Don’t Jump to Faulty Conclusions: Using the Synthetic Control Method to Evaluate the Effect of a Counterfactual Lockdown in Sweden

By
Jonas Herby*

Keywords: COVID-19, lockdown, non-pharmaceutical interventions, mortality, Synthetic Control Method
JEL codes: I18

Abstract*

Several studies based on the Synthetic Control Method (SCM) link the many COVID-19 deaths in Sweden during the spring of 2020 to Sweden’s decision not to lock down. I show that the estimated effect of lockdown varies greatly between four hypothetical Swedens depending on the spread of COVID-19 in societies before lockdown. This suggests that the effect of lockdowns found in earlier studies using Sweden as a control are likely to be (at least partially) driven by unobserved variables such as the spread of COVID-19 in societies before lockdown, and other misspecifications. Including a proxy for the spread of COVID-19 before lockdown does not improve the results. Suggesting that the very short data series on the spread of COVID-19 before lockdown limits the SCM ability to accurately estimate the effect of lockdowns.

* Jonas Herby, Economist (MSc.), Copenhagen
E: jherby@gmail.com // P: +45 2728 2748

1. Introduction

In the spring of 2020, Sweden, led by state epidemiologist Anders Tegnell, opted against a mandatory lockdown, a decision that drew significant international attention and academic scrutiny. Despite reporting low excess mortality overall during the pandemic, Sweden experienced substantial COVID-19-related deaths early on.

The case of Sweden has fueled a body of research, and several studies have tried to estimate the effect of lockdowns using Sweden as a control group. These studies, which mostly focus on the first pandemic wave, generally conclude that a Swedish lockdown would have prevented thousands of deaths in Sweden in the spring of 2020.

One group of studies employs the Synthetic Control Method (SCM) to examine the effect of a counterfactual lockdown in Sweden. The SCM is a statistical technique used to estimate the causal effects of interventions in observational studies, particularly when randomized control trials are not feasible. This method constructs a synthetic version of the treated unit by optimally weighing a combination of units from a donor pool that did not receive the intervention but share similar pre-intervention characteristics and outcomes. The goal of SCM is to create a counterfactual scenario that closely approximates what would have happened to the treated unit in the absence of the intervention, thereby isolating the intervention's effect on the outcome of interest.

Based on a 24 day pre-intervention window, Conyon and Thomsen (2021) conclude that »stricter lock-down policies are associated with fewer COVID-19 infections and deaths«, and Cho (2020) uses a 25 day pre-intervention period and find that »lockdown measures would have been associated with a lower excess mortality rate in Sweden by 25 percentage points«. Born et al. (2021) employ a 13 day pre-intervention period and find that »a 9-week lockdown in the first half of 2020 would have reduced [...] deaths by about 38 %«, and Latour et al. (2022) find that »a lockdown would have had sizable effects within one week« based on a 15 day pre-intervention window. Recently Ege et al. (2023) found that »not imposing a mandatory lockdown resulted in [...] a substantial increase in mortality« using excess mortalities over a 20 week pre-intervention window, although their estimate appears to lack statistical significance.[1] The estimated effects in all these SCM-

1. Based on the results of their placebo test, Ege et al. conclude that the »estimated impact of not implementing a mandatory lockdown in Sweden is not likely to be observed by chance and is clearly driven by the treatment in question.« However, according to Abadie et al. (2010) and Abadie et al. (2015), if a significant proportion of control units exhibit outcomes similar to the treated unit, this suggests that the estimated effects of the intervention might be due to chance, akin to a statistical p-value. Since 4-5 out of the 29 donor countries (14 %-17 %) show the same or higher excess mortality by the end of the study period, the result in Ege et al. appears to lack statistical significance.

studies are in the lower thousands for a population the size of Denmark[2] and thus significantly smaller than those predicted by epidemiological studies (see for example Ferguson et al. (2020)) or by studies that cannot distinguish between the effect of voluntary behavioral changes and lockdown measure (Flaxman et al. (2020)).

The SCM method has also been applied in other locations. Herby et al. (2024) show that these studies tend to focus on outliers with notably high COVID-19 mortality rates during the first wave. Herby et al. (2024) report that at least seven out of eleven eligible SCM studies targeted regions that experienced substantial deaths early in the pandemic.[3] In contrast, Mader and Rüttenauer (2022), who use the Generalized Synthetic Control Method (GSCM) covering 169 countries to avoid selection bias, 'do not find substantial and consistent mitigating effects of any NPI under investigation on COVID-19-related deaths per capita.'

The SCM-studies examining a counterfactual lockdown in Sweden predominantly rely on a limited pre-intervention window and tend to focus on a very short post-intervention window neglecting the Swedish health authorities' emphasis on designing their public health strategy for the long run. In this study I seek to qualify the use of the SCM in relation to examining the effect of a lockdown in Sweden.

First, I use historical mortality rates since 2016 to extend the pre-intervention period significantly. Abadie (2021) underscores the significance of a comprehensive pre-intervention period when designing synthetic control studies, stating that »it is of crucial importance to collect information on the affected unit and the donor pool for a large pre-intervention window«. The term »large pre-intervention window« is inherently ambiguous; however, asserting that a pre-intervention window of merely a few weeks qualifies as »large« is, under any circumstances, a contention that warrants critical examination.

I also extend the post-intervention period considerably compared to abovementioned studies. Through the pandemic, Swedish health authorities emphasized that their public health strategy was designed for the long run. One key reason for avoiding strict lockdowns was to prevent fatigue among the population, which could potentially lead to higher mortality rates over time. This approach can be interpreted as a strategic trade-off, accepting a higher number of deaths in

2. For example, Born et al. (2021) estimate that »a lockdown would have reduced the number of deaths by 38 %, from 5,795 to 3,573« (2,222 avoided deaths), Ege et al. (2023) estimate that »up to about 4411« deaths could have been avoided, and Conyon and Thomsen (2021)'s estimate corresponds to app. 3,000 avoided deaths. These numbers translate to approximately 1,250 to 2,500 avoided deaths in Denmark with a population almost half of that in Sweden. The difference-in-difference studies estimate similar effects of lockdowns. In comparison, influenza causes an excess mortality of 1,644 deaths in the 2017/18 influenza season according to Statens Serum Institut (2018).
3. See Supplementary file2 in Herby et al. (2024).

the short run to reduce the risk of greater mortality in the long run. Thus, evaluating the Swedish COVID-19 policy on the short run may lead to wrong conclusions.

I find that, in the short run (by week 26 of 2020), a counterfactual lockdown would have reduced mortality rates by approximately 22 deaths per 100,000. However, in the longer run (by week 26 of 2021), the same lockdown scenario would have increased mortality by about 19 deaths per 100,000. Crucially, both of these estimates are far from statistically significant.

I then examine the impact of the spread of COVID-19 before lockdown. I use the timing of the winter holiday as an instrument for the spread of COVID-19 in different regions of Sweden at the time when other countries implemented lockdowns. Arnarson (2021) and Björk et al. (2021) illustrate how the timing of the winter holiday significantly influenced the spread of the virus in Europe. Their research indicates that regions which scheduled their winter holidays later, in weeks 9 or 10 rather than weeks 6 to 8, experienced disproportionately high COVID-19 case numbers during the initial wave.[4] This was largely due to virus transmission by ski tourists returning from the Alps. Specifically, regions with their winter holiday in week 9 (which was the case for approximately one-third of the Swedish population, including the heavily affected Stockholm region) had, ceteris paribus, almost twice (90 %) the COVID-19 cases in March 2020 compared to those with winter holiday in week 7.

I find large discrepancies in the effect estimates across four hypothetical Swedens with different timing of their winter holiday suggesting that differences in excess mortality are less a result of the no-lockdown policy and more a result of unobserved variables such as the spread of COVID-19 before lockdown. This result indicates that the effect of lockdowns found in other studies using Sweden as a control are at least partially driven by unobserved variables.

To account for the pre-lockdown spread of COVID-19, I use excess death rates in the weeks following the lockdown as a proxy for pre-lockdown transmission. Although this approach narrows the range of the estimated counterfactual lockdown effects among the four hypothetical countries, the resulting span remains unacceptably large. This finding suggests that the SCM model does not fully capture all relevant factors – most likely due to the limited pre-intervention window for COVID-19.

Rhetorically, one might ask if researchers would ever base multi-week forecasts of business turnover on merely 15–20 days of noisy data, or whether this approach is peculiar to the COVID-19 context.

4. Week 6: February 3-9, Week 7: February 10-16, Week 8: February 17-23, Week 9: March 2-8, Week 10: March 9-15.

2. Data & Methodology

2.1. Methodology

The SCM is a specialized variant of the difference-in-difference analysis which serves as an econometric tool for policy evaluation. Noted for its interpretability and transparency, SCM is particularly valuable in constructing counterfactual scenarios to assess the impact of interventions without randomized control trials, cf. Abadie (2021). In this study, SCM is employed to create a »synthetic« (also called »doppelgänger«), which hypothetically adheres to lockdown measures. This enables a robust comparative analysis between the actual pandemic trajectories observed in Sweden and those that would have potentially occurred under lockdown conditions.

The construction of the synthetic control involves selecting a donor pool of potential control units (countries) and then determining a set of non-negative weights that minimize the difference between the treated unit and the synthetic control in terms of pre-intervention characteristics and outcomes. This weighted combination forms a synthetic control that serves as a benchmark for comparison during the post-intervention period. The causal effect of the intervention is estimated by comparing the post-intervention outcomes of the treated unit with those of the synthetic control.

