

The benefits of the Synthetic Control Method help tackle the issues presented in the empirical research. For starters, the requirements for larger samples and large amounts of observed variables are not necessary to conduct an accurate Synthetic Control Method regression, which, compared to traditional regression methods, is a significant advantage. Unlike techniques such as Difference-in-Difference, the Synthetic Control Method requires a relatively more minor pool of controls and can produce consistently better and more accurate results at the expense of computational time ONeill et al. (2016). Another benefit is what Abadie (2021) calls "Transparency of the Fit". The Synthetic Control Method provides detailed information about the co-factors of the synthetic unit, the actual treated unit, the control group chosen to build the synthetic model, and the complete data set used. Finally, the Synthetic Control Method only requires data up until the occurrence of an event for the treated variable. Unlike other regression methodologies such as OLS, ML, Difference-in-Difference, or any time-series models, no observations of the control group are needed past the treated period. In fact, as a robustness exercise, it is common only to take part in the pre-treatment periods to obtain synthetic weights and use the rest of the pre-treatment to verify fit. This again highlights that while large data sets are clearly desired, the Synthetic Control Method can work in situations where this is not the case.

This method is not without any shortcomings. As Abadie (2021) mentions, inference testing is probably the most significant limitation of this method. Compared to traditional parametric regression methods, inference for Synthetic Control Method regressions relies heavily on control group selection. While successful regressions can be conducted with smaller control pools, there is a higher chance of rejection relative to traditional statistical inference. Another issue is the ease of use/replication of the method. As of the date of writing of this paper, very few libraries exist that support Synthetic Control Method, which limits researchers' ability to use this technique.

Another standard limitation in a lot of the empirical research is that the Synthetic Control Method can be biased if there are expectations of policy changes, such as the announcement of the implementation of policies ahead of time. One proposed solution is to set the treatment period to the announcement date, not implementation. In the case of COVID policies, this is not necessary. First, announcements of policies being implemented were usually expected. Second, there was little evidence mentioned in the literature about self-selection in measures such as school closings, work-from-home orders, restrictions on gatherings, or restrictions on public events. Even so, assuming that this expectation exists, compared to traditional methods, the Synthetic Control Method offers an elegant solution to this issue. An easy way of handling this self-selection bias is backdating the treatment periods and assuming that it starts from the date of announcement of a policy, or for even more robustness, that it starts when individuals notice a sharp increase in death rates since that could be interpreted as a signal to engage in social distancing measures. The researcher then conducts the regression with that period as their treatment period. In this way, the Synthetic Control Method circumvents this major pitfall of the previous empirical methods. For this research, I conduct both backdating events to ensure no bias in the synthetic model results.

2.3 Other pandemic effects

In this section, I summarize the literature that has explored other areas impacted by the pandemic. Much research has been done so far, and it is almost impossible to pack all of it comprehensively. For each topic, I highlight the earlier research that was done and complement that with more recent literature in the hopes of showcasing the progress that had been made from the start of the pandemic. What is of relevance to my research are the models and econometric techniques used in these different papers, as some of them use the SIR model. In contrast, others have used specific econometric techniques, including Difference-in-Difference and the Synthetic Control Method. I cover the impact of the pandemic on GDP, firm production, and household consumption across countries, the impact on the labor force and how containment policies affected unemployment, and finally, how financial markets reacted to the pandemic and containment measures.

2.3.1 Impact on GDP, firms, and households

Fiscal stimulus was an essential policy tool to mitigate the pandemic's potential effects. Auerbach et al. (2021) find that the restrictions on economic activity severely impact the economy's supply side, reducing the effectiveness of government spending and the size of the fiscal multiplier. They conclude that fiscal policy should not be a simple reaction to the economy's contraction. However, they should consider the channels through which this downturn affected the economy, in this case, worldwide supply lines. Similarly, Baldwin and Freeman (2021) tackled the solution provided to this issue at the beginning of the pandemic, which was shorter and more domestic supply chains. They find that global supply chains did not simply cause economic shortages. The design of these chains, while partly responsible for the economic downturn, was also why vaccines and medical supplies were deployed rapidly worldwide.

In another paper, Lu et al. (2021) develop an agent-based model for Idaho's potato supply chains. They show that producing versatile input (such as potatoes) can stabilize prices and reduce the impact of supply chain disruptions by about half. Industries with such input did not experience severe disruptions in their production, lessening the impact on GDP. More specifically they talk about food differentiation which can alleviate risks in the food supply chain.

Using a panel Vector Auto Regression model , Ludvigson, Ma, and Ng (2020) study the economy's response to a significant natural disaster shock by calibrating this shock to the March 2020 effect of the COVID-19 virus on the U.S. economy. They find that such a shock leads to a monthly decrease in GDP equal to 12% per month and that in the mildest case, they expect a drop of 20% in production over 12 months. They also find that persistent adverse shocks increase the duration of macro uncertainty by the length of the shock. However, they highlight caveats in their research. The large shocks selected for predictions in their model did not impact production as much as the pandemic did and the policy responses to these were more localized. They Baker et al. (2020) use a Vector Auto Regression model calibrated to large disasters to assess the impact that the uncertainty caused by COVID has had on GDP. They use an empirical model with moments calibrated to the COVID-19 shock. They find that a year-over-year contraction of 11 percent is expected at the peak, two quarters after the pandemic's start. They find that more than half is actually induced by uncertainty. This model considers the different restrictions put in place and finds that these restrictions on their own cannot explain the contraction in GDP. This is, in fact, relevant when discussing policies to be implemented and comparing the trade-off between loss of life and contraction of GDP. Disaster models are well fitted to analyze the impact of the pandemic, and the Synthetic Control Method is a suitable model to calculate such shocks.

Similarly, Jones, Philippon, and Venkateswaran (2020a) study the macroeconomic differences between large economic disasters such as wars or earthquakes and pandemics. More specifically, these disasters tend to be short-term relative to the length of a pandemic which can span several years, necessitating different responses from policymakers. While large disasters destroy capital, pandemics do not, reducing the return on investment as labor becomes scarce. Their analysis suggests that the pandemic would have reduced the natural interest rate under typical situations. However, given the nature of the disease, it is likely that this rate will go up precisely due to aggressive fiscal expansion. They contrast this to traditional natural disasters, more specifically in poorer countries.

2.3.2 Impact on labor force

There is no denying that the pandemic has caused severe disruption to the labor market. Unemployment from the closure of businesses is a significant concern, but another critical topic is the transformation of initial work-from-home mandates from temporary solutions to permanent realities. On top of this, one of the more controversial containment policies that countries worldwide implemented was wage compensation or wage subsidies. The biggest fear was that it would incentivize workers not to work, leading to high levels of unemployment and lower production. While the exact details of these policies differed between states and countries, the main objective behind such policies was to offer immediate compensation to workers who lost jobs due to the pandemic and potentially lower the pandemic's impact on consumer spending. In some cases, there was also an aim to provide a stopgap for workers until work from home was established as an alternative to on-site work. Amongst economists, there is a general agreement that working from home will be a part of the labor market moving forward (Brueckner, Kahn, and Lin, 2021) Eberly, Haskel, and Mizen, 2021.

In their paper, Bai et al. (2021) show that the transition to a more digital economy has made jobs and firms more resilient to shocks, using a difference-in-difference framework. They build a work-from-home index for firms and find that this index increases uniformly across firms after the pandemic, signaling a higher ability to conduct work from home even after the pandemic is over. This result is in line with work from other authors such as Eberly, Haskel, and Mizen (2021) and Brueckner, Kahn, and Lin (2021).

On a different note, Rojas et al. (2020) measure how school closures impacted the job market conditions of states in the United States, more specifically, unemployment. The key takeaway from their paper is that this dramatic increase in unemployment was due to the response to the pandemic by households, not the mitigation strategies themselves. In other words, policies did not result in increases in unemployment. Local anti-contagion policies (specifically school closures) seemed to have little impact on the labor market.

Using register data from Norway, Alstadster et al. (2020) find that the majority, or around 90%, of layoffs were temporary. They find that the rise of unemployment overstates the loss of output by about a third but that the majority of the impacted population is the most vulnerable population, i.e., less-educated workers or women with children were more likely to be laid off during the pandemic.

Substantial work has also been done to understand the impact of governments' responses to the pandemic. The consensus is that these programs did not increase unemployment. The findings of Bartik et al. (2020) or Holzer, Hubbard, and Strain (2021) suggest that there was no correlation between governmental programs and unemployment. Results that were tested by authors such as Finamor and Scott (2021) using a standard regression model.

Dube (2021) uses difference-in-differences to estimate the employment effects of the United States unemployment benefits program offered during the pandemic. He finds similar results to previous work and finds that this policy did not have any substantial and long-term adverse effect on the job market, even after its expiry. Moreover, while Marinescu, Skandalis, and Zhao (2021) found that these welfare payments decreased competition in the labor market between applicants, and there was a decline in job applications being sent. Their results, however, still find that these payments did not decrease employment.

2.3.3 Impact on financial markets

An important component of modern markets are financial assets, and a lot of work has tried to understand the impact that the pandemic had on asset prices such as Davis, Liu, and Sheng (2021), Alfaro et al. (2020), and Ramelli and Wagner (2020).

Davis, Liu, and Sheng (2021) find that asset prices fell between 20 to 50 percent in countries around the world. This drop was followed by a drop in economic activity about three weeks later. This was the trend in 32 out of 35 countries in their panel, with the notable exceptions being South Korea and Taiwan, where in the former, there was only a minor drop in economic activity, and in the latter saw none at all. China was the only country that saw a contemporaneous dip in economic activity and asset prices. The key finding of this paper is that asset prices served as a strong indicator of incoming economic activity decline.

A key finding of the paper of Davis, Liu, and Sheng (2021) is that the most efficient policy responses to the pandemic involved rapid implementation of virus containment efforts but not necessarily strict lock-downs on economic and social activity. This means that mask mandates, social distancing, or work from home requirements had little impact on economic activity, but as per the authors this part of their research was not extensive, and only looked at groupings of lock-down measures, but not individual measures, and in fact, does not account for the varying effects of individual measures.

This work improved previous work done by Alfaro et al. (2020). Alfaro et al. (2020) try to understand the link between infection rates and asset prices throughout the COVID pandemic in the United States and the SARS pandemic in Hong Kong. Using a standard model of Infectious diseases (SIR), they find that unanticipated changes in predicted values of infection rates across the U.S. forecast next-day stock returns. More precisely, they find that if predicted values of infections double, stock returns drop by 4 to 10 percent, and vice versa. Their work mirrors that of Ramelli and Wagner (2020), who found that the most significant indicators for firms' stock prices during the pandemic were debt ratios and cash reserves.

Eichenbaum et al. (2020) find that older consumers reduced spending more than younger consumers when the pandemic hit. A pandemic is a low-probability event with significant consequences (similar to a disaster). They find that their result is consistent with the theory around natural disasters. The theory states that as people age, they become more risk-averse in the face of disasters. This is important in formulating the appropriate containment policies for different age groups in a country.

Another crucial financial market asset is life insurance. These assets play an important role in an individual's lifetime savings and behavior and are tools that are highly affected by containment policies. Since higher mortality risks are generally linked to lackluster containment policies, it is crucial to understand how these policies affect life insurance contract prices. Harris, Yelowitz, and Courtemanche (2021) find that these prices did not fluctuate during or after the pandemic. The mortality rates being relatively low compared to the population and the rapid deployment of containment policies across different states in the U.S. resulted in a relatively small increase in life insurance prices.

Finally, Acharya, Engle III, and Steffen (2021) study the crash of bank stock prices in the early onset of the pandemic. They find that while all bank stocks were heavily impacted during the pandemic because of an increase in lending risk, banks with higher liquidity performed relatively better, similar to the 2007 mortgage crisis. Markets still perceive this as a binding constraint to their activities. The quick fiscal stimulus response served as a dampener on this crash relative to the 2007 crisis.

2.4 Conclusion

Very few events have garnered as much attention from researchers as the COVID-19 pandemic has. Two years later, and economists are still working on identifying potential repercussions of different anti-contagion policies, and the pandemic itself, not only to better prepare for future mutations of the virus, but also any similar events that might arise such as large and natural disasters from climate change.

This chapter asks three questions about anti-contagion policies and the pandemic. First, what can be concluded from the current research, and what is still missing to better understand the effectiveness of different policies on lowering death rates from the COVID-19 virus? The second question asked is how are researchers currently tackling the gaps in the research and what are some potential avenues that can be explored. Finally, how do these policies impact different economic variables, and what is the trade-off between human and economic costs.

These questions are answered to an extent in this chapter. First, the current body of work shows the importance of anti-contagion policies in reducing death rates, and that the economic trade-off of these policies is minimal, especially in the long run. There is still however gaps surrounding the impact of individual policies. This gap is being tackled through the use of different models, and the Synthetic Control Method is a powerful tool to conduct such research. The first chapter of this thesis discusses the Synthetic Control Method's ability to quantify the impact of policies in handling large crises. The rest of the thesis tackles the first two questions posed in this chapter that have not been explored in great detail in the literature around the pandemic. The first is the targeted impact of anti-contagion policies on deaths in the United States. While much research has worked on the combined impact of these anti-contagion policies, or just the economic subsidy policy, I focus on unravelling the individual policies themselves, thereby providing policymakers with more accurate data on the impact of such policies. Secondly, I focus on the cumulative effect of these policies throughout the pandemic, and I propose a novel approach to the Synthetic Control Method. Much of the literature has discussed the differences between large economic disasters (financial or natural) and the pandemic and has used the Difference-in-Difference policy to conduct such analysis. As mentioned by Abadie, Diamond, and Hainmueller (2010a) and Li and Shankar (2020b) the Synthetic Control Method is often a better tool for studying policies than Difference-in-Difference for studying the pandemic since it can be used to analyze more events with a looser restriction on available data. While the Synthetic Control Method is a great tool for policy analysis, it is not always suitable for quantifying the total impact of anti-contagion policies in its basic form. I propose changes to this tool that could be applied to other similar responses.

Chapter 3

Assessing the effectiveness of anti-contagion policies in The United States using a synthetic control method approach

Abstract

The COVID-19 pandemic is the most studied event of the last ten years. With several viruses already spreading since the original variant and potentially more to come, understanding the impact of anti-contagion and social distancing policies is crucial. I present an empirical model that tackles the impact of each policy on curbing death rates from COVID-19. The studied period is over 490 days and covers seven policies implemented in 50 U.S. states. The results show the difference in the effectiveness of each policy. For example, closing schools resulted in a potential reduction of death rates from COVID by 12%. On the other hand, restrictions on gathering on average reduced death rates by up 3.86 times their predicted value, or by 4 cases per 100,000. This was also the most efficient policy having a significant impact almost 80% of the time 30 days following the implementation of the policy. The 20% gap is most likely due to lack of enforcement of such a policy, since backdating was conducted on all events to ensure that self-selection was not a possibility. At the same time, closing schools was shown to be the least effective in reducing death rates. I find that with at least 70% of the population vaccinated, death rates are not significantly affected when removing anti-contagion policies.

3.1 Introduction

The COVID-19 pandemic proved just how unprepared countries were for such an event in terms of the ability of their healthcare system to handle the shock and their policy responses. Actions taken during the pandemic to curb the death rate (or flatten the curve) varied wildly between and within countries, mainly due to the plethora of policy options that policymakers and scientists proposed. These include, but are not limited, to policies such as imposing the use of masks in confined spaces (Rancourt, 2020), lockdowns (Acemoglu et al., 2021), social distancing in public areas, restrictions on gatherings (Deb et al., 2020), or work-from-home orders (CDC, 2021).

Initially, without much data or proper understanding of the impact of the different policies, policymakers imposed and removed different restrictions with various claims about the effectiveness of these measures thrown around. There was literature evidence backing many of these claims, and the variance in the stringency of these policies varied wildly. The literature surrounding COVID-19 anti-contagion policies has since exploded, with many authors trying to understand the impact of anti-contagion policies. Some have measured the effect of stringier contamination policies on death rates, such as Hallas, Hatibie, Koch, et al. (2021a), World Health Organization (2020), and Hsiang et al. (2020). Others have looked at the long-term effect of the pandemic on economic growth (Famiglietti and Leibovici, 2021), (Acemoglu et al., 2021). From the start of the pandemic in February 2020, different policies have been implemented worldwide with varying outcomes and not always with the desired impact (Deb et al., 2020). There are many difficulties in separating the impact of individual policies, such as the random and often overlapping implementation of anti-contagion policies, lack of data, or model restrictions. Chapter 2 discusses these in more detail.

This chapter contributes to the research by identifying the effectiveness of anti-COVID measures separately in reducing death rates. This research aims to provide more insightful recommendations for policymakers for future waves of COVID-19 or similar events by understanding the impact of individual anti-contagion policies. Removing policies is also a contested topic, and this chapter identifies the effects of removing policies on death rates. Anti-contagion policies were always meant to be temporary. However, given how quickly the virus can spread and mutate, it was essential to identify what conditions were necessary to ensure the safe removal of these policies and not increase death rates. This chapter is an ex-post performance review of handling the pandemic for different U.S. states.

The overlapping implementation of these policies means it is difficult to disentangle the singular effect of each policy. Using the traditional Synthetic Control Method (Synthetic Control Method) proposed by (Abadie, Diamond, and Hainmueller, 2010a), I study different anti-contagion policies, such as movement restrictions, curfews, and workplace closing, as well as one economic policy, namely, a wage subsidy. Chapter 2 of this thesis showed precisely why this methodology is more suited to analyze the effect of these policies than traditional policy analysis methods. Another benefit of understanding the impact of each policy on reducing death rates is that it provides a springboard for economists to tackle the trade-off of implementing a policy on death rates versus economic impact. A policymaker is incentivized to apply the policy that reduces the death tolls the most while having a negligible impact on the market.

The results show that all policies were significant in reducing death rates from the virus, but they did not have the same level of impact. In other words, they did not all reduce death rates the same way and for the same duration. The least efficient was closing schools, reducing death rates on average by 0.1 per 100,000 during the same period. The results indicate that imposing these policies reduced deaths by 60,000 to 200,000 from March 2019 to June 2020. On the other hand, removing restrictions on a population that is not well vaccinated leads to an increase in death tolls. When the population is sufficiently vaccinated, removing restrictions does not impact the death rates. Sufficiently, in this case, refers to vaccine rates above 75%, where removing restrictions with such conditions had little to no impact on death rates. This is discussed in the results section of this chapter.

The rest of the paper is structured as follows. Section 3.2 explores the most up-to-date literature surrounding the pandemic and the Synthetic Control Method. Section 3.3 details the data used and the methods of selection for the data. Section 3.4 develops the model used in this chapter and sections 3.5 and 3.6 describe the results and the conclusion of the paper.

3.2 Literature

Various authors have voiced fears of variants of COVID-19 or other viruses from the same family replicating the pandemic's initial impact (Wong et al., 2020), (Wise, 2021). Understanding how the current response to the pandemic has affected infections and death rates is vital in ensuring quick recovery in the future and reducing the impact that this virus has had on human lives and the economy. This trade-off was highlighted in Chapter 2, and it can even be extended to other viral events that may not necessarily be of the same magnitude but are more localized.

As the virus spread in U.S. states, anti-contagion policy implementation was varied between these states. However, it almost always included gathering restrictions, closure of public transport, work from home orders, restrictions on movement, curfews, and international travel restrictions. The stringency of these policy responses has also fluctuated across states since the start of the pandemic as measured by the Oxford COVID-19 Government Response Tracker indicators. The average of the stringency index across states spiked early in the U.S. with a low variance to this index, indicating strict anti-contagion policy implementation for all states. There was a precipitous drop in the average stringency index April 2020, and it steadily went down over that period. Over the same time period, the variance of the index remained high and steady, with only a tiny drop in April 2020, and remained at consistently high levels after June 2020. This means that while U.S. states had initially decided to lower restrictions, a sizeable portion of them decided to increase their restrictions throughout the pandemic to curb death rates further. This also indicates that there was very high variance between states in terms of strictness of these policies. The peak amount of policies implemented was April 2020, with an average of 5.8 policies implemented by the 50 U.S. states. Since then, fewer and fewer states have had active closures and containment policies. Figures 3.1 and 3.2 show the mean and variance of this index. By July 2021, only 31 states had more than one restriction actively enforced.

