Thesis Project

Historical Persistence: Examining the Effects of the 1918 Influenza Pandemic on COVID-19 Vaccination Patterns

Juan José Ramírez Cruz
ANR 616980
SNR 2106296

BSc Thesis for Economics
Tilburg School of Economics and Management
Tilburg University
Supervisor: Antonio De Araujo Laplana
Number of words: 7.952
June 12th, 2024

Contents

1 Introduction

Historical persistence is an important factor to understand how our institutions, former policies, and historical events shaped our societal constrains nowadays; our present-day attitudes; and how our economic interactions will be shaped in the long run. The identification of persistence in interdisciplinary sectors can have important implications for policy-making in future strategies, making these decisions more effective and targeted at our objectives by researching how former approaches shaped contemporary responses.

The motivation for this study is the impact of former historical events and policies such as the 1918 Great Influenza, and the interventions to mitigate it, on present-day decisions concerning public health issues, specifically COVID-19 vaccination rates. By interpreting how these historical events shaped current behaviours and institutional frameworks, it enhances the opportunity to design efficient crisis mitigation plans which can address disparities across demographics and geography, and it is an alternative point of view for policy-makers and researchers to understand inherited human behaviour and societal resilience for future research and more accurate governmental interventions.

Therefore, the research question regarding this topic is *"How mortality and former governmental interventions during the 1918 influenza pandemic persisted and influenced COVID-19 vaccination rates in cities across the United States?"* This question has the aim to examine if the 1918 Influenza Pandemic and its measures to mitigate it have a persistent effect and an impact on COVID-19 vaccination patterns.

Most of the literature regarding the 1918 Great Influenza is focused on the effects of the latter on its economic consequences such as economic activity (Barro et al., 2020) and innovation (Berkes et al., 2023), while few of them emphasise its impact on social behaviour and its development in the long-run (Aassve et al., 2021; Aksoy et al., 2020). This paper would enhance the knowledge on how societal attitudes on public health concerns are shaped due to former policies, differ from one region to another, and it crosses generations.

The COVID-19 pandemic generated a buzz to study the long-lasting consequences of former pandemics in society's economic and social attitudes. Aassve et al. (2021) explores the effects of the 1918 Great Influenza on lower social trust for the survivors descendants while Guiso et al. (2016) shows how historical events, such as the self-governance of Italian cities in the Middle Ages, can be a factor for long-term persistence in development and economic growth. These are some of reviewed papers that evidence the repercussions of historical events and why its importance for further research development.

In order to develop this study, three main variables were used with the motivation to find long-term policy persistence for COVID-19 vaccination rates as Guiso et al. (2016) and Tabellini (2008) found developmental and cultural persistence affecting civic capital and economic development respectively: Firstly, COVID-19 vaccination rates per county in the United States will be used in order to identify where people have been vaccinated the most and it has been determined as the outcome variable of this research. This data is retrieved from the United States Centers for Disease Control and Prevention (CDC). Secondly, a dataset of the weekly total deaths caused by the 1918 Great Influenza. This data, by Markel et al. (2007), reports weekly the total deaths in 43 cities across the United States during the pandemic's second wave, between 31st August 1918 until 15th March 1919, and further expanded to 46 cities by Correia et al. (2022), which is the dataset used for this paper. These records have been used many times on this pandemic research by other authors such as Barro et al. (2020) and Berkes et al. (2023). Finally, the last dataset used is Non-Pharmaceutical Interventions (NPIs) enacted during the pandemic. NPIs refer to measures implemented to mitigate the spread of infectious diseases when pharmaceutical options (e.g. vaccines and antibiotics) are unavailable, such as mask-wearing mandates, quarantines and public event bans. This data collection was primarily retrieved from Markel et al. (2007) as the paper compacts all the NPI measures taken in 43 cities across the United States during the 1918 influenza pandemic, and further expanded by Berkes et al. (2023) and Correia et al. (2022) to 53 cities which is the dataset used for this research. In comparison with Correia's dataset, which complies NPI actions as categorical data, this paper quantifies these NPIs as weekly indices similar to the ones used to measure governmental actions during the COVID-19 pandemic (Hale et al., 2020).

With these datasets, I conducted different panel data regression analyses in order to measure the impact of the 1918 events and interventions on COVID-19 vaccination rates. The results show that higher death cases increase the percentage of COVID-19 vaccination rates; the NPIs from 1918 have a negative effect on vaccination rates in line with the literature (Troesken, 2019); and the interaction between death cases and

NPIs have a positive effect on vaccination rates i.e., a higher number of death cases and stricter/more NPIs increase the percentage of vaccinated population.

The paper is divided into eight sections: Section 2 shows relevant literature on the topic to sustain this paper's expectations. Section 3 comprises a framework with important topics which will be used for the remainder of the paper such as the 1918 Great Influenza Pandemic and Non-Pharmaceutical Interventions to understand further definitions and the development of this study. Section 4 comprises data description where the data gathering and cleanup is further explained as well as the summary statistics from this paper. Section 5 discusses the empirical methods used for the methodology. Section 6 reflects the empirical results of this study with their robustness checks. Section 7 includes the limitations and further policy implications of this research. Finally, Section 8 concludes with the main results.

2 Literature Review

As explained by Acemoglu et al. (2005), institutions determine the constraints on a society and how their economic interactions will be shaped in the long run; it also shapes social relationships and thus, trust (Algan & Cahuc, 2010); culture (Tabellini, 2008, 2010); financial development (Guiso et al., 2004); economic performance (Acemoglu et al., 2001); as well as a wide range of other societal factors.

Historical Persistence is a consequence of it and indicates how historical events, policies, and former institutions have influenced current attitudes and societal transactions even when this institutions do not exist anymore. These types of persistence can be seen in papers such as Voigtländer & Voth (2012) which reflects on the persistence of medieval anti-Semitic attitudes during the 14th Century and higher violent behaviour against Jews in Nazi Germany; the current higher levels of civic capital in Italian northern cities which were independent during the Middle Ages presented by Guiso et al. (2016); and lower levels of trust in African regions affected by slave trade (Nunn & Wantchekon, 2011). This thesis mainly contributes to existing literature by opening a new area, public health issues, in which historical persistence also plays an important role in order to explain social attitudes regarding vaccination rates.

Historical events also raised the interest in current literature on the impact of pandemics, i.e., COVID-19 and the 1918 Influenza and its consequences for economic, social and human capital for future generations (Arthi & Parman, 2021; Bridgman & Greenaway-McGrevy, 2023). The COVID-19 pandemic already has persistent effects economically (Barrett et al., 2021) and psychologically (Heitzman, 2020); however, it is relatively a recent historical event. Therefore, it is only possible to hypothesise the persistent effects of COVID based on past pandemics (Furceri et al., 2022; Emmerling et al., 2021). Further research in the medium and long-run can observe more accurately the COVID-19 historical persistence.

One of the main contributing papers for this research is *'Non-Pharmaceutical Interventions Implemented by U.S. Cities During the 1918-1919 Influenza Pandemic'* by Markel et al. (2007). This paper shows how the use of Non-Pharmaceutical Interventions mitigate the consequences and the excess death rates of the 1918 Great Influenza across the United States and its results on lower mortality rates in cities that implemented earlier NPIs during the 1918 pandemic even if these interventions were short-lived. Combinations of NPIs such as school closures and restrictions on public gatherings had the most significant effect on weekly excess deaths, also since they were the most implemented measures of the affected cities. Also, its database is a clear standpoint for this paper and for future studies which analyse the effects of the Great Influenza in multiple research areas. Berkes et al. (2023) estimate the effects of NPIs -mostly related to lockdowns and workplace closures- on innovation during 1918 extending Markel's database to 50 cities and indicating no decline on patenting rates for cities with NPI policies, higher aggregate productivity and possible positive long-run economic persistence.

The study of NPIs with Markel's contribution was further developed by Correia et al. (2022) to study the effects of those measures on economic activity and its impact on mortality. As also estimated by Markel et al. (2007), it states that the NPIs were mostly successful if they were quickly enacted. Furthermore, they find that information availability to policymakers (which was constrained and related to geographical factors) was an important driver of NPIs as well as local preferences. The trade-off between economic activity and reducing disease transmission depended on the groups affected by the interventions, but NPIs mainly mitigate the most disruptive shocks created by the pandemic such as labour supply shortages and illness, so the main source of disruption for these groups and economic activity were not the NPIs but the pandemic itself.

The 1918 Influenza happened at a moment of global turmoil and affected millions of people which endure aftermath consequences. Barro et al. (2020) present the macroeconomic impact caused by the Pandemic in the U.S. while Beach et al. (2022), even if it takes also into account the economic effects of the Influenza, they mainly show its effects regarding global and American health effects; short and medium-run economic effects; fertility; migration; and political economy. It is important for persistence studies, and as it was shown in these two papers, to be cautious of the difficulty to differentiate the effects caused by the Influenza and likely confounding effects for these historical studies, such as World War I, wartime production, air pollution, and female's labour participation.

For the measure of Non-Pharmaceutical Interventions in 1918, the *Oxford COVID-19 Government Response Tracker* (OxCGRT) by Hale et al. (2020) is essential to measure how strict were the NPIs enacted by each city that enacted NPIs in the U.S in 1918. This project records and represents in various indices how strict were the governmental policies during COVID regarding school and workplace closures; quarantines; cancellation of public events and public gatherings; restrictions on internal movement; public transportation restrictions, international controls; testing and contact policies; informational campaigns; and mask mandates.

The interventions used to reduce COVID-19 infections did barely change in comparison with those used during 1918: the use of this measures, e.g. quarantines, have been used since the 14th Century in order to control the spread of other diseases (Beach et al., 2022) while hand washing and mask wearing were "new" measures that were understood to decrease tuberculosis transmission before the antibiotic revolution in the mid-twentieth century (Tomes, 1999, 2000). Therefore, this metrics were clear and possibly comparable in order to compact each city NPIs into an Stringency and Containment Index during the 1918 pandemic.

Some other papers such as Correia et al. (2022) have also used NPI Intensity to measure its relationship of peak excess mortality, but this paper has the unique approach of using these indices as a way to make comparative applications between past and present measures to see if the effect has worked before and makes sense to enact it in future situations while also evaluating the effect that these NPIs have on pandemic mortality.