SCM has been widely applied in various fields, including economics, political science, and public health, due to its ability to handle complex longitudinal data and provide transparent and intuitive results. For instance, Abadie and Gardeazabal (2003) utilized SCM to assess the economic impact of terrorism in the Basque Country by comparing it with a synthetic counterpart composed of other Spanish regions. Similarly, Abadie et al. (2010) employed this method to evaluate the effects of California's tobacco control program by constructing a synthetic control from other U.S. states. Additionally, Grier and Maynard (2016) used SCM to analyze the economic consequences of Hugo Chavez's policies in Venezuela, creating a synthetic control from similar countries to isolate the impact of Chavez's administration on Venezuela's economy, demonstrating significant negative effects.

Ege et al. (2023) use cumulative excess mortality to construct their synthetic controls. However, this approach raises several issues. First, there is little reason to believe that the initial spread of COVID-19 during the early days of the pandemic correlates with cumulative excess mortality. Even if deaths from influenza or other respiratory viruses in a typical winter season spread to Europe in a way similar to COVID-19 (e.g. via ski tourists in the Alps, as Arnarson (2021); Björk et al. (2021) show was the case with COVID-19), this would manifest itself in data as mortality, not (cumulative) excess mortality. Second, since the COVID-19 pandemic in Europe began in March, it is not comparable to earlier seasons where mortality typically peaked months before COVID-19 deaths peaked. Third, using cumulative excess mortality makes the synthetic control calculation very sensitive to the specific starting point of the analysis, specifically the choice of when the cumulative count is set to zero. Figure 8 in Appendix A shows that the results in Ege et al. (2023) change dramatically, if one set the zero point for accumulation to week 11 of 2020 rather than week 43 of 2019.

Hence, I construct the synthetic controls using historical mortality rates. Consequently, my synthetic control reflects the expected mortality in the post-intervention period – rather than the expected excess mortality, as in Ege et al. (2023). Therefore, the difference between observed mortality in Sweden and the synthetic control can be interpreted as the excess mortality.

I use mortality rates for each season (week 27 year X to week 26 year X+1) in the pre-intervention window to account for trends. For the 2019/20 season I stop at week 8 in 2020 well before the first COVID-19 death in Sweden. Besides seasonal mortality rates, I also employ hospital beds per 1,000 people, urban population as share of total population, GDP per capita (PPP), share of population aged 65 and above, and the migrant share of the population as control variables.

2.2. Data

I employ three primary data sources.

First, I source weekly mortality data for donor countries from the Short-Term Mortality Fluctuations (STMF) data provided by mortality.org (accessed July 2, 2024). The STMF collection is derived from official national statistics and adjusted for completeness and consistency, ensuring a high level of reliability. This comprehensive dataset facilitates cross-national comparisons and temporal analyses that are essential for public health planning, policy-making, and epidemiological research. I calculate the weekly mortality rates as the yearly mortality rate in STMF (variable »Rtotal«) divided by 52.[5]

Second, I source data for hospital beds per 1,000 people, urban population as share of total population, GDP per capita (PPP), share of population aged 65 and above, and the migrant share of the population from the World Bank's World Development Indicators database (accessed June 30, 2024).[6] I use simple forward/backward fill to deal with missing data.

Third, I utilize weekly deaths and population data from SCB Statistics Sweden for detailed regional analysis within Sweden (accessed June 30, 2024).[7] Importantly, the data from SCB Statistics for the entirety of Sweden corresponds precisely

5. See https://www.mortality.org/File/GetDocument/Public/STMF/DOC/STMFNote.pdf p. 6.
6. Indicator id's SH.MED.BEDS.ZS, SP.URB.TOTL.IN.ZS, NY.GDP.PCAP.PP.KD, SP.POP.65UP.TO.ZS, and SM.POP.TOTL.ZS.
7. We use the tables »Döda efter region, vecka och ålder« and »Folkmängd efter region, år och ålder«. Links to all datasets are available in our R-code provided as supplementary material.

with the mortality data provided by mortality.org, ensuring consistency and accuracy in the analysis.

I exclude the United Kingdom (UK) from the donor pool because its approach to the pandemic was similar to Sweden's in the early weeks. Additionally, I remove Italy – where lockdowns in Lombardy and the Veneto region began as early as February – from the donor pool. Finally, I restrict the donor pool to countries in the Northern Hemisphere.[8]

3. Sweden's COVID-19 policy

I expand on three critical aspects of the SCM, which have significant implications for evaluating the consequences of Sweden's decision not to implement lockdown measures, as well as for interpreting results from other studies that employ similar methodologies.

3.1. The impact of a short pre-intervention window

Most existing SCM studies examining Sweden's COVID-19 policy are based on very short pre-intervention – often just a few weeks long. This approach appears to contrast sharply with the guidelines suggested by Abadie (2021), who emphasized the importance of using a »large pre-intervention window« to gather sufficient data on the affected unit and the donor pool. Compared to previous studies, by using mortality rates from mortality.org, I can extend the pre-intervention period considerably, as STMF, SCB and the World Bank supply data for 36 countries going back to 2016.[9]

3.2. The impact of a short post-intervention window

The mentioned SCM studies all end their analysis in the summer or fall of 2020 focusing solely on the first part of the pandemic. While some studies are limited by data availability at the date of publication, Ege et al. (2023) ends their analysis arguing that »by November, synthetic Sweden is no longer an appropriate counterfactual for Sweden due to policy changes in the donor countries.« I see potential flaws in this argument.

8. Abadie (2021) advises selecting a donor pool of moderate size that consists of units resembling the treated unit in both observable and unobservable characteristics. Given the pandemic's staggered global onset, I therefore conducted the SCM analysis using a more focused donor pool, restricted to European countries (excluding CAN, ISR, KOR, TWN, and USA). This narrower donor pool produced only minor changes in my results and did not affect my overarching conclusions.
9. Russia is not part of the donor pool as there are no mortality data after week 53 of 2020. There are data for 26 countries going back to 2007. Extending the pre-intervention period to 2007 only have a marginal effect on my results and does not change my conclusions.

Swedish health authorities repeatedly stressed a long-term public health strategy, notably aiming to avoid population fatigue – an approach intended to prevent potentially higher mortality over time. In essence, this strategy represents a deliberate trade-off: accepting somewhat elevated short-term mortality to avert even larger losses in the future. Supporting this, , Mulligan and Arnott (2022) report increases in non-COVID excess deaths following lockdowns, indicating that a narrow, short-term focus may be insufficient when assessing the broader impact of these measures.

On the other hand, if Sweden's comparatively lenient lockdown in early 2020 allowed COVID-19 to spread more widely during the first wave, it may have started later waves at a higher baseline of infections, as posited by Arnarson (2021), and consequently required more stringent interventions further down the line. However, despite implementing additional mandatory measures during the 2020/2021 winter, Sweden's restrictions remained considerably more lenient than those in other (donor) countries (contrary to what Ege et al. (2023) suggest). For instance, Sweden did not close elementary schools, and although bars and restaurants faced certain limitations, they were not required to shut down. Moreover, hairdressers, fitness centers, and shopping malls remained open, and stay-at-home orders were never imposed.

Overall, choices made during the pandemic's first wave likely shaped subsequent policy responses. Examining outcomes through the summer of 2021 – by which time vaccines had been widely distributed to vulnerable populations – can thus offer insights into the longer-term ramifications of the paths chosen in spring 2020. However, it should be emphasized that this is not a straightforward comparison of »lockdown versus no lockdown« under identical conditions; rather, it reveals how early decisions may have set a path-dependent course for the pandemic's evolution.

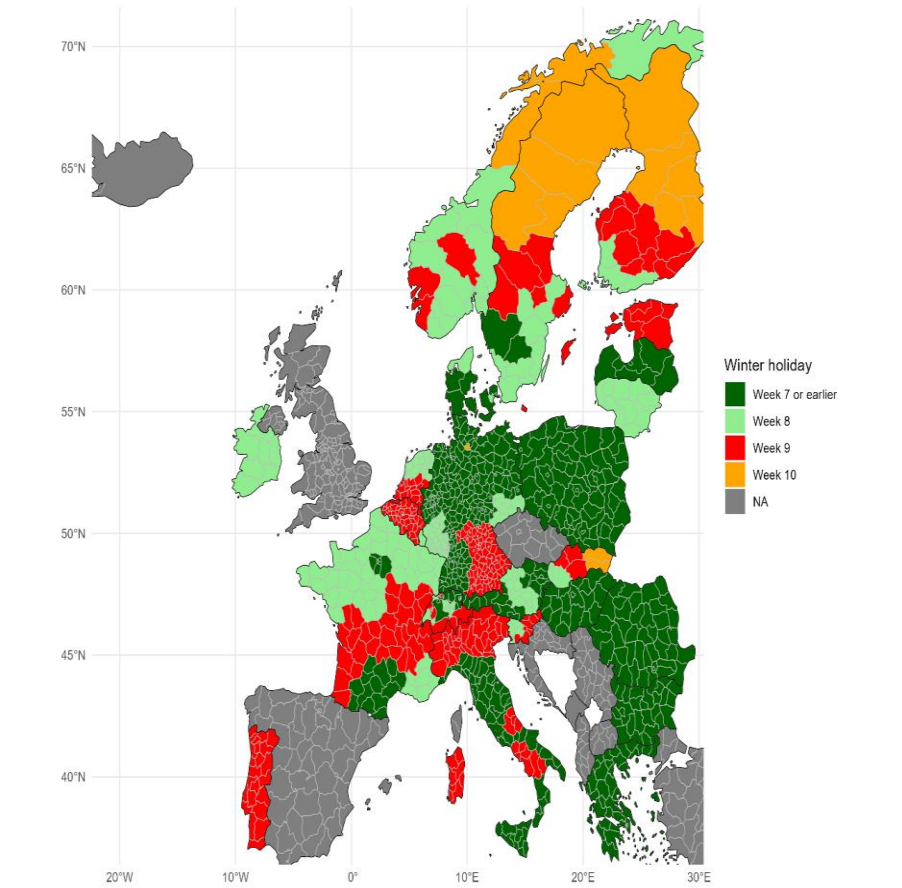
3.3. The spread of COVID-19 before lockdown and the impact of the timing of the winter holiday

The timing of the winter holiday emerges as a substantial factor in the spread of COVID-19 before lockdowns, particularly among regions that observed their winter holidays later (weeks 9 or 10) compared to regions with earlier winter holidays (week 6, 7 or 8). Arnarson (2021) and Björk et al. (2021) demonstrate that these regions experienced markedly higher COVID-19 case rates during the first wave, likely exacerbated by the return of ski tourists from the Alps. For instance, Arnarson (2021) shows regions that had their winter holiday in week 9 reported almost twice as many COVID-19 cases in March 2020 compared to those with winter holiday in week 7.