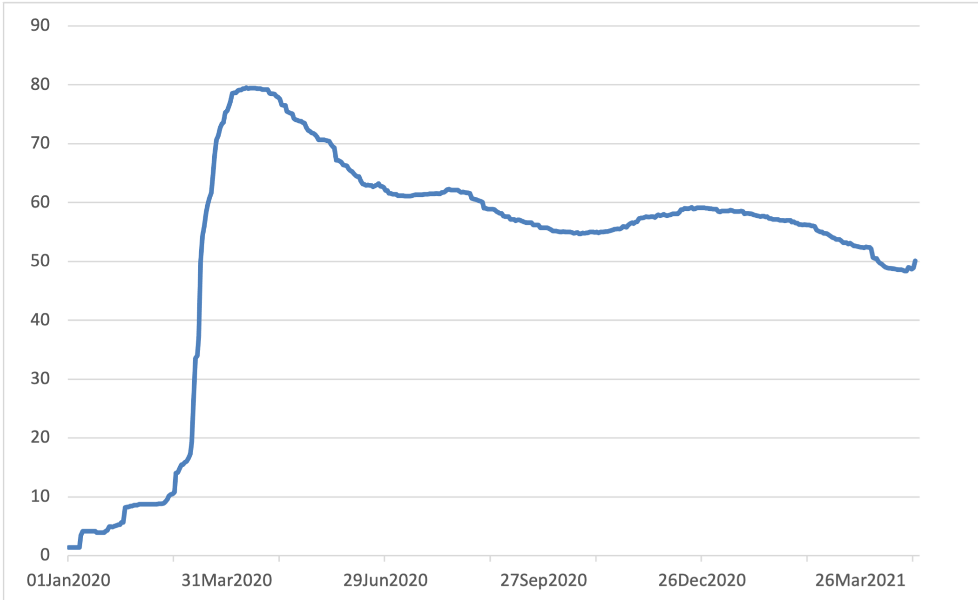


Figure 3.1: Average stringency index across U.S. states
1

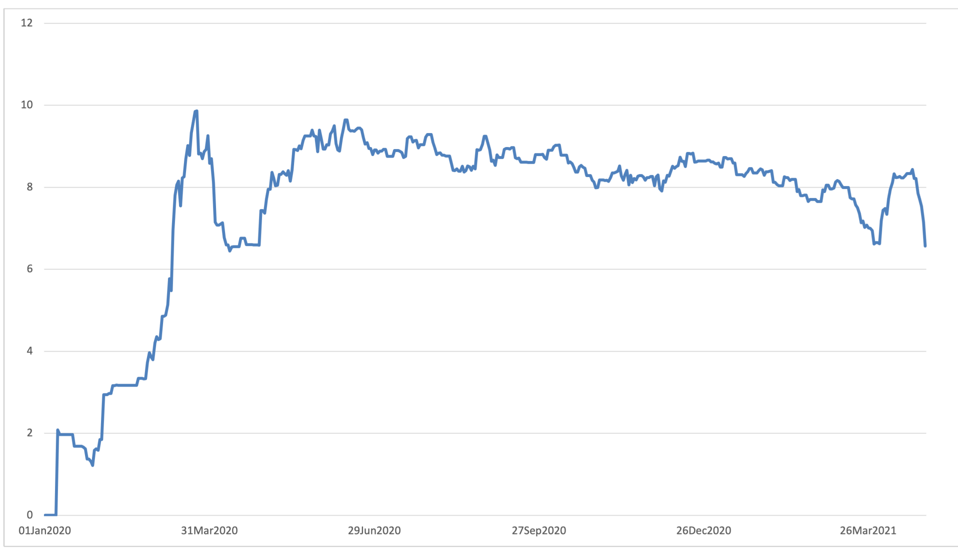


Figure 3.2: Standard deviation of stringency index across U.S. states
2

this chapter also provides a performance review of policymakers in different U.S. states. A descriptive analysis of U.S. response to the pandemic Hallas, Hatibie, Koch, et al. (2021a) found that for a panel of states, the regional and political variation in stringency was significant throughout the pandemic and even widened within states, with Democrat-led states having the most stringent policy responses and Republican-led states having the least. Targeted geographic policies continued to exist even when states lifted state-wide policies. In Republican-led states, policies were more stringent in cities that leaned more towards the Democrats than in rural areas, while the same was not valid for Democratic-led states. In their work, however, Hallas, Hatibie, Koch, et al. (2021a) fail to account for the relative importance of the population density of these cities in the decision-making process of removing restrictions. Political considerations often overshadowed the realities of the data in other countries. Abbasi (2020) criticizes the United Kingdom's response as inefficient and unstructured in terms of the implemented policies, with politics and optics overshadowing proper policy making. He mentions that the reality of deaths in hospitals and care homes looked entirely disconnected from government COVID-19 briefings.

In a seminal paper concerning the impact of the pandemic, Hsiang et al. (2020) use a simple susceptible-infected-recovered (SIR) disease model to evaluate the effect of anti-contagion policies on the growth rate of infections and death rates. They find that these policy actions significantly slowed the growth rates of infections across a panel of six countries, China, South Korea, Italy, Iran, France, and the United States, and prevented or delayed around 519 million cases of COVID-19. They do not tackle death rates, a shortcoming in their paper. Chinazzi et al. (2020) find that an international travel ban resulted in a 77% reduction in the potential spread of new cases from China to neighboring countries.

Jones, Philippon, and Venkateswaran (2020b) extend a basic representative-agent framework to understand the optimal containment policy in a pandemic. They propose extending the neoclassical model to include contagion dynamics such as social distancing and working from home. They find that relative to the incentives of private agents, a planner wishes to significantly front-load mitigation strategies and that the prospect of mitigation, together with the possibility of agents working from home, gives quantitatively meaningful reductions in the spread of disease.

The consensus in the research is that these anti-contagion policies were critical in reducing either case rates or deaths. However, little has been done to identify each policy's impact on its own. In their paper, Famiglietti and Leibovici (2021) investigate the impact of containment policies on economic activity using a VAR model. They find that these measures successfully curbed the spread of the virus, with only a transitory impact on economic activity, while also highlighting the importance of economic policies as a complement to health policies. In their model, they use exports of every state as a measure of economic activity. Finally, using the OxCGRT stringency index for a panel of 32 countries and using an updated version of the model proposed by Hsiang et al. (2020), Dergiades, Milas, and Panagiotidis (2020) found that anti-contagion policies were effective in the full list of countries studied and that the higher stringency index values equated to bigger drops in both case rates and death rates. Specifically, in certain countries, confirmed cases and deaths were reduced by more than 90 percent relative to the underlying country-specific path in the absence of anti-contagion measures.

In order to study these individual effects, I use the Synthetic Control Method as outlined initially by Abadie, Diamond, and Hainmueller (2010a). This chapter adds to this method by using the improvements outlined in chapter 1, loosening the restrictions on the weights of the control regions in a two-step synthetic control method proposed by Li and Shankar (2020b), and by using different inference methods outlined by Abadie (2021), to ensure the robustness of the results. These improvements will be discussed in section 3.4. There are several arguments for using the Synthetic Control Method to conduct this analysis. As mentioned in chapter 2 and earlier subsections, the different timings of implementation and removal of policies are suitable for this model as it allows for the study of different policies in isolation. In their paper comparing different econometric tools for policy evaluation, Athey and Imbens (2017) referred to the Synthetic Control Method as "[...]*arguably the most important innovation in the policy evaluation literature in the last 15 years*".

Other authors, such as McClelland and Gault (2017) also test changes to the original Synthetic Control Method method used by Abadie, Diamond, and Hainmueller (2010a) and argue that the method is robust to a slew of alterations and modifications. For this chapter, an advantage of the method is the length of the significance of any particular event. McClelland and Gault (2017) find that the while the standard Synthetic Control Method model loses some predictive power for more extensive time series, or in other words, the accuracy of the predicted value is high for smaller series, depending on the quality of the data used in predictions, up to $T < 100$. Due to the limitation of the data set used in this chapter, only the first sixty days after the occurrence of an event are considered in the analysis and results, meaning the results from any significant events would be valid over the entire period of consideration. Finally, to avoid the issue of multiple treated units (regions experiencing the same event simultaneously), the improvements proposed by Kreif et al. (2016) are considered. More specifically, optimal weights are relaxed to be negative or more than one possibly. There is no significant loss of accuracy when these restrictions are relaxed.

There has been a wealth of literature surrounding the pandemic, tackling different topics and policies as mentioned earlier in chapter 2. To my knowledge, there is no research that uses the Synthetic Control Method to study the impact on mortality rates in the U.S. of different containment policies to such an extent.

3.3 Data Sources

Using the proper co-factors is essential in conducting accurate Synthetic Control Method regressions to ensure that the model predictions are significant. These covariates are crucial in finding the proper weights for the synthetic model. They are identified from the literature mentioned in section 3.2 surrounding the COVID-19 pandemic. The treated variable in this chapter is death rates per 100,000; however, as case rates impact death rates, any research that tackles both of these is considered when selecting covariates.

Deb et al. (2020) find that containment measures were more effective in reducing case rates in countries with lower temperatures, lower population density, and where the share of the population aged 65 or older was larger. These variables are selected as co-factors. They use several control variables for their regression, such as air quality and mobility score, which are added to the list. Health system quality as measured by Harris, Yelowitz, and Courtemanche (2021) is also added to the list, as it is found to be significant in impacting death rates. The complete list of covariates is as follows:

- **Unemployment:** unemployment data was obtained from the FRED database. The data is monthly for all 50 states. This co-variate underlines possible transmission channels coupled with on-site work, as lower unemployment with no work-from-home restrictions could lead to more transmissions (Chernozhukov, Kasahara, and Schrimpf, 2021).

- **Density score:** a variation on traditional density measures. Instead of using state-wide density, I build a composite that I call the density score. For every state, I find the three largest cities in that particular state in terms of population, and I build a weighted average of their density based on their share of the total population. The data is calculated on a monthly basis. Since certain metropolitan areas can stretch across states, the population of the city considered is restricted to the size of the city, and not the metropolitan area it represents to ensure consistency for the calculation of this index.

- **Vaccination rate:** vaccination rate per 100,000 person per state, obtained from the CDC database. I make no distinction between the different vaccines used. Since vaccination only started halfway through the data set I am working with, I dropped this co-factor for any simulations done before the beginning of vaccinations.

- **Old age population share:** the share of people aged 65 and above, obtained from the U.S. census bureau. This population is more vulnerable to the disease, and therefore a state with an older population will experience more deaths (Mueller, McNamara, and Sinclair, 2020).

- **Testing:** testing data, which includes number of confirmed cases, was obtained from the Johns Hopkins Centers for Civic Impact[3](Johns Hopkins University, 2021), and it includes data from March 2020 to July 2021.

- **Healthcare quality index:** a composite index, obtained from the Agency for Healthcare Research and Quality as part of the department of the U.S. Health and Human services. It provides a healthcare quality index per U.S. state, based on insurance costs, access to healthcare, quality of healthcare, and surveys from healthcare professionals. I consider the complete index, and the index for access to healthcare separately. The data covers all 50 U.S. states and is available for 2020.

- **Air quality:** I obtain this data from the Air Quality Open Data Platform, which includes

[3]https://civicimpact.jhu.edu/

daily data for the most populous city per state. Deb et al. (2020) show that better air quality allows for more outdoor events, leading to lower virus transmission rates.

- **Mobility trends:** data mobility trends obtained from both Google and Apple (Apple, 2021) and (Google, 2021) for the two most populous cities per state including walking and driving within the cities.

- **Exports:** aggregated per month, I obtain monthly data from the U.S. census bureau in U.S. dollars.

- **Stringency index:** to control for the impact of potential policies already implemented for a given region, the 14-day average of the stringency index is considered.

The key variable in this chapter is deaths caused by COVID-19, and daily data from the Johns Hopkins Centers for Civic Impact(Johns Hopkins University, 2021) is used. It includes data from March 2020 to July 2021. The data is a 7-day moving average of the number of deaths per 100,000.

3.3.1 COVID Policies

The stratification of the implementation of anti-contagion policies by different U.S. states means a significant pool of controls and placebo variables is available to conduct statistically significant regressions. Enough states exist for any event chosen from which to build proper counterfactuals. The data obtained from OxCGRT details several policies, their implementation date, and their severity. A benefit of this model is that there is no inherent bias if the policy was announced before its implementation, especially since the average duration between an announcement and the decision being made is less than a week (Deb et al., 2020). The inclusion of mobility co-variates controls for the importance of mobility restrictions in reducing death rates.

The data for the policies implemented in each state was obtained from OxCGRT's COVID policy tracker (Hale et al., 2021b). It includes daily data about the policy implemented and its respective levels of intensity for all U.S. states. There are three main policy categories: containment, economic, and health. this chapter uses all the containment policies and one economic policy in the analysis, namely economic stimulus. This is because health policies did not fluctuate enough during the pandemic to be included in the research (Hale et al., 2021b). The economic stimulus has been added because it was the most significant policy implemented throughout the pandemic compared to the rest of the economic policies. The data is available from March 1st, 2020, until July 31st, 2021. Depending on the policy indicator, the values range between 0 and 4, where 0 means no restrictions are in place, 1 is a recommendation, and 4 is the most severe level of stringency for that policy. Descriptions of these indicators are provided below.

- *School closing (sc, 1-3)*: records the closing of schools and universities, where the max value indicates the closing of all levels of education.
- *Workplace closing (wc, 1-3)*: records the closing of workplaces, where the max value records the closing of everything except essential services, such as hospitals.
- *Cancelling public events (cp, 1-2)*: records the cancelling of public events. The max value records cancelling all events.
- *Restrictions on gathering (rg,1-4)*: records putting limits on gatherings, where the highest value records restricting gatherings on less than 10 people.
- *Closing public transportation (pt,1-2)*: records closing public transportation where the highest value prohibits any public transportation use.
- *Stay at home requirements (sh, 1-3)*: records staying at home requirements where the highest value requires to only leaving home for emergencies.
- *Internal movement restrictions (mr, 1-2)*: records restriction on movement between cities and region, where the highest value records no movement in and out of state.
- *International travel controls (ir,1-4)*: records international restriction where the maximum value records total international border closure.
- *Income support (es,1-2)*: records government providing cash support to those who have lost jobs due to the pandemic, where the maximum value records the government replacing 50% or more of worker's salaries.

3.4 Model

This model was discussed in previous chapters, however a brief refresher is provided in this section. In a standard Synthetic Control Method setting, we start with a set of regions I where $I = \{1, \ldots, N\}$ are the fifty states of the United States and $N = 50$, with no particular order given to the states. One of these states is exposed to the "treatment", such as the implementation or removal of an anti-contagion policy, where $N-1$ states are not treated and a state $i = tr$ is the state treated. For the sake of simplicity, for any given simulation, the treated stated is $i = 1$. The outcome variable in the model is deaths per 100,000. I assume $y_{i,t}$ is the outcome variable of state i at time t. The outcome variable is observed over T periods. At a point $t = T_0 < T$, the treatment occurs, but only for the affected region $i = tr$, leaving $T - T0$ of treated periods moving forward, meaning the treatment is uninterrupted. It is assumed that

$$y_{i,t} = \hat{y}_{i,t} + \alpha_{i,t} D_{i,t}$$

which can be rewritten as:

$$y_{1,t} = \sum_{i=2}^{N} w_i y_{i,t}$$

where $t \leqslant T$. To solve this model, I build a set of positive weights w_i where $i = 2, \ldots, I$ such that $\sum\limits_{i=2}^{N} w_i = 1$. There are ideal weights w_i^* such that

$$\sum_{i=2}^{N} w_i^* y_{i,t} = y_{tr,t} \;\; \forall t \in \; T \quad \text{and} \quad \sum_{i=2}^{N} w_i^* z_i = z_{tr} \;\; \forall t \in T$$

and $z_{i,t}$ is a vector of covariates chosen to find the appropriate w^* We can then use

$$\hat{\alpha}_{i,t} = y_{i,t} - \sum_{i=2}^{N} w_i^* y_{i,t}$$

as a way to estimate $\alpha_{i,t}$ where $t \in \{T_0 + 1, \ldots, T\}$. This will be the deviation of the synthetic predicted model from the actual data.

Seeing as restrictions typically lasted for weeks or even months, it would be expected for many states to have a particular restriction already in effect when another state adopts theirs. This means that for every event or treatment, the control group included states already being treated. An essential benefit of the Synthetic Control Method is that it allows finding a weighted average of control units, which leaves room for a larger pool of possible comparable units. To avoid the issue with multiple treated units, the weights are chosen in such a way that the synthetic alternative closely tracks the performance of the treated variable before the treatment, and no restrictions are imposed on the values of the weights. By allowing weighted averages of other states to act as matches, which Difference-in-Difference (DiD) does not, Synthetic Control Method expands the pool of possible comparators, which is valuable when the number of untreated states is limited.

Another requirement for a properly fitted Synthetic Control Method is that there cannot be spillovers in terms of the treated variable between the treated region and regions in the donor pool. In the case of the pandemic, this is not necessarily an issue since, due to lockdown policies, there was little movement between U.S. states during the pandemic, at least after the initial implementation of the policies, resulting in little to no spillover in terms of the treated variable (deaths per 100,000).

The weights are obtained in such as a way as to minimize a penalty function defined as follows:

$$\underset{W^*}{argmin}||\boldsymbol{z}_1 - W\boldsymbol{z}_0|| = \sqrt{(\boldsymbol{z}_1 - \boldsymbol{z}_0 W)' V (\boldsymbol{z}_1 - \boldsymbol{z}_0 W)}$$

where $\boldsymbol{z}_1$ is a vector of pre-treatment variables and co-variates relevant for the treated region and $\boldsymbol{z}_0$ is the same vector of variables for the non-treated regions, and V is a positive semi-definite matrix. The vector $\boldsymbol{z}$ is defined as

$$\boldsymbol{z}_i = [x_{i,1}, \dots, x_{i,m}, y_{i,0}, \dots, y_{i,n}]'$$

where i refers to the region, $x_{i,m}$ is a covariate variable for the region i. In our case the $m = 9$ and $y_{i,n}$ is a set of pre-treatment outcome variables where $n \leqslant T_0$. n could be chosen in such a way that the vector $\boldsymbol{z}$ can include all outcome variables up until the treatment period, however, for this model, $n = \frac{T_0}{2}$ is rounded down, as a way of training the model. This means that only half of the outcome variables are used prior to the treatment to find the optimal weights, and the other half is used as a validation for the accuracy of my model.

In section 3.3 the density score was mentioned as an alternative to state density as it represents a more accurate co-variate for this analysis. This score is a weighted average of the total density of a state, especially since cities are more likely to be impacted by anti-contagion policies compared to rural areas (Hallas, Hatibie, Koch, et al., 2021a). First, the three largest cities in a state in terms of population are found and sorted. Only the city is considered but not the metropolitan area, as some metropolitan areas extend over several states. The respective population and size are subtracted from that of the total population and size of the state. A new density is obtained for each state i where the remaining population is divided by the remaining land size. Assuming cities are ranked from 1 to 3 (1 being the most populous), the density of state i is then

$$\text{dens}_i = \frac{Pop_i - pop_{1,i} - pop_{2,i} - pop_{3,i}}{Size_i - size_{1,i} - size_{2,i} - size_{3,i}}$$

where $pop_{n,i}$ is the population of city n in state i and $size_{n,i}$ is size in km^2 of city n in state i. The weights of the cities are found as the ratio of their population to total population of the state or

$$w_{n,i} = \frac{pop_{n,i}}{Pop_i}$$

The density score is calculated as

$$ds_i = w_{1,i}d_{1,i} + w_{2,i}d_{2,i} + w_{3,i}d_{3,i} + \left(1 - \sum_{n=1}^{3} w_{n,i}\right) \text{dens}_i$$

where $d_{n,i}$ is the respective density of the largest three cities and dens_i is the density of the state excluding these largest cities. This density score provides a more accurate representation of the density of a state when conducting the Synthetic Control Method. Data availability limits the density score to be a monthly variable. A potential problem that may occur is population movement once the pandemic had started, which could change the value of the score. While indeed there was movement between cities, trends in mobility were not heavily affected in major cities as noted by Deb et al. (2020). In fact, when looking at mobility trends from both Apple and Google, there is no significant dip or increase after the declaration of the national pandemic in March of 2020, indicating little change in city densities.