3 Institutional Framework

3.1 The 1918 Great Influenza Pandemic

The 1918 Influenza Pandemic was the deadliest pandemic in humanity and the last global event with similar shocks to the COVID-19 pandemic. It began in the early 1918, probably in Kansas, China or France (Beach et al., 2022), until 1920. It is possible that the military had a role spreading the virus across the United States, from their crowded military bases and their return home from World War I in Europe (Beach et al., 2022; Taubenberger & Morens, 2006).

The pandemic comprises three main infectious waves. The second wave, from September 1918 until November 1918, was the pandemic's peak and the deadliest one (Beach et al., 2022). In the United States alone, where this research is focus on, 0.66 percent of the population or 675.000 people perished (Correia et al., 2022); a significant number but far from other affected countries such as India, were 5.2 percent of its population or 16.7 million people died (Barro et al., 2020). Nevertheless, there is no precise number of how many people were infected or died because of the disease worldwide, but estimates predict that between 20 and 100 million people perished to the "Spanish Flu" (Johnson & Mueller, 2002; Beach et al., 2022; Taubenberger & Morens, 2006; Barro et al., 2020). The deviations between the literature is mainly caused due to underreporting in densely populated areas, World War I casualties, political censorship due to the war, etc. (Beach et al., 2022). Therefore, it is impossible to calculate the real magnitude of this historical event.

Figure 1 comprises the weekly deaths during the pandemic's second wave. The second wave killed on average 10.000 people per week in the United States, peaking at 28.000 people on the week of October 26th, 1918 (Markel et al., 2007) as it is highlighted in Figure 1. During the pandemic, not all regions of the country were affected the same. Figure 2 indicates the deaths per week by each region of the country. It shows that the Northeast of the United States was the most affected region during the entire pandemic. This region was the first to be affected by the second wave during the last week of August near Boston, Massachusetts; started to spread quickly to the neighbouring states of New York and Pennsylvania, and reached most to the Western region by mid-September (Beach et al., 2022). Influenza death cases increased so quickly that,

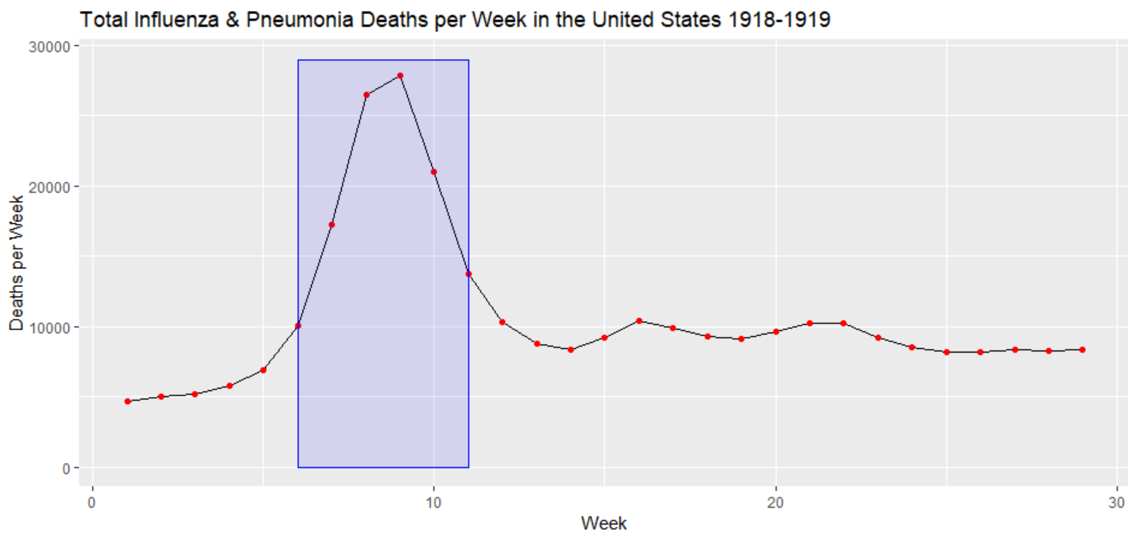

Figure 1: Total Influenza & Pneumonia Deaths per Week during the second Influenza wave in the United states. 1918-1919. The first week with recorded deaths cases is August 31st 1918 until March 15th, 1919. The highlighted region comprises the wave's peak, from October 5th, 1918 to November 9th, 1918. Weekly data retrieved from Markel et al. (2007).

by August 31st, 1918, most the largest counties in the country reported death cases related to pneumonia or influenza, marking the beginning of the second wave (see Figure 3).

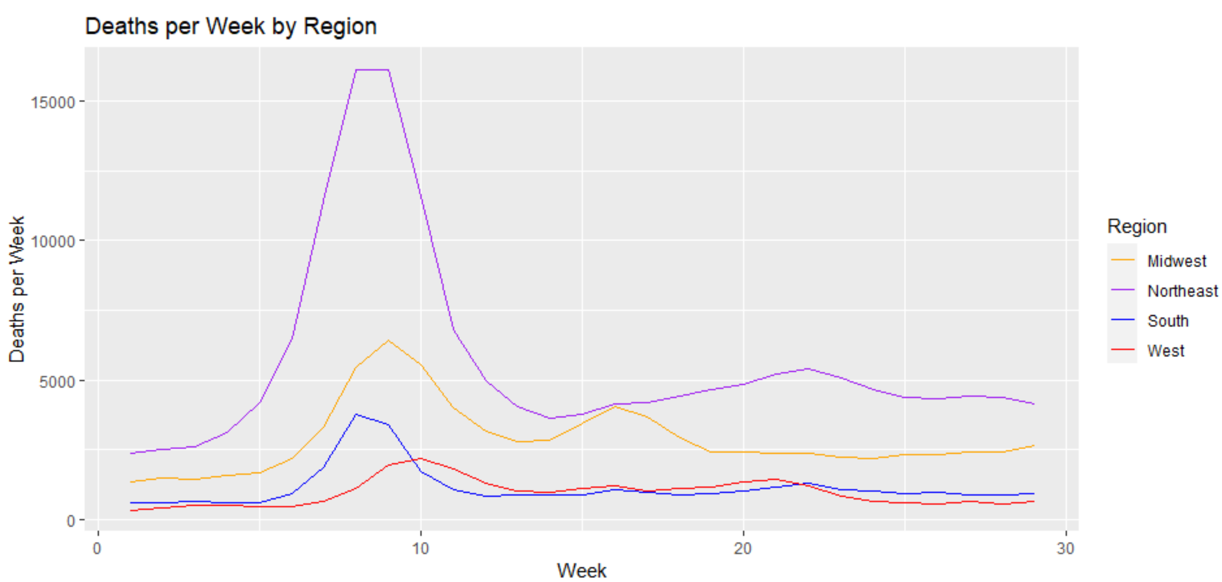

Figure 2: Deaths per Week by Region 1918-1919. The first week with recorded deaths cases is August 31st 1918 until March 15th, 1919. Weekly data retrieved from Markel et al. (2007).

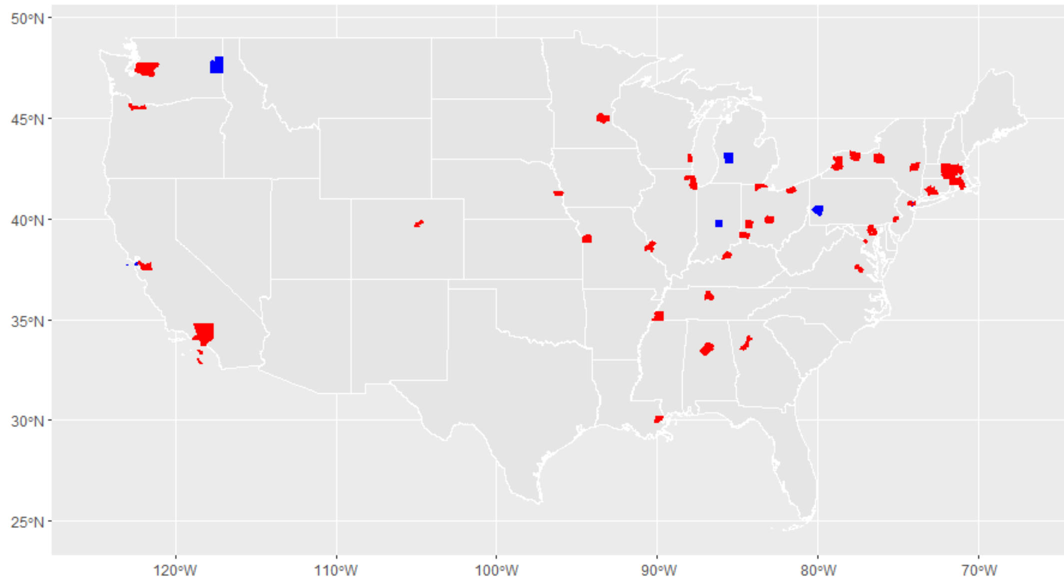

Figure 3: Counties with confirmed Influenza death cases in 1918. Counties in red had its first death cases during the first week of the pandemic's second wave (August 31st, 1918). Counties in blue had its first death cases the week after (September 7th, 1918). Data retrieved from Markel et al. (2007).

3.2 Non-Pharmaceutical Interventions (NPIs)

A pandemic certainty shocks different sectors in society and it must be mitigated as fast as possible in order to minimise the number of infections and, most importantly, deaths. Therefore, when pharmaceutical assets i.e., vaccines or antibiotics are not available, non-pharmaceutical interventions or NPIs play an important role on the pandemic's earliest stages (Ferguson et al., 2005). These interventions, in comparison to the ones issued by the pharmaceutical sector, can be applied easily for the general population, such as mask wearing and individual isolation.

Each of these NPIs have different effects on influenza containment and infection rates (Ferguson et al., 2006). From this measures, we can divide the types of NPIs mainly into two types: Containment and Stringency measures. This division is also used by Hale et al. (2020) to calculate government responses and mitigate new COVID-19 cases and each come with different practices. Containment measures are those which primarily focus on preventing outbreaks and the spread of the disease, this includes interventions such as quarantines; school and workplace closures; informational campaigns; social event bans; testing; and contact tracing. Stringency measures, on the other hand, not only takes in consideration the prevention of outbreaks, but also the severity of these measures. This type of response reflects how much economic activities and daily life is affected.