Figure 1 below illustrates the timing of the winter holiday across European regions and highlights several notable aspects. Firstly, it seems to confirm the findings from Arnarson (2021) and Björk et al. (2021) as all regions in Belgium – the

European country with the highest number of COVID-19 deaths per capita in spring 2020 – had their winter holiday in week 9. However, the figure also demonstrates that the timing of the winter holiday is not deterministic in terms of COVID-19 outcomes, as evidenced by Norway and Finland. Both countries have regions where the winter holiday occurs in week 9, yet, unlike Sweden, they experienced very few COVID-19 deaths during the pandemic.

Figure 1
Timing of the winter holiday in European regions



Source: Data from Arnarson (2021). The red color of the regions in week 9 indicates that this week had the greatest impact on the pandemic, as the outbreak was not discovered until week 10.

However, an important difference exists between these countries. According to Anders Tegnell's recent book, approximately 1 million Swedes returned from international travel during the winter holidays, predominantly visiting the Alps during weeks 9 and 10. In contrast, Norwegians and Finns are less likely to travel to the Alps during these weeks, typically preferring vacations within their own countries, Tegnell (2023).

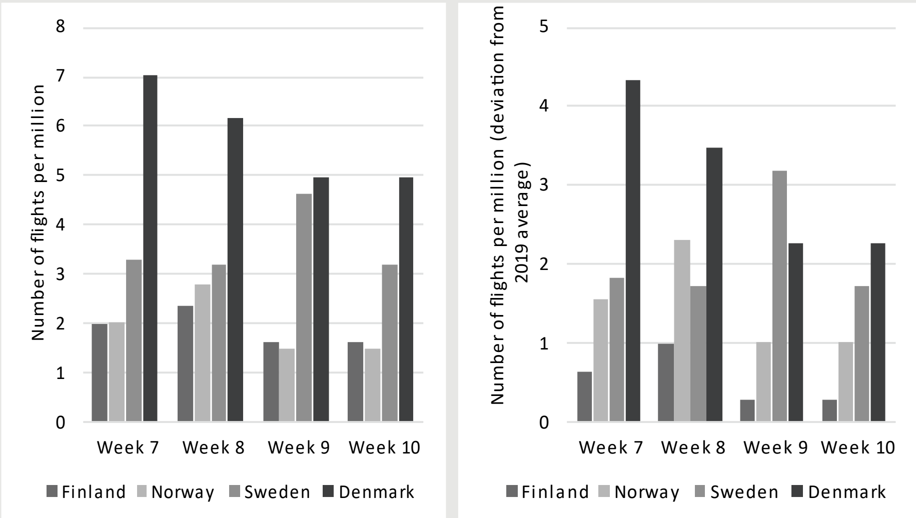
This behavior is corroborated by official flight data, as illustrated in Figure 2. Panel A displays the number of arrivals from the Alps by country and week in 2020, and Panel B highlights the deviations in these numbers from the 2019 average. The data reveal that Swedish arrivals from the Alps were approximately three times higher than those of Norwegians and Finns when accounting for differences in population sizes. This corresponds to data provided by Arnarson (2021) which shows that Swedes were more than twice as likely to spend a night in Austria during February 2020 compared to Norwegians and Finns.[10]

Interestingly, while many Danes also traveled to the Alps during this period, Denmark's COVID-19 death rate during the first wave was about double that of Norway and Finland, but only about one-fifth compared to Sweden. This indicates that while the timing of the winter holiday was a significant factor, it is far from being the sole determinant of the differences in COVID-19 outcomes.

Figure 2
Skiing in the Alps

Panel A: Number of flight arrivals from the Alps in 2020 by week and country

Panel B: Number of flight arrivals from the Alps in 2020 by week and country (deviation from 2019-average)



Source: https://zenodo.org/records/7923702

10. See supplementary information for Arnarson (2021), Table C1

The data presented in Figure 2 indicates that Swedes were more likely than Norwegians and Finns to travel to the Alps during the weeks coinciding with an undisclosed COVID-19 outbreak. This was particularly true for week 9, during which Arnarson (2021) identified the most substantial impact of the virus.

4. Results for Sweden (naïve specification)

I begin by analyzing the potential impact of a Swedish lockdown in weeks 11 or 12 of 2020, aligning with the timeframe when most European countries imposed similar measures.[11] I utilize a pre-intervention period starting from week 27 of 2016 and extending to a post-intervention period ending in week 26 of 2021, when vaccines were widely distributed to vulnerable residents. By selecting week 27 in 2016 as the commencement of the pre-intervention period, rather than week 1, I aim to avoid the influence of random peaks in mortality rates that could occur during the 2015/16 winter, which might otherwise introduce variability depending on whether they fall in 2015 or 2016. This approach ensures that the synthetic control remains unaffected by such random fluctuations. Based on these data boundaries the available donor pool consists of 30 countries.[12]

Table 1 compares the predictor variables for Sweden and its synthetic counterpart prior to the lockdown. As the table shows, Synthetic Sweden aligns closely with Sweden on most included control variables, with the notable exception of hospital beds, where Synthetic Sweden has substantially more. The table also indicates that while Sweden's mortality rates have been declining since the 2016/17 season, those of Synthetic Sweden remain more or less unchanged.

11. According to Born et al. (2021), the majority of European nations locked down in week 12 or 13. For example, Finland banned gatherings, closed museums, and announced school closures on Monday, March 16 (week 12), with closures taking effect on Wednesday, March 18. In Denmark, the lockdown was announced on the evening of Wednesday, March 11 (week 11), but most restrictions, including school closures, did not begin until Monday, March 16. Business closures started on Wednesday, March 18 (week 12), following an announcement on Tuesday, March 17.
12. AUT, BEL, BGR, CAN, CHE, CZE, DEU, DNK, ESP, EST, FIN, FRA, GRC, HRV, HUN, ISL, ISR, KOR, LTU, LUX, LVA, NLD, NOR, POL, PRT, SVK, SVN, SWE, and USA.

Table 1
Comparing Sweden to Synthetic Sweden

Predictor	Sweden	Synthetic Sweden	Absolute Difference	Relative Difference
Hospital beds per 1,000	2.2	3.1	0.9	43 %
Urban population share	87.4	86.6	-0.8	-1 %
GDP per cap (PPP)	61,054	63,135	2,080	3 %
Aged 65+ share	19.9	18.9	-0.9	-5 %
Migrant share	16.8	14.2	-2.6	-15 %
Mortality 2016/17	17.3	16.7	-0.6	-3 %
Mortality 2017/16	16.9	16.8	-0.1	-1 %
Mortality 2018/19	16.1	16.5	0.4	2 %
Mortality 2019/20*	16.1	16.6	0.5	3 %

Note: The table shows the average values of the predictor variables in the pre-treatment window for Sweden and Synthetic Sweden.
* *Only week 27 of 2019 to week 8 of 2020 are included.*

Figure 3 below displays the results of the SCM. Panel A shows the full pre- and post-intervention window of the analysis while Panel B focuses on the post-intervention period. Panel C shows the (cumulative) excess deaths for the post-intervention period based on the weekly differences between actual and synthetic Sweden. The donors are Finland (25 %), Canada (29.2 %), Netherlands (31.6 %), Denmark (10.4 %), and Luxembourg (3.7 %) (donor weights are available in Appendix B).

Figure 3 reveals some interesting findings. First, Panel A shows Synthetic Sweden's ability to track Actual Sweden prior to lockdown. In general, Synthetic Sweden tracks Actual Sweden closely.

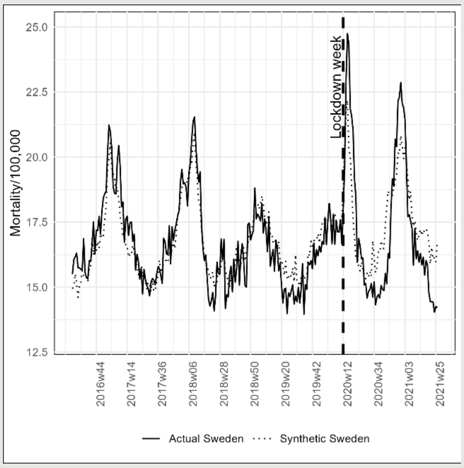
Second, Panel B shows that the mortality rate in Actual Sweden is higher compared to Synthetic Sweden during the first and second wave of the pandemic, but lower between and after the waves.

Third, Panel C shows that my cumulative results in the short run (the summer of 2020) are comparable to earlier findings. After the first wave (week 26 of 2020), the cumulative mortality in Actual Sweden is approximately 22 deaths per 100,000 higher than in Synthetic Sweden corresponding to approximately 2,300 avoided deaths in total. In comparison, both Conyon and Thomsen (2021) and Ege et al. (2023) estimate an effect of approximately 30 deaths per 100,000 (3,100 deaths in total), and Born et al. (2021) estimate that a lockdown would have reduced deaths by 38 % corresponding to approximately 2,000 avoided deaths. In a Danish context, my results correspond to approximately 1,300 avoided deaths with a population of 5.8 million.