The choice to use 3 cities is used mainly for consistency of the score. In terms of co-factor, using the density score versus state density value results in an average reduction of pre-treatment Mean Square Predicted Errors by 1.8%. There is no loss of accuracy around the use of the density score, and therefore this score is preferred over total density. The existence of better co-factors is a possibility to be explored in subsequent research.

Once the predicted values $\hat{y}_{i,t}$ are obtained, inference analysis is conducted by subtracting the predicted value from the actual value, thereby calculating the alphas where

$$\alpha_{i,t} = y_{i,t} - \hat{y}_{i,t}$$

Since the treated variable is death rate per 100,000, the value $\alpha_{i,t}$ is then the deviation in death rates. If an event is significant, and its implementation (or removal) has an impact on the outcome variable, the alphas are expected to reflect this result. The assumption is that imposing restrictions will reduce death rates to be significant, and removing them will increase death rates over the actual rate. A significant synthetic region where a restriction was imposed is then expected to have a higher synthetic value than the actual region, meaning the alphas are expected to be negative, since $y_i < y_i^n$. The inverse can be applied to removing restrictions, meaning the alphas are expected to be positive. This will be the terminology used then in referring to events, *negative* for imposing restrictions, and *positive* for removing restrictions throughout the rest of the paper and chapter 4.

The method proposed by Cavallo et al. (2013) is used to conduct inference testing:

1. The placebo effect is computed for every event by simulating the outcome variables for the available controls for the corresponding event date.

2. At every point in time following the occurrence of an event (called leads), all the placebo alphas are calculated and then averaged.

3. The actual lead is ranked in the distribution of placebo averages.

4. The lead specific p-value is given by the following formula.

$$p\text{-value} = \frac{\sum_{np=1}^{N_{pl}} I(\bar{\alpha}_l^{pl(np)} < \bar{\alpha}_i)}{N_{pl}}$$

   where $\bar{\alpha}_i$ is the effect of the restriction on the state in question and $\bar{\alpha_i^p}$ is the average placebo effect. This is only the case if α_i is expected to be positive (imposing restriction). If α_i is expected to be negative then the alternative p-value formulation is considered, more specifically.

$$p\text{-value} = \frac{\sum_{np=1}^{N_{pl}} I(\bar{\alpha}_l^{pl(np)} > \bar{\alpha}_i)}{N_{pl}}$$

Two modifications to the treatment of the alphas that affect the analysis of the inference tests are considered. First, because of how the virus interacts with the human body and its impact on people affected, only periods ten days after treatment are considered for the inference test. The virus does not kill immediately upon infection and takes at least a week before the worst effect, i.e., death occurs. Transmission is not always immediate, and adherence to restrictions is not always immediate. As mentioned earlier, the incubation duration of the virus is around 14 days, after which the most severe effects of the virus are usually mitigated (World Health Organization, 2020).

According to the World Health Organization, while COVID-19 can last up to three months in the human body, most people who develop symptoms improve without treatment in 2-6 weeks. A person can transmit the virus 48 hours before developing symptoms (World Health Organization, 2020). Table 3.1 provides a summary by the CDC of the three stages of a typical COVID-19 infection. The CDC also states that moderate illness appears ten days after symptoms appear and that severe illness can develop 20 days after symptoms appear (CDC, 2021). Therefore, when considering the impact of the policies, it does not make sense to look simply at the immediate impact on death rates but rather a specific period after. The data itself also backs this change.

When looking at events where restrictions were imposed in Table 3.3b, it is clear that prior to the 10-day mark, there are minor changes to the outcome variable, with a sizeable jump happening around the 10-day mark. As mentioned in chapter 2, the effects of imposing policies tend to peak and slow down after the 50th day from the implementation date.

Symptoms	**Time of occurrence**
No Symptoms	Up to 10 days from the moment of infection
Mild illness	From 10 to 20 days after symptoms show up. Death is a possibility during this stage
Severe illness/death	20 days after symptoms, highest possibility of death occurs during this period

Table 3.1: Effects of Covid 19
4

Different ranges are considered for the inference tests, mainly for robustness. Due to the nature of the data, and since only events that were uninterrupted by any other event for at least 50 days were considered, any range up to that point can be used for the tests to ensure comparable results between simulations. The following ranges are considered when calculating the p-value, where T_0 is the date of implementation or removal of a policy:

* From $T_0 + 10$ to: $T_0 + 25$, $T_0 + 40$ and $T_0 + 50$

* From $T_0 + 15$ to: $T_0 + 30$, $T_0 + 40$ and $T_0 + 50$

A policy is considered significant if the p-value is sufficiently low in four out of the six periods mentioned above. Around 94% of the total events considered were significant for 4 out of the 6 ranges specified.

The second addition to the inference tests is the use of the percentage deviation as well as the simple deviation in absolute value for the p-value calculations. Rather than the simplified version of α_i mentioned above, an additional value of α is considered:

$$\alpha_i^p = \frac{y_{i,t} - y_{i,t}^n}{y_{i,t}}$$

The same process for the inference tests conducted above is used for α^p. The result of these two procedures is two p-values for every event. An event is considered significant if both p-values are below 0.05. All but one event are significant in both scenarios as well.

3.4.1 Event Selection process

The selection process for events is crucial for the integrity and accuracy of the model, and it is summarized in this subsection. The event period spans from March 1st, 2020, to June 31st, 2021, a total of 487 days. The selection process for events is as follows. For every state i, a daily data set of policies in place at any given time is constructed. To do this, for every policy p, I find the full list of dates where state i either exited or entered policy p. I consider a policy "in-place" if the policy flag is two or above. If the flag is 1 or 0, then the policy to not active. The main reason I do this is that a flag of 1 refers to a recommendation for implementation or a *suggestion*, and 0 means the policy is not in place. Since there is no way to see a policy's enforceability level, I use two or above to proxy an effective policy. For 50 states, I find that out of 24,350 possible date/state combinations (487 days multiplied by 50 states), 20,571 of these had at least one policy being actively implemented.

For every state, I find all the dates where the absolute change in policy numbers is 1 (either 1 policy was implemented or removed). I then select all these dates as potential treatment periods. I then filter out these potential periods according to the following requirements:

- In the first day of a change in policy number, I find what policy was changed, implemented or removed.

- If a policy was implemented, I find the number of states where this policy was not implemented at the time, and vice versa. I add these states to my list of controls. If the list is more than 20, I proceed through the remaining steps. Otherwise I ignore this event.

- For every event selected, I check whether there was another change in policy 50 days after that particular event. If there is not, I proceed to the next step, otherwise I ignore this event.

- With the control states selected, I filter out the policies implemented in the previous 14 days, and build the covariate data of current policies in place. This serves as a control for the impact of other events.

After an event is selected, I sort it into a list of implemented or removed anti-contagion policies. The total amount of events selected is 256.

The model loses significance the more extended treatment lasts. Policies that are significant ten days after the start of an event will remain significant 40 days after. There is a slight drop-off from the 40th to the 50th day after the policy. After the 50th day, and for the policies where such data was measured, the drop in significance is noticeable, with around 30% of events being insignificant 50 days after treatment.

3.4.2 Improvements to model

Two improvements are considered to the traditional Synthetic Control Method. The first is the selection of the V^* vector shown in chapter 1. The same method proposed in that chapter is used, where regions are grouped according to the treated variable in different clusters. A V vector is obtained for every region as an initial guess for each simulation conducted. The V in each cluster that minimizes the Mean Square Predicted Errors is used as a guess for any synthetic regression, including those conducted for the inference. The main benefit of this change is that it improves accuracy by reducing pre-treatment Mean Square Predicted Errors between predicted and actual data at the cost of higher simulation time.

The second improvement is the removal of the restriction placed on the weights. As proposed by Doudchenko and Imbens (2016), the removal of these restrictions implies the possibility of negative weights for each region. This allows for greater flexibility in selecting the donor group and reduces the need for large control groups for the regression. These changes were among several proposed by Doudchenko and Imbens (2016). Other changes include adding an intercept to the model; however, testing this led to no change in the accuracy of the Synthetic Control Method regressions.

The restriction that $w_i \geqslant 0$ for $i \in \{2, \ldots, I\}$ is no longer binding, while still maintaining the restriction $\sum_{i=2}^{I} w_i = 1$. In order to test whether this change is valid, I use the method proposed by

Li and Shankar (2020b) to compare traditional Synthetic Control Method to the amended version of the model. In total, 9,000 simulations or synthetic regressions were conducted with a dataset that includes 4.5 million distinct data points used in this chapter, including the regressions done for inference.

3.5 Results from the baseline model

Two types of events are considered in this model, imposing restrictions and removing restrictions. In total, 256 events were analyzed. Of these, 88 were for imposing restrictions, and 168 were for removing restrictions. The Synthetic Control Method is run on these events, and significance is set at a p-value of 0.05 for each one of these events.

This results in 42 (47.7%) significant *negative (implementing)* events and 62 (36.9%) *positive (removing)* events . While there are more significant events where restrictions were removed, in absolute terms, they are less significant as a percentage of the total events in their respective sample for two reasons. The first is the vaccination rate when removing restrictions. A fair assumption is that policymakers would remove restrictions when a population is significantly vaccinated, meaning infections are less likely to lead to deaths. The second reason that might explain this difference is that when specific sanitary conditions are implemented, removing restrictions will lead to less virus transmission. These conditions include mask mandates, social distancing rules, or proper sanitization of surfaces. The impact of the former reason is tested in an OLS regression in subsection 3.5.2. Tables 3.2 and 3.3 show the distribution of these events by policy.

Policy	**Number selected**	**Significant**	**Percentage**
Closing public transport	6	3	50
Restrictions on gathering	4	2	50
Movement restrictions	13	10	76.9
Stay at home requirements	11	6	54.5
Cancelling public events	10	5	50
Economic Subsidy	24	13	54.1
School Closing	5	1	20
Workplace closing	15	6	40

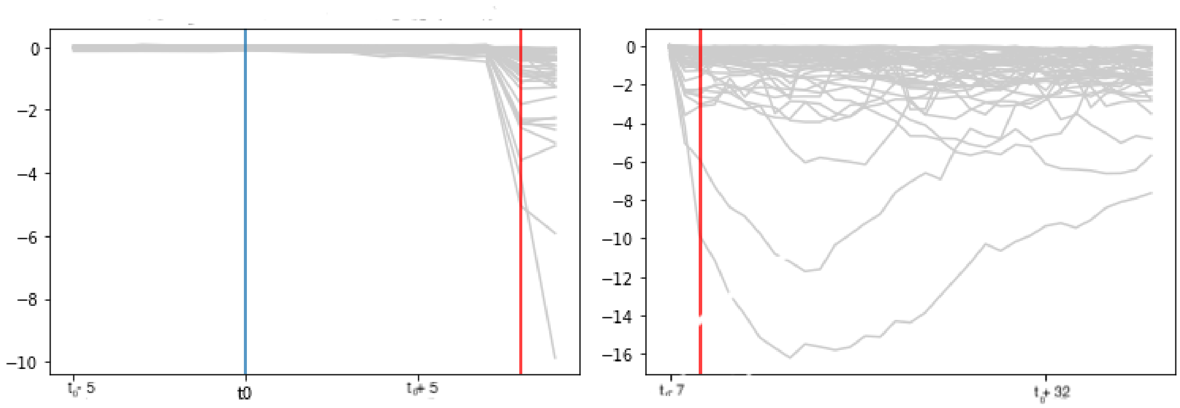
Table 3.2: Number of significant policies when Imposing restrictions

Policy	Number selected	Significant	Percentage
Closing public transport	13	0	0
Stay at home requirements	27	12	44.4
Workplace closing	30	13	43.3
Movement restrictions	15	5	33.3
School Closing	22	10	45.4
Restrictions on gathering	9	4	44.4
Cancelling public events	20	4	20
Economic Subsidy	32	14	43.8

Table 3.3: Number of significant policies when removing restrictions

Out of 13 events where restrictions on public transport were removed, there was not a single statistically significant event that led to increased death rates. Several reasons could lead to this. First, in most cities, public transportation was heavily sanitized, and people just avoided using it as a whole, even after restrictions were removed. Second, vaccine rates were highest in cities where public transportation systems are most active. Finally, work-from-home requirements meant less congested public transportation methods.

When imposing restrictions, there is a sharp drop off of the impact of the restriction from its peak, with an average decrease of around 45% 50 days after the implementation of the anti-contagion policy. Figures 3.3a and 3.3b show the absolute deviation from actual death rates (deaths per 100,000) when imposing restrictions. When implementing restrictions, we see in figure 3.3a that five days before and up to 10 days after the event, there is little change in death rates, and the actual change starts occurring around ten days after the implementation of a policy. This is important as inference is conducting by looking at deviations 10 days after the implementation of a policy, and not immediately after. Figure 3.3b shows a sort of inverted U-shaped curve is observed for such policies, with a drop-off around the 40-day mark after the implementation of the treatment. The red lines on the graphs are for the period of $T_0 + 10$.



(a) Deviation pre $T_0 + 10$ (b) Deviation post $T_0 + 10$

Figure 3.3: Percentage deviation of death rates for all significant negative policies

To quantify the impact of these events, the average deviation is calculated for every policy across all treatments for every day after the implementation of a treatment. Tables 3.4 and 3.5 show the predicted percentage change of implementing (Table 3.4) or removing (Table 3.5) a restriction. The values represent the average percentage change of death rates within a specific time frame. Four of these time frames are chosen that would be significant given how the virus affects humans. The first thing that can be noticed is that the absolute average effect of removing a restriction is significantly lower than the absolute average effect of imposing a restriction. This is particularly important because it highlights several points made earlier in this chapter about the conditions for removing a policy. More specifically, when removing restrictions, policymakers anticipated a certain level of immunization or public safety standard such that the death rate would not be heavily altered. Individuals and businesses adapted to public sanitization policies, which could also help explain this absolute difference. This is especially clear when considering relatively more enforceable policies such as workplace closing or restrictions on public transportation, where there is more accountability from proper sanitization to wearing masks and adhering to public safety recommendations like social distancing.

Policy	T_{0+10} to T_{0+30}	T_{0+10} toT_{0+50}	T_{0+20} to T_{0+40}	T_{0+20} to T_{0+50}
gathering restrictions	-381.1	-333.8	-360.8	-320.5
closing public transport	-62.2	-72.0	-73.7	-79.4
movement restrictions	-115.8	-123.9	-120.7	-124.9
stay-at-home	-144.6	-124.2	-134.0	-128.8
cancelling public events	-58.5	-77.6	-101.2	-81.3
economic subsidy	-166.0	-245.3	-198.7	-275.9
closing schools	-12.7	-14.8	-12.1	-14.4
workplace closing	-63.1	-59.0	-60.2	-57.4

Table 3.4: Average percentage deviation from actual death rate when implementing a restriction

Policy	T_{0+10} to T_{0+30}	T_{0+10} toT_{0+50}	T_{0+20} to T_{0+40}	T_{0+20} to T_{0+50}
gathering restrictions	47.1	41.1	46.3	42.9
closing public transport	-	-	-	-
movement restrictions	38.7	41.7	44.8	44.3
stay-at-home	35.0	34.7	35.8	35.1
cancelling public events	33.9	35.9	37.5	38.7
economic subsidy	30.7	30.3	32.0	31.1
closing schools	12.0	14.8	15.7	16.2
workplace closing	8.2	7.8	7.5	7.4

Table 3.5: Average percentage deviation from actual death rate when removing a restriction

Imposing restrictions on public transportation also had one of the lowest effects on reducing death rates, indicating the effectiveness of cleanliness and sanitization procedures. It is interesting to note that school closing was the least significant restriction regarding actual event count and effect. Finally, economic subsidies were also very effective, and their effect only grew stronger the longer they were implemented. This could be explained by individuals adhering to other safety measures, as they felt their income was not jeopardized by staying home. In fact, for most policies, the longer the time frame, the lower the effect the policy had. The only other exception to this was public transportation closures.

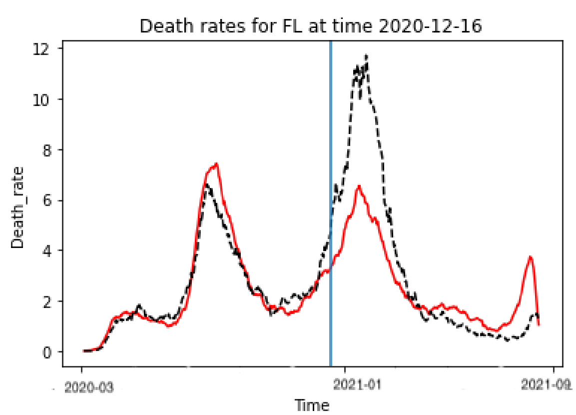
3.5.1 Estimation of impact on the death toll

To check the reduction in deaths from these policies, the significant events where an anti-contagion policy was implemented are selected per state. The 30, 40, and 50-day averages for events are selected. Since the outcome variable is death rates per 100,000, each policy's impact was matched to the population of the state where it was implemented.

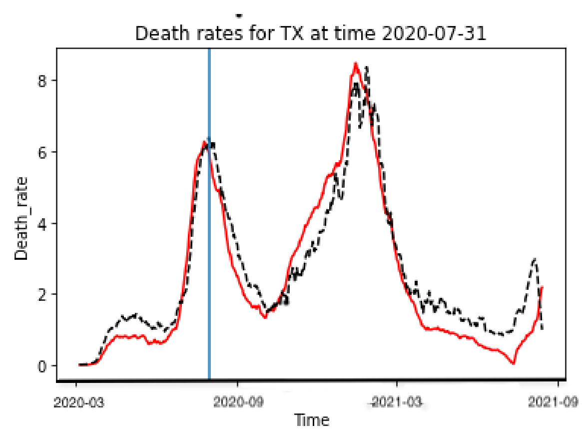
$$dr_{p,i,T} = \sum_{i=1}^{I} \sum_{t=T_0+10}^{T} \frac{\alpha_{p,i,t} * pop_{i,t}}{100,000}$$

Where p is a given policy out of the significant policies. For each policy p the sum of deviations in values (αs) across states i are calculated and multiplied by the state's population in year t, where these policies were significant for at least 50 days. The results are summarized in table 3.6. The last column represents the average death reduced per policy from $T_0 + 10$ to $T_0 + 50$ (or 50 days after the policy is implemented).

The results show the scale and importance of these events in reducing death rates. It is important to note that while restrictions on gatherings were not necessarily the most impactful policy in terms of percentage change in death rates, it was by far the most impactful in terms of absolute change. I attribute this to states that imposed restrictions on gatherings during peak death rates of COVID as highlighted by Figures 3.4a and 3.4b.



(a) Gathering restriction implementation



(b) Gathering restriction implementation

Figure 3.4: Examples of states imposing restrictions

Unfortunately, the sample of restrictions on gatherings and school closings was relatively small due to the selection process and the requirement for an event to occur in isolation (without any event alongside it). A significant analysis of the impact of implementing such measures cannot be made with confidence.

Policy	$T_{0+10} to T_{0+30}$	$T_{0+10} to T_{0+40}$	$T_{0+10} to T_{0+50}$	Average
public transport restrictions	$1,138$	$2,430$	$3,644$	$1,215$
stay-at-home	$8,930$	$18,396$	$31,910$	$5,318$
workplace closing	$5,743$	$10,367$	$15,979$	$2,663$
movement restrictions	$14,091$	$32,847$	$50,859$	$5,085$
school closing	120	497	621	621
gathering restrictions	$9,335$	$27,040$	$48,871$	$24,435$
cancelling public events	$2,973$	$7,527$	$8,819$	$1,763$
economic subsidy	$18,149$	$33,652$	$46,857$	$3,604$
Total	$60,481$	$132,758$	$207,563$	$5,189$

Table 3.6: Total estimated deaths avoided

Table 3.6 shows each policy's total estimated deaths reduced. The most inefficient policy was school closing regarding reduced deaths and significant events. With 621 deaths reduced on average, it is 83 times less effective than the most effective policy regarding total death reduction. The two most effective measures for reducing deaths were movement restrictions (mr) and restrictions on gatherings (rg). The model estimates that they reduced deaths by around 50,000 over 50 days of implementation. However, restrictions on gatherings were five times more effective at reducing death rates, while movement restrictions were the second most effective. Economic subsidies, stay-at-home requirements, and workplace closing follow in terms of effectiveness. I suspect that this effect difference is due to the enforceability of the actual restrictions. Restrictions on gatherings and closing public transportation would be the easiest to implement.