Because of this, stringency also includes all the containment measures (excluding testing and contact tracing), but it represents a broader spectrum to calculate the severity of NPIs. Following the indices created by Hale et al. (2020), and the NPI measures gathered by Markel et al. (2007), it is possible to create a Containment and Stringency Index during the 1918 influenza. As shown in Figure 4, NPI measures began by mid-September, two weeks after the first increase of death cases during the pandemic's second wave. Regions with stricter NPI measures had lower mortality rates and were able to reduce these measures quicker than other regions (Correia et al., 2022). However, since the disease travelled from East to West, it is also possible that more distant regions had more time to prepare, as the information about the virus was clearer for both policymakers and the general population (Correia et al., 2022; Beach et al., 2022).

These approaches in the United States were done mostly in a local-level without any coordinated federal

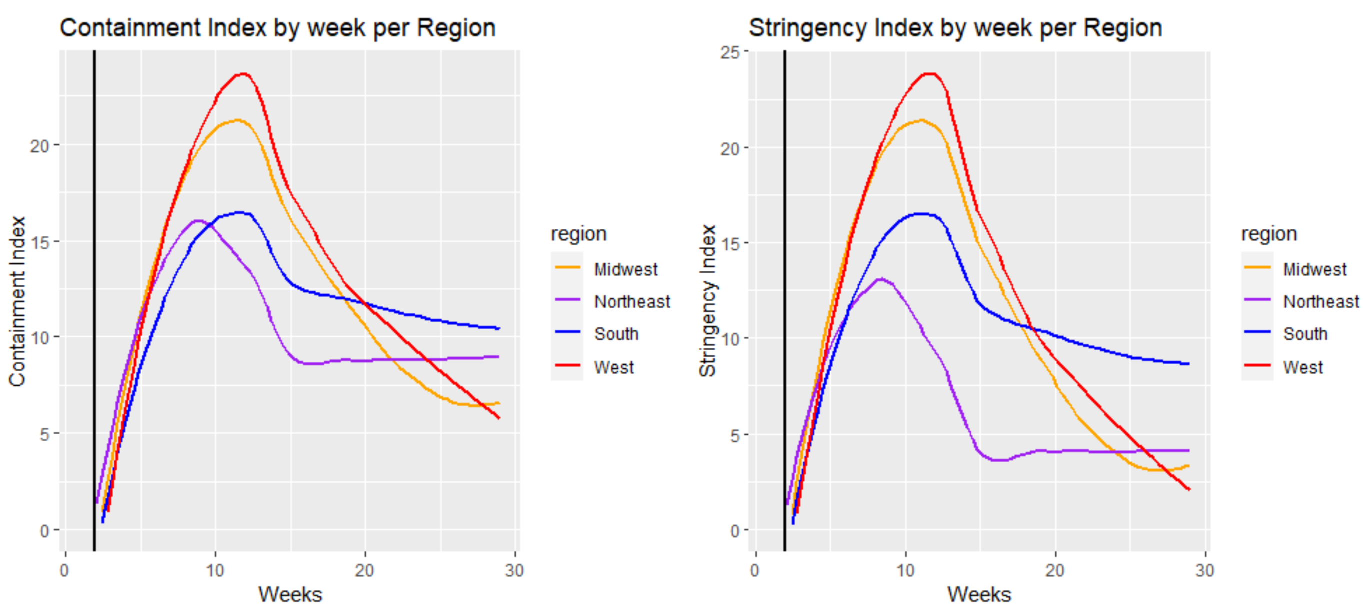

Figure 4: Containment & Stringency Index per Region in the US. 1918-1919. NPIs retrieved by Markel et al. (2007) and Government Response Tracker retrieved by Hale et al. (2020).

framework (Correia et al., 2022). These limited government and federalist approach have mixed effects on coordinated decision-making regarding public health issues and its trade-off with high economic growth: Troesken (2019) states that, due to high property rights protection; institutions promoting economic liberty;

and the decentralisation of high-priority policies from the federal government, the country benefited communities to develop water network infrastructure at low interest rates; however, the country also had higher levels of waterborne diseases and smallpox, this latter due to the unwillingness for universal vaccination schemes.

During the 1918 Pandemic, multiple cities in the U.S. enacted a variety of NPIs to lessen the infection and mortality rate across the country with different levels of compliance from different societal groups. As shown by Markel et al. (2007), most of the 34 out of 43 cities used NPI combinations from this category, school closures and public gathering bans. These measures had a mixed response among the population: groups in San Francisco protested against city mask mandates, businesses and unions lobbied against workplaces closures in some cities, and the clergy opposed or defied church closures and mass cancellations (Correia et al., 2022).

After the second wave by end-November (week 14), most of the interventions and stricter measures were lifted gradually in the affected cities (see Figure 4). Since the third wave was relatively mild in comparison with its predecessor, no new NPIs were implemented and the pandemic was considered over by the mid-1920 (Beach et al., 2022).

4 Data Description

4.1 Data Gathering

For this research, I employed three main datasets to include all the variables which are needed to comprehend the effect of persistence. As a foundation for this analysis, the datasets provided by Markel et al. (2007) and further updated by Correia et al. (2022), which track historical weekly death cases and enacted NPIs across cities in the United States during the 1918 Influenza pandemic, are a standpoint to assess the effectiveness of this measures amid the pandemic, as well as to observe any kind of policy or historical persistence on COVID-19 vaccination rates. Additionally, I also include data from the Centers for Disease Control and Prevention (2023), which details information on county-level COVID-19 vaccination rates for those counties where every city of this research is located.

At the same time, the Containment and Stringency Indices were included after analysing the NPI categorical data from Markel et al. (2007) and Correia et al. (2022), and quantifying it based on the index by Hale et al. (2020) used to measure the metrics enacted during the COVID-19 pandemic and how they have evolved over time. This measure creates a score between 0 and 100, being 100 the strictest governmental response. With the guidebook provided to create these indices, all historical NPIs were transformed into quantitative data, assessed, and divided into the established categories in order to calculate their scores for each indicator and the subsequent creation of the Containment and Stringency Indices. The next subsection will provide an insight of the NPIs used for each index thoroughly and which categories compose them. Finally, the 1918 city population was retrieved from the Integrated Public Use Microdata Series of the United States (IPUMS USA) and also confirmed with the city population retrieved by Markel et al. (2007).

4.2 Data Cleanup

The cleanup process for the weekly death cases and NPIs were quite similar since it was retrieved from the same two authors: Markel et al. (2007) and Correia et al. (2022). The main process was to group 29 weeks from August 31st, 1918 to March 15th, 1919 into a single week column on a panel dataset which comprise the weekly deaths from each city from which I had data. Additionally, the Federal Information Processing Standard (FIPS) county code, a unique code that defines each county in the United States, was included to identify the county's location of each city and make it consistent with modern data, as with CDC's dataset on vaccination rates.

The division of each NPI category was manually evaluated each week for all the cities provided by Correia et al. (2022) and then added to the panel datasets for the regression which will be using those indices. Out of 53 cities, 46 of them have records of providing any type of governmental intervention. Additionally, these cities were also divided per Regions as defined by the Census Bureau[1]: 16 of them are from the Midwest; 20 are from the Northeast; 9 from the South; and 8 are from the West of the United States. This was done as some cities were outliers and underestimated the coefficients of the entire dataset. Therefore, the division between geographical regions can enhance this research by getting more precise estimates by region.

[1]U.S. Census Bureau (2010). *Census Regions and Divisions with State FIPS Codes.* U.S. Department of Commerce, Economics and Statistics Administration.

The NPIs used for the measurement of the indices are according to the same metrics used by Hale et al. (2020) which are called “Containment and Closure Metrics” and “Health System Metrics”. The “Containment and Closure Metrics” measures are: School closure (1); Workplace closure (2); Public event cancellations (3); Restrictions on Gatherings (4); Public transportation restrictions (5); Stay at home requirements (6); Restrictions on internal movement (7); and Restrictions on international Travel (8). The “Health System Metrics” measures are: Public information campaigns (9); Testing policy (10); and Contract tracing (11). While the Containment Index takes into account all the already named measures (1-11), the Stringency Index does not take in consideration Restrictions on international Travel; Testing policy; and Contract tracing (1-7, 9).

Finally, the cleanup process for the COVID-19 vaccination rates by county from the Centers for Disease Control and Prevention (2023) was to keep variables such as the percentage of people who completed their vaccination scheme and a subset of 65+ population that also completed their vaccination scheme. Since the data was last updated in 2023, those who wanted to complete their vaccination programme already did it and those who did not want to get vaccinated still are not; therefore, using this variable can prevent any over-, underestimation of the coefficients in comparison with other determinants such as first vaccine takeover without completing it or booster shots.

4.3 Summary Statistics

For the summary statistics, three tables are provided, all of them with columns that reports the minimum value, the maximum value, the mean, and the number of observations per each variable which, in total, are four of them: *Weekly Death Cases*; *Containment Index*; *Stringency Index*; and *COVID-19 Vaccination*.

Table 1 has overall information which comprises all the measured regions in the United States. Table 2 splits into the four regions defined by the Census Bureau[2] to account for possible outliers amid regions and showing a summary about weekly death cases, their city population as well as the total deaths from the pandemic’s second wave. Table 3 accounts for a summary per region of its Containment and Stringency Index and the COVID-19 vaccination rates of the evaluated places.

[2]U.S. Census Bureau (2010). *Census Regions and Divisions with State FIPS Codes.* U.S. Department of Commerce, Economics and Statistics Administration.

Table 1: General Summary Statistics

Variable	Minimum	Maximum	Mean	N
Weekly Death Cases				
Reported Weekly Deaths	0	6.783	200,7	1537
City Population 1918	104.432	5.482.949	465.728	53
Total I&P Deaths per city	0	69.111	5.821	53
Containment Index (0-100)				
Index	0.000	61.458	10.804	1537
Stringency Index (0-100)				
Index	0.000	63.750	8.656	1537
COVID-19 Vaccination (%)				
Complete Vaccination	55,70	93,40	73,17	53
Total Number of Deaths				
Total I&P Deaths		308.509		

Firstly, Weekly Death Cases contains, in absolute numbers, the reported death cases per week in every city. The number of cities in this research is 53 and this, by the number of weeks which comprises the second wave of the pandemic (29 weeks), accounts for 1.537 observations in total. Furthermore, it also details the total city population in 1918 and the total number of deaths related to Influenza and Pneumonia (I&P) during the second wave.

Containment and Stringency Index reports weekly the number of measures taken to contain the disease and the severity of those measures respectively. Both indices have a score from 0 to 100, 100 being the strictest governmental response. This indices have 1.537 observations.

Finally, COVID-19 Vaccination accounts for the percentage of the county population where the evaluated cities are located with a complete vaccination schedule against the disease. This variable has 53 also observations.