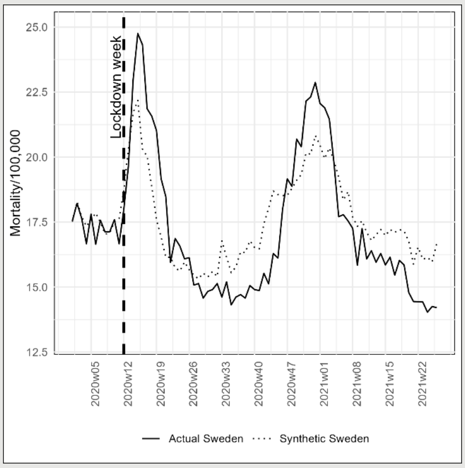
Fourth, Panel C indicates that the effect reverses over the longer term. By the end of the post-intervention period, the cumulative mortality in Actual Sweden is 19 deaths per 100,000 lower than in Synthetic Sweden – equivalent to nearly 2,000 fewer deaths, or about 1,100 fewer deaths in a population the size of Denmark's.

Figure 3
Results from the SCM for Sweden (naïve specification)

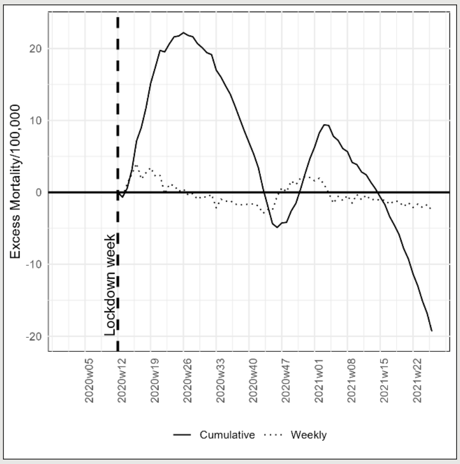
Panel A: Full pre- and post-intervention window



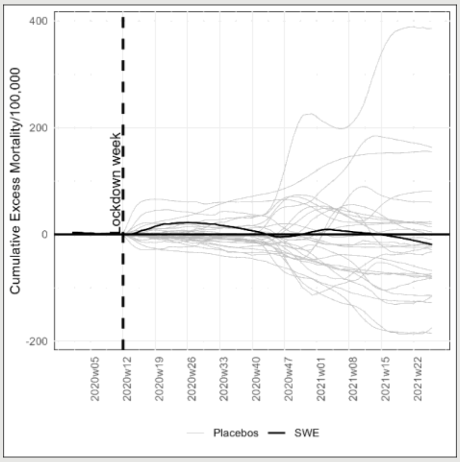
Panel B: Panel A zoomed to post-intervention window



Panel C: Post-intervention excess mortality



Panel D: Placebo test



Note: The vertical dashed line illustrates the timing of the lockdown in synthetic Sweden in week 12 of 2020. Cumulative excess mortalities have been normalized to zero at the time of the lockdown. Donor weights can be found in Appendix B.

4.1. Placebo test

Panel D presents the results of the placebo test, a robustness check in the SCM that compares the estimated treatment effect for Sweden to the distribution of effects in untreated units. In this approach, synthetic controls are created for donor countries that did not receive the intervention (i.e., no lockdown), and the resulting placebo effects are compared with Sweden's actual estimated outcome. If the observed effect in Sweden is substantially larger (or smaller) than the placebo effects, it suggests that the estimated impact is not merely a result of chance or unobserved confounders. Following Abadie et al. (2010) and Abadie et al. (2015), the fraction of donor units that produce outcomes similar to the treated unit serves as a proxy for a statistical p-value.

Given 30 donors, an effect is broadly »significant« at the 5 % level if at most one or two placebo effects exceed Sweden's estimated impact. Here, this threshold is not met. In every week following the lockdown, several placebos show higher cumulative mortality than Sweden, indicating that Sweden's results are not statistically significant.

This finding is not entirely unexpected. High mortality rates were by no means unique to Sweden, which did not implement a strict lockdown. For instance, Belgium enacted a stringent lockdown around the same time as Denmark yet recorded about 60 % more COVID-19 deaths than Sweden. Ireland and the Netherlands also imposed lockdown measures but still experienced relatively high COVID-19 mortality rates, not markedly different from Sweden's trajectory.[13]

5. The lockdown effect in Four hypothetical Swedens

I now turn to examine how the timing of the winter holiday could affect the spread of COVID-19 prior to any lockdown measures. To clarify, I am not primarily interested in the direct effect of the holiday timing itself; rather, I use it as an instrument to capture variations in pre-lockdown viral transmission across different Swedish regions (län). Specifically, I divide Sweden into four hypothetical countries – »Sweden w7,« »Sweden w8,« »Sweden w9,« and »Sweden w10« – based on each region's winter holiday schedule. Table 2 provides descriptive statistics for these four hypothetical countries.

13. Source: https://ourworldindata.org/explorers/covid?zoomToSelection=true&time=2020-06-30&facet=none&country=DNK~FIN~NOR~SWE~DEU~European+Union~ITA~ESP~BEL~NLD~GBR&pickerSort=asc&pickerMetric=location&hideControls=true&Metric=Confirmed+deaths&Interval=Cumulative&Relative+to+population=true

For context, the populations of Denmark, Norway, and Finland in 2020 were about 5.8 million, 5.5 million, and 5.5 million, respectively. Thus, three of the four hypothetical countries are slightly smaller than their Scandinavian neighbors but still exceed the populations of, for instance, the Baltic states.

Table 2
Basic statistics for hypothetical countries

Hypothetical country	Population 2020	Cumulative COVID-19 mortality per 100,000 by week 26, 2020
Sweden w7	2.100.000	44
Sweden w8	3.790.000	35
Sweden w9	3.590.000	79
Sweden w10	900.000	30
Sweden	10.380.000	52

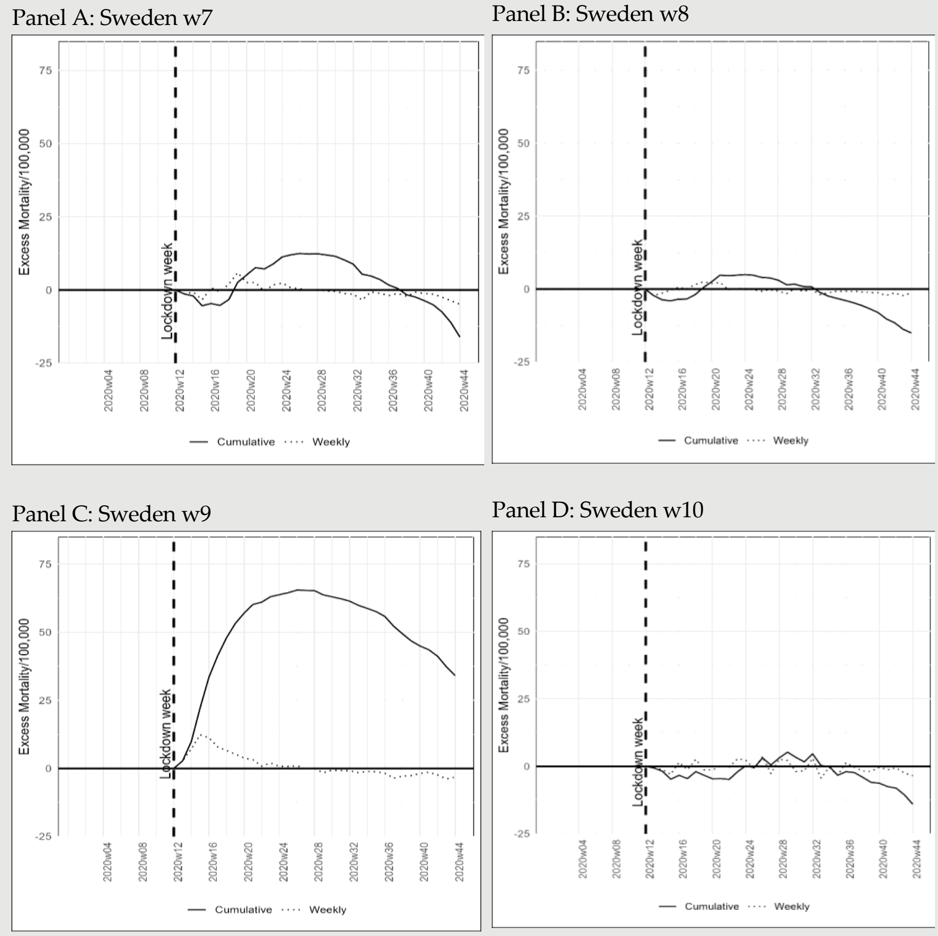
Source: SCB Statistics and winter holiday data from Arnarson (2021)

All four hypothetical countries maintained the same COVID-19 policies, notably abstaining from lockdowns. If the decision against lockdowns were an important cause of excess mortality, we would expect similar excess mortalities in all four hypothetical countries. Contrary, if lockdowns had little effect and the differences in pandemic outcomes were driven by unobserved spread of COVID-19 in societies before lockdown, we would expect to find large differences in the estimated effect with a large effect in Sweden w9 and possibly Sweden w10 and smaller or no effects in Sweden w7 and Sweden w8.

As I lack specific data on variables such as hospital beds and urban population for each of the four hypothetical countries, I apply the corresponding Swedish data uniformly to all four cases. Consistent with my main analysis, the pre-intervention period begins in week 27 of 2016; however, I concentrate on the short-run period prior to the second COVID-19 wave and therefore end the post-intervention analysis in week 44 of 2020.