3.5.2 Vaccinations and restrictions

Using the Synthetic Control Method model results, specifically the removal of policies, this chapter explains why many events were insignificant in changing the death rates. As mentioned in sections 3.22 and 3.3, the number of significant events is expected to be lower when removing restrictions. Since a possible reason behind this is that removing a restriction could result from sufficient immunization, in this subsection, this hypothesis is tested using a simple OLS model. When policymakers feel that restrictions have run their course, there is no reason to keep these restrictions, and removing them should not result in any change to the death rates, or at least any significant one.

Out of the 62 events where removing restrictions was significant and ended up increasing the death rates, there was no vaccination done at the time of removing the restrictions in 51 of these events, or 82.2% of total events. The highest vaccination percentage out of the remaining events was 45% (Florida, April 13th, 2021), and the average was 8%. Out of 106 events that were not significant, 85 had a positive vaccination rate, with the average being 50% and the maximum being 120%.

Several models are considered, with two different data sets of events. The first set is the full list of 186 events where restrictions were removed, while the second only includes events where restrictions were removed, but the vaccination rate was positive (> 0). There is no restriction set on only considering significant events. The models are defined as follows

$$Y_i = \beta X_i + \gamma Z_i + u_i$$

Y_i is the mean deviation between the actual and counterfactual. Different mean values are tested, as shown in table 3.4, but no significant difference is found between these means in terms of the results from the regression. The mean deviation between counterfactual and actual data between T_{0+10} to T_{0+50} is the one that is presented as it has been the one used the most in this chapter so far. The results of the OLS regressions for the other three means are shown in the appendix. X_i is a vector of explanatory variables that include the ratio of the vaccinated population, as mentioned in the previous section, which is the sum of all vaccines given over the total population. This ratio goes from 0 to 2, where 2 is the entire population double vaccinated. Finally, the vector Z_i contains control variables, such as the density score, case average, share of the population that is 65 years old, and quality of healthcare. No significance is lost between using the single-dose vaccination ratio or the total vaccination ratio.

The tables below summarize the result of the regressions.

	Model 1	Model 2	Model 3	Model 4	Model 5
Vaccination rate	-61.7848**	-70.4107 *	-63.7416*	-59.1924**	-64.0156 *
	(-13.1493)	(-16.7438)	(-20.4534)	(-16.5785)	(-28.9196)
Case average		0.0076	0.0024	0.0027*	0.0025*
		(0.0055)	(0.0058)	(0.0058)	(0.0058)
Density score			-1.2205**	-1.3713**	-0.8802*
			(-0.4868)	(-0.6131)	(-0.3177)
Health score				0.0037	0.0033
				(0.0040)	(0.0041)
Over 65 pop. share					-0.5458
					(0.9761)
N	186	186	186	186	186
R-squared	0.057	0.144	0.153	0.159	0.1605
R-squared Adj.	0.048	0.034	0.034	0.034	0.0292
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 3.7: Regression with full dataset

	Model 1	Model 2	Model 3	Model 4	Model 5
Vaccination rate	-63.4592**	-61.0383**	-68.4777**	-60.7184*	-62.5412*
	(-15.3418)	(-15.7294)	(-13.2690)	(-23.9518)	(-19.8612)
Case average		0.0035*	0. 0042*	0.0027	0.0041*
		(0.0011)	(0.0028)	(0.0028)	(0.0021)
Density score			-1.512*	-1.825**	-1.422*
			(-0.86)	(-0.91)	(-0.95)
Health score				0.5622	0.6311
				(0.5729)	(0.5837)
Over 65 pop. share					-0.423**
					(-0.0825)
N	48	48	48	48	48
R-squared	0.0754	0.0772	0.0845	0.0966	0.1072
R-squared Adj.	0.0557	0.0471	0.0466	0.0463	0.0461
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 3.8: Regressions using events with non-zero vaccination rates

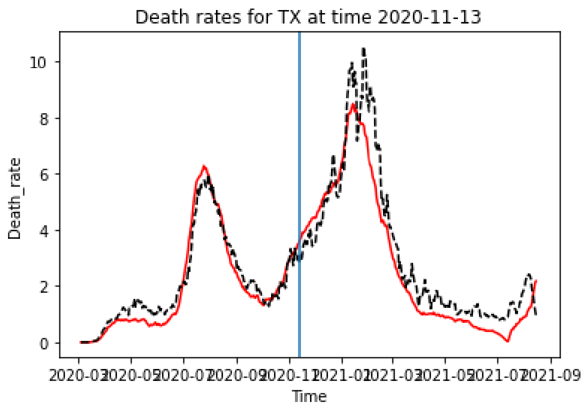
The vaccination rate is significant in both cases, even when controlling for other variables. The density score and the share of the population over 65 seem significant in terms of death rate changes. Table 3.8 shows that when the vaccination ratio goes up by 1, i.e., the population is fully vaccinated, differences in death rates from removing restrictions are reduced by around 60 to 70 percent from their predicted value.

3.5.3 Robustness checks

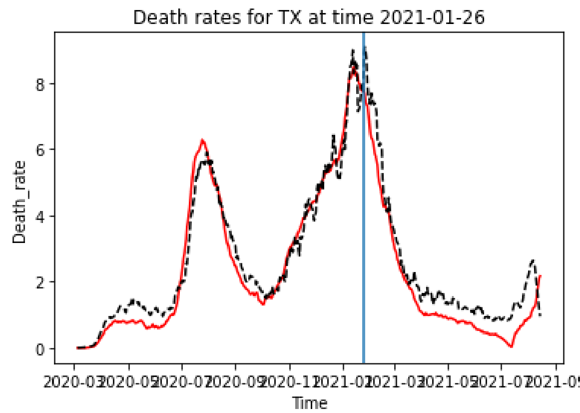
Several robustness tests are conducted to check the Synthetic Control Method results' validity as Abadie (2021) proposed. The first check is backdating, where the start of a treatment period is backdated to a random point, and the deviations from the backdated event are compared to the actual event. The second robustness test is changing the composition of the co-factor vector. When backdating, for every significant event, and where possible, an Synthetic Control Method regression is done ten days earlier than the event's occurrence. The average from T_{0+20} to T_{0+60} is computed, which I call $\bar{\alpha_{ro}}$ and subtracted from the T_{0+10} to T_{0+50} average deviation, which was obtained earlier. For the list of events, the following null hypothesis is proposed:

$$H_0 : \bar{\alpha}^2_{ro,i} = \bar{\alpha}^2_i \quad H_a : \bar{\alpha}^2_{ro,i} \neq \bar{\alpha}^2_i$$

where $\bar{\alpha}_{ro,i}$ is the average deviation from the robust Synthetic Control Method for an event, and the null is rejected with a confidence level of 95%. This robustness technique checks whether the treated period is indeed the period where events occur, and corrects for potential self-selection that could happen from the announcement of a policy implementation/removal. For example, taking the case of Texas, and for the January 26th, 2021 date, the state of Texas imposed a restriction (Movement restriction). Since this event is significant, it is expected that the predicted value of the Synthetic Control Method will be above the actual data. Assuming the event happens before January 26th, for example, on January 16th, the changes to the death rate should still occur around February 7th. As an additional robustness exercise, further periods were used to check for the results' validity. The following figure shows the case of selecting October 26th as a placebo for the event that occurred on January 26th. The null is not rejected for the panel of events studied for every event selected, and the results from this robustness test are summarized in Appendix B.



(a) Placebo simulation



(b) Actual simulation

Another robustness check is conducted on selecting vaccination rate as a co-factor since the CDC calculates three different rates of vaccinations. This covariate could change the predictive power of the model. For a random pool of states and events from the list of events considered, a simple Synthetic Control Method is run for each one of the selected events, using three different vaccination values as co-factors.

- Single dose vaccinations: percentage of the population that has received one dose of the vaccine.
- Double dose vaccination: percentage of the population that has received two doses of the vaccine.
- Total vaccination ratio: ratio of total vaccines administered over total population. This ratio has an upper value of 2, assuming every individual has been vaccinated).

The average Mean Square Predicted Errors (mean square predictive error) for the pre-treatment treated value is calculated for each permutation of event and vaccination rate used. The Mean Square Predicted Errors associated with the total vaccination rate is the lowest, but the difference between the first dose percentage and the total vaccination ratio is negligible. The formula used to calculate the Mean Square Predicted Errors is:

$$MeanSquarePredictedErrors_N = \frac{1}{Pr}\sum_{pr=0}^{Pr}\left(\frac{\sum_{n=0}^{N}\sum_{t=0}^{T_0-1}\alpha_{n,t,pr}^2}{N \times T}\right)$$

Where i is the region, t is the period-specific to that region, and P is an arbitrary number of permutations. For simplicity, 100 possible permutations are chosen for the three different values of N. The data is summarized in Table 3.9.

	Single dose	Double dose	Total vaccination ratio
$N = 3$	0.22	0.42	0.27
$N = 5$	0.20	0.33	0.17
$N = 10$	0.18	0.45	0.13

Table 3.9: Mean Square Predicted Errors results

Therefore using the single dose or total vaccination ratio does not impact the model's accuracy.

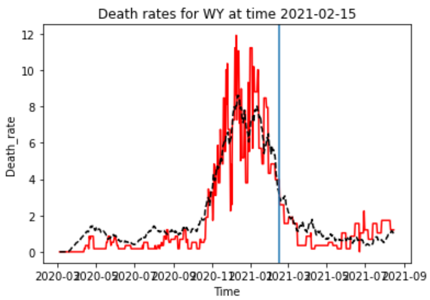
3.6 Conclusion

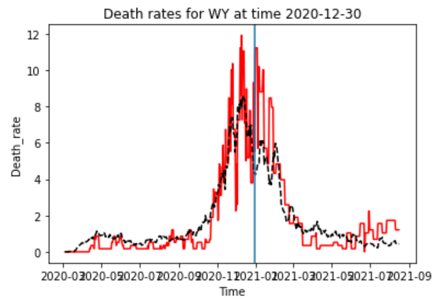
From February 2020 to June 2021, deaths from the COVID-19 pandemic had reached around 700,000, making this virus the highest-killing disease during that period, eclipsing almost every other deadly disease. These numbers do not account for under-reporting deaths from the virus. Restrictions are still being constantly implemented and removed, and this chapter provides a more precise idea of the effectiveness of these anti-contagion policies. This chapter estimates that by implementing anti-contagion policies, U.S. states reduced deaths by around 200,000 or 30% of the total deaths during the same period. A caveat is that these values are not estimates of the total number of deaths that were reduced by all policies, but only the ones selected for this chapter. The model described in this chapter requires that no other policies were implemented in a 50-day time frame, meaning several periods were excluded from the regressions. The final list consists of 256 potential events for analysis, spanning from March 2020 to June 2021.

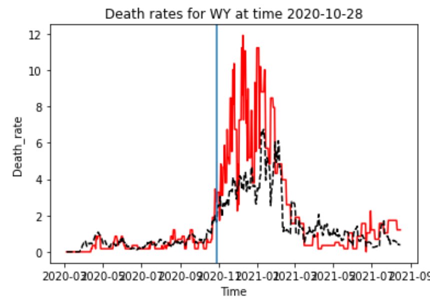
The intricacies of different policies, the stringency of their application, and the multitude of policies implemented simultaneously make it difficult to assess the efficiency of these policies properly. The contribution provided by this chapter is disentangling those effects from each other. This chapter finds that workplace closures, stay-at-home requirements, and restrictions on gatherings were the most efficient. This could be due to the ability of a policy maker to impose those restrictions compared to something like limiting mobility. It is also important to note that closing schools was the least efficient in reducing death rates. This chapter highlights the importance of properly securing and sanitizing a public environment to minimize disease transmission.

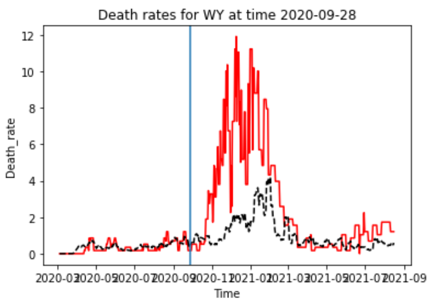
This chapter defines the vaccine ratio as the number of vaccines administered within a state over the population eligible to receive vaccines in that state. Removing restrictions with an under-vaccinated population (vaccine ratio of less than one) leads to an increase in the death rates by around 40%. It is also estimated that when a population is at least 66% double vaccinated (or vaccine ratio equal to 1), death rates are reduced by around 60% after removing restrictions. Additional research could tackle the initial conditions that could have impacted the effectiveness of these policies.