Table 2 and 3 shows the same variables as Table 1, but it disaggregates the data into regions. Each section now shows the summary of each defined region: the Midwest region comprises 16 cities and 464 observations for Weekly Reported Deaths and Containment and Stringency Index; the Northeast includes 20 cities and 580 observations; the South 9 cities and 261 observations; and finally, the West region, comprises 8 cities and 232 observations. With this divided section, it is possible to see how the disease travelled from East to West and when the NPIs were enacted.

Table 2: Weekly Death Cases per Region: Summary

Variable	Minimum	Maximum	Mean	N
Weekly Death Cases				
Midwest				
Reported Weekly Deaths	0	3.137	179,84	464
City Population 1918	119.436	2.617.088	506.623	16
Total I&P Deaths per city	0	31.231	5.215	16
Total Number of Deaths - Midwest				
Total I&P Deaths		83.445		
Northeast				
Reported Weekly Deaths	0	6.783	282,77	580
City Population 1918	108.952	5.482.949	603.205	20
Total I&P Deaths per city	0	69.111	8.200	20
Total Number of Deaths - Northeast				
Total I&P Deaths		164.009		
South				
Reported Weekly Deaths	0	1.691	127,4	261
City Population 1918	117.110	704.378	252.121	9
Total I&P Deaths per city	0	10.961	3.695	9
Total Number of Deaths - South				
Total I&P Deaths		33.258		
West				
Reported Weekly Deaths	0	804	119,81	232
City Population 1918	104.432	527.715	280.551	8
Total I&P Deaths per city	0	7.591	3.475	8
Total Number of Deaths - West				
Total I&P Deaths		27.797		

The peak week of the disease was reached first in the South during the 8th week (see Figure 7) of the second wave; followed by the Midwest and the Northeast by week 9th (see Figure 5 & 6) and later by the West on week 10th (see Figure 8). All regions had a delayed surge of NPIs during the acceleration weeks of the disease and continuous decrease as the worst part of the second wave was comes to an end: the South and the Midwest had their first measures during the 5th week of the wave; the Northeastern region enacted their first NPIs by week 6 which at a time when their death cases were already increasing exponentially. The regions also experienced a second peak not as deadly as the first one, and it caused no further increase on the already decreasing NPIs.

Additionally, the regions experienced the divergence between Containment and Stringency Index as the second wave was claiming less victims with the exception of the Northeast (see Figures 5-8), meaning that even if the NPIs were still in place, they were less severe for the general population. This opened the

Table 3: Containment / Stringency Index and COVID-19 Vaccination per Region: Summary

Containment Index (0-100)				
Midwest				
Index	0.000	55.208	11.348	464
Northeast				
Index	0.000	55.903	9.915	580
South				
Index	0.000	52.430	10.810	261
West				
Index	0.000	61.458	11.931	232
Stringency Index (0-100)				
Midwest				
Index	0.000	56.250	9.923	464
Northeast				
Index	0.000	50.417	6.304	580
South				
Index	0.000	52.917	9.995	261
West				
Index	0.000	63.750	10,50	232
COVID-19 Vaccination (%)				
Midwest				
Complete Vaccination	55,70	78,30	66,71	16
Northeast				
Complete Vaccination	70,60	93,40	79,08	20
South				
Complete Vaccination	56,30	76,10	65,79	9
West				
Complete Vaccination	64,50	88,00	79,60	8

possibility for small gatherings; staged business schedules; and recommendations rather than prohibitions based on the evolution of NPIs gathered by Markel et al. (2007) during the second wave period.

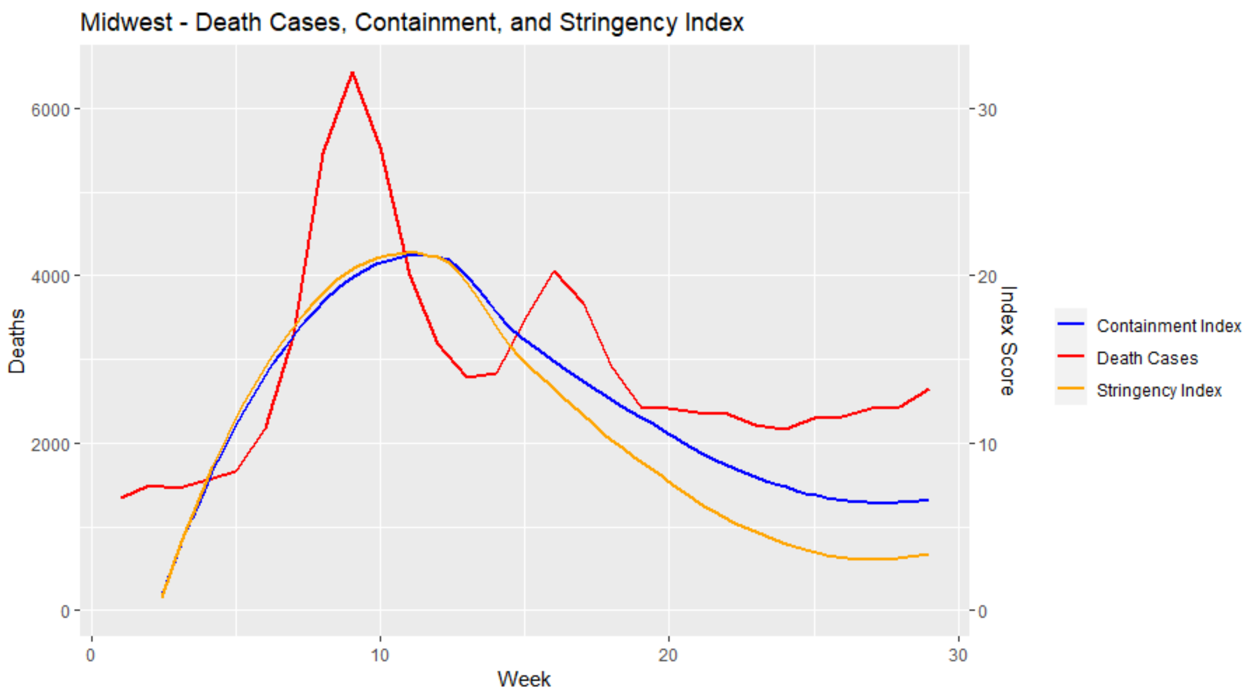

Figure 5: Reported weekly death cases, Stringency, and Containment Index in the Midwest region. 1918-1919. NPIs retrieved by Markel et al. (2007) and own measured 1918 Index using Hale et al. (2020) guidebook.

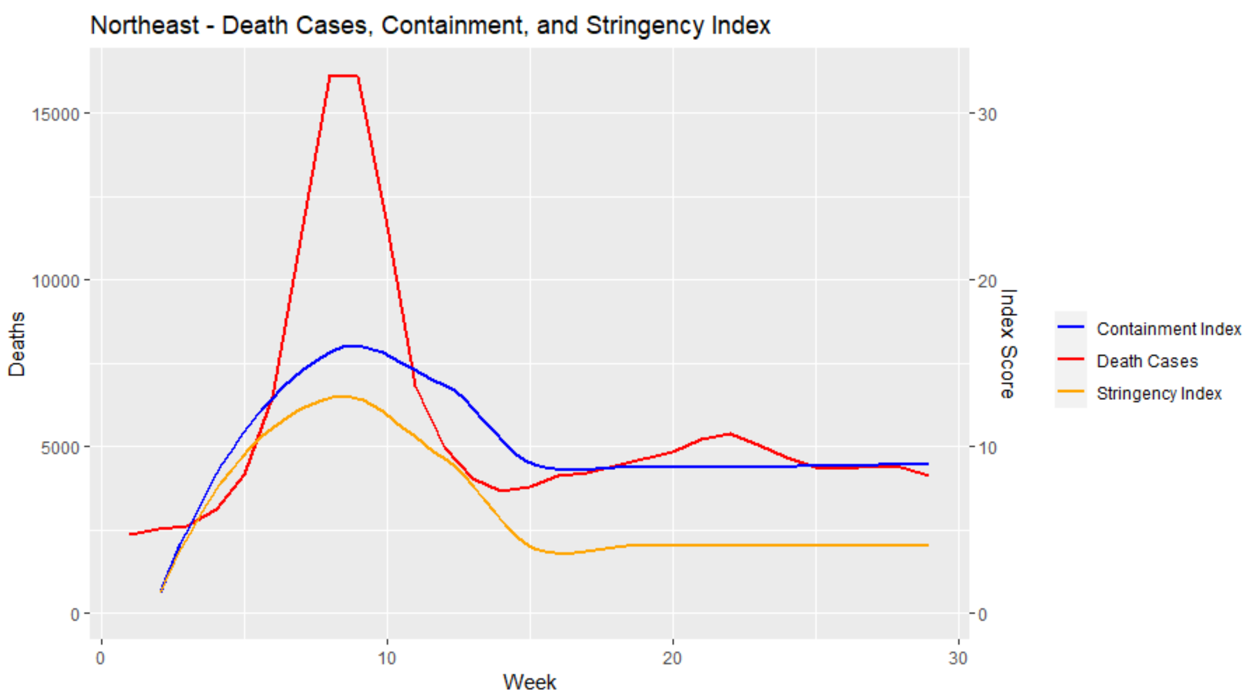

Figure 6: Reported weekly death cases, Stringency, and Containment Index in the Northeast region. 1918-1919. NPIs retrieved by Markel et al. (2007) and own measured 1918 Index using Hale et al. (2020) guidebook.