Figure 4 presents the outcomes of the analysis, demonstrating that the estimated effect of a Swedish lockdown varies considerably depending on the timing of the winter holiday. In the short run, the estimated effect is substantial in Sweden w9, but notably smaller in Sweden w7, Sweden w8, and Sweden w10.

Figure 4
Post-intervention excess mortalities for hypothetical countries



Note: The vertical dashed line illustrates the timing of the lockdown in synthetic Sweden. Excess mortalities have been normalized to zero in week 12 of 2020. Donor weights can be found in Appendix B.

The findings in Figure 4 suggest that the results from my analysis of Sweden as a whole (and from other SCM-based studies) are likely influenced by unobserved factors, such as the extent of COVID-19 transmission prior to lockdown. The seemingly large effect in Sweden w9 appears to confirm the significance of the winter holiday's timing, as demonstrated by Arnarson (2021) and Björk et al. (2021).

Another contributing factor is the significant variation in donor weights across different regions. For instance, Belgium contributes exclusively to the synthetic versions of Sweden w7 (21.9 %) and Sweden w10 (20.5 %) – not to Sweden w8 and Sweden w9. In my work, I have observed that seemingly minor adjustments to the model, such as altering the pre-intervention window, markedly shift donor weights. Given that Belgium by far experienced the highest death rate during the first wave among donor countries, these seemingly arbitrary and unstable donor weight assignments can have a large impact on the estimates. This underscore fundamental problems with SCM.

6. Controlling for the spread of COVID-19 before lockdown

The distinct estimates obtained for the four hypothetical countries suggest that pre-lockdown COVID-19 transmission may be a critical factor in understanding the effectiveness of lockdown measures. Including this factor in the analysis could potentially improve the model's accuracy.

To capture pre-lockdown spread, I use a simple measure of excess deaths in the weeks immediately following the lockdown.[14] As detailed by Herby et al. (2024)[15], the average time from infection to death is approximately 20–26 days, with 29–42 % of deaths occurring in the third week after infection (day 15–21) and roughly 25 % in the fourth week (day 22–28). Consequently, the excess death rate in the weeks after lockdown may serve as a proxy for pre-lockdown COVID-19 transmission.[16]

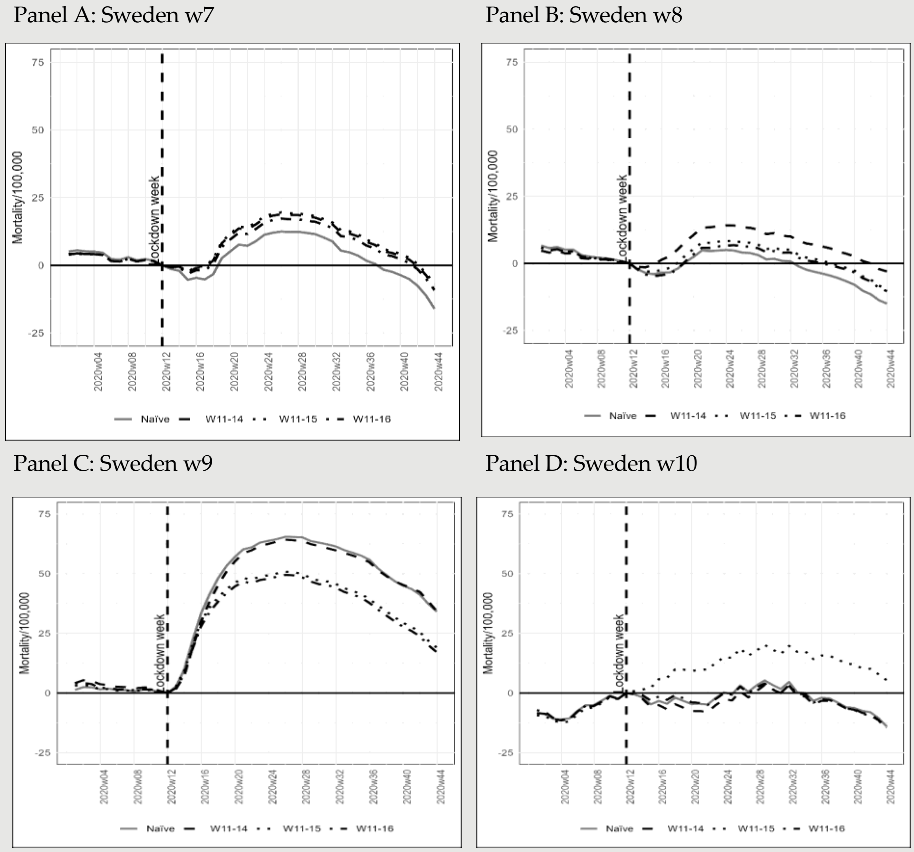
I define the excess death rate in a given week X as the death rate in week X minus the average death rate in weeks 8 to 10. Because of the variability in the infection-to-death lag, multiple post-lockdown weeks can arguably be used as a proxy for the pre-lockdown spread of COVID-19. Accordingly, I run the SCM under three different specifications, using weeks 11–14, 11–15, and 11–16 of 2020 as proxies.

14. Another way to do this would be to include the timing of the winter holiday in the model. However, mortality data at the NUTS3 level, available only annually from Eurostat, lacks the necessary granularity for this purpose. Furthermore, although the timing of the winter holiday appears to significantly influence outcomes in Sweden, as previously discussed for Norway and Finland, it may not reliably proxy the pre-lockdown spread of COVID-19 in other countries.
15. See the appendix in Supplementary file2.
16. The time lag between infection and death appears to be misrepresented in parts of the literature. For example, Ege et al. (2023) suggest that their »results are comparable to more recent studies using COVID-19 deaths as outcomes, where effects are seen within a week after the introduction of lockdown in synthetic Sweden.« However, it is implausible for lockdowns to yield visible impacts on mortality within such a short timeframe. A more likely explanation is that the underlying model is misspecified.

Figure 5 presents the estimated effects of a counterfactual lockdown for the four hypothetical countries when including the proxy for pre-lockdown COVID-19 transmission. Overall, incorporating this proxy does not substantially alter the estimated lockdown effects, and the magnitude of these estimates still varies across the hypothetical countries.

Although these proxies do narrow the range of estimated effects – shifting from 3–65 avoided deaths per 100,000 in the naïve model (without the proxy) to 8–51 avoided deaths per 100,000 using weeks 11–15 as the proxy – the remaining wide range indicates that the SCM model continues to omit key factors.

Figure 5
Post-intervention excess deaths for hypothetical countries using different proxies for the spread of COVID-19 before lockdown.



Note: Each line in the figure represents the results from one specification – the naïve model and the three models that use different proxy measures. The vertical dashed line illustrates the timing of the lockdown in synthetic Sweden. Excess mortalities have been normalized to zero in week 12 of 2020. Donor weights can be found in Appendix B.

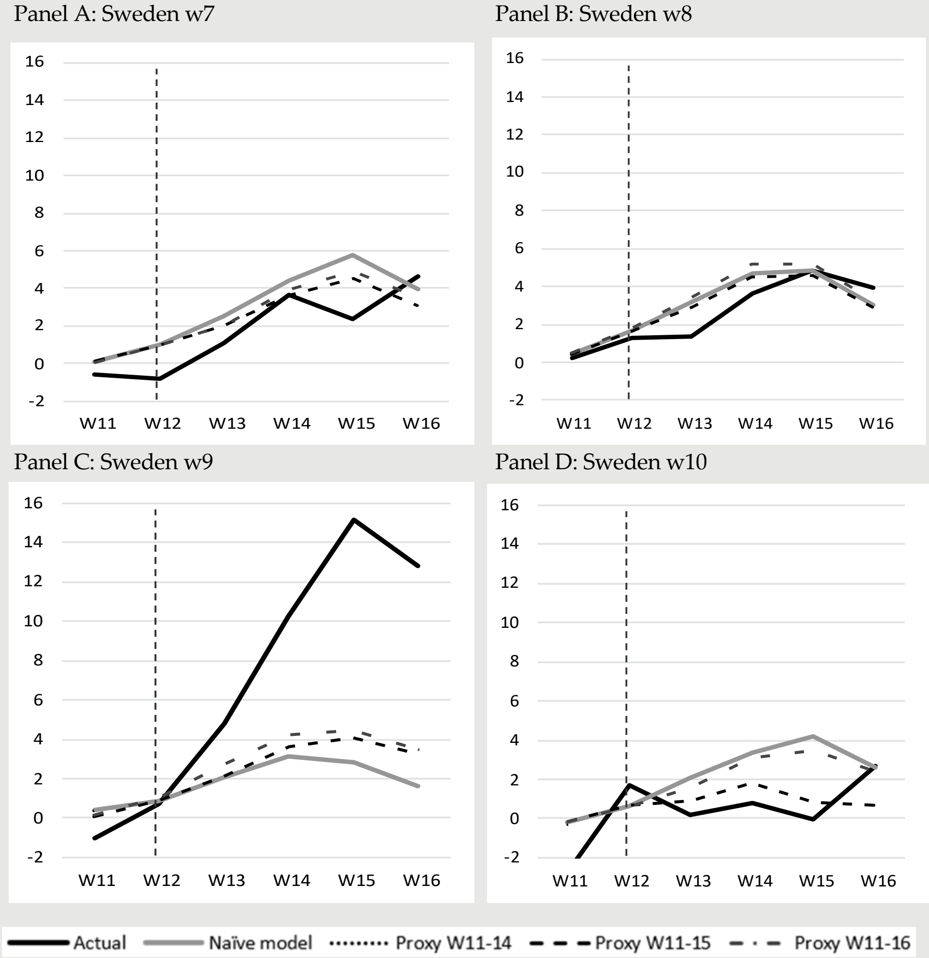
The failure of the SCM models to consistently estimate lockdown effects may stem from various factors. One possibility is that my proxy for pre-lockdown COVID-19 transmission is simply inadequate. Alternatively, other unobservable elements – such as the organization of care homes or differences in how authorities communicate about the virus – could also be critical.