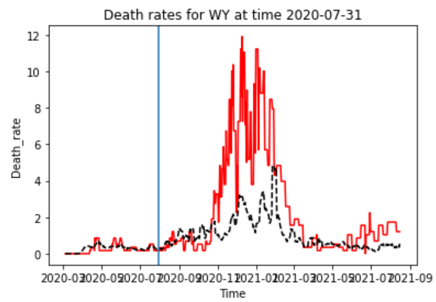
Appendix A

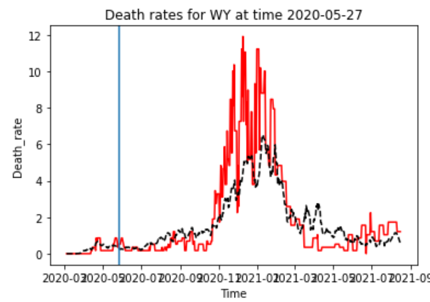

Death rates for WY at time 2021-02-15
Death_rate
Time

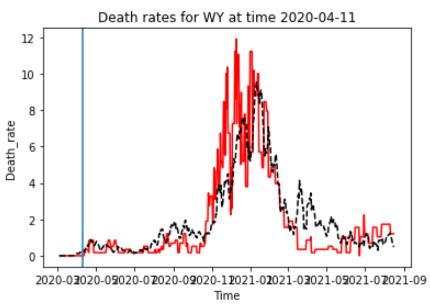

Death rates for WY at time 2020-12-30
Death_rate
Time

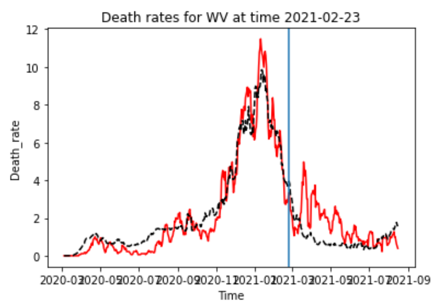

Death rates for WY at time 2020-10-28
Death_rate
Time

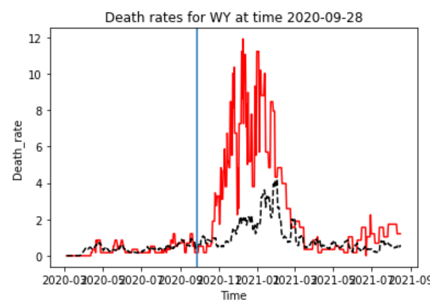

Death rates for WY at time 2020-09-28
Death_rate
Time

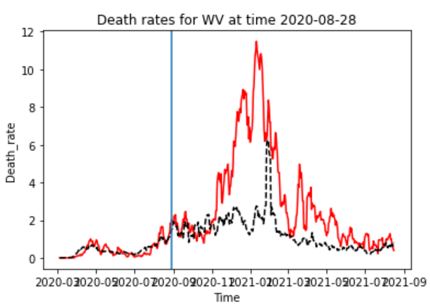

Death rates for WY at time 2020-07-31
Death_rate
Time

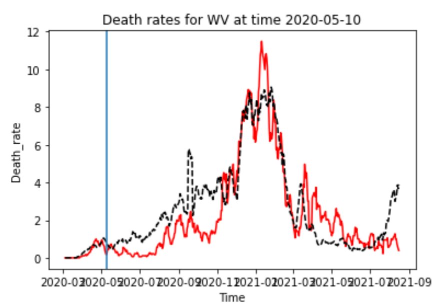

Death rates for WY at time 2020-05-27
Death_rate
Time

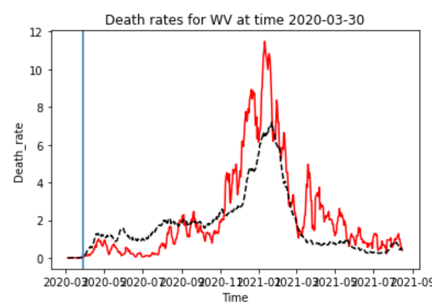

Death rates for WY at time 2020-04-11
Death_rate
Time

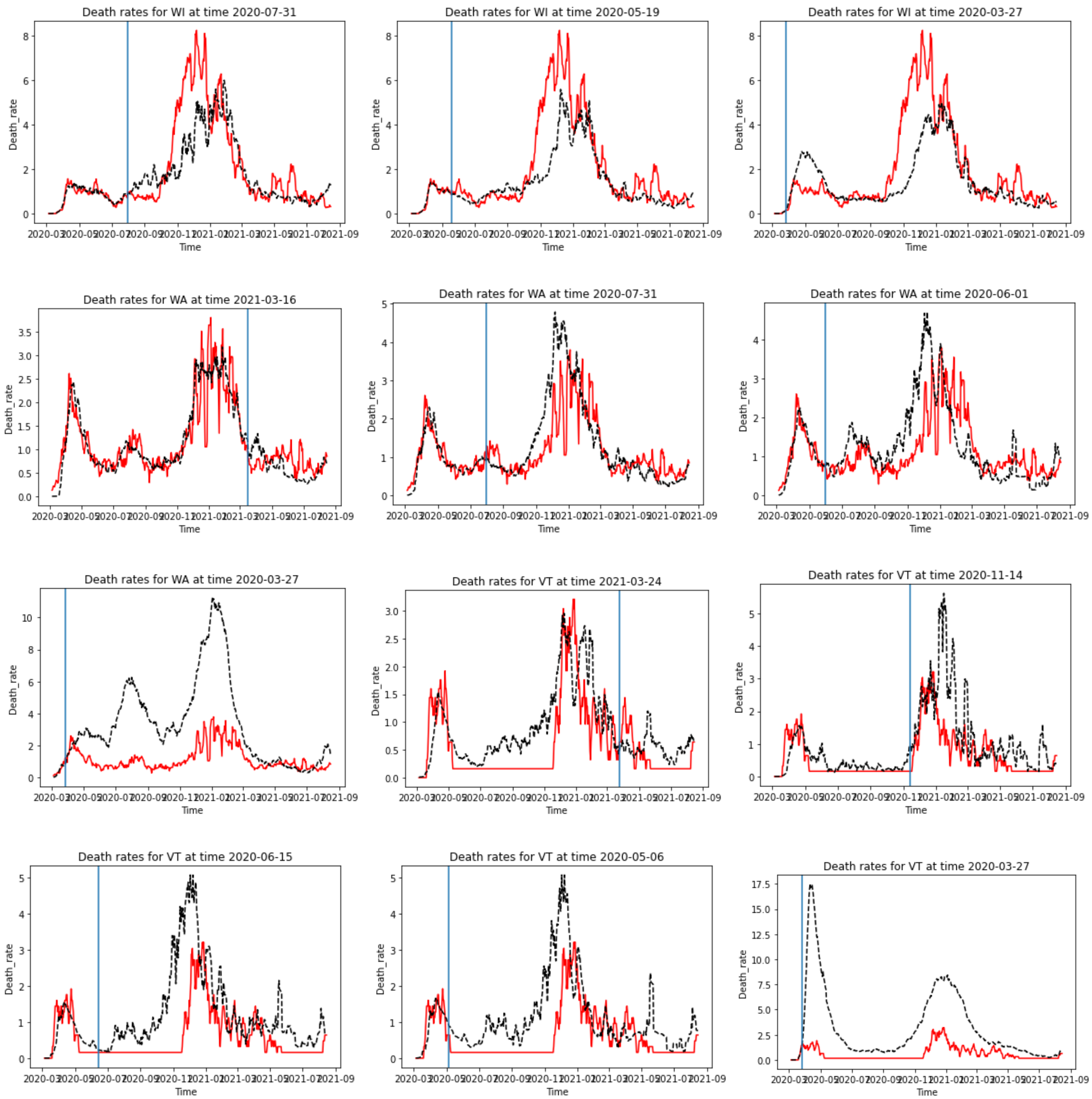

Death rates for WV at time 2021-02-23
Death_rate
Time

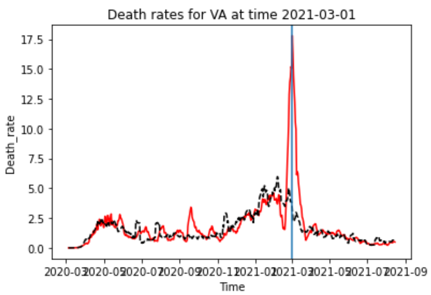

Death rates for WY at time 2020-09-28
Death_rate
Time

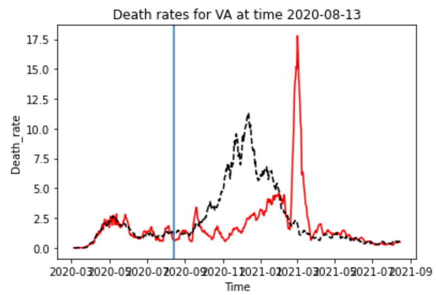

Death rates for WV at time 2020-08-28
Death_rate
Time

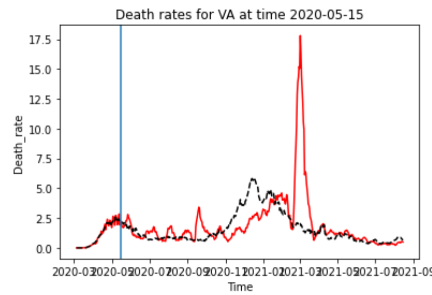

Death rates for WV at time 2020-05-10
Death_rate
Time

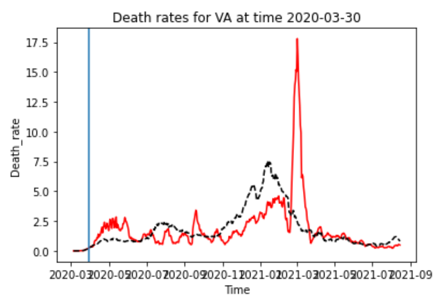

Death rates for WV at time 2020-03-30
Death_rate
Time

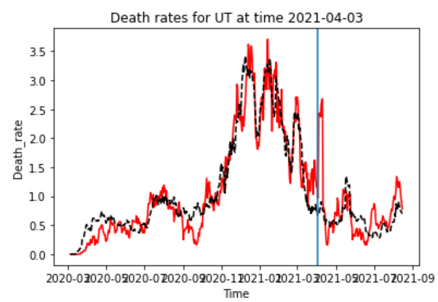

Death rates for WI at time 2020-07-31
Death rates for WI at time 2020-05-19
Death rates for WI at time 2020-03-27
Death rates for WA at time 2021-03-16
Death rates for WA at time 2020-07-31
Death rates for WA at time 2020-06-01
Death rates for WA at time 2020-03-27
Death rates for VT at time 2021-03-24
Death rates for VT at time 2020-11-14
Death rates for VT at time 2020-06-15
Death rates for VT at time 2020-05-06
Death rates for VT at time 2020-03-27
Death_rate
Time

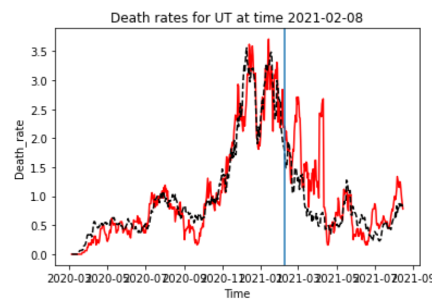

Death rates for VA at time 2021-03-01
Death_rate
Time

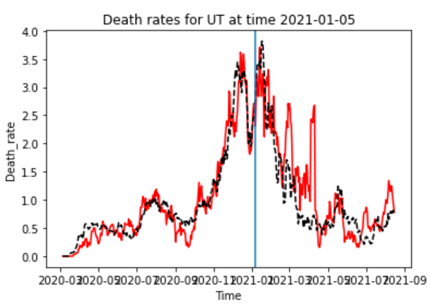

Death rates for VA at time 2020-08-13
Death_rate
Time

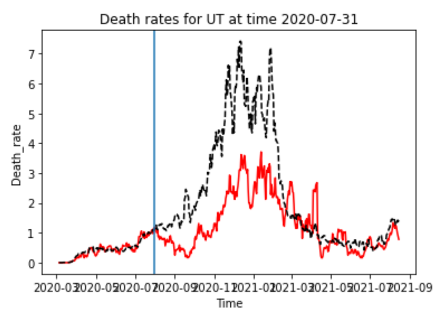

Death rates for VA at time 2020-05-15
Death_rate
Time

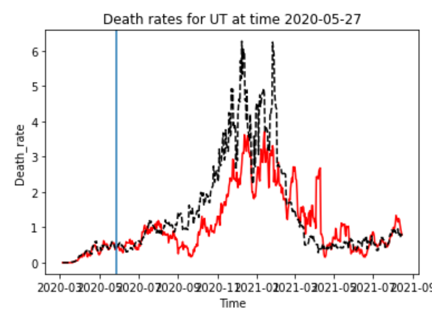

Death rates for VA at time 2020-03-30
Death_rate
Time

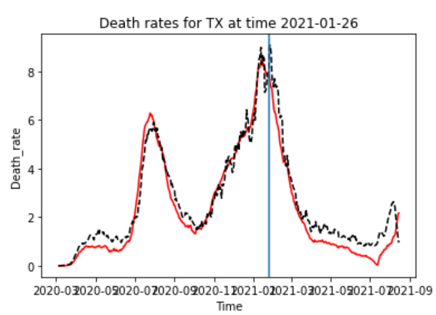

Death rates for UT at time 2021-04-03
Death_rate
Time

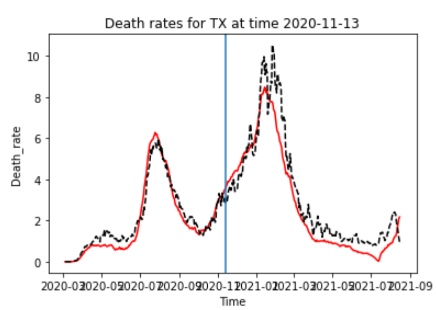

Death rates for UT at time 2021-02-08
Death_rate
Time

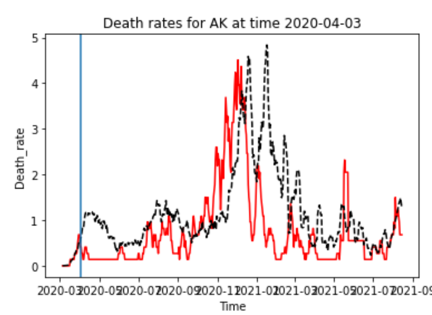

Death rates for UT at time 2021-01-05
Death_rate
Time

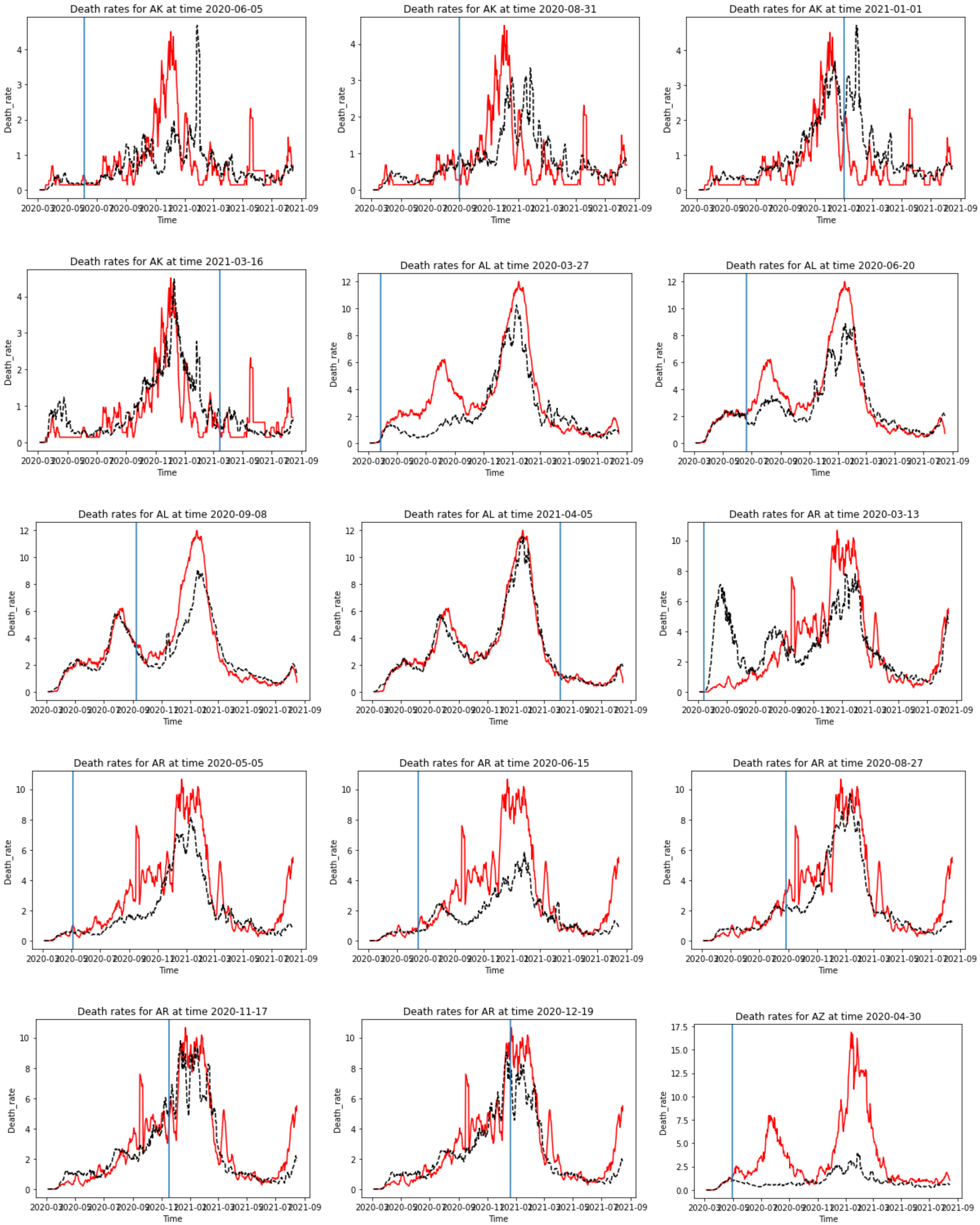

Death rates for UT at time 2020-07-31
Death_rate
Time


Death rates for UT at time 2020-05-27
Death_rate
Time


Death rates for TX at time 2021-01-26
Death_rate
Time


Death rates for TX at time 2020-11-13
Death_rate
Time


Death rates for AK at time 2020-04-03
Death_rate
Time


Death rates for AK at time 2020-06-05
Death rates for AK at time 2020-08-31
Death rates for AK at time 2021-01-01
Death rates for AK at time 2021-03-16
Death rates for AL at time 2020-03-27
Death rates for AL at time 2020-06-20
Death rates for AL at time 2020-09-08
Death rates for AL at time 2021-04-05
Death rates for AR at time 2020-03-13
Death rates for AR at time 2020-05-05
Death rates for AR at time 2020-06-15
Death rates for AR at time 2020-08-27
Death rates for AR at time 2020-11-17
Death rates for AR at time 2020-12-19
Death rates for AZ at time 2020-04-30
Death_rate
Time

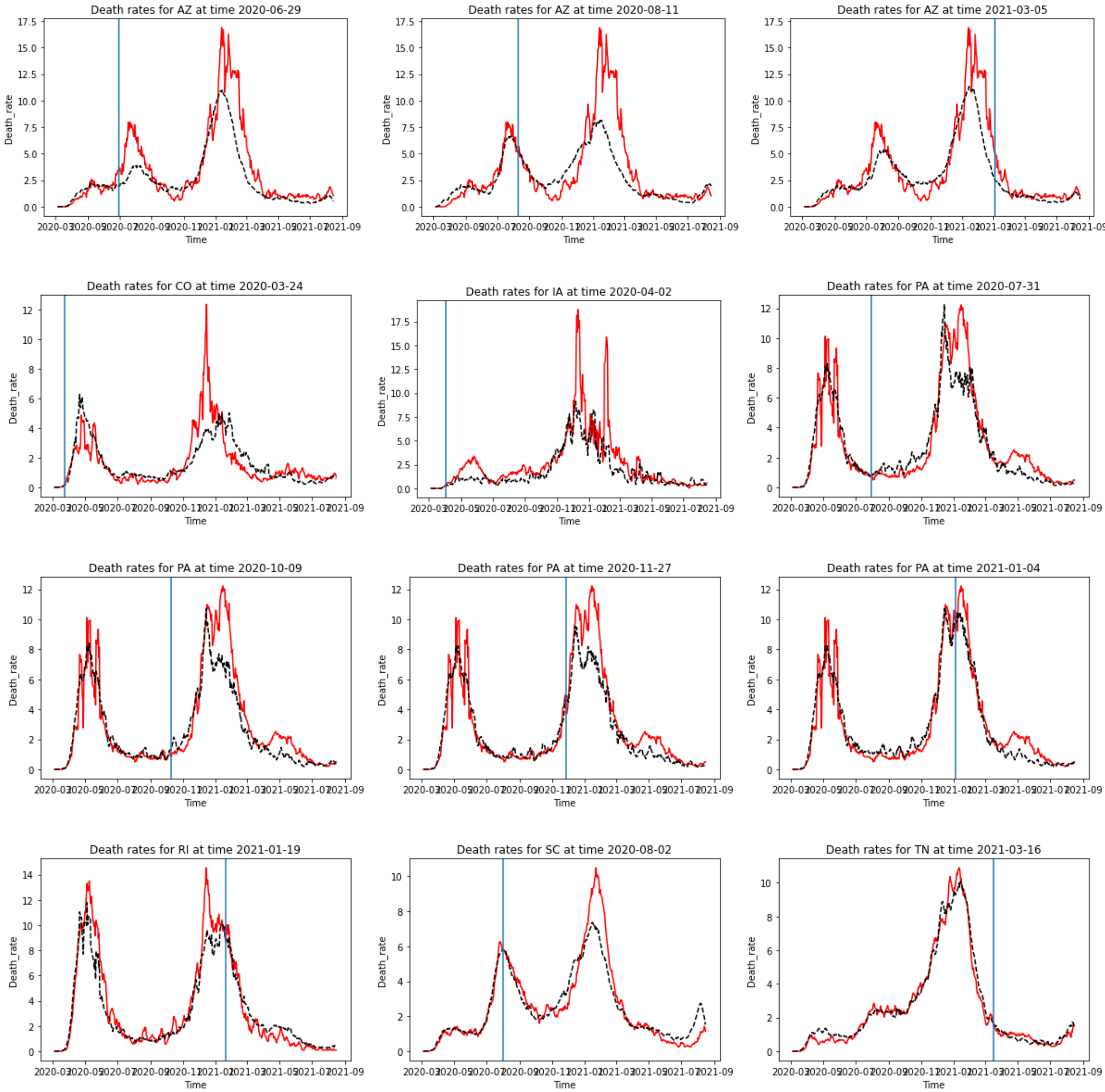


Figure 3.-5: Simulations obtained from Synthetic Control Method

Appendix B

FDS is the full data set and NZV is non-zero vaccination rates. Below are the results from the regressions on the different means of average deviations or $\delta_{FDS,T}$ and $\delta_{NZV,T}$ T_{0+10} to T_{0+30},T_{0+20} to T_{0+40}, and T_{0+20} to T_{0+50}

	R 1	R 2	R 3	R 4	R 5
Vaccination rate	-55.84**	-58.35**	- 57.44**	-58. 97**	-60.31**
	(-22.41)	(-19.47)	(-21.34)	(-28.33)	(-21.97)
Case average		0.008*	0.007*	0.007	0.006*
		(0.005)	(0.006)	(0.0007)	(0.0004)
Density Score			-1.41**	-1.33**	-1.82*
			(-0.43)	(-0.48)	(-1.03)
Health Score				0.21	0.11
				(1.18)	(1.42)
Over 65 pop.share					-0.77
					(-0.98)
R-squared	0.0681	0.0775	0.102	0.105	0.111
R-squared Adj.	0.0554	0.0521	0.0511	0.0510	0.0507
N	155	155	155	155	155
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 3.10: OLS regression (FDS) for T_{0+10} to T_{0+30}

Table 3.11: OLS regression (FDS) for T_{0+10} to T_{0+40}

	R 1	R 2	R 3	R 4	R 5
Vaccination rate	-62.41***	-60.22***	- 59.31***	-63. 93**	-61.48**
	(-20.33)	(-24.81)	(-18.40)	(-22.81)	(-19.85)
Case average		0.007*	0.007*	0.007	0.007*
		(0.004)	(0.004)	(0.0006)	(0.0005)
Density Score			-1.32**	-1.38**	-1.33*
			(-0.94)	(-0.98)	(-1.03)
Health Score				0.35	0.48
				(2.18)	(1.02)
Over 65 pop.share				-0.37**	
					(-0.18)
R-squared	0.0803	0.0825	0.122	0.125	0.141
R-squared Adj.	0.0753	0.0729	0.0688	0.0681	0.0667
N	155	155	155	155	155
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 3.12: OLS regression (FDS) for T_{0+20} to T_{0+40}

	R 1	R 2	R 3	R 4	R 5
Vaccination rate	-62.41***	-60.22***	- 59.31***	-63. 93**	-61.48**
	(-20.33)	(-24.81)	(-18.40)	(-22.81)	(-19.85)
Case average		0.007*	0.007*	0.007	0.007*
		(0.004)	(0.004)	(0.0006)	(0.0005)
Density Score			-1.32**	-1.38**	-1.33*
			(-0.94)	(-0.98)	(-1.03)
Health Score				0.35	0.48
				(2.18)	(1.02)
Over 65 pop.share				-0.37**	
					(-0.18)
R-squared	0.0803	0.0825	0.122	0.125	0.141
R-squared Adj.	0.0753	0.0729	0.0688	0.0681	0.0667
N	155	155	155	155	155
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 3.13: OLS regression (FDS) for T_{0+20} to T_{0+50}

Chapter 4

A multi-stage synthetic control approach: assessing the cumulative effect of COVID-19 anti-contagion policies

Abstract

The synthetic control method is a potent policy analysis tool. An application of this tool is to study the impact of anti-contagion policies during the COVID-19 pandemic on various economic and health indicators. A shortcoming of this method in its basic form is that a treated event needs to be uninterrupted after its treatment. I propose a modified version of this model, with the proper inference tools, that considers multiple interruptions of treatments over long periods. I test this model for various U.S. states and found that this model can account for interruptions in events and provides predictions for cumulative treatment effects. Results from the multi-staged Synthetic Control Method show that in the case of multiple events occurring within short periods, the Synthetic Control Method understates the impact of anti-contagion policies by around 200%.

4.1 Introduction

Chapter 3 discussed how implementing and removing different anti-contagion policies impacted the death rates and total deaths in various states in the United States. The effects of specific policies were quantified by using the synthetic control method. For this, synthetic control regressions were applied at specific times and for unique events, where only one policy was implemented or removed at a time. A significant restriction for event selection was that anti-contagion policies had to run uninterrupted for 50 days or more. Since restrictions were often put in place or removed in overlapping sequences, analysis often failed to capture or understand the cumulative impact of anti-contagion policies.