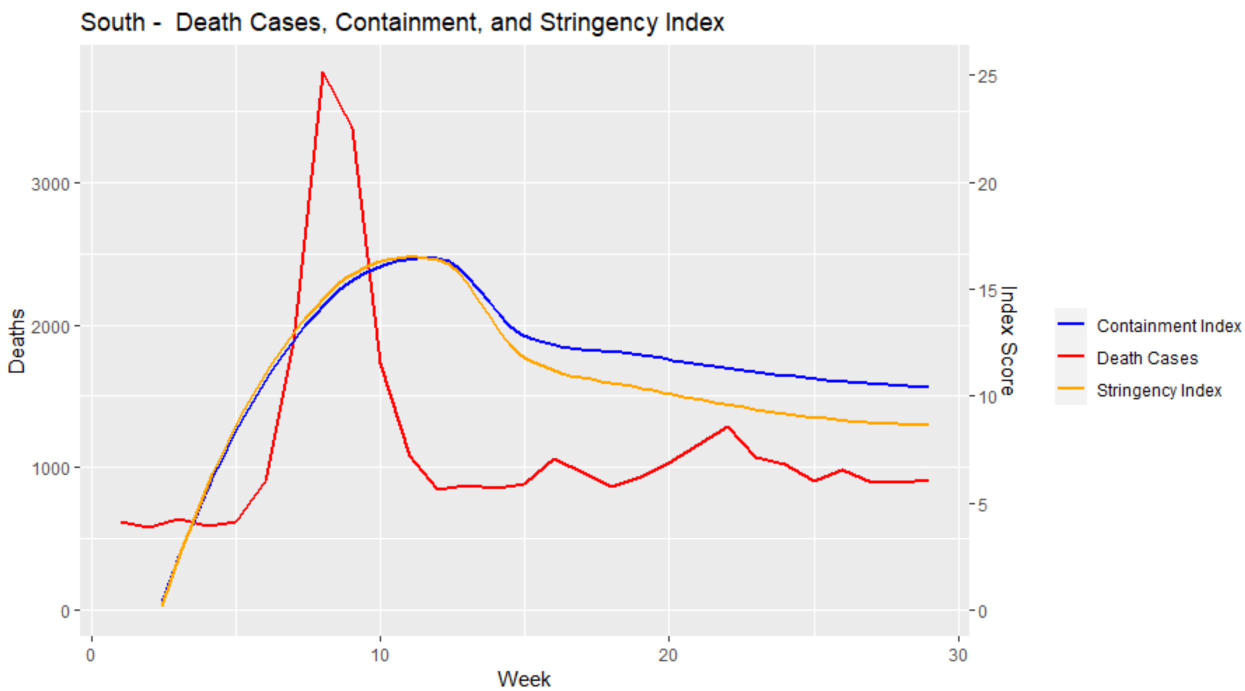

Figure 7: Reported weekly death cases, Stringency, and Containment Index in the South region. 1918-1919. NPIs retrieved by Markel et al. (2007) and own measured 1918 Index using Hale et al. (2020) guidebook.

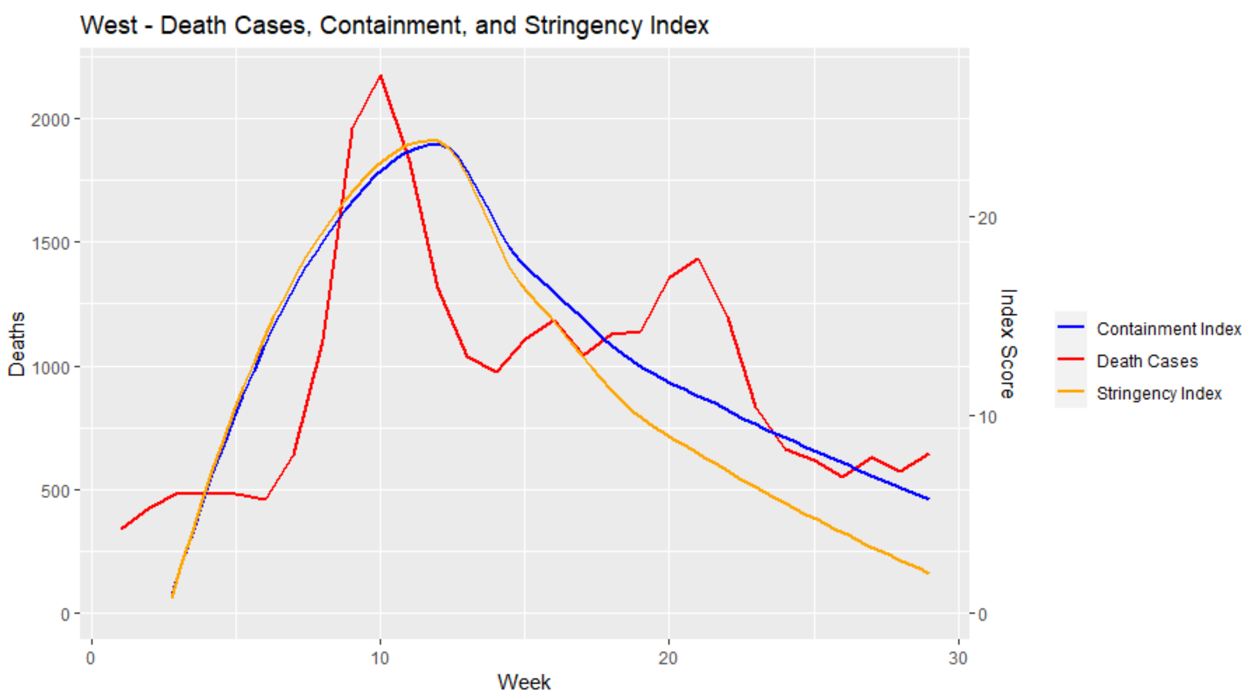

Figure 8: Reported weekly death cases, Stringency, and Containment Index in the West region. 1918-1919. NPIs retrieved by Markel et al. (2007) and own measured 1918 Index using Hale et al. (2020) guidebook.

5 Empirical Method

To identify the persistence shocks of the 1918 Influenza Pandemic in COVID-19 Vaccination Rates, five Panel Data regressions were conducted to measure this relationship. The first regression goes as follows:

$$\text{COVID-19 Vaccination Rate}_i = \beta_0 + \beta_1 Deaths_{1918i} + \beta_2 \text{Region} + \varepsilon \tag{1}$$

Where *COVID-19 Vaccination Rate* takes in consideration vaccination rates from city i from which we have also data from 1918; $Deaths_{1918}$ has the weekly deaths from those cities during the 1918 pandemic. The variable is log-transformed due to the exponential nature of deaths during a pandemic in order to linearlise the relationship; and *Region* will measure the effects comparing regions regarding vaccination rates.

This regression is a Random Effects Panel Data regression. This model was seen as more appropriate in order to have a larger number of observations with time-varying independent variables. Additionally, a Hausman Test for Panel Models was conducted and hence, Random Effects was seen more convenient.

The second and third regressions go as follows:

$$\begin{split} Deaths_{1918_{i,t}} = \beta_0 + \beta_1 \text{ NPI Containment Index}_{1918,it} \\ + \beta_2 \text{ NPI Containment Index}_{it+1} + \beta_3 \text{City Population}_{1918} + \alpha_i + \varepsilon \end{split} \tag{2}$$

$$\begin{split} Deaths_{1918_{i,t}} = \beta_0 + \beta_1 \text{ NPI Stringency Index}_{1918,it} \\ + \beta_2 \text{ NPI Stringency Index}_{it+1} + \beta_3 \text{City Population}_{1918} + \alpha_i + \varepsilon \end{split} \tag{3}$$

Where $Deaths_{1918it}$ are the log-transformed deaths of the measured cities i in time period t; *NPI Containment Index* measure the amount of measures taken enacted by each city i during time period t. Both indices also follow a non-linear relationship and thus, they will be also be log-transformed for this research. Furthermore, since NPIs can have an effect in later weeks, $NPI\ Containment\ Index_{i,t+1}$ will measure the Containment Index of city i during time period *t+1* to catch the delayed NPI effect of the amount of containment NPIs while also addressing possible endogeneity concerns. Additionally, $City\ Population_{1918}$

from each of the measured cities is taken as a control variable to control for the effect of the population size across them and ensure that the effect of the Indices are not confounded by the population size of the cities. These regressions, in comparison with regression 1, were divided per regions due to the deviation between them, which at the end could underestimate the results; nevertheless, the overall results are also included in Appendix (see Section 9). α_i represent an instrumental variable to control for factor Week due to the time-specific effects that the peak of the pandemic and its progression in all cities; and ε accounts for the error term.

In comparison with the first regression, regression 2 is a Fixed Effect Panel Data regression. By using this model, it is possible to account for individual-specific effects while also controlling to deviations in deaths during the pandemic; any time-invariant factors; or unobserved heterogeneity. As well as with the first regression, a Hausman Test was conducted in order to choose this Fixed Effect model.

The third regression takes the same variables from the second regression, with the difference of using the log-transformed Stringency Index of each city i in time period t rather than the Containment Index used in the latter and represent the level of strictness of this NPI measures. Similarly, this regression will use the Stringency Index of period *t+1* to catch the delayed NPI effect of the severity of this measures. The instrumental variable α_i represent the same IV during all the paper accounting for Week. In the same way as the other regressions, a Hausman test was conducted which shows that a Fixed Effect regression was also more suitable.

Finally, the fourth and fifth regressions go as follows:

$$\begin{aligned}\text{COVID-19 Vaccination Rate}_i = \beta_0 + \beta_1 \text{ NPI Containment Index}_{1918t} + \beta_2 Deaths_{1918} + \\ \beta_3 \text{ NPI Containment Index}_{1918t} * Deaths_{1918} + \alpha_i \ + \ \varepsilon\end{aligned} \tag{4}$$

$$\begin{aligned}\text{COVID-19 Vaccination Rate}_i = \beta_0 + \beta_1 \text{ NPI Stringency Index}_{1918t} + \beta_2 Deaths_{1918} + \\ \beta_3 \text{ NPI Stringency Index}_{1918t} * Deaths_{1918} \ + \ \alpha_i + \varepsilon\end{aligned} \tag{5}$$

In addition to the already explained variables from regression 1 to 3, regression 4 and 5 will include an interaction term between their indices and the deaths attributed to the 1918 influenza in order to measure the joint effects between those on COVID-19 vaccination rates. Both regressions are Random Effects Panel Data regressions which were considered suitable by running a Hausman test.

By using these regressions, it is possible to measure how those deaths persisted and have an effect on COVID-19 vaccination rates even at a minimum level; the effects that these NPIs had at the immediate moment to tackle the 1918 pandemic shocks; and how these governmental policies and historical events persisted and have a current social behaviour effect regarding public health issues such as vaccination rates which can also differ substantially between regions. As well as in regression 2 and 3, these regressions were conducted by regions. The nature of the pandemic to spread in unexpected ways could also have different effects depending on the region the cities are located; however, the results of regressions 2-5 without dividing by regions are also included in the Appendix section of this paper (see Section 9).

