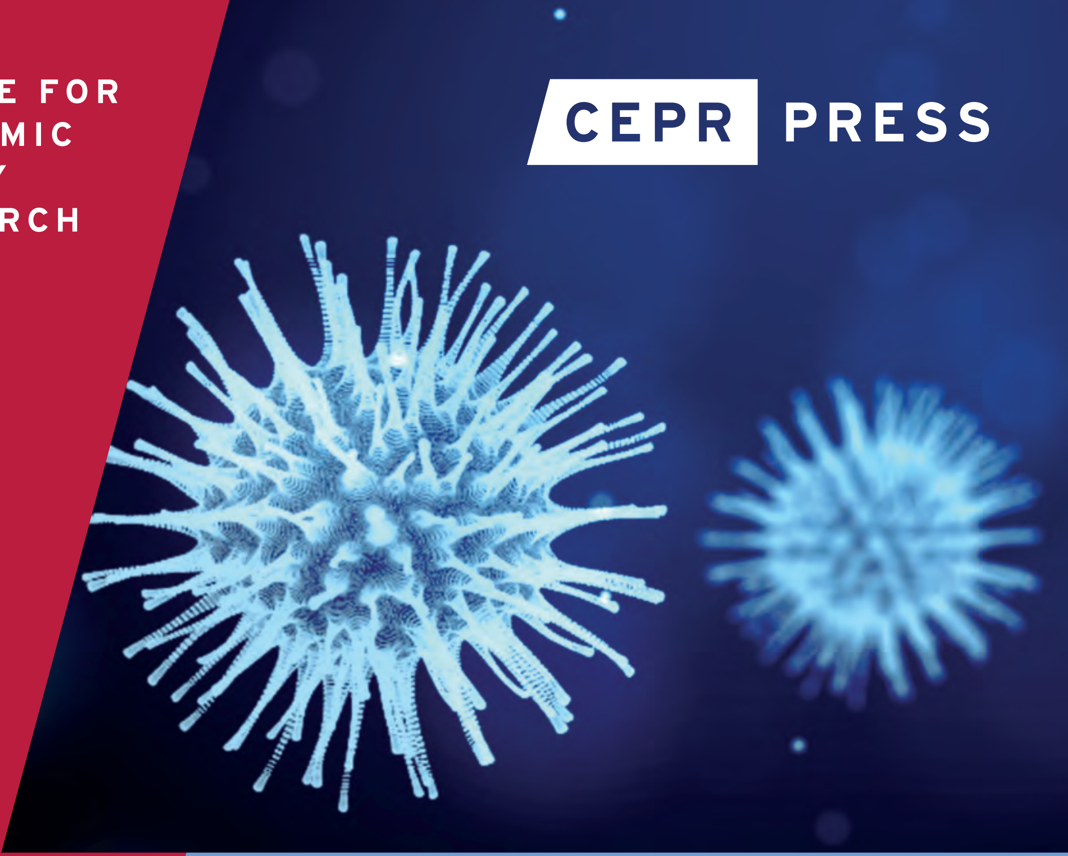
CENTRE FOR
ECONOMIC
POLICY
RESEARCH



CEPR PRESS

COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 35
7 JULY 2020

WILL THE VACCINE BE USED?
Linda Thunström, Madison Ashworth, David Finnoff and Stephen C. Newbold

LIQUIDITY OF FIRMS
Fabiano Schivardi and Guido Romano

THE SWEDISH EXPERIMENT
Sang-Wook (Stanley) Cho

HOW TO MEASURE THE CPI?
Pascal Seiler

THE IMPACT OF MASKS, POLICIES AND BEHAVIOR
Victor Chernozhukov, Hiroyuki Kasahara and Paul Schrimpf

TRANSITIONS IN HEALTH STATUS
Antoine Djogbenou, Christian Gourieroux, Joann Jasiak, Paul Rilstone and Maygol Bandehali

TASK CONTENT AND JOB LOSSES
Filippos Petroulakis

WEATHER AND SOCIAL DISTANCING
Daniel J. Wilson

Covid Economics Issue 35, 7 July 2020

Covid Economics 35, 7 July 2020: 70-95

Quantifying the impact of non-pharmaceutical interventions during the COVID-19 outbreak: The case of Sweden[1]

Sang-Wook (Stanley) Cho[2]

Date submitted: 29 June 2020; Date accepted: 1 July 2020

This paper estimates the effect of non-pharmaceutical intervention (NPI) policies on public health during the recent COVID-19 outbreak by considering a counterfactual case for Sweden. Using a synthetic control approach, I find that strict initial lockdown measures played an important role in limiting the spread of the COVID-19 infection and that Swedish policymakers would have eventually reduced the infection cases by more than half had they followed those policies. As people dynamically adjust their behavior in response to information and policies, the impact of NPIs becomes visible with a time lag of around 5 weeks. An alternative difference-in-differences research design that allows for changes in behavioral patterns also confirms the effectiveness of a strict lockdown policy. Finally, extending the analysis to excess mortality, I find that the lockdown measures would have lowered excess mortality in Sweden by 23 percentage points, with a steep age gradient of more than 30 percentage points for the most vulnerable elderly cohort. The outcome of this study can help policymakers lay out future policies to further protect public health, as well as facilitate an economic plan for recovery.

1 I thank the editor Charles Wyplosz as well as Nicola Aravecchia, Hansoo Choi and Julián P. Díaz for their constructive feedback and comments.

2 Senior Lecturer, University of New South Wales.

1. Introduction

Coronavirus disease 2019 (COVID-19) is a viral respiratory illness caused by a new coronavirus, first reported in Wuhan, Hubei Province, China in November 2019. Over the next few months, the illness rapidly spread to almost every country. In response, the WHO declared COVID-19 a pandemic on March 11, 2020. As vaccines or medicines for COVID-19 have yet to be available, most countries around the world resorted to non-pharmaceutical interventions (NPIs), or community mitigation strategies, to help slow the spread of the illness. Some of the NPIs involve government measures to close schools and workplaces, canceling and restricting public events and gatherings, shutting down public transport and stay-at-home requirements, as well as restrictions on domestic and international travel, not to mention general public information campaigns. By late March, nearly every country in Europe have implemented these policies basically putting themselves into a nationwide lockdown. These government policies remained in place until late May with a gradual easing of some of the harshest measures. One country, however, stood out for its decision to remain open: Sweden. In fact, Swedish officials chose not to implement a nationwide lockdown, trusting that people would voluntarily do their part to stay safe. For example, while high schools and universities have switched to distance learning, elementary and preschools have remained open. In addition, while the government recommended people to stay at home, many non-essential businesses such as restaurants, gyms and bars were still open, while gatherings up to 50 people were allowed. Given this divergence in the policy measures between Sweden and the rest of Europe, I study the public health impact of NPIs by asking how the trajectories of the COVID-19 infection and mortality would have evolved had Sweden opted for more stringent lockdown measures.

In order to study this counterfactual scenario, I first employ the synthetic control method (SCM) pioneered by Abadie and Gardeazabal (2003) and analyze how a parallel (or "synthetic") version of Sweden would have evolved had it enforced a mandatory lockdown policy. This parallel version of Sweden is first constructed through a data-driven process with weights assigned to all possible donor countries that would best approximate the pre-lockdown characteristics of Sweden (our "treatment" unit). Once the policy intervention takes place, we can trace its effect with the evolution of the untreated synthetic control unit to assess the counterfactual situation corresponding to the parallel regime where strict lockdown measures were in place. The causal effect of the lockdown is measured by the post-intervention difference in infection rates of the treatment and the synthetic control unit. It has been shown that the synthetic control method offers several advantages over traditional difference-in-differences or fixed-effect models as not only is the procedure a transparent data-driven one but also it allows the effect of unobservable country heterogeneity to vary over time as discussed by

Covid Economics 35, 7 July 2020: 70-95

Covid Economics 35, 7 July 2020: 70-95

Abadie, Diamond and Hainmueller (2010) and Imbens and Wooldridge (2009). I further quantify the causal effect of counter-COVID measures by using a difference-in-differences (DD) research design that allows for additional variables regarding people's behavior. This would enable us to understand how much of the observed infection rate dynamics is attributed to the effect of NPIs by itself relative to voluntary changes in people's behavior for fear of infection.

The key findings from regression analysis are as follows. I find that the lockdown measures played an important role in limiting the spread of the COVID-19 infection and that Swedish policymakers would have contained the infection cases by more than half had they followed similar policies implemented elsewhere. I also find that as people dynamically adjust their behavior in response to information and policies, the impact of NPIs does not manifest immediately but only with a time lag of approximately five weeks. Profiling excess mortality for the synthetic Sweden, I find that the excess mortality rate in Sweden would have been reduced by approximately 23 percentage points had the policymakers followed strict counter-COVID measures. The effectiveness in death prevention becomes disproportionately higher by age, with more than a 30 percentage point reduction in the excess mortality rate for the elderly cohort aged 85 and above.

This paper contributes to the ongoing discussion on the effectiveness of NPI policy response to the COVID-19 shock, see Chen and Qiu (2020); Gonzalez-Eiras and Niepelt (2020); Ullah and Ajala (2020); Goodman-Bacon and Marcus (2020); Chernozhukov, Kasahara and Schrimpf (2020) and the contributions in the volume by Baldwin and di Mauro (2020). Empirically, this paper extends cross-country experiences in the policy effectiveness. Castex, Dechter and Lorca (2020) shows that the effectiveness of NPIs differ by various socio-economic and public health systems, and the effectiveness of lockdown policies is declining with GDP per capita, population density and surface area; and increasing with health expenditure and proportion of physicians in the population. In terms of scope and methodology, the paper is closest in spirit to Born, Dietrich and Müller (2020), hereinafter BDM, that conducts a similar counterfactual lockdown scenario for Sweden using the synthetic control method. Documenting infection dynamics of one month post-lockdown, they find that the counterfactual Sweden did not differ from actual infection dynamics observed in Sweden. In their discussion, they attribute this outcome to the voluntary precautions taken by the general public that essentially had the same impact as a mandatory lockdown.

This paper extends BDM in the following aspects. First, I consider post-lockdown period extending for two months, which completely covers the time horizon during which the initial lockdown measures were fully in place outside Sweden. Consistent with BDM, I also find that during the first half, the infection dynamics in the counterfactual Sweden was not lower than those in actual Sweden. However, over time, the counterfactual Sweden shows a significant slowdown in the infection rate, which demonstrates that the lockdown measures would eventually have a containment effect in the longer horizon. Second, using a difference-in-differences approach, I formally control for the behavioral changes using Google Mobility Tracker and show that the mandatory lockdown measures would have significantly reduced the infection rate in comparison to a voluntary social distancing scenario.

The rest of the paper is organized as follows. Section 2 describes the methodology and data for the synthetic control approach. Section 3 presents the main estimation results and robustness checks. Section 4 extends the analysis to mortality and discusses the role of voluntary social distancing, followed by a difference-in-differences estimation in Section 5. Finally, conclusion is provided in Section 6.

Covid Economics 35, 7 July 2020: 70-95

2. Data and Methodology

In this section I describe the synthetic control method (SCM) proposed by Abadie and Gardeazabal (2003), and later developed in Abadie, Diamond and Hainmueller (2010) and Abadie, Diamond and Hainmueller (2015). The SCM is a popular approach for comparative case studies, which has also been used to quantify the economic effects of shocks or policy interventions. [1]

Under the synthetic control approach, we can generate a counterfactual designed to capture how the infection rates would have evolved in Sweden had it followed a similar policy approach (or a mandatory lockdown) taken by other European countries. This counterfactual (or synthetic control) unit would track the actual path of infection rates in Sweden (our treatment unit) as closely as possible prior to the policy intervention. After the policy intervention, the control unit followed a path of mandatory lockdown measures while Sweden did not. As such, the notion of policy intervention in our setting refers to the absence of mandatory lockdown measures, or no changes in government policy. Due to difficulties in picking individual countries that satisfy these criteria, we resort to a weighted average of potentially comparable countries that best resemble the characteristics of Sweden prior to the policy intervention. Any discrepancy in the infection dynamics between the two units after the policy intervention can be interpreted as an outcome of the policy or the treatment effect.

As the SCM exploits the pre-intervention data to form better counterfactual values, it is often preferred over other program evaluation methods such as difference-in-differences in comparative case studies.

[1] See Abadie (2020) for a broader overview of the methodology.

2.1. Data

The outcome variable of interest is the infection dynamics as measured by cumulative infection per million population. As for potential donor pool, I select 29 countries consisting of the European Union members (excluding Malta due to lack of data) as well as Iceland, Israel, Norway and Switzerland. As the infection dynamics varied across countries, we normalize the time unit such that "Day 1" refers to the day on which the infection per million exceeds one. For country-specific characteristics, epidemiological studies suggest demographic factors such as population size and the rate of urbanization to be crucial to understand the infection dynamics. I also include population density as it has been found to catalyze the spread of COVID-19 by Rocklöv and Sjödin (2020). The latest available figures for all three country-specific covariates were taken from the World Development Indicators (WDI).

Table 1: COVID-19 and demographic characteristics

Variables	Sweden	All donors (n=29)
COVID-19 dynamics		
- Day 1	29 February	4 March
- Case per million on Day 1	1.18	1.51
- Lockdown day		28 March
- Pre-lockdown duration (days)		23.9
- Case per million on Lockdown day	199.6*	471.4
- Stringency Index (SI) on Lockdown day	32.4*	82.4
Demographics		
- Population (million)	10.1	18.5
- Urban population fraction (%)	87.4	75.1
- Population density	24.7	146.7

Note: Day 1 refers to the date on which the infection per million exceeds one. Lockdown day refers to the date on which the SI index reached the maximum. For reference, * denote the numbers for Sweden on 24 March, 24 days since Day 1.

Next, for the policy intervention, lockdown measures consist of various socioeconomic measures including school and workplace closing, cancellation of public events, restrictions on gatherings, closing of public transport, stay at home requirements, restrictions on domestic/international travels, as well as public info campaigns. As these measures took place over different time with varying magnitudes, we resort to an all-inclusive index measure. The OxCGRT data[2] provides a Government Response Stringency Index (Stringency Index, SI), which ranges from 0 to 100 with each additional government response leading to a higher

[2] https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker.

Covid Economics 35, 7 July 2020: 70-95

Covid Economics 35, 7 July 2020: 70-95

index value.[3] The Stringency Index varies across jurisdictions and time on a daily basis. I pick the date at which this stringency index reached its peak in each country to pinpoint the timing of our policy intervention. Table 1 summarizes the COVID-19 dynamics as well as country specific characteristics for Sweden and the simple average of all 29 donors.

It's worth noting that the COVID-19 infection started a few days earlier in Sweden compared to the average of all donor countries. For the latter group, it took around 24 days for a full lockdown measure was in place when the Stringency Index reached 82.4. For Sweden, on the other hand, the index remained at 32.4 around the same time and never reached higher than 55 during the whole period of our analysis.[4] For the demographic covariates shown in the last three rows, we note that Sweden is characterized by a smaller population with one of the highest urbanization rate and a significantly low population density than the average of the donor group.

Next, we proceed to find the weighted average of the countries in the donor pool which will generate the synthetic control unit for Sweden. The weights are assigned by minimizing the distance between Sweden and the synthetic control unit along all three demographic covariates as well as the average infection rates in the first 20 days since Day 1. Including lagged terms of the dependent variable often helps mitigate the problem of omitting important predictor effects as suggested by Athey and Imbens (2006).

3. Results

Table 2 summarizes the predictor variables for the synthetic Sweden, which is constructed as a weighted average of Finland, France and Norway, with the largest weight assigned to Finland and followed by Norway.[5] Compared to the simple average of all countries in the donor pool (as shown in Table 1), the synthetic control unit provides a much better matched profile of Sweden along the predictors. In other words, the weighted selection of countries seem more appropriate as a control unit than taking a simple average of all countries in the donor pool. In the synthetic control approach, the root mean square prediction error (RMSPE) measures the gap between the variable of interest for the treated country and its synthetic counterpart. The last row of Table 2 reports the RMSPE for the pre-intervention period.

[3] Hale, Webster, Petherick, Phillips and Kira (2020) provides detailed information on the construction of the stringency index.

[4] No single country in the donor pool had the Stringency Index peaking below this maximum value for Sweden.

[5] In the Appendix, Table A1 breaks down the demographic and pre-intervention epidemiological profile while Figure A1 shows the dynamics of the SI index for each country comprising the synthetic unit.

Covid Economics 35, 7 July 2020: 70-95

Table 2: Predictor variables and RMSPE

	Sweden	Synthetic Sweden
		Finland (0.643), France (0.076), Norway (0.281)
Predictors		
- Population (million)	10.099	10.047
- Urban population fraction (%)	87.431	84.126
- Population density	24.718	25.041
- Case per million (first 20 days)	42.830	42.843
RMSPE		8.605

Note: For countries with positive weights, the weights are shown in brackets. All other countries in the donor pool receive zero weight.

Next, in the upper panel of Figure 1, I show the profile of infection dynamics for the synthetic Sweden together with the actual Sweden. I consider a period of 75 days which roughly corresponds to the entire months of March and April as well as the first half of May. As many countries started to gradually ease some of the lockdown measures in late May, the period under observation covers the full period of the initial lockdown. The policy intervention takes places on Day 24—as indicated by the dashed vertical line—which falls on the midpoint of lockdown dates of the three countries comprising the synthetic control unit.

For the first two weeks upon the policy intervention, the cumulative infection cases in the synthetic Sweden follow the actual Sweden quite closely or even higher than the latter.[6] After this period of incubation, there is a divergence in which the actual Sweden follows a much steeper path than its synthetic counterpart. By the end of our sample period on Day 75[7], or roughly 7 weeks after the lockdown intervention, the infection case in Sweden reaches around 2,700. On the other hand, the figures for the synthetic Sweden reaches slightly below 1,300. In other words, Swedish policymakers would have reduced the infection cases by more than half had they followed similar policies implemented elsewhere, which signifies the important role of the lockdown measures in limiting the spread of the COVID-19 infection.