Many improvements to the Synthetic Control Method have been made since Abadie, Diamond, and Hainmueller (2010a) first introduced the concept. Most notably and recently was the augmented synthetic control method proposed by Ben-Michael, Feller, and Rothstein (2021), which corrects any bias that might result from a lousy pre-treatment fit of the data by correcting for the covariates in the regression. Kreif et al. (2016) presented a different version of the synthetic control method that considers multiple treated units receiving the same treatment. This technique was something that was used previously in this thesis. Both Ben-Michael, Feller, and Rothstein (2021) and Abadie (2021) highlight that in its basic form, the Synthetic Control Method cannot possibly find the impact of events that are interrupted after their occurrence. This was also mentioned by Cavallo et al. (2013). In his research, the selection of disasters chosen was limited by the Synthetic Control Method's inability to deal with multiple disasters that would occur within short periods.

This chapter proposes a novel addition to the synthetic control method to tackle such a gap, an augmented version of this model that accounts for more than one treatment prior to the last occurrence of uninterrupted treatment. This change allows the user to understand the cumulative effect of multiple overlapping policy changes. The validity of this method is checked by conducting several robustness tests using established data and results. In a situation where treatment is interrupted by the occurrence of another treatment or the end of that treatment, there is a compounding effect that a traditional Synthetic Control Method regression is not able to capture. On the other hand, applying a multi-stage Synthetic Control Method to an event where no significant events occurred before it does not change the results. In other words, when policies are implemented in a stratified way, where they are not all implemented at the same time, it is possible to distinguish the effect of each policy and quantify the cumulative effect of all the policies at a final point.

Using a selection of 8 states from the U.S. with varying policy implementation schedules, I find that there is a stacking effect that is not captured by simple Synthetic Control Method regressions. The simple Synthetic Control Method for one-time events was understated by up to 150% for certain states. Second, when applied to the sample chosen in this chapter, the states that removed policies too early do not experience a compounding effect, even if the implementation of policies significantly reduced death rates. Despite many treatments, I find that if a state removed policies implemented too early, there is no evidence of a lasting compounding effect. These changes to the Synthetic Control Method are robust to various checks when conducting a multi-stage Synthetic Control Method (Multi Synthetic Control Method) regression on events where there is no prior impactful policy change. The results are similar to the standard Synthetic Control Method, with no significant distinction between the two.

An essential contribution of this chapter to previous work in this dissertation and the literature is the ability to track the cumulative effects of policies that overlap with other policies. The Multi Synthetic Control Method regressions involve an iterative process for different events in the same region. I supplement this with a change to the inference method to test this new proposal. The Synthetic Control Method is already an excellent tool for this type of policy analysis, and the additions mentioned in this chapter expand the possible scenarios where this tool can be used. More specifically, it allows researchers to understand the impact of events that do not run uninterrupted, unlike the basic Synthetic Control Method regression. When considering how anti-contagion policies were frequently implemented and interrupted in cycles, this improvement allows me to tackle this topic more efficiently.

As far as my knowledge goes, this chapter provides a novel approach to the Synthetic Control Method that has not been explored in this way by other authors so far. The rest of the chapter is organized as follows. In section 4.2, I discuss the data used to test out this model and the sources I used for this data. In section 4.3, I explain the changes to the baseline model of the Synthetic Control Method and highlight how these changes are implemented, including inference and robustness testing. I test this updated model on benchmark Synthetic Control Method examples. In section 4.4, I show the results of my analysis. Finally, in sections 4.5 and 4.6, I run through the robustness exercises mentioned earlier, and I discuss my conclusion and potential further improvements to this model.

4.2 Data and data selection

4.2.1 Data sources

Data sources are similar to those in chapter 3. However, data selection is different. There is no restriction on whether treatments are uninterrupted. Additional co-factors are introduced for more robust and accurate simulations. Air quality was removed as it was insignificant in the previous chapter in the regressions. The covariates chosen for this model are

- **Unemployment:** I use unemployment data obtained from the FRED database. The data is monthly for all 50 states. This covariate underlines possible transmission channels coupled with on-site work, as lower unemployment with no work-from-home restrictions could lead to more transmissions.

- **Density score:** a variation on traditional density measures. Instead of using state-wide density, I build a composite that I call the density score. For every state, I find the three largest cities in that particular state in terms of population, and I build a weighted average of their density based on their share of the total population.

- **Vaccination rate:** I use the vaccination rate per 100,000 persons per state, obtained from the CDC database. I make no distinction between the different vaccines used. Since vaccination only started halfway through the data set. This co-factor is dropped for any simulations done before the beginning of vaccinations.

- **Old age population share:** the share of people aged 65 and above, obtained from the U.S. census bureau.

- **Testing:** testing data, which includes the number of confirmed cases, was obtained from the Johns Hopkins Centers for Civic Impact (Johns Hopkins University, 2021), and it includes data from March 2020 to July 2021.

- **Healthcare quality index:** I use this composite index, obtained from the Agency for Healthcare Research and Quality, part of the department of the U.S. Health and Human Services. It provides an index on the quality of healthcare per U.S. state, based on insurance costs, access to healthcare, quality of healthcare, and surveys from healthcare professionals. I consider the complete index, and the index for access to healthcare separately. The data covers all 50 U.S. states and is available for 2020.

- **Mobility trends:** data mobility trends obtained from both Google and Apple (Apple, 2021 Google, 2021) for the two most populous cities per state, including walking and driving within the cities.

The treated variable considered is deaths caused by COVID-19, and I obtained daily data from the Johns Hopkins Centers for Civic Impact[1](Johns Hopkins University, 2021). It includes data from March 2020 to September 2021, a total of 518 days. The data is a 7-day moving average of the number of deaths per 100,000. An essential benefit of the Synthetic Control Method compared to traditional policy analysis tools such as difference-in-difference is the ability to select from a wide variety of control groups, even if the control group is relatively small. When conducting their panel regression on the impact of anti-contagion policies in different countries using DiD, this was one of the struggles mentioned by Deb et al. (2020).

4.2.2 Data selection

Similar to chapter 3, the history of all "strict" policies implemented in all 50 U.S. states from March 2020 to September 2021 is recorded for any policy where the flag was more than 2. The

[1]https://civicimpact.jhu.edu/

selection process for events is as follows:

1. Every time there is a treatment, only one treatment must have occurred. In other words, when a restriction was removed or added, only one restriction must have been done at that particular time. There is no requirement on time between treatments, as long as it is more than one day.

2. When treatment occurs, I calculate how many states have also received that treatment in the 15 days prior. I do not consider any event where more than ten states (20%) of my regions received such a treatment within the past 15 days. This is done to ensure a large enough pool of donors not affected by this treatment. This can be relaxed, however it does not affect the results of the research. It however increases computational time.

These two criteria are set in such a way as to ensure comparability between the Synthetic Control Method and the Multi Synthetic Control Method. The Synthetic Control Method cannot untangle the effect of multiple events occurring simultaneously. Considering the states where the daily death rates were somewhere between the 20th and the 80th percentile of all states allows for the selection of a group without any outliers, removing the need for relaxing assumptions about the Synthetic Control Method.

A daily rank is assigned for each state, between 1 and 50, where one refers to the state with the highest death rate at time t, and 50 is the lowest death rate. Any state ranking anywhere between 1 and 10 or 40 and 50 for more than 200 days was eliminated from the selection of states. The pool of possible states to choose events from is reduced to 14. They are highlighted in Table 4.1.

state	days	events	state	days	events
IN	393	13	IA	412	11
KY	465	15	MD	399	13
MN	443	7	NM	405	12
NC	496	5	OH	440	4
TX	436	17	VA	395	11
ID	382	16	NE	321	17
MA	340	10	UT	302	5

Table 4.1: Potential states

The percentiles chosen for outliers could be changed with some loss to computational speeds if they are expanded or if the assumption on the weights is relaxed, namely $w_i \geqslant 0$ not holding; however, the results would not be significantly different. These assumptions are relaxed in Section 4.5, and this analysis is done for a panel of 20 events. There is no loss in accuracy in terms of the deviations obtained from the Synthetic Control Method, but on average, the regressions take twice as much time to finish.

After the initial selection of states from the initial filters, all policy changes over the 518 days are identified. Then the states where the timing between policy changes was at most sixty days for consecutive implementation or removal of policies were selected. Results from different authors mentioned in chapter 2 and the author's results in chapter 3 show that the impact of anti-contagion policies on death rates drops around 50 days after their implementation but does not entirely disappear. Eight states are left for the final Multi Synthetic Control Method regressions. Considering that the Synthetic Control Method can accurately present counterfactual values for one-hundred periods after an event, this assumption can be relaxed for any other research.

The total number of events considered after selecting states is 113. Since eight states were selected, the first event for each state is discarded in the multi-stage analysis as it would be the same result as the traditional Synthetic Control Method. The different events for the eight states are also not equally distributed between them. Out of these 105 events, 41 were for events where restrictions were added, and 64 were events where restrictions were removed. Table 4.2 shows the distribution of these 105 events listed below and the average duration of implemented events in days. The inference is made on the standard Synthetic Control Method and Multi Synthetic Control Method for each of these treatments.

State	# implemented	# removed	Avg. duration (days)
Iowa (IA)	5	6	52
Kentucky (KY)	9	6	31
Indiana (IN)	7	6	32
Texas (TX)	10	7	35
Idaho (ID)	7	9	40
Nebraska (NE)	9	9	21
Maryland (MD)	7	6	57
Massachusetts (MA)	6	4	63

Table 4.2: States selected and policy distribution and duration

4.3 Multi-stage Synthetic Control Method model

The basic Synthetic Control Method model was explained thoroughly in chapters 1 and 3, but a brief refresher is provided in this section to highlight the additions of the multi-stage regression. The changes to the model are two-fold. The first change is the regression itself, where the vector of co-factors that was used in obtaining the optimal weights is changed for the synthetic model. Second is the inference methods used in testing the validity of the results.

In a standard Synthetic Control Method setting, the user of the model starts with a set of regions i where $i = \{1, \ldots, I\}$, in this case, the fifty states of the United States and $I = 50$, with no particular order given to the states. One of these states is exposed to the "treatment", such as the implementation or removal of an anti-contagion policy, where $I - 1$ states are not treated and a state $i = tr$ is the state treated. This state is treated at time t. For simplicity, I assume that the treated region is $i = 1$. The model's outcome or treated variable is the 7-day moving average deaths per 100,000. $y_{i,t}$ is the actual outcome variable of state i at time t. The outcome variable is observed over T periods. At a point $t = T_0 < T$, the treatment occurs, but only for the affected region $i = tr$, leaving $T - T0$ of treated periods moving forward, meaning the treatment is uninterrupted (in other words, there are no more treatments). The treated variable will then look as follows

$$y_{i,t} = \hat{y}_{i,t} + \alpha_{i,t} D_{i,t}$$

which can be rewritten as:

$$\hat{y}_{1,t} = \sum_{i=2}^{I} w_i y_{i,t}$$

where $t \leqslant T$. With $t \leqslant T_0$, $D_{i,t} = 0$, and$D_{i,t} = 1$ otherwise. To solve this model, I build a set of positive weights w_i where $i = 2, \ldots, I$ and $i \neq tr$, such that $\sum_{i=2}^{I} w_i = 1$. There are ideal weights w_i^* such that

$$\sum_{i=2}^{I} w_i^* Y_{i,t} = Y_{tr,t} \forall t \in T \quad \text{and} \quad \sum_{i=2}^{I} w_i^* Z_i = Z_{tr} \forall t \in T$$

where $Z_{i,t}$ is a vector of covariates chosen to find the set of appropriate W^*. Therefore the choice of covariates is crucial in properly identifying the weights. We can then use

$$\hat{\alpha}_{i,t} = y_{i,t} - \sum_{i=2}^{I} w_j^* Y_{j,t}$$

as a way to calculate $\alpha_{i,t}$ where $t \in \{T_0 + 1, \ldots, T\}$. This will be the deviation of the synthetic predicted model from the actual data. In this model, $\hat{y}_{i,t}$ then refers to the predicted treated variable.

4.3.1 Multi-stage changes

In the multi-stage synthetic control method proposed in this chapter, there is no longer a single treatment period for a treated event. For every region $i = tr$, there are now several treated periods defined as T_k. $\tilde{y}^i_{t,k}$ being the synthetic counterpart of the treated variable, for region i, time t, and event k. In the multi-stage synthetic control method, the aim is to estimate:

$$\tilde{\alpha}^i_{t,k} = y^i_{t,k} - \tilde{y}^i_{t,k}$$

This deviation is the cumulative deviation caused by consecutive policies, where $\tilde{y}$ reflects the estimated outcome variable after a series of events had occurred. In the simple Synthetic Control Method, $\hat{y}^i_{t,k}$ is the synthetic counterpart of an outcome variable that had been only affected with one treatment. $\tilde{y}^i_{t,k}$ is the synthetic counterpart of the same variable, having been affected by several treatments. The difference between these two variables is defined as

$$\delta^i_{t,k} = \hat{y}^i_{t,k} - \tilde{y}^i_{t,k}$$

which is the deviations caused by the consecutive events from the traditional Synthetic Control Method. By definition, since the α from the Synthetic Control Method already reflect the impact that a certain policy had at time t, $\delta^i_{t,k}$ captures the effect of prior events. If $\delta^i_{t,k} = 0$, this means that the multiple treatments did not impact the outcome variable as much as the treatment implied by $\hat{y}^i_{t,k}$

In order to capture this cumulative effect $\tilde{\alpha}$, the vector Z, which is the co-factor vector, is changed to reflect past effects of treatments. Recall that in the Synthetic Control Method the optimal weights are obtained in such as a way as to minimize a penalty function defined as follows:

$$||Z_1 - WZ_0|| = \sqrt{(Z_1 - Z_0W)'V(Z_1 - Z_0W)}$$

where Z_1 is a vector of pre-treatment variables and covariates relevant for the treated region and Z_0 is the same vector of variables for the non-treated regions. V is a positive semi-definite matrix that defines the relative importance of every co-factor in Z. In a traditional Synthetic Control Method $Z_1 = [x_1, \dots, x_m, Y_{1,0}, \dots, Y_{1,t}]'$ where x_m is covariate m and $Y_{i,t}$ is a linear combination of the treated variable at time t prior to the treatment. For simplicity and without any loss of accuracy, a valid combination is $Y^1_t = y^1_t$. In other words, the complete pre-treatment set of treated variables is included in the co-factor vector. In chapters 1 and 3, I had included only half of the pre-treatment set as an additional test for the accuracy of the model. This is not feasible in the Multi Synthetic Control Method.

The changes implemented in this chapter on the vector Z are the following. For region i, with treatment periods $k = 1, \dots, K$ with K being the total number of treatments for every treatment I consider treated variables up until the previous treatment. Each treatment starts at period $T_{0,k}$

where each event k lasts tor $T_{0,k+1} - T_{0,k}$ periods. I can then define

$$Z_K^i = [x_{i,1}, \ldots, x_m^i, \tilde{Y}_{1,1}^i, \ldots, \tilde{Y}_{T_{0,0}}^i, Y_{1,2}^i, \ldots, Y_{T_{0,K},K}^i]'$$

where i refers to the region, x_m^i is a covariate variable for the region i, and $\tilde{Y}_{t,k}^i = \hat{y}_{t,k}^i$, where $\hat{y}_{t,k}^i$ is the predicted value obtained from a standard Synthetic Control Method. This however can be changed to any other combination of the actual treated variable value and simple Synthetic Control Method predicted value. For this chapter, the Synthetic Control Method predicted value is used for its simplicity. I propose three propositions about the updates to the co-factor vector.

- The results of the first step of the multi-stage Synthetic Control Method and the traditional Synthetic Control Method are the same since, up until that point, no other events had occurred.
- There is no difference between using the predicted values from the traditional Synthetic Control Method or the multi-stage Synthetic Control Method when conducting the second stage of the multi-stage Synthetic Control Method. The predicted treated variable was obtained from this regression.
- For any treated period after the second period, the compounded effect of different policies is captured by the inclusion of the value of the predicted treated variable from the Synthetic Control Method as long as that value is significant.

The procedure used to run the regressions of the Multi Synthetic Control Method is as follows. For every region (or state), a traditional Synthetic Control Method regression is done for every event chosen; the values are tested for significance and stored if significant. These are then used for any subsequent regressions of the multi-stage regression. When calculating the appropriate weights, the vector that contains the covariates will change depending on the date chosen. The two main assumptions of the Synthetic Control Method do not necessarily have to hold. These assumptions are

$$\sum_{i=0}^{I-1} w_i = 1$$

$$0 \leqslant w_i \leqslant 1 \;\; for\{0, \ldots, I-1\}$$

When conducting the Multi Synthetic Control Method, the counterfactual treated variables could potentially become outliers. These counterfactuals would be used in the different stages of the regression, meaning a convex set of weights is no longer possible to find even if the chosen states and events are not necessarily outliers themselves initially. Outlier states will have fewer comparable states in their control pools. As pointed out by Li and Shankar (2020b), under certain circumstances, relaxing the assumptions mentioned above is possible without loss of accuracy to the regression. Increasing the bounds of the weights allows for more accurate Synthetic Control Method regressions if the treated variable is an outlier.

Once the values from the Synthetic Control Method are calculated and tested for significance, the second round of regressions are conducted and the observations from these regressions of the treated variable $\tilde{y}^i_{t,k}$ where $t \geqslant T_{0,K}$ are stored. They are then subtracted from the actual observation $y^i_{k,t}$ to get the new set of deviations, $\tilde{\alpha}$

4.3.2 Inference in the multi-stage Synthetic Control Method

Inference in the Multi Synthetic Control Method requires two steps. In the first step, inference is run on all the treated periods selected for the multi-stage. Any treated insignificant event is removed from the list of events chosen for the multi-stage list. This is done before to obtaining the results from the second stage of the Multi Synthetic Control Method. In the second step and after obtaining the predicted values $\tilde{y}$ from the Multi Synthetic Control Method, an additional inference test is conducted for each event in the multi-stage list that remains. There is only a slight deviation from the routine inference procedures of the Synthetic Control Method. Before doing this, though a test is also conducted on the difference between the Synthetic Control Method's and the Multi Synthetic Control Method's deviations. The null hypothesis for this test and every event k is

$$H_0 : \bar{\tilde{\alpha}}_{i,k} = \bar{\hat{\alpha}}_{i,k} \quad H_a : \bar{\tilde{\alpha}}_{i,k} \neq \bar{\hat{\alpha}}_{i,k}$$

Since this is a simple average of values, a t-test is sufficient to conduct hypothesis testing. Rejecting the null means that the deviations from the Synthetic Control Method are not equal to the deviations of the Multi Synthetic Control Method, and it is then possible to proceed with the rest of the inference procedure which is similar to that of the standard model.

1. I compute a placebo effect for every event by conducting an Multi Synthetic Control Method regression on the outcome variables for the available controls for the corresponding event. It is called a placebo effect because the control regions are selected so that they are not subjected to the same treatment as the region in question at the date chosen.

2. At every point in time following the occurrence of the disaster (called leads), I compute all the placebos alphas and then take the average across all placebos.

3. The actual lead is ranked in the distribution of placebo averages

4. The following formula gives the lead-specific p value.

$$p\text{-value} = \frac{\sum_{np=1}^{N_{pl}} I(\bar{\alpha}_l^{pl(np)} < \bar{\alpha}_i)}{N_{pl}}$$

   where $\bar{\alpha}_i$ is the effect of the restriction on the state in question and $\bar{\alpha}_i^p$ is the average placebo effect. This is only the case if α_i is expected to be positive (imposing restriction). If α is expected to be negative, then I consider the alternative p-value formulation; more specifically.

$$p\text{-value} = \frac{\sum_{np=1}^{N_{pl}} I(\bar{\alpha}_l^{pl(np)} > \bar{\alpha}_i)}{N_{pl}}$$

4.3.3 Testing the updated model

To test the validity of the changes, the model is applied to two established applications of the synthetic control method. These models are the benchmarks used in most updates to the Synthetic Control Method (Li and Shankar, 2020b), (Ben-Michael, Feller, and Rothstein, 2021), and (Kreif et al., 2016). The objective is to see if the additions proposed in this chapter would change the results obtained in these benchmark papers. The first is Abadie, Diamond, and Hainmueller, 2010a (**P1**), which finds the impact of California's Tobacco Control program on the consumption of Tobacco in the state. The second is Abadie, 2021 (**P2**),where the authors test the impact of the German reunification on the GDP of West Germany. In order to test the model, three, four, and five events are randomly selected prior to the concurrence of the event described in each of the papers. The post-treatment deviations obtained from each selection of the Multi Synthetic Control Method are compared to the results from the papers. 10 post treatment period averages are used for each test. The final results are summarized in table 4.3.