6 Results

6.1 Regression 1

Table 4: Regression (1) Results: COVID-19 Vaccination Rate & 1918 Deaths

	COVID-19 Vaccination Rate			
	(1)	(2)	(3)	(4)
Deaths_{1918}	0.677***	0.677***	0.036	0.036*
	(0.084)	(0.025)	(0.045)	(0.020)
Northeast	12.457***	12.457***	2.197***	2.197***
	(0.395)	(0.042)	(0.209)	(0.004)
South	−0.872*	−0.872***	−1.127***	−1.127***
	(0.490)	(0.036)	(0.260)	(0.003)
West	13.030***	13.030***	2.082***	2.082***
	(0.510)	(0.031)	(0.270)	(0.085)
Constant	63.860***	63.860***	91.211***	91.211***
	(0.460)	(0.085)	(0.244)	(0.085)
Robust Standard Errors		X		X
Population 65+			X	X
Observations	1,537	1,537	1,537	1,537
R^2	0.518	0.518	0.137	0.137
Adjusted R^2	0.517	0.517	0.135	0.135
F Statistic	1,647.038***	1,647.038***	243.456***	243.456***
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table 4 shows the results of Regression 1 showing the effect of the 1918 Influenza deaths on COVID-19 vaccinations[3]. The findings reveal that indeed the 1918 Influenza deaths had a small but persistent and significant effect on COVID-19 vaccination rates across the United States. A 1% increase on Influenza & Pneumonia Deaths (I&P Deaths) has a 0.677% increase on completed vaccination schemes against COVID-19. It is also important to remark the role of the regions for the willingness to get vaccinated against COVID-19: the West and Northeast cities affected by the influenza are more prone to get vaccinated against the virus by 13.03% and 12.45% respectively, while the Southern region (-0.872%) is less likely to do it. The use of percentages was chosen as it is more clear to see the effects. As deaths during a pandemic does not

[3]Dependent variable: COVID-19 Vaccination Rates. Column (2) and (4) present standard errors robust to heteroskedasticity.

follow a linear relationship, the log-transformation of the variables gives a more clear and robust coefficients. The appendix present the same table without log-transformation (see Table 9). Although the coefficients are still significant without a log-transformation, the interpretation of the variable *Deaths* is not as direct as with the log-transformation presented in table 4, therefore the log-transformed variables was considered better suited for this regression as well as for regression 4 and 5 that consider the same dependent variable.

In order to prevent confounding effects, age was also taken in consideration on column (3) and (4) since older population might have higher vaccination rates due to increased vulnerability to COVID-19: Columns (3) and (4) take the same variables into account but instead the sample of COVID-19 vaccination rates was subset for people of 65 years or older. On these columns, the effect of influenza deaths is reduced considerably but is still significant. Additionally, the results per region decrease sharply, but with an increased percentage of older people vaccinated as the constant coefficient increases to 91.21%. The effect of deaths is relatively small, but by having in consideration the period of time that has passed between these two events can remark the Influenza's historical legacy that has impacted societal behaviour on public health concerns.

6.2 Regression 2

Table 5: Regression (2) Results: Deaths & Containment Index

	Midwest		Northeast		South		West	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lag_{t+1}	-2.984**	-2.984*	-1.112***	-1.112*	0.011	0.011	-0.454	-0.454**
	(1.420)	(1.624)	(0.399)	(0.699)	(0.044)	(0.019)	(0.281)	(0.225)
Index_t	4.291***	4.291**	2.054***	2.054**	0.226***	0.226***	0.712***	0.712***
	(1.586)	(2.030)	(0.402)	(0.921)	(0.047)	(0.051)	(0.254)	(0.237)
Population_{1918}	0.111	0.111	1.148***	1.148***	0.332	0.332***	0.657***	0.657***
	(0.288)	(0.403)	(0.088)	(0.191)	(0.390)	(0.056)	(0.124)	(0.136)
Robust Standard Errors		X		X		X		X
Observations	464		580		253		225	
R^2	0.076		0.416		0.233		0.298	
Adjusted R^2	0.053		0.407		0.201		0.278	
F Statistic	19.062***		452.308***		134.276***		13.824***	
Note:								*p<0.1; **p<0.05; ***p<0.01

Table 5 accounts for the effect of Containment NPI on Influenza Deaths during the 1918 pandemic[4]. All the variables are log-log transformed since the relationship is non-linear as already explained in the empirical method (see Chapter 5) and the results are divided by regions due to the disparities between them. The appendix has the overall results for all cities (see Table 10). The results show that Containment NPIs have a delayed and significant effect on reducing deaths during the second wave of the pandemic in all regions but the South; the lag effect of the NPIs enacted the previous week is higher in the Midwest and the Northeast with -2.98% and -1.11% respectively and less in the Western region with a negative effect of -0.454%. This shows that the larger the amount of measures taken, the sharper the decline of death cases will be during the next week, showing that the effectiveness of this measures usually takes time in reducing deaths and highlights the importance of early NPIs' implementations to contain the disease. Additionally, we see that the variable Index_{1918} interestingly shows that increases in NPIs should increase death cases in the same week, which should sound quite counterintuitive; however, this positive response could reflect a reactive approach and not direct causal effect from the policymakers during the second wave. It implies that containment measures are enacted mainly due to rising death cases in the same week rather than increasing the number of deaths. With this interpretation, it makes sense that the regions with the higher coefficients, the Midwest (4.29%) and the Northeast (2.05%), were also the regions with higher death tolls relative to their containment NPIs as it was necessary to react quickly.

Furthermore, population also plays an important role on death cases: higher city population also increase death cases in all regions, although not significant in the Midwest. The Northeast, the most populated and dense region in the United States, shows a 1.15% increase in death cases with a 1% increase in its city population. This coefficients are lower but significant in the West (0.66%) and also the South (0.33%). One explanation for population not being significant in the Midwest could be the concentration of death cases in a few highly populated cities and other big cities with less death cases: this is the case of Chicago, the city accounted for 37,4% of all death cases in the Midwest while Detroit, the second largest city, have 0 death cases during the second wave.

[4]Dependent variable: 1918 Influenza death cases. Columns (2), (4), (6), and (8) present standard errors robust to heteroskedasticity.

6.3 Regression 3

Table 6: Regression (3) Results: Deaths & Stringency Index

	Midwest		Northeast		South		West	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lag_{t+1}	0.837*	0.837	1.669**	1.669	-0.035	-0.035	0.077*	0.077***
	(0.472)	(0.837)	(0.846)	(1.221)	(0.047)	(0.026)	(0.040)	(0.024)
$Index_t$	0.265	0.265	-0.808	-0.808	0.273***	0.273***	0.208***	0.208***
	(0.447)	(0.907)	(0.800)	(0.932)	(0.050)	(0.043)	(0.041)	(0.028)
$Population_{1918}$	0.076	0.076	1.360***	1.360***	0.354	0.354***	0.812***	0.812***
	(0.120)	(0.376)	(0.074)	(0.288)	(0.386)	(0.057)	(0.080)	(0.143)
Robust Standard Errors		X		X		X		X
Observations	450		564		253		225	
R^2	0.316		0.465		0.252		0.574	
Adjusted R^2	0.299		0.456		0.222		0.560	
F Statistic	104.799***		538.119***		97.320***		20.446***	
Note:								$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 6 shows the results of the Stringency Index on 1918 Influenza deaths.[5] with the same variables as the latter as well as the division per regions. The appendix includes the overall result from all cities (see table 11). Population is still significant and similar variable as seen in Regression 2: the Northeast is also the more impacted region with 1.36% increase in death cases by a 1% increase in its city population, followed by the West (0.81%), the South (0.35%) and the Midwest (0.076%, although not significant). In comparison to the effects of Lag and Index shown in Table 5, we observe that few of these results are statistically significant and positive, providing the opportunity to interpret about the strictness of the NPIs during the pandemic. The lag variable see a decrease in death cases in the South (although not significant) and an increase of death cases in the Midwest, the Northeast and the West region (with the West being the only statistically significant result). Although a positive coefficient might be interpreted as a direct positive effect on death cases, there are possible explanations to it: firstly, the measures were not as sophisticated and strict as the interventions taken to contain the COVID-19 virus which could, despite the possibility of reducing the number of increasing death cases, these might be not strict enough to make a change: the measures taken

[5]Dependent variable: 1918 Influenza death cases. Columns (2), (4), (6), and (8) present standard errors robust to heteroskedasticity.

were mainly focused on social distancing without closing non-essential businesses or public transportation, allowing to the disease to still spread but at a lower pace. These interventions can also be validated from the NPIs overview by Correia et al. (2022). Besides, we can also observe a divergence between the Containment and Stringency Indices in all regions (see Figures 5-8), this also implies that, although the measures were still in place during the second wave of the pandemic, the measures were more flexible, which could also be counterproductive to contain the virus as seen on the variables Lag and Index. Secondly, the variation across cities could also be an explanation. Since the there was no coordinated federal response, local governments were implementing what they found plausible to contain the disease with complete different responses with some cities and groups debating the efficacy of mask wearing, the economic consequences of the NPIs, and church closures. An interesting study shown also by Correia et al. (2022) from the Twin Cities of Minneapolis and St. Paul in the Midwest with different responses: Minneapolis had a quick response to control the disease in early October (and shows a maximum score of 36.25 according to the Stringency Index) while St. Paul waited until November to take those measures (and with a maximum stringency of 20 points according to the Stringency Index). Lastly, another reason for this finding can be policy compliance: the Stringency Index measures the strictness of the interventions, but not how those measures were enforced to by the public. Although few cities tried to enforce those rules via possible arrests or heavy fines, these punishments were in reality not used by the government (This can be seen from the NPI event dataset retrieved from Markel et al. (2007)).

6.4 Regression 4

Table 7 now takes in consideration the results of Regression 4 regarding the effects of Influenza Deaths, the Containment Index, and the interaction term between these two[6].

The results show that indeed deaths caused by the influenza in 1918 have an effect on COVID-19 vaccination rates depending on the affected region[7]. The most affected regions by the pandemic have a positive and significant effect on vaccinations rates with a 1% increase in Influenza & Pneumonia Deaths: 3.259% for the Midwest and 1.293% for the Northeast. The South also has an effect of 2.752% while the West has

[6]Dependent variable: COVID-19 Vaccination Rates. Columns (2), (4), (6), and (8) present standard errors robust to heteroskedasticity.

[7]The overall effect including all cities can be seen in the Appendix (see Table 12).