It may also be the case that the pre-lockdown spread of COVID-19 is the primary driver of the observed differences. In my current SCM specification, factors such as GDP per capita, historical mortality rates, and hospital bed availability are treated as equally important as the pre-lockdown spread. Although these aspects are certainly relevant, they may be far less important than the actual level of transmission, which varied substantially across countries at the time of lockdown. For instance, in week 15 of 2020 – before lockdowns could substantially affect mortality – ten times as many people died of COVID-19 in Belgium and two-and-a-half times as many in Sweden compared to Denmark, placing Belgium and Sweden in an entirely different situation unrelated to their lockdown decisions.[17] Thus, if the pre-lockdown spread is the most influential factor, the unsatisfactory SCM results could simply reflect the synthetic controls' inability to match the genuine transmission dynamics in each hypothetical setting.

Figure 6 illustrates these proxies for the real and synthetic versions of the four hypothetical Swedens. While the proxies align reasonably well for Sweden w7, Sweden w8, and Sweden w10, they diverge considerably for Sweden w9 – the only hypothetical country for which the SCM estimates a notable lockdown effect. This sizable gap suggests that the estimated effect for Sweden w9 is not a true consequence of the counterfactual lockdown; rather, it appears to stem from the SCM's failure to adequately account for pre-lockdown COVID-19 spread.

17. https://ourworldindata.org/explorers/covid?tab=chart&time=2020-04-10&pickerSort=asc&pickerMetric=location&hideControls=true&Metric=Confirmed+deaths&Interval=Weekly&Relative+to+population=true&country=DNK~SWE~BEL

Figure 6
Comparing proxies for spread of COVID-19 before lockdown for actual and synthetic hypothetical countries



Note: The figure shows the proxies for the pre-lockdown spread of COVID-19 – measured as 'excess deaths' in the weeks following the lockdown – for both the actual hypothetical Swedens and each synthetic hypothetical Sweden. The vertical dashed line indicates the timing of the counterfactual lockdown in synthetic Sweden.

One might attempt to adjust the SCM to place greater emphasis on the pre-lockdown spread of COVID-19, for instance, by incorporating daily COVID-19 deaths as done by Born et al. (2021). However, this approach reintroduces the primary issue we aim to avoid: the reliance on only a few weeks of noisy data to construct the synthetic control.

The challenge of a very short pre-intervention window also brings to the forefront a seemingly simple yet complex question: »When did the lockdown in donor countries begin?« Taking Denmark as an example, the timeline of lockdown measures is multifaceted. Governmental indoor cultural institutions, libraries, and leisure facilities were closed by Friday, March 13th. Schools followed by closing on Monday, March 16th, and bars and restaurants were closed and public gatherings limited to 10 people by Wednesday 18th, at 10:00 AM. Consequently, pinpointing the exact start date of the Danish lockdown – whether it is Friday the 13th, Monday the 16th, or Wednesday the 18th – is non-trivial. Given the very short pre-intervention window, the choice of lockdown start date can significantly influence the SCM estimates.

7. The limits of the SCM in evaluating the effect of a counter factual lockdown in Sweden

Earlier studies assessing the impact of a counterfactual lockdown in Sweden using COVID-19 data, such as positive cases or deaths, are constrained by very short pre-intervention windows. This limitation arises because countries in the donor pool typically implemented lockdowns within weeks after their first national COVID-19 case was reported. The short pre-intervention periods severely restrict the effectiveness of the SCM.

To address the challenge of limited pre-intervention data, some studies, including the present one, utilize historical mortality rates instead of COVID-19-specific measures. However, this approach has its own drawbacks. Historical mortality rates may not accurately predict how countries were impacted by COVID-19 during the initial weeks of the pandemic. For example, in Arnarson (2021) and Björk et al. (2021) demonstrate that the timing of winter holidays significantly influenced the spread of COVID-19 in Europe, whereas such timing is unlikely to have a meaningful effect on pre-pandemic mortality rates. –

Leveraging the timing of winter holidays as an instrument for pre-lockdown COVID-19 spread, I divide Sweden into four hypothetical Swedens – »Sweden w7,« »Sweden w8,« »Sweden w9,« and »Sweden w10« – based on their respective winter holiday weeks (7, 8, 9, and 10). The analysis reveals that the estimated effect of a counterfactual lockdown varies substantially across these hypothetical regions. This large discrepancy suggests that differences in excess mortality are less attributable to the absence of a lockdown policy and more to unobserved variables, such as the extent and spread of COVID-19 in societies prior to lockdown.

To further investigate, I expand the naïve SCM model by incorporating excess deaths after lockdown as a proxy for pre-lockdown COVID-19 spread. While this adjustment reduces the discrepancy in the estimated effects—shifting the range from 3–65 avoided deaths per 100,000 in the naïve model to 8–51 avoided deaths per 100,000 when using weeks 11–15 as the proxy—the remaining wide range indicates that the SCM model still fails to account for all relevant factors. This inadequacy is likely due to either unobserved variables or the critical role of pre-lockdown COVID-19 spread, for which only a few weeks of data are available.

These findings imply that the estimated effects of lockdowns in other SCM studies using Sweden as a control are at least partially driven by unobserved variables and the short pre-intervention matching periods. Consequently, the reliability of SCM estimates in this context is questionable, highlighting the need for more robust methods or additional data to accurately assess the impact of lockdown policies.

7.1. Causality problems are minor

Another potential limitation is the potential causality problem. If increasing infection rates lead governments to introduce lockdown policies, and declining infection rates subsequently lead them to ease lockdowns, the estimated association between policy stringency and mortality may be biased. It is essential for the SCM that the treatment is the only important factor that differs between treatment and control group in the post-treatment period. This may not be satisfied if there is a great degree of reverse causality and may explain why I do not find any effect of lockdowns.

However, for a number of reasons I believe that this potential bias is likely to be very small.

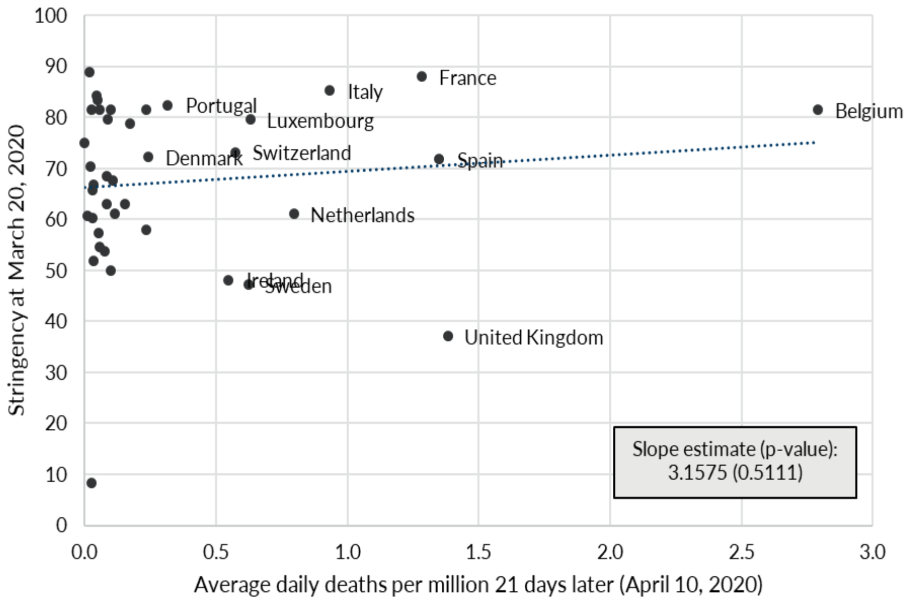
In the beginning of the pandemic there was very little information available to governments because of very limited testing capacities. This means that governments could not react based on information about infections. Rather, Sebhatu et al. (2020) show that government policies were strongly driven by the policies initiated in neighboring countries rather than by the severity of the pandemic in their own countries. In short, Sebhatu et al. show that it is not the severity of the pandemic that drives the adoption of lockdowns, but rather the propensity to copy policies initiated by neighboring countries.[18] Similar results are found by Engler et al. (2021) and Mistur et al. (2023), while Bjørnskov and Voigt (2022) show that the use of lockdown was influenced by the additional emergency powers granted by the constitution.

These insights are illustrated in Figure 7 below which shows the correlation between the spread of infection on 20 March 2020 (proxied by the daily death toll three weeks later) and the degree of lockdown. The figure shows that

18. Sebhatu et al. (2020) finds a statistically significant correlation between COVID-19 mortality and the degree of lockdown in a country, but the magnitude is modest, explaining only 2.1 stringency points on average. In comparison, the gap between the strictest and most lenient lockdowns in Europe was between 67 and 92 stringency points in the period from 16 March to 15 April 2020), c.f. Herby et al. (2024).

there was a positive correlation, so that countries that had a high spread of infection also shut down more strictly. But the correlation is very small and by no means statistically significant.

Figure 7
There was no correlation between the spread of infection and the degree of lockdown in the spring of 2020



Source: Our World in Data (2022)
Note: The figure shows the correlation between the strictness of the lockdown on March 20, 2020 measured by the Stringency index and the average number of deaths per day 21 days later (as a proxy for the spread of infection on March 20, 2020). The average number of deaths is calculated as a seven-day average.

According to Sebhatu et al. (2020), this kind of imitation is common among decision-makers when they are unsure about the effects of a decision. By doing what everyone else does, decision-makers can avoid being criticized for being too slow or not responding at all (which Sweden received a lot of criticism for). In addition, politicians can adopt rules that are similar to those of other countries to signal that they are part of the »normal«.