Method	Results from **P1**	Results from **P2**
3-stage Synthetic Control Method	-0.03***	0.08***
4-stage Synthetic Control Method	-0.03***	0.05**
5-stage Synthetic Control Method	-0.12**	0.06**
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 4.3: Results from testing Synthetic Control Method

Both of these events were unique and there were no significant events prior to these that had affected the treatment variable. It is expected that the results from the Multi Synthetic Control Method should not deviate from the traditional Synthetic Control Method. Table 4.3 shows that for all the different Multi Synthetic Control Method variations, the null from the test is not rejected. The values displayed in the two right columns are the differences between the average from the Multi Synthetic Control Method and the papers studied.

4.4 Results from models

This paper presents an application of this model, however, it is computationally intensive. For the 113 events selected and the follow up inference tests, a total of 4,381 regressions were conducted. Out of these events, 74 were significant, and these events are the ones selected for the second step of regressions. I then start the work's second step, but only using those 74 in my multi-stage regressions. With inference testing, an additional 3,126 regressions were conducted.

Means	IA	IN	ID	TX	MD	NE	MA	KY
T_{0+15}-T_{0+30}	-31.7	-32.5	-33.8	9.8	-15.8	34.7	-131.5	3.3
T_{0+15}-T_{0+45}	-41.9	-76.9	-37.3	14.0	-25.0	39.9	-156.4	5.2
T_{0+15}-T_{0+60}	-24.1	-98.3	-37.5	13.3	-34.8	49.0	-92.9	-9.7

Table 4.4: Average percentage deviation for the last event using simple Synthetic Control Method

4.4.1 Standard Synthetic Control Method regressions

In this subsection, I discuss the results from the Synthetic Control Method regressions. The discrepancy between the length of the different policies means that I cannot observe the same means for these policies. In other words, for some policies, the maximum length of time available until the next set of policies was implemented was relatively small, and while the event itself was significant, the follow up event occurred within a time frame where in such a I am unable to compare 50 or 60 day averages across events. However, for the entirety of the regions considered, the final treatment or event was uninterrupted for a total of 60 days. In fact, for all eight states, the last policy considered was the removal of one or all restrictions, which was a significant event for 7 out of the 8 states.

Out of the 39 insignificant events, the majority were for removing restrictions, a total of 28 out of 39. These results are similar to those found in chapter 3. The only exceptions are Kentucky and Nebraska, where the only insignificant events were the removal of restrictions. For the last event selected for each state, only the event that was selected for the state of Kentucky was insignificant. Table 4.4 displays the synthetic results of several averages of the predicted treated variables on the vaccination rates only for the standard Synthetic Control Method and for the last event in the list. These values will then be subtracted from the Multi Synthetic Control Method for those events.

4.4.2 Multi-stage regression

Given the objective of this chapter, I provide regression results only on the last treatment period of the multi-stage Synthetic Control Method for every state. No two treatment periods start on the same day for any of the eight states. However the predicted values 60 days after the treatment are obtained regardless. The treatment, in this case, being the last date where a major policy change was recorded for any of these states within the time frame mentioned earlier. The latest treatment in this case occurred on April 5th, 2021; in Kentucky. This event was the removal of all policies except restrictions on public transportation.

After conducting inference testing on all eight states, the differences between the Synthetic Control Method and the Multi Synthetic Control Method were significant in 7 different cases, the exception being the state of Kentucky. This difference is calculated as

$$\delta_{i,t} = \frac{\hat{y}_{i,t} - \tilde{y}_{i,t}}{\hat{y}_{i,t}}$$

Zhere i refers to region i and $t \geqslant T_{0,K}$ refers to the treatment period for event k. Figure 4.1 shows these differences from T_0 to $T_0 + 90$. The blue bands around the lines are the 95% confidence levels. Any line that intersects with the blue 0 lines indicates that $\delta_{i,t}$ is 0. We can see that for Kentucky, there is no noticeable significant cumulative effect of previous policies. This is not to say that the last policy was insignificant, but that the difference between the Multi Synthetic Control Method and the Synthetic Control Method is negligible. When conducting the same type of inference on the remaining seven states (and seven events) I find that they are all significant at a 95% confidence level.

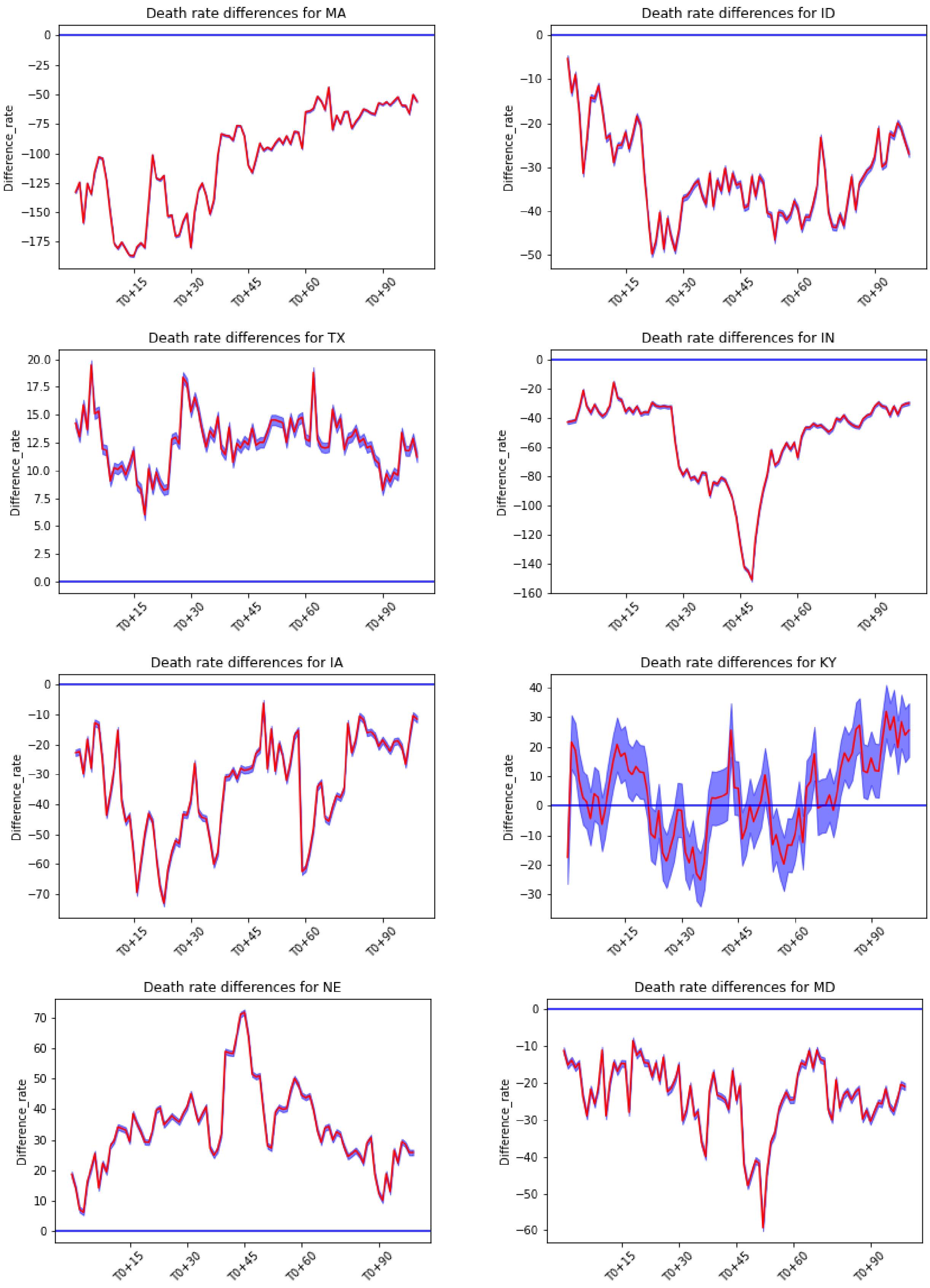


Figure 4.1: Deviations, $\delta_{i,t}$ 90 days after period $T_{0,K}$

Values for several averages consistent with the impact of the COVID-19 virus on humans in terms of the incubation period, and symptoms, including death, are recorded based on results obtained from earlier research. These means are displayed in table 4.5, and they represent the percentage difference between the two sets of deviations $\hat{\alpha}_{i,t}$, and $\tilde{\alpha}_{i,t}$. The most significant deviation was for the state of Massachusetts. On average, the deviations obtained from the multi-stage Synthetic Control Method are 150% stronger. In other words, the predicted values obtained from conducting a simple Synthetic Control Method for that treated period are, on average under-predicted, and in fact, the cumulative effect of all the policies is almost twice as strong as the effect predicted. I then study what could potentially explain these variations in these predicted values. For the panel of states shown, I find that the predicted values from the Multi Synthetic Control Method are between 10 and 150% larger than those predicted by the traditional Synthetic Control Method.

Means	IA	IN	ID	TX	MD	NE	MA	KY
15-30 mean	-55.7	-32.5	-33.8	9.8	-15.8	34.7	-131.5	3.3
30-45 mean	-41.9	-76.9	-37.3	14.0	-25.0	39.9	-156.4	5.2
45-60 mean	-24.1	-98.3	-37.5	13.3	-34.8	49.0	-92.9	-9.7
60-75 mean	-27.3	-42.5	-33.6	12.4	-22.2	29.8	-65.6	-0.2

Table 4.5: Average percentage deviation between Synthetic Control Method and Multi Synthetic Control Method for different states

The deviation peak was around the 40th day after the policy had been implemented for most events, with a slight dip after this day. It is worth noting that when computing various means, the 45 to 60-day average is usually higher in absolute value than the 15 to 30-day average or the 30 to 45-day average.

In order to explain this difference between the Synthetic Control Method value and the Multi Synthetic Control Method values of the treated variable, I conducted an OLS regression with my dependent variable being the 15-60 day average. This was chosen because it provides a well-rounded representation of this difference without being limited by data availability.

$$\bar{\delta}_i = \beta_i X_i + \gamma_i Z_i u_i$$

where $\bar{\delta}_i$ is the 15- 60 day average and X_i is a vector of explanatory variables. The vector includes the mean and variance of the stringency index as calculated by the Oxford COVID-19 Government Response Tracker indicators, the number of policies implemented, and number of policies removed for each state. Vector Z_i is a vector of controls that includes the density factor calculated earlier in chapter 3, percentage of the population over 65, and health quality. The results are in Table 4.6.

	M1	M2	M 3	M4	M5
Stringency Mean	10.21**	8.05 **	7.31**	7.65**	7.28**
	(2.18)	(2.43)	(2.45)	(2.31)	(2.53)
Stringency Variance	0.51	0.78	0.91	-0.21	1.63
	(0.31)	(0.34)	(0.45)	(0.62)	(1.21)
Policies imp.		1.95**	1.89**	1.86**	1.92**
		(0.41)	(0.60)	(0.58)	(0.55)
Policies rem.			-0.54 *	-0.51*	-0.65
			(-0.21)	(-0.22)	(-0.66)
Over 65 pop.				0.22 **	0.18**
				(0.07)	(0.06)
Density score					0.13*
					(0.08)
N	105	105	105	105	105
R-squared	0.3415	0.3489	0.3672	0.3724	0.3753
R-squared Adj.	0.3348	0.3341	0.3331	0.3323	0.3398

Table 4.6: OLS Regression results

Table 4.6 shows the results of the OLS regressions around the determinants of the differences between the Synthetic Control Method and Multi Synthetic Control Method. The mean of the stringency is significant in determining the percentage difference between the two models. Table 4.6 shows that for every point increase in the average stringency index, the Multi Synthetic Control Method deviates by 8 to 10% more than the Synthetic Control Method . As states imposed more restrictions, and as they increased the severity of the restrictions, the effect of these policies increased over time. The deviations between the Multi Synthetic Control Method and the Synthetic Control Method for the last policy increased by 2% for each policy imposed on average. The results show that the Synthetic Control Method under-predicts the true impact of a policy if a series of significant events preceded it.

4.5 Robustness testing

Two robustness tests are conducted to test the validity and accuracy of the results obtained. These tests are augmented versions of robustness tests conducted for traditional Synthetic Control Method regressions. The first test is the placebo test. As Abadie (2021) highlighted, when conducting a placebo test, a treatment period is arbitrarily chosen prior to the actual treatment period T_0. A synthetic model is constructed using the arbitrarily chosen period, and then the inference tests are conducted on the synthetic treated variables of the actually treated period.

In this chapter, this test is altered to accommodate the changes in the model. IA number of events are randomly picked, specifically, three random dates where the variable of interest had not been affected by any policies prior to that point for a significantly long period (either not at all or more than 60 days). For each event, simple Synthetic Control Method regression is conducted for the treated variable on that date, and the deviations from the actual value of the treated variable are recorded.

For the Multi Synthetic Control Method, I pick any random date prior to the event in question and any number of consecutive events. I choose to conduct this experiment with a 2-stage, 3-stage, and 4-stage Synthetic Control Method, with at least ten days between each date. Data obtained from the multi-stage regressions is compared to the standard model. If there are no significant differences between the Synthetic Control Method and the Multi Synthetic Control Method, then the multi-stage regression is the same as the Synthetic Control Method.

Three periods are chosen due to data limitations. For each event used, a minimum of 50 observations is needed to conduct inference, without any interruptions or policy changes. This does not mean that no policies present at the time but that no policies were changed for 50 days prior to the event was selected. The cumulative effects were unlikely to have any effect. The events chosen are for the following states on the following dates: a)South Dakota, 2020-07-31, b)Alabama, 2020-09-07, c)Nevada, 2021-02-21. The results of the multi-stage deviation regression for the three events chosen are compared to the one-period Synthetic Control Method specifically, the percentage change.

Regression used	Event 1	Event 2	Event 3
Synthetic Control Method	19.82%**	34.14% **	-70.52% *
2-stage Multi Synthetic Control Method	19.79% *	33.98% *	-70.31%*
3-stage Multi Synthetic Control Method	19.63% *	34.34%**	-71.28%*
4-stage Multi Synthetic Control Method	19.37% **	34.22%**	-70.48% *
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 4.7: Robustness check on random events

The results are summarized in Table 4.7, which displays the average 30-day percentage difference in death rates for each of the three events, in both the multi-stage and traditional Synthetic Control Method in their respective columns from the actual value. The signs represent the direction of change. A negative sign means that the predicted value is more than the actual value (implementation of a policy), and a positive sign indicates that the predicted value is less than the actual value (removal of a policy). When conducting inference using Synthetic Control Method, the events are significant at a 95% confidence level. From this table, I can see that there is no difference between these events, meaning that in a scenario where these cumulative effects are not present, the Multi Synthetic Control Method delivers predicted values similar to the Synthetic Control Method, meaning we can use the Multi Synthetic Control Method without any loss of accuracy.

Another robustness exercise involves changing the values of the treated variable in the vector Z when attempting to obtain counterfactuals. In a traditional Synthetic Control Method this is usually to only use half of the pre-treatment treated values to train the model. This chapter adapts this technique to the model defined above. For every event after the first event in the multi-stage regression, I only use half of the predicted values from the Synthetic Control Method. In other words, the vector Z_1 becomes

$$Z_{i,K} = [x_{i,1}, \dots, x_{i,m}, \tilde{Y}_{i,1,1}, \dots, \tilde{Y}_{i,\frac{T_{0,0}}{2}}, Y_{i,1,2}, \dots, Y_{i,\frac{T_{0,K}}{2},K}]$$

For the set of 7 significant final treatment periods, there is no noticeable change in the means highlighted in table 4.5 if the co-factors are changed.

4.6 Conclusion

Policymakers responded to the pandemic in the United States by imposing various anti-contagion policies at different times. These policies would often overlap, and this made analyzing their impact difficult. While a popular tool in its standard form, the Synthetic Control Method can only quantify the impact of policies that are uninterrupted after their occurrence by any other policy. This chapter presents a modified version of this method that relies on changing the co-factor vector and multiple iterations of the Synthetic Control Method to solve this issue.

The findings indicate that when a series of significant anti-contagion policies precede an anti-contagion policy, a synthetic control regression can underestimate the impact of the last policy in this series by up to 150% in absolute value. On average, each restriction imposed can add 2% to the total deviation between the Synthetic Control Method and the Multi Synthetic Control Method. The more stringent these policies the higher the deviation, almost 10% more for every point increase in stringency.

These findings are robust when tested against events that were not interrupted or were not preceded by significant events. Using key examples from the Synthetic Control Method literature, I find no significant difference between the Synthetic Control Method and the Multi Synthetic Control Method results. In these cases, the event studied was not preceded by any significant events. The Multi Synthetic Control Method predicted similar results to that of the Synthetic Control Method with several tests conducted on the average post-event differences. The Multi Synthetic Control Method results were within a 99% confidence interval of the Synthetic Control Method results. Future work can add further modifications to the model that can be undertaken. The model could be used to focus on more concentrated and specific treatments or for more extended treatment periods.

Appendix A

The results of robustness checks are in table 4.8. In this case, for every event, 4 random stages are used, and the 45 day average is taken as a benchmark. Increasing the number of stages, or taking a different average does not change the results.

Regression	ID	IN	IA	TX
SCM	19.82%**	34.14% **	-70.52% ***	10.71 ***
2-stage M-SCM	19.79% **	33.98% *	-70.31%***	13.22 ***
3-stage M-SCM	19.63% **	34.34%**	-71.28%***	8.43***
4-stage M-SCM	19.37% **	34.22%**	-70.48% ***	9.71 ***
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 4.8: Robusteness checks for selected significant events-positive values

Regression	MD	NE	MA	KY
SCM	-12.32***	33.41***	-122.33***	0.33***
2-stage M-SCM	-15.20***	42.22***	-146.92***	5.87***
3-stage M-SCM	-27.42*	28.47***	-125.87***	2.10***
4-stage M-SCM	-15.04***	29.04***	-161.55***	4.22***
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 4.9: Robusteness checks for selected significant events- negative values

Bibliography

Abadie, Alberto (2021). "Using synthetic controls: Feasibility, data requirements, and methodological aspects". In: *Journal of Economic Literature* 59.2, pp. 391–425.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010a). "Synthetic control methods for comparative case studies: Estimating the effect of Californias tobacco control program". In: *Journal of the American Statistical Association* 105.490, pp. 493–505.

— (2010b). "Synthetic control methods for comparative case studies: Estimating the effect of Californias tobacco control program". In: *Journal of the American statistical Association* 105.490, pp. 493–505.

— (2015). "Comparative politics and the synthetic control method". In: *American Journal of Political Science* 59.2, pp. 495–510.

Abbasi, Kamran (2020). *COVID-19: a public inquiry in hard times?*

Abel, Andrew B and Stavros Panageas (2021). "Social distancing, vaccination and the paradoxical optimality of an endemic equilibrium". In: *Vaccination and the Paradoxical Optimality of an Endemic Equilibrium (April 12, 2021).*

Abuin, Pablo et al. (2020). "Characterization of SARS-CoV-2 dynamics in the host". In: *Annual Reviews in Control* 50, pp. 457–468.

Acemoglu, Daron, Simon Johnson, and James A Robinson (2001). "The colonial origins of comparative development: An empirical investigation". In: *American economic review* 91.5, pp. 1369–1401.

Acemoglu, Daron et al. (2021). "Optimal targeted lockdowns in a multigroup SIR model". In: *American Economic Review: Insights* 3.4, pp. 487–502.

Acharya, Viral V, Robert F Engle III, and Sascha Steffen (2021). *Why did bank stocks crash during COVID-19?* Tech. rep. National Bureau of Economic Research.

Akao, KenIchi and Hiroaki Sakamoto (2018). "A theory of disasters and long-run growth". In: *Journal of Economic Dynamics and Control* 95, pp. 89–109.