Table 7: Regression (4) Results: COVID-19 Vaccination, Deaths, & Containment Index

	Midwest		*Northeast*	
	(1)	(2)	(3)	(4)
$Deaths_{1918}$	3.259***	3.259***	1.293***	1.293*
	(0.277)	(0.009)	(0.249)	(0.767)
$Index_{1918}$	−5.217***	−5.217***	−8.986***	−8.986***
	(0.688)	(0.016)	(0.773)	(1.580)
$Deaths_{1918}$ x $Index_{1918}$	0.914***	0.914***	1.069***	1.069***
	(0.161)	(0.003)	(0.175)	(0.360)
Constant	55.618***	55.618***	80.394***	80.394***
	(2.311)	(0.048)	(1.824)	(2.392)
Robust Standard Errors		*X*		*X*
Observations	464		580	
R^2	0.444		0.352	
Adjusted R^2	0.441		0.349	
F Statistic	365.504***		253.346***	

	South		*West*	
	(5)	(6)	(7)	(8)
$Deaths_{1918}$	2.752	2.752	1.762***	1.762
	(2.122)	(9.332)	(0.505)	(1.992)
$Index_{1918}$	6.959*	6.959	0.383	0.383
	(3.874)	(15.325)	(2.729)	(2.925)
$Deaths_{1918}$ x $Index_{1918}$	−1.397	−1.397	0.099	0.099
	(0.940)	(3.867)	(0.510)	(1.232)
Constant	52.585***	52.585	71.080***	71.080***
	(8.588)	(36.391)	(1.296)	(2.498)
Robust Standard Errors		*X*		*X*
Observations	261		232	
R^2	0.011		0.169	
Adjusted R^2	0.001		0.158	
F Statistic	11.631***		60.007***	

Note: *p<0.1; **p<0.05; ***p<0.01

an effect of 1.762%, although these last two are not significant after adjusting for robust standard errors. This divergence between regions have several explanations: the Midwest and Northeast were among the most populated areas in the country as well as having the first influenza cases which then spread towards the West of the country, which at the time was more rural and included a few cities which could be possible infectious areas during the pandemic (See Table 2). Furthermore, as the Influenza travelled from East to West, the virus could have become weaker (Correia et al., 2022). Although the affects are quite small, it clarifies the persistence of this event in the collective memory and it is reflected in higher vaccination rates.

The following variable, Index, shows a sharp negative and significant relationship on vaccination rates also for the Midwest (-5.217%) and the Northeast (8.986%). The negative effect of the Containment Index might indicate a lasting perception of distrust and avoidance to governmental interventions which passed through generations as also seen in the literature (Aassve et al., 2021; Beach et al., 2022). As also explained for the deaths variable, as the effects of the influenza went from East to West, the main shocks were first perceived in the Midwest and Northeast; with more time to prepare for the arrival of the flu, population in the West and South had more time to prepare (Correia et al., 2022), making it less reliant on local policies as the effects in the South are positive (6.959%) and in the West (0.383%) although both are not significant.

The interaction term between Deaths and Index also shows a significant and positive effect for the Midwest (0.914%) and Northeast regions (1.069%) as well as for the West (0.099%), although not significant for the latter. These results can be interpreted as a small persistence effect on preventive health behaviours and historical memory: places with a higher number of containment NPIs and higher number of deaths follow a positive relationship that has a persistent effect on COVID-19 vaccination rates. On the other hand, the interaction term in the South is negative (-1.397%) and not significant. This effect can also follow relevant literature about the difference between this region and the other ones. The mixed effects that NPIs and Deaths have on population behaviour and public health issues are important in a country where limited government and personal liberties are the backbone of American society (Beach et al., 2022; Troesken, 2019). Because of this scopes, governmental intervention can be seen as an abuse from the federal government which, at the end, undermines the effect of the NPIs as well as their interaction terms.

6.5 Regression 5

Finally, the results Table 8 accounts for the effects of Influenza Deaths, the Stringency Index, and the interaction term between these two[8].

[8]Dependent variable: COVID-19 Vaccination Rates. Columns (2), (4), (6), and (8) present standard errors robust to heteroskedasticity.

Table 8: Regression (5) Results: COVID-19 Vaccination, Deaths, & Stringency Index

	Midwest		*Northeast*	
	(1)	(2)	(3)	(4)
Deaths_{1918}	3.592***	3.592***	0.974***	0.974
	(0.255)	(0.923)	(0.171)	(0.736)
Index_{1918}	−5.486***	−5.486***	−10.189***	−10.189***
	(0.819)	(2.016)	(1.268)	(3.438)
Deaths_{1918} x Index_{1918}	0.944***	0.944***	1.344***	1.344*
	(0.161)	(0.138)	(0.252)	(0.688)
Constant	54.181***	54.181***	79.517***	79.517***
	(2.468)	(3.437)	(1.869)	(2.819)
Robust Standard Errors		X		X
Observations	464		580	
R^2	0.429		0.303	
Adjusted R^2	0.425		0.299	
F Statistic	357.464***		190.938***	

	South		*West*	
	(5)	(6)	(7)	(8)
Deaths_{1918}	3.948	3.948	3.522***	3.522*
	(3.298)	(12.556)	(0.715)	(2.006)
Index_{1918}	9.280	9.280	−8.491**	−8.491
	(6.370)	(22.490)	(3.806)	(5.823)
Deaths_{1918} x Index_{1918}	−2.039	−2.039	0.957	0.957
	(1.545)	(5.628)	(0.640)	(2.115)
Constant	48.429***	48.429	71.058***	71.058***
	(13.292)	(49.007)	(1.452)	(2.531)
Robust Standard Errors		X		X
Observations	261		232	
R^2	0.003		0.110	
Adjusted R^2	-0.009		0.099	
F Statistic	6.403*		53.494***	

Note: *p<0.1; **p<0.05; ***p<0.01

The results of the 1918 Influenza deaths are, in comparison with Table 7, quite similar[9]. There is a positive effect on vaccination rates by the increase of 1% in *Deaths* during the pandemic; 3.592% for the Midwest; 0.974% for the Northeast (although not significant); the South (3.948%, also not significant); and

[9]The overall effect including all cities not dividing by region can be seen in the Appendix (see Table 13).

the West (3.522%). The death effects were already explained for Table 7 and they are still small, but they still generate a positive influence on vaccination rates. The Stringency Index also shows the same negative behaviour in the Midwest (-5.486%), the Northeast (-10.189%), and the West (-8.491%) and a positive not significant effect in the South (9.280%). This negative effect is sharper in the Midwest and Northeast than the results reflected on table 7, this can be explained due to the perception of these interventions as ineffective, lower compliance due to longer and harsher interventions as it happened during COVID, and also from the population that protested and lobbied against it (Correia et al., 2022). As seen from Figure 5-8, the Stringency Index diverged from the Containment Index after the second wave's peak but death cases were still being recorded in all regions. This negative intervention experience could be transmitted through communities that impact current attitudes towards vaccination, as seen by Aassve et al. (2021) and lower social trust.

The interaction term is also similar and significant from table 7 for the Midwest (0.944%) and the Northeast (1.344%) while it is not significant for the South (-2.039%) and the West (0.957%). As also explained in the last regression, the interaction from the severity of the pandemic due to the number of death cases in these regions, and the importance of containment of these diseases by stringent NPIs, results in a small persistence effect that leads to higher vaccination rates nowadays. Although the effect can be classified as almost negligible, there is still a collective behaviour effect from a historical event, mainly in the regions that suffered the second wave the most.

7 Discussion

According to the results, this research is in line with the current literature review on how former institutions and historical events influenced social attitudes nowadays. However, I also believe that the regression may have omitted variable issues that could be enhanced for further approaches.

Despite using age to identify potential confounding effects for regression 1, there might be other possible omitted variables from a contemporary and historical perspective which could enhance this project. Ethnicity is an important factor to take into account: in spite of being the most affected populations by the Influenza, there is few city records of influenza death cases of People of Colour, mainly in segregated cities at the time[10]. Other variables could be Influenza and Pneumonia death cases attributed to others factors rather than the Influenza such as air pollution caused by coal power plants as Correia et al. (2022) controls for.

For Regression (2) and (3), I believe that other factors can also be taken into account possible confounding effects and omitted variables for the death variable: air pollution, for example, could be a possible variable that can affect influenza and pneumonia death cases. However, as stated for Correia et al. (2022), their use of air pollution (coal-power plant prevalence) as a control variable left the effects of the NPIs unchanged. Nonetheless, the inclusion of this variables can enhance this research, its validity and robustness as well.

Furthermore, regarding contemporary variables, political attitudes and institutional trust regarding the COVID-19 pandemic can play role on affecting its vaccination rates. For this research, institutional trust was also taken in consideration in order to measure the effects of the pandemic on trust and thus, on COVID-19 vaccination rates. Unfortunately, there was no reliable data to measure this type of trust firsthand as there is not available surveys or trust measurement post-pandemic or even from the XIXth Century.

This research results have also policy implications for future pandemics as its persistence can change attitudes concerning public health issues: as seen in subsection 6.2, the amount of NPIs can reduce the number of death cases in order to stop the transmission of the disease, however its implications in the long-run can be detrimental for the government as it can reduce its trust, making other governmental interventions such as COVID-19 vaccinations, less efficient (see Table 7). Furthermore, in subsection 6.3, the NPIs have to be stringent enough in order to be effective in reducing disease transmission, otherwise, and as the findings on Table 6 show, the interventions can be inefficient in reducing cases and it will also reduce the efficacy of other governmental interventions in the long run due to the lack of population compliance (see Table 8).

8 Conclusion

This research suggest that there is a small institutional and historical persistence from Non-Pharmaceutical Interventions and the 1918 Influenza deaths on COVID-19 vaccination rates: The effects of the 1918 Influenza deaths on COVID-19 vaccination rates show a positive effect where regions most affected by the Influenza

[10] Few cities from the CDC Mortality Statistics (1920) have records of People of Color's Death Rate from Influenza per 100,000 population. Out of 66 cities, only 15 cities have segregated data between Whites and People of Colour.

have higher levels of vaccinated population overall and for older population as well. While analysing the effects of the NPIs on Influenza deaths, this has a lagged effect on reducing death cases as it catches the effect and highlights the importance of early NPI implementations to prevent higher transmission and death cases. Nonetheless, the stringency of these NPIs diverged and was lower than than the amount of NPIs in the same period, which did not improve the reduction of death cases, implying that the strictness of these measures was not high enough or followed by the community as much as expected by the policy-makers.

Finally, the findings show that the NPIs from 1918 have a negative effect on COVID-19 vaccination rates; this goes in line with the literature which shows that federalised communities are unwilling to adopt universal vaccination schemes and federal containment measures (Beach et al., 2022; Troesken, 2019).

Overall, the historical persistence effects are small, but significant, robust and in line with current literature. This persistence approach can be useful for public health policy-makers, as their decisions could have unintended long-term effects by changing social attitudes for future pandemics.

9 Appendix

Table 9: Regression (1) Results: COVID-19 Vaccination Rate & 1918 Deaths. Linear relation

	COVID-19 Vaccination Rate			
	(1)	(2)	(3)	(4)
Deaths_{1918}	0.004***	0.004***	0.0001	0.0001***
	(0.0003)	(0.001)	(0.0002)	(0.00003)
Northeast	11.958***	11.958***	2.185***	2.185***
	(0.388)	(0.045)	(0.210)	(0.003)
South	−0.703	−0.703***	−1.125***	−1.125***
	(0.480)	(0.063)	(0.260)	(0.001)
West	13.139***	13.139***	2.079***	2.079***
	(0.499)	(0.031)	(0.270)	(0.002)
Constant	65.972***	65.972***	91.349***	91.349***
	(0.294)	(0.099)	(0.160)	(0.005)
Robust Standard Errors		X		X
Population 65+			X	X
Observations	1,537	1,537	1,537	1,537
R^2	0.539	0.539	0.137	0.137
Adjusted R^2	0.538	0.538	0.135	0.135
F Statistic	1,791.076***	1,791.076***	242.872***	242.872***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 10: Regression (2) Results: Deaths & Containment Index - Overall results

	1918 Influenza Deaths	
	(1)	(2)
Lag_{t+1}	-1.754^{***}	-1.754^{***}
	(0.380)	(0.753)
$Index_t$	2.766^{***}	2.766^{**}
	(0.400)	(1.013)
$Population_{1918}$	0.833^{***}	0.833^{***}
	(0.086)	(0.256)
Robust Standard Errors		X
Observations	1,522	
R^2	0.188	
Adjusted R^2	0.171	
F Statistic	344.785^{***}	
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 11: Regression (3) Results: Deaths & Stringency Index - Overall results

	1918 Influenza Deaths	
	(1)	(2)
Lag_{t+1}	0.269	0.269
	(0.284)	(0.484)
$Index_t$	0.684^{**}	0.684
	(0.276)	(0.596)
$Population_{1918}$	1.045^{***}	1.045^{***}
	(0.048)	(0.292)
Robust Standard Errors		X
Observations	1,492	
R^2	0.397	
Adjusted R^2	0.384	
F Statistic	916.339^{***}	
Note:	*p<0.1; **p<0.05; ***p<0.01	

.

Table 12: Regression (4) Results: COVID-19 Vaccination, Deaths & Containment Index - Overall results

	COVID-19 Vaccination Rate	
	(1)	(2)
$Deaths_{1918}$	1.264^{***}	1.264
	(0.182)	(1.004)
$Index_t$	-5.112^{***}	-5.112^{**}
	(0.533)	(2.492)
$Deaths_{1918}$ x $Index_{1918}$	0.708^{***}	0.708
	(0.124)	(0.583)
Constant	70.036^{***}	70.036^{***}
	(1.658)	(3.372)
Robust Standard Errors		X
Observations	1,537	
R^2	0.147	
Adjusted R^2	0.145	
F Statistic	252.197^{***}	
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 13: Regression (5) Results: COVID-19 Vaccination, Deaths & Stringency Index - Overall results

	COVID-19 Vaccination Rate	
	(1)	(2)
$Deaths_{1918}$	1.533^{***}	1.533^{*}
	(0.135)	(0.921)
$Index_t$	-3.750^{***}	-3.750
	(0.759)	(3.572)
$Deaths_{1918}$ x $Index_{1918}$	0.321^{***}	0.321
	(0.156)	(0.741)
Constant	69.267^{***}	69.267^{***}
	(1.633)	(3.572)
Robust Standard Errors		X
Observations	1,537	
R^2	0.127	
Adjusted R^2	0.126	
F Statistic	204.278^{***}	
Note:	*p<0.1; **p<0.05; ***p<0.01	

10 References

Aassve, A., Alfani, G., Gandolfi, F., & Le Moglie, M. (2021, apr). Epidemics and trust: The case of the Spanish Flu. *Health Economics*, *30*(4), 840–857. Retrieved from https://doi.org/10.1002/hec.4218 doi: https://doi.org/10.1002/hec.4218

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American economic review*, *91*(5), 1369–1401.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Institutions as a fundamental cause of long-run growth. *Handbook of economic growth*, *1*, 385–472.

Aksoy, C. G., Eichengreen, B., & Saka, O. (2020). The political scar of epidemics. *NBER Working Paper*(w27401).

Algan, Y., & Cahuc, P. (2010). Inherited Trust and Growth. *American Economic Review*, *100*(5), 2060–2092. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.100.5.2060 doi: 10.1257/aer.100.5.2060

Arthi, V., & Parman, J. (2021). Disease, downturns, and wellbeing: Economic history and the long-run impacts of covid-19. *Explorations in Economic History*, *79*, 101381.

Barrett, M. P., Das, M. S., Magistretti, G., Pugacheva, E., & Wingender, M. P. (2021). *After-effects of the covid-19 pandemic: Prospects for medium-term economic damage*. International Monetary Fund.

Barro, R. J., Ursúa, J. F., & Weng, J. (2020, March). *The coronavirus and the great influenza pandemic: Lessons from the "spanish flu" for the coronavirus's potential effects on mortality and economic activity* (Working Paper No. 26866). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w26866 doi: 10.3386/w26866

Beach, B., Clay, K., & Saavedra, M. (2022). The 1918 Influenza Pandemic and Its Lessons for COVID-19. *Journal of Economic Literature*, *60*(1), 41–84. Retrieved from https://www.aeaweb.org/articles?id=10.1257/jel.20201641 doi: 10.1257/jel.20201641

Berkes, E., Deschênes, O., Gaetani, R., Lin, J., & Severen, C. (2023, feb). Lockdowns and Innovation: Evidence from the 1918 Flu Pandemic. *The Review of Economics and Statistics*, 1–30. Retrieved from `https://doi.org/10.1162/rest{_}a{_}01289` doi: 10.1162/rest_a_01289

Bridgman, B., & Greenaway-McGrevy, R. (2023). The economic impact of social distancing: Evidence from state-collected data during the 1918 influenza pandemic. *Explorations in Economic History*, *90*, 101531. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0014498323000256` doi: https://doi.org/10.1016/j.eeh.2023.101531

Centers for Disease Control and Prevention. (2023). *COVID-19 Vaccinations in the United States, County.* Data provided by IISInfo. Retrieved from `https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh/about_data` (Public Domain U.S. Government)

Correia, S., Luck, S., & Verner, E. (2022). Pandemics Depress the Economy, Public Health Interventions Do Not: Evidence from the 1918 Flu. *The Journal of Economic History*, *82*(4), 917–957. Retrieved from `https://www.cambridge.org/core/product/5902CC4230B6B10805C89E374BBE89A6` doi: DOI:10.1017/S0022050722000407

Emmerling, J., Furceri, D., Monteiro, F. L., Loungani, M. P., Pizzuto, P., & Tavoni, M. (2021). *Will the economic impact of covid-19 persist? prognosis from 21st century pandemics.* International Monetary Fund.

Ferguson, N. M., Cummings, D. A., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., ... Burke, D. S. (2005). Strategies for containing an emerging influenza pandemic in southeast asia. *Nature*, *437*(7056), 209–214.

Ferguson, N. M., Cummings, D. A., Fraser, C., Cajka, J. C., Cooley, P. C., & Burke, D. S. (2006). Strategies for mitigating an influenza pandemic. *Nature*, *442*(7101), 448–452.

Furceri, D., Loungani, P., Ostry, J. D., & Pizzuto, P. (2022). Will covid-19 have long-lasting effects on inequality? evidence from past pandemics. *The Journal of Economic Inequality*, *20*(4), 811–839.

Guiso, L., Sapienza, P., & Zingales, L. (2004). The role of social capital in financial development. *American economic review*, *94*(3), 526–556.

Guiso, L., Sapienza, P., & Zingales, L. (2016, dec). Long-Term Persistence. *Journal of the European Economic Association*, *14*(6), 1401–1436. Retrieved from `https://doi.org/10.1111/jeea.12177` doi: 10.1111/jeea.12177

Hale, T., Angrist, N., Kira, B., Petherick, A., Phillips, T., & Webster, S. (2020). *Variation in government responses to COVID-19 VO - RT - Unpublished Work* (No. WP -). University of Oxford.

Heitzman, J. (2020). Impact of covid-19 pandemic on mental health. *Psychiatr Pol*, *54*(2), 187–198.

Johnson, N. P., & Mueller, J. (2002). Updating the accounts: global mortality of the 1918-1920" spanish" influenza pandemic. *Bulletin of the History of Medicine*, 105–115.

Markel, H., Lipman, H. B., Navarro, J. A., Sloan, A., Michalsen, J. R., Stern, A. M., & Cetron, M. S. (2007, aug). Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic. *JAMA*, *298*(6), 644–654. Retrieved from `https://doi.org/10.1001/jama.298.6.644` doi: 10.1001/jama.298.6.644

Nunn, N., & Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American economic review*, *101*(7), 3221–3252.

Tabellini, G. (2008). Institutions and culture. *Journal of the European Economic association*, *6*(2-3), 255–294.

Tabellini, G. (2010, jun). Culture and Institutions: Economic Development in the Regions of Europe. *Journal of the European Economic Association*, *8*(4), 677–716. Retrieved from `https://doi.org/10.1111/j.1542-4774.2010.tb00537.x` doi: 10.1111/j.1542-4774.2010.tb00537.x

Taubenberger, J. K., & Morens, D. M. (2006). 1918 influenza: the mother of all pandemics. *Revista Biomedica*, *17*(1), 69–79.

Tomes, N. (1999). *The gospel of germs*. Harvard University Press.

Tomes, N. (2000). The making of a germ panic, then and now. *American Journal of Public Health*, *90*(2), 191.

Troesken, W. (2019). *The pox of liberty: how the constitution left Americans rich, free, and prone to infection.* University of Chicago Press.

Voigtländer, N., & Voth, H.-J. (2012, aug). Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany*. *The Quarterly Journal of Economics*, *127*(3), 1339–1392. Retrieved from `https://doi.org/10.1093/qje/qjs019` doi: 10.1093/qje/qjs019