The findings of Sebhatu et al. (2020), Engler et al. (2021), and Mistur et al. (2023) indicate that politicians and the general public did not primarily respond to the pandemic's progression within their own countries but instead based their actions on shared international knowledge. Countries such as Denmark, Sweden, Norway, and other Northern European nations accessed similar information around the same time, which explains why their governments chose to implement lockdowns and why their citizens became more cautious almost simultaneously, despite being at different stages of the pandemic curve. For example, Norway and Belgium reported their first COVID-19 deaths and initiated lockdowns around the same period, with a comparable number of cases per 100,000 people. However, three weeks later, Belgium had recorded nearly ten times as many COVID-19 deaths per capita as Norway, highlighting that the circumstances during the lockdown were markedly different between the two countries.

The reliance on international knowledge complicates efforts to align pandemic curves by matching countries based on specific milestones, such as the occurrence of one COVID-19 death per 100,000 people, as employed by Born et al. (2021). Shifting the pandemic curves in this manner can distort the shared international knowledge, leading to discrepancies between countries. For instance, Born et al. (2021) align March 11, 2020, in Sweden with March 30, 2020, in Finland. However, the information and understanding available to the populations of these two countries on those dates were significantly different, undermining the validity of such temporal alignments.

8. Conclusion

Contrary to Denmark, Finland, and Norway, Sweden experienced significant COVID-19 mortality in the spring of 2020 without enforcing a mandatory lockdown. These contrasts led many to conclude that the absence of a Swedish lockdown resulted in a larger than necessary spread of the virus and consequent excess mortality. Numerous studies utilizing the SCM have supported this initial assumption. However, as argued herein, these studies may possess critical flaws, notably the short pre-intervention windows and the unaccounted spread of COVID-19 prior to lockdown measures.

These methodological shortcomings have largely been overlooked, despite the fact that the results of many SCM-based studies should have raised significant concerns. For instance, Ege et al. (2023) observe that »excess deaths in Sweden and synthetic Sweden start to diverge immediately« following the lockdown. This finding is epidemiologically implausible, given that the period from infection to death typically spans at least two to three weeks (Herby et al. (2024).[19] Such immediate divergence suggests potential model misspecification.

In the present study, I address two potentially critical flaws in existing SCM literature: the reliance on very short pre-intervention periods (sometimes limited to a few weeks) and the exclusive focus on short-term lockdown effects. Specifically, I extend the pre-intervention window substantially, utilizing mortality rates

19. See the appendix in Supplementary file2.

from week 27 of 2016 to week 12 of 2020. This period is significantly longer than those used in prior studies. Additionally, I expand the post-intervention window to week 26 of 2021, encompassing the period when vaccines were distributed to the most vulnerable populations, thereby effectively concluding the severe phase of the pandemic. This extended post-intervention period allows for a more comprehensive analysis of the potential trade-offs between short-term benefits and long-term costs of lockdowns.

In my naïve model I estimate that a counterfactual lockdown would have reduced mortality rates by approximately 22 deaths per 100,000 in the short run (by week 26 of 2020), a figure comparable to previous studies employing similar methodologies. Conversely, in the longer run (by week 26 of 2021), a counterfactual lockdown would have increased mortality by over 19 deaths per 100,000. Importantly, both the short-term and long-term estimates are statistically insignificant.

Using the timing of the winter holiday as an instrument for the spread of COVID-19 prior to lockdown I find large discrepancies in the estimated effects across regions with different holiday timings. These discrepancies suggest that differences in excess mortality are less attributable to the absence of a lockdown policy and more to unobserved variables, such as the extent and spread of COVID-19 before lockdown measures were implemented. This finding indicates that the lockdown effects identified in other SCM studies using Sweden as a control are at least partially driven by unobserved variables.

To proxy the spread of COVID-19 before lockdown, I utilize excess death rates in the weeks following the (counterfactual) lockdown. While this adjustment slightly improves the consistency of the results, the remaining large discrepancies across regions with different winter holiday timings indicate inherent problems with the SCM when applied to assessing the effects of a counterfactual lockdown in Sweden. These issues are likely related to the very short pre-intervention period available for capturing the spread of COVID-19.

Consequently, one of the critical takeaways from this study is the need for caution when employing the SCM to evaluate the effects of lockdowns. Although SCM is a powerful and intuitive tool, it may inadvertently fail to account for critical omitted variables in a COVID-19 context, thereby skewing the interpretation of the true impact of lockdown measures.

Many SCM studies tend to focus on outliers with notably high COVID-19 mortality rates during the first wave. Herby et al. (2024) report that at least seven out of eleven eligible SCM studies targeted regions that experienced substantial deaths early in the pandemic.[20] This selective focus likely introduces selection bias, which could lead to misleading conclusions about the general effectiveness of

20. See Supplementary file2 in Herby et al. (2024).

interventions. In contrast, Mader and Rüttenauer (2022) utilized the Generalized Synthetic Control Method (GSCM) across a broader dataset of 169 countries, effectively minimizing such biases. Their findings reveal no significant or consistent mitigating effects of non-pharmaceutical interventions (NPIs) on COVID-19-related deaths per capita. This disparity underscores the importance of employing diverse methodologies and comprehensive datasets to avoid skewed interpretations and to better understand the true impact of public health measures.

Despite these limitations, several scholars have employed SCM to estimate the effect of a counterfactual lockdown in Sweden. This raises a pertinent question: would researchers be willing to generate multi-week forecasts for other time series, such as business turnover, based on merely 15–20 days of noisy data with considerable uncertainty regarding when businesses altered their strategies? The analogy underscores the potential fragility and questionable reliability of SCM estimates when based on limited and noisy pre-intervention data.

9. Appendices

9.1. Appendix A – Using cumulative excess mortalities in Ege et al. (2023)

The analysis in Ege et al. (2023) of a counterfactual lockdown in week 11 of 2020 starts in week 44 of 2019, where the cumulative excess mortality equals the excess mortality for that week, effectively setting week 43 as the zero point for accumulation. Figure 8 below presents the results from Ege et al., alongside an identical SCM where the only change is setting week 11 of 2020 as the zero point for accumulation. This adjustment is made by subtracting the value in week 11 of 2020 from all prior and subsequent weeks by country.

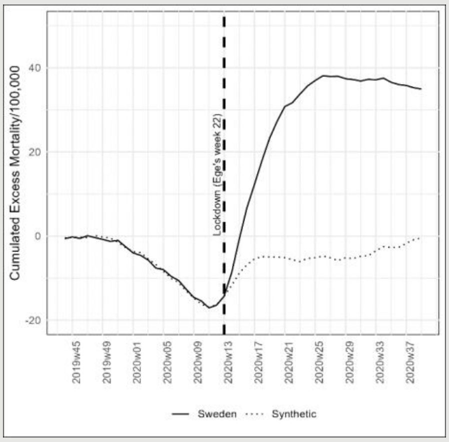
The top panels of Figure 8 illustrates the result when the lockdown week is set to week 13 of 2020 (the main specification in Ege et al.), while the bottom panels row shows the result when the lockdown week is set to week 11 of 2020 (the secondary specification). The first column, panel A, presents the original results from Ege et al., and the second column, Panel B, shows the results when the zero point for accumulation is shifted to week 11 of 2020.

Figure 8 demonstrates that adjusting the zero point for accumulation significantly alters the results markedly. In the main specification (Panel A), the estimated effect of the lockdown is almost halved, and in the secondary specification (Panel B), the effect of the lockdowns disappears entirely.

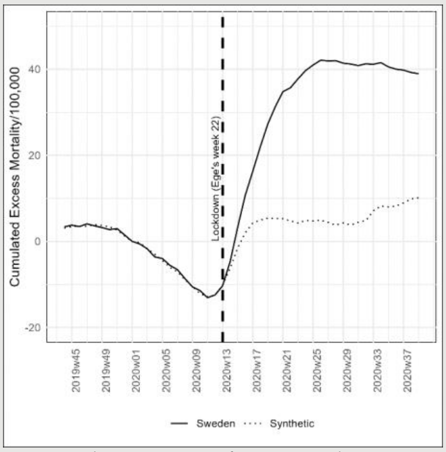
Figure 8

Selecting Other Weeks as Zero Points for Cumulative Excess Mortality in Ege et al. (2023)

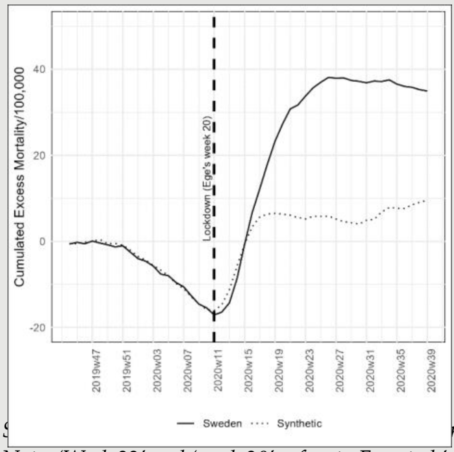
Panel A1: 'Week 22' result from Ege et al. (2023) with week 43 of 2019 as the zero point for accumulation



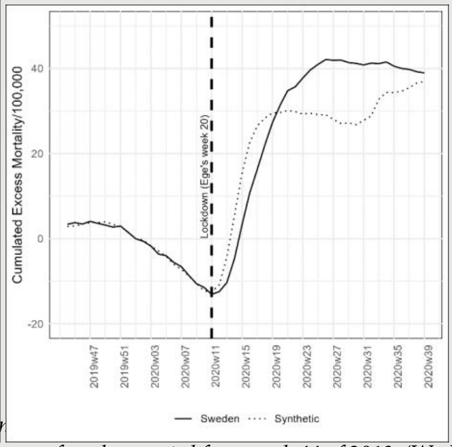
Panel B1: 'Week 22' result when week 1 of 2020 is the zero point for accumulation



(2023) with week 43 of 2019 as the zero point for accumulation



2020 is the zero point for accumulation



Note: 'Week 22' and 'week 20' refers to Ege et al.'s use of weeks counted from week 44 of 2019. 'Week 22' corresponds to week 13 of 2020 while 'week 20' corresponds to week 11 of 2020. The results in Panel B1 differ from the results presented in Ege et al. (2023) Figure 1(e). It is not clear why there is a discrepancy between the figure in the paper and the figure created by the R-code provided by Ege et al. My points hold, nevertheless.

9.2. Appendix B –Donor weights from the SCM model

Table 3
Donor weights for the counterfactual synthetics in the naïve SCM model

Donor country	Sweden	Sweden w7	Sweden w8	Sweden w9	Sweden w10
FIN	25.0 %	29.1 %	23.0 %	37.6 %	21.3 %
CAN	29.2 %	41.8 %	21.3 %		
NLD	31.6 %		31.6 %	14.8 %	
DNK	10.4 %		23.6 %		23.3 %
BEL		21.9 %			20.5 %
ISR				25.7 %	
BGR					25.0 %
LUX		5.2 %			7.8 %
CHE				15.6 %	
Others	3.8 %	2.0 %	0.5 %	6.3 %	2.1 %

Note: Only weights larger than 5 % are shown in the table.

Table 4
Donor weights for the counterfactual synthetics using mortality rates in week 11 to 14 as proxies for the spread of COVID-19 before lockdown

Donor country	Sweden w7	Sweden w8	Sweden w9	Sweden w10
FIN	25.4 %	21.1 %	24.7 %	14.5 %
CAN	33.1 %	9.2 %		14.4 %
DNK	15.2 %	37.8 %		12.7 %
CHE		11.4 %	47.4 %	6.4 %
BEL	17.1 %			24.4 %
NLD		15.3 %		
BGR				26.9 %
KOR			18.7 %	
LUX		5.3 %		
ESP			6.6 %	
Others	9.2 %	5.2 %	9.2 %	0.7 %

Note: Only weights larger than 5 % are shown in the table.

Table 5

Donor weights for the counterfactual synthetics using mortality rates in week 11 to 15 as proxies for the spread of COVID-19 before lockdown

Donor country	Sweden w7	Sweden w8	Sweden w9	Sweden w10
CAN	31.5 %	11.3 %	65.1 %	
DNK	15.6 %	33.2 %		36.8 %
NLD		30.4 %	22.0 %	
FIN	26.4 %	19.5 %	8.2 %	
BGR				30.0 %
BEL	15.0 %			
CHE				28.1 %
ISL		5.6 %		
Others	11.5 %	0.0 %	4.7 %	5.1 %

Note: Only weights larger than 5 % are shown in the table.

Table 6

Donor weights for the counterfactual synthetics using mortality rates in week 11 to 15 as proxies for the spread of COVID-19 before lockdown

Donor country	Sweden w7	Sweden w8	Sweden w9	Sweden w10
CAN	38.2 %	9.9 %	60.5 %	20.8 %
FIN	23.4 %	15.6 %	9.2 %	13.1 %
NLD		36.8 %	19.7 %	
DNK	13.4 %	32.4 %		9.9 %
BEL	18.1 %			14.9 %
BGR				22.0 %
ISL		5.3 %	7.0 %	
PRT				8.7 %
LVA				5.3 %
Others	6.9 %	0.0 %	3.6 %	10.6 %

Note: Only weights larger than 5 % are shown in the table.

10. Supplementary material

R-code available at https://github.com/JonasHerby/Dont-Jump-to-Faulty-Conclusions

References

Abadie, Alberto. 2021. »Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects.« *Journal of Economic Literature* 59 (2):391–425. https://doi.org/10.1257/jel.20191450.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. »Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program.« *Journal of the American Statistical Association* 105 (490). Taylor & Francis:493–505. https://doi.org/10.1198/jasa.2009.ap08746.

– 2015. »Comparative Politics and the Synthetic Control Method.« *American Journal of Political Science* 59 (2). John Wiley & Sons, Ltd:495–510. https://doi.org/10.1111/ajps.12116.

Abadie, Alberto, and Javier Gardeazabal. 2003. »The Economic Costs of Conflict: A Case Study of the Basque Country.« *American Economic Review* 93 (1):113–32. https://doi.org/10.1257/000282803321455188.

Arnarson, Björn Thor. 2021. »How a School Holiday Led to Persistent COVID-19 Outbreaks in Europe.« *Scientific Reports* 11 (1):24390. https://doi.org/10.1038/s41598-021-03927-z.

Björk, Jonas, Kristoffer Mattisson, and Anders Ahlbom. 2021. »Impact of Winter Holiday and Government Responses on Mortality in Europe during the First Wave of the COVID-19 Pandemic.« *European Journal of Public Health* 31 (2):272–77. https://doi.org/10.1093/eurpub/ckab017.

Bjørnskov, Christian, and Stefan Voigt. 2022. »This Time Is Different? – On the Use of Emergency Measures during the Corona Pandemic.« *European Journal of Law and Economics* 54 (1):63–81. https://doi.org/10.1007/s10657-021-09706-5.

Born, Benjamin, Alexander M. Dietrich, and Gernot J. Müller. 2021. »The Lockdown Effect: A Counterfactual for Sweden.« *PLOS ONE* 16 (4). Public Library of Science:e0249732. https://doi.org/10.1371/journal.pone.0249732.

Cho, Sang-Wook (Stanley). 2020. »Quantifying the Impact of Nonpharmaceutical Interventions during the COVID-19 Outbreak: The Case of Sweden.« *The Econometrics Journal* 23 (3):323–44. https://doi.org/10.1093/ectj/utaa025.

Conyon, Martin J., and Steen Thomsen. 2021. »COVID-19 in Scandinavia.« https://doi.org/10.2139/ssrn.3793888.

Ege, Florian, Giovanni Mellace, and Seetha Menon. 2023. »The Unseen Toll: Excess Mortality during Covid-19 Lockdowns.« *Scientific Reports* 13 (1):18745. https://doi.org/10.1038/s41598-023-45934-2.

Engler, Sarah, Palmo Brunner, Romane Loviat, Tarik Abou-Chadi, Lucas Leemann, Andreas Glaser, and Daniel Kübler. 2021. »Democracy in Times of the Pandemic: Explaining the Variation of COVID-19 Policies across European Democracies.« *West European Politics* 44 (5–6). Routledge:1077–1102. https://doi.org/10.1080/01402382.2021.1900669.

Ferguson, Neil M, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, et al. 2020. »Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID- 19 Mortality and Healthcare Demand«, March, 20.

Flaxman, Seth, Swapnil Mishra, Axel Gandy, H. Juliette T. Unwin, Thomas A. Mellan, Helen Coupland, Charles Whittaker, et al. 2020. »Estimating the Effects of Non-Pharmaceutical Interventions on COVID-19 in Europe.« *Nature* 584 (7820):257–61. https://doi.org/10.1038/s41586-020-2405-7.

Grier, Kevin, and Norman Maynard. 2016. »The Economic Consequences of Hugo Chavez: A Synthetic Control Analysis.« *Journal of Economic Behavior & Organization* 125 (May):1–21. https://doi.org/10.1016/j.jebo.2015.12.011.

Herby, Jonas, Lars Jonung, and Steve H. Hanke. 2024. »Were COVID-19 Lockdowns Worth It? A Meta-Analysis.« *Public Choice*, November. https://doi.org/10.1007/s11127-024-01216-7.

Latour, Chiara, Franco Peracchi, and Giancarlo Spagnolo. 2022. »Assessing Alternative Indicators for Covid-19 Policy Evaluation, with a Counterfactual for Sweden.« *PLOS ONE* 17 (3). Public Library of Science:e0264769. https://doi.org/10.1371/journal.pone.0264769.

Mader, Sebastian, and Tobias Rüttenauer. 2022. »The Effects of Non-Pharmaceutical Interventions on COVID-19 Mortality: A Generalized Synthetic Control Approach Across 169 Countries.« *Frontiers in Public Health* 10 (April): 820642. https://doi.org/10.3389/fpubh.2022.820642.

Mistur, Evan M., John Wagner Givens, and Daniel C. Matisoff. 2023. »Contagious COVID-19 Policies: Policy Diffusion during Times of Crisis.« *Review of Policy Research* 40 (1):36–62. https://doi.org/10.1111/ropr.12487.

Mulligan, Casey B., and Robert D. Arnott. 2022. »The Young Were Not Spared: What Death Certificates Reveal about Non-Covid Excess Deaths.« *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 59 (January). SAGE Publications Inc:00469580221139016. https://doi.org/10.1177/00469580221139016.

Our World in Data. 2022. »COVID-19 Data Explorer.« Our World in Data. 2022. https://ourworldindata.org/coronavirus-data-explorer.

Sebhatu, Abiel, Karl Wennberg, Stefan Arora-Jonsson, and Staffan I. Lindberg. 2020. »Explaining the Homogeneous Diffusion of COVID-19 Nonpharmaceutical Interventions across Heterogeneous Countries.« *PNAS*, August, 202010625. https://doi.org/10.1073/pnas.2010625117.

Statens Serum Institut. 2018. »EPI_NYT Uge 23/24-2018.« June 13, 2018. https://www.ssi.dk/aktuelt/nyhedsbreve/epi-nyt/2018/uge-23_24-2018.

Statsministeriet. 2020. »Pressemøde om COVID-19 den 11. marts 2020.« https://www.stm.dk/presse/pressemoedearkiv/pressemoede-om-covid-19-den-11-marts-2020/.

Tegnell, Anders. 20023. *Tankar efter en pandemi: Och lärdomarna inför nästa*. Natur & Kultur Digital.