Albi, Giacomo, Lorenzo Pareschi, and Mattia Zanella (2021). "Control with uncertain data of socially structured compartmental epidemic models". In: *Journal of Mathematical Biology* 82.7, pp. 1–41.

Alfano, Vincenzo, Salvatore Ercolano, and Lorenzo Cicatiello (2021). "School openings and the COVID-19 outbreak in Italy. A provincial-level analysis using the synthetic control method". In: *Health Policy* 125.9, pp. 1200–1207.

Alfaro, Laura et al. (2020). *Aggregate and Firm-Level Stock Returns During Pandemics, in Real Time*. Working Paper 26950. National Bureau of Economic Research. DOI: 10.3386/w26950. URL: http://www.nber.org/papers/w26950.

Alstadster, Annette et al. (2020). *The First Weeks of the Coronavirus Crisis: Who Got Hit, When and Why? Evidence from Norway*. Working Paper 27131. National Bureau of Economic Research. DOI: 10.3386/w27131. URL: http://www.nber.org/papers/w27131.

Alvarez, Fernando E, David Argente, and Francesco Lippi (2020). *A simple planning problem for covid-19 lockdown*. Tech. rep. National Bureau of Economic Research.

Apple (2021). *Mobility Trends Reports*. URL: https://covid19.apple.com/mobility.

Athey, Susan and Guido W Imbens (2017). "The state of applied econometrics: Causality and policy evaluation". In: *Journal of Economic Perspectives* 31.2, pp. 3–32.

Atkeson, Andrew (2020). *What Will Be the Economic Impact of COVID-19 in the US? Rough Estimates of Disease Scenarios*. Working Paper 26867. National Bureau of Economic Research. DOI: 10.3386/w26867. URL: http://www.nber.org/papers/w26867.

Atkeson, Andrew, Karen Kopecky, and Tao Zha (2020). *Estimating and Forecasting Disease Scenarios for COVID-19 with an SIR Model*. Working Paper 27335. National Bureau of Economic Research. DOI: 10.3386/w27335. URL: http://www.nber.org/papers/w27335.

Auerbach, Alan J et al. (2021). *Fiscal Multipliers in the COVID19 Recession*. Working Paper 29531. National Bureau of Economic Research. DOI: 10.3386/w29531. URL: http://www.nber.org/papers/w29531.

Baez, Javier, Alejandro De la Fuente, and Indhira Vanessa Santos (2010). "Do natural disasters affect human capital? An assessment based on existing empirical evidence". In: *IZA Discussion Paper*.

Bai, John et al. (2021). *Digital Resilience: How Work-From-Home Feasibility Affects Firm Performance*. Working Paper 28588. National Bureau of Economic Research. DOI: 10.3386/w28588. URL: http://www.nber.org/papers/w28588.

Baker, Scott R et al. (2020). *COVID-Induced Economic Uncertainty*. Working Paper 26983. National Bureau of Economic Research. DOI: 10.3386/w26983. URL: http://www.nber.org/papers/w26983.

Baldwin, Richard and Rebecca Freeman (2021). *Risks and global supply chains: What we know and what we need to know*. Working Paper 29444. National Bureau of Economic Research. DOI: 10.3386/w29444. URL: http://www.nber.org/papers/w29444.

Barone, Guglielmo and Sauro Mocetti (2014). "Natural disasters, growth and institutions: a tale of two earthquakes". In: *Journal of Urban Economics* 84, pp. 52–66.

Barro, Robert J (2001). "Human capital and growth". In: *American economic review* 91.2, pp. 12–17.

— (2006). "Rare disasters and asset markets in the twentieth century". In: *The Quarterly Journal of Economics* 121.3, pp. 823–866.

— (2009). "Rare Disasters, Asset Prices, and Welfare Costs". In: *American Economic Review* 99.1, pp. 243–64. DOI: 10.1257/aer.99.1.243. URL: http://www.aeaweb.org/articles?id=10.1257/aer.99.1.243.

Bartik, Alexander W et al. (2020). *Measuring the labor market at the onset of the COVID-19 crisis*. Working Paper 27613. National Bureau of Economic Research. DOI: 10.3386/w27613. URL: http://www.nber.org/papers/w27613.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein (2021). “The augmented synthetic control method”. In: *Journal of the American Statistical Association* 116.536, pp. 1789–1803.

Benson, Charlotte and Edward J Clay (2000). “Developing countries and the economic impacts of natural disasters”. In: *Managing disaster risk in emerging economies*, pp. 11–21.

Bouttell, Janet et al. (2018). “Synthetic control methodology as a tool for evaluating population-level health interventions”. In: *Journal of Epidemiol Community Health* 72.8, pp. 673–678.

Broek-Altenburg, Eline van den and Adam Atherly (2021). “Adherence to COVID-19 policy measures: Behavioral insights from The Netherlands and Belgium”. In: *PloS one* 16.5, e0250302.

Brueckner, Jan, Matthew E Kahn, and Gary C Lin (2021). *A New Spatial Hedonic Equilibrium in the Emerging Work-from-Home Economy?* Working Paper 28526. National Bureau of Economic Research. DOI: 10.3386/w28526. URL: http://www.nber.org/papers/w28526.

Casares, Miguel, Paul Gomme, and Hashmat Khan (2022). “COVID-19 pandemic and economic scenarios for Ontario”. In: *Canadian Journal of Economics/Revue canadienne d’économique* 55, pp. 503–539.

Cavallo, Eduardo, Ilan Noy, et al. (2011). “Natural disasters and the economya survey”. In: *International Review of Environmental and Resource Economics* 5.1, pp. 63–102.

Cavallo, Eduardo et al. (2013). “Catastrophic natural disasters and economic growth”. In: *Review of Economics and Statistics* 95.5, pp. 1549–1561.

CDC, Center for Disease Control (2021).

Chernozhukov, Victor, Hiroyuki Kasahara, and Paul Schrimpf (2021). “Causal impact of masks, policies, behavior on early covid-19 pandemic in the US”. In: *Journal of Econometrics* 220.1, pp. 23–62.

Chilarescu, Constantin (2008). “An analytical solutions for a model of endogenous growth”. In: *Economic Modelling* 25.6, pp. 1175–1182.

Chinazzi, Matteo et al. (2020). “The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak”. In: *Science* 368.6489, pp. 395–400.

Cho, Sang-Wook (2020). “Quantifying the impact of nonpharmaceutical interventions during the COVID-19 outbreak: The case of Sweden”. In: *The Econometrics Journal* 23.3, pp. 323–344.

Cucinotta, Domenico and Maurizio Vanelli (2020). "WHO declares COVID-19 a pandemic". In: *Acta Bio Medica: Atenei Parmensis* 91.1, p. 157.

Davis, Steven J, Dingqian Liu, and Xuguang Simon Sheng (2021). *Stock Prices and Economic Activity in the Time of Coronavirus*. Working Paper 28320. National Bureau of Economic Research. DOI: 10.3386/w28320. URL: http://www.nber.org/papers/w28320.

Deb, Pragyan et al. (2020). "The effect of containment measures on the COVID-19 pandemic". In: *CEPR Discussion Paper No. DP15086*.

Dergiades, Theologos, Costas Milas, and Theodore Panagiotidis (2020). "Effectiveness of government policies in response to the COVID-19 outbreak". In: *Available at SSRN* 3602004.

Donohue, John J, Abhay Aneja, and Kyle D Weber (2019). "Right-to-carry laws and violent crime: A comprehensive assessment using panel data and a state-level synthetic control analysis". In: *Journal of Empirical Legal Studies* 16.2, pp. 198–247.

Doudchenko, Nikolay and Guido W Imbens (2016). *Balancing, regression, difference-in-differences and synthetic control methods: A synthesis*. Tech. rep. National Bureau of Economic Research.

Drury, A Cooper and Richard Stuart Olson (1998). "Disasters and political unrest: An empirical investigation". In: *Journal of Contingencies and Crisis Management* 6.3, pp. 153–161.

Dube, Arindrajit (2021). *Aggregate employment effects of unemployment benefits during deep downturns: Evidence from the expiration of the Federal Pandemic Unemployment Compensation*. Tech. rep. National Bureau of Economic Research.

Eberly, Janice C, Jonathan Haskel, and Paul Mizen (2021). *Potential Capital, Working From Home, and Economic Resilience*. Working Paper 29431. National Bureau of Economic Research. DOI: 10.3386/w29431. URL: http://www.nber.org/papers/w29431.

Eichenbaum, Martin S, Sergio Rebelo, and Mathias Trabandt (2021). "The macroeconomics of epidemics". In: *The Review of Financial Studies* 34.11, pp. 5149–5187.

Eichenbaum, Martin S et al. (2020). *How do people respond to small probability events with large, negative consequences?* Tech. rep. National Bureau of Economic Research.

EMDAT, Center for Research on the Epidemiology of disasters (2020). *The International Disaster Database*. URL: https://www.emdat.be/.

Erosa, Andres, Tatyana Koreshkova, and Diego Restuccia (2010). "How important is human capital? A quantitative theory assessment of world income inequality". In: *The Review of Economic Studies* 77.4, pp. 1421–1449.

Famiglietti, Matthew and Fernando Leibovici (2021). "The Impact of Health and Economic Policies on the Spread of COVID-19 and Economic Activity". In: *FRB St. Louis Working Paper* 2021-005.

Fernández-Villaverde, Jesús and Charles I Jones (2020). *Macroeconomic outcomes and covid-19: A progress report*. Tech. rep. National Bureau of Economic Research.

Finamor, Lucas and Dana Scott (2021). "Labor market trends and unemployment insurance generosity during the pandemic". In: *Economics Letters* 199, p. 109722.

Fu, Zhengqing, Goulin Liu, and Lanlan Guo (2019). "Sequential quadratic programming method for nonlinear least squares estimation and its application". In: *Mathematical Problems in Engineering* 2019.

George, Bert et al. (2020). "A guide to benchmarking COVID-19 performance data". In: *Public Administration Review* 80.4, pp. 696–700.

Gevertz, Jana L et al. (2021). "A novel COVID-19 epidemiological model with explicit susceptible and asymptomatic isolation compartments reveals unexpected consequences of timing social distancing". In: *Journal of theoretical biology* 510, p. 110539.

Google (2021). *Community Mobility Reports*. Reports created 2021-08-14. URL: `https://www.google.com/covid19/mobility/`.

Gourio, Francois (2012). "Disaster risk and business cycles". In: *American Economic Review* 102.6, pp. 2734–66.

Hale, Thomas et al. (2021a). "A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker)". In: *Nature Human Behaviour* 5.4, pp. 529–538.

— (2021b). "A global panel database of pandemic policies (Oxford COVID19 Government Response Tracker)". In: *Nature Human Behaviour* 5.4, pp. 529–538.

Hall, Robert E and Charles I Jones (1999). "Why do some countries produce so much more output per worker than others?" In: *The quarterly journal of economics* 114.1, pp. 83–116.

Hallas, L, A Hatibie, R Koch, et al. (2021a). *Variation in US states COVID-19 policy responses.*

— (2021b). *Variation in US states COVID-19 policy responses.*

Harris, Timothy F, Aaron Yelowitz, and Charles Courtemanche (2021). "Did COVID-19 change life insurance offerings?" In: *Journal of Risk and Insurance* 88.4, pp. 831–861.

Heersink, Boris, Brenton D Peterson, and Jeffery A Jenkins (2017). "Disasters and elections: Estimating the net effect of damage and relief in historical perspective". In: *Political Analysis* 25.2, pp. 260–268.

Heger, Martin, Alex Julca, and Oliver Paddison (2008). *Analysing the impact of natural hazards in small economies: the Caribbean case.* 2008/25. WIDER Research Paper.

Holzer, Harry J, R Glenn Hubbard, and Michael R Strain (2021). *Did Pandemic Unemployment Benefits Reduce Employment? Evidence from Early State-Level Expirations in June 2021.* Tech. rep. National Bureau of Economic Research.

Hsiang, Solomon et al. (2020). "The effect of large-scale anti-contagion policies on the COVID-19 pandemic". In: *Nature* 584.7820, pp. 262–267.

Hsiang, Solomon M and Amir S Jina (2014). *The causal effect of environmental catastrophe on long-run economic growth: Evidence from 6,700 cyclones.* Tech. rep. National Bureau of Economic Research.

Huber, Martin and Henrika Langen (2020). "Timing matters: the impact of response measures on COVID-19-related hospitalization and death rates in Germany and Switzerland". In: *Swiss Journal of Economics and Statistics* 156.1, pp. 1–19.

Ikefuji, Masako and Ryo Horii (2012). "Natural disasters in a two-sector model of endogenous growth". In: *Journal of Public Economics* 96.9-10, pp. 784–796.

Jamison, Julian C et al. (2021). "Comparing the impact on COVID-19 mortality of self-imposed behavior change and of government regulations across 13 countries". In: *Health services research* 56.5, pp. 874–884.

Jaramillo, Christian R (2009). "Do natural disasters have long-term effects on growth?" In: *Documento CEDE* 2009-24.

Johns Hopkins University, JHU (2021).

Jones, Callum J, Thomas Philippon, and Venky Venkateswaran (2020a). *Optimal mitigation policies in a pandemic: Social distancing and working from home*. Tech. rep. National Bureau of Economic Research.

— (2020b). *Optimal mitigation policies in a pandemic: Social distancing and working from home*. Tech. rep. National Bureau of Economic Research.

Jong-A-Pin, Richard (2009). "On the measurement of political instability and its impact on economic growth". In: *European Journal of Political Economy* 25.1, pp. 15–29.

Kajitani, Yoshio and Hirokazu Tatano (2018). "Applicability of a spatial computable general equilibrium model to assess the short-term economic impact of natural disasters". In: *Economic Systems Research* 30.3, pp. 289–312.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi (2010). "The worldwide governance indicators: Methodology and analytical issues". In: *World Bank policy research working paper* 5430.

Kermack, William Ogilvy and Anderson G McKendrick (1927). "A contribution to the mathematical theory of epidemics". In: *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115.772, pp. 700–721.

Kleven, Henrik Jacobsen, Camille Landais, and Emmanuel Saez (2013). "Taxation and international migration of superstars: Evidence from the European football market". In: *American economic review* 103.5, pp. 1892–1924.

Kreif, Noémi et al. (2016). "Examination of the synthetic control method for evaluating health policies with multiple treated units". In: *Health economics* 25.12, pp. 1514–1528.

Li, Kathleen and Venkatesh Shankar (2020a). "Estimating the Causal Effect of A Digitally Native Retailer Opening a New Store: A New Two-Step Synthetic Control Method". In: *Available at SSRN 3628589*.

— (2020b). "Estimating the Causal Effect of A Digitally Native Retailer Opening a New Store: A New Two-Step Synthetic Control Method". In: *Available at SSRN 3628589*.

Lu, Liang et al. (2021). *Demand Shocks and Supply Chain Resilience: An Agent Based Modelling Approach and Application to the Potato Supply Chain*. Tech. rep. National Bureau of Economic Research.

Lucas Jr, Robert E (1988). "On the mechanics of economic development". In: *Journal of monetary economics* 22.1, pp. 3–42.

Ludvigson, Sydney C, Sai Ma, and Serena Ng (2020). *COVID-19 and the macroeconomic effects of costly disasters*. Tech. rep. National Bureau of Economic Research.

Marinescu, Ioana, Daphne Skandalis, and Daniel Zhao (2021). "The impact of the federal pandemic unemployment compensation on job search and vacancy creation". In: *Journal of Public Economics* 200, p. 104471.

Matta, Samer, Michael Bleaney, and Simon Appleton (2022). "The economic impact of political instability and mass civil protest". In: *Economics & Politics* 34.1, pp. 253–270.

McClelland, Robert and Sarah Gault (2017). "The synthetic control method as a tool to understand state policy". In: *Washington, DC: The Urban Institute*.

Mills, Melinda C and Tobias Rüttenauer (2022). "The effect of mandatory COVID-19 certificates on vaccine uptake: synthetic-control modelling of six countries". In: *The Lancet Public Health* 7.1, e15–e22.

Montreal Gazette, MG (2021). URL: `https://montrealgazette.com/business/local-business/none-of-this-makes-sense-business-owners-push-back-against-health-restrictions`.

Mueller, Amber L, Maeve S McNamara, and David A Sinclair (2020). "Why does COVID-19 disproportionately affect older people?" In: *Aging (albany NY)* 12.10, p. 9959.

Mulligan, Casey B and Xavier Sala-i Martin (1992). *Transitional Dynamics in Two-Sector Models of Endogenous Growth*. Working Paper 3986. National Bureau of Economic Research. DOI: `10.3386/w3986`. URL: `http://www.nber.org/papers/w3986`.

ONeill, Stephen et al. (2016). "Estimating causal effects: considering three alternatives to difference-in-differences estimation". In: *Health Services and Outcomes Research Methodology* 16.1, pp. 1–21.

Parker, Jonathan A et al. (2022). *Household Spending Responses to the Economic Impact Payments of 2020: Evidence from the Consumer Expenditure Survey*. Tech. rep. National Bureau of Economic Research.

Pindyck, Robert S (2020). *COVID-19 and the welfare effects of reducing contagion*. Tech. rep. National Bureau of Economic Research.

Pindyck, Robert S and Neng Wang (2013). "The economic and policy consequences of catastrophes". In: *American Economic Journal: Economic Policy* 5.4, pp. 306–39.

Public Broadcast Service, PBS (2021). URL: `https://www.pbs.org/newshour/nation/u-s-governors-push-back-against-mandates-closures-amid-omicron-surge`.

Raddatz, Claudio (2007). "Are external shocks responsible for the instability of output in low-income countries?" In: *Journal of Development Economics* 84.1, pp. 155–187.

Ramelli, Stefano and Alexander F Wagner (2020). "Feverish stock price reactions to COVID-19". In: *The Review of Corporate Finance Studies* 9.3, pp. 622–655.

Rancourt, Denis G (2020). "Masks Dont Work A review of science relevant to COVID-19 social policy". In: *American Journal of Infection Control* 37.5, pp. 417–419.

Rojas, Felipe Lozano et al. (2020). *Is the cure worse than the problem itself? Immediate labor market effects of COVID-19 case rates and school closures in the US*. Tech. rep. National Bureau of Economic Research.

Ruiter, Marleen C de et al. (2020). "Why we can no longer ignore consecutive disasters". In: *Earth's future* 8.3, e2019EF001425.

Sadeghi, Mahdiar, James M Greene, and Eduardo D Sontag (2021). "Universal features of epidemic models under social distancing guidelines". In: *Annual reviews in control* 51, pp. 426–440.

Skidmore, Mark and Hideki Toya (2002). "Do natural disasters promote long-run growth?" In: *Economic inquiry* 40.4, pp. 664–687.

Stock, James H (2020). *Data gaps and the policy response to the novel coronavirus*. Tech. rep. National Bureau of Economic Research.

UNDRR, United Nations Office for Disaster Relief (2017). *THIS IS NOT A DRILL: HOW 1985 DISASTER TAUGHT MEXICO TO PREPARE FOR EARTHQUAKES*. URL: `https://www.`

preventionweb.net/news/not-drill-how-1985-disaster-taught-mexico-prepare-earthquakes.

Uzawa, Hirofumi (1965). "Optimum technical change in an aggregative model of economic growth". In: *International economic review* 6.1, pp. 18–31.

Vogler, Justin (2010). *Chile: politics of an earthquake*. Accessed on 2020-12-13. URL: https://www.opendemocracy.net/en/chile-politics-of-earthquake-0/.

Wedel, Michel and Rik Pieters (2017). "A review of eye-tracking research in marketing". In: *Review of marketing research*, pp. 123–147.

Wise, Jacqui (2021). *Covid-19: The E484K mutation and the risks it poses*.

Wong, Yik Chun et al. (2020). "Natural transmission of bat-like SARS-CoV-2 PRRA variants in COVID-19 patients". In: *Clinical Infectious Diseases*.

World Health Organization, WHO (2020). *Long term effects of COVID-19*. accessed on August 2021. URL: https://www.who.int/docs/default-source/coronaviruse/risk-comms-updates/update-36-long-term-symptoms.pdf?sfvrsn=5d3789a6_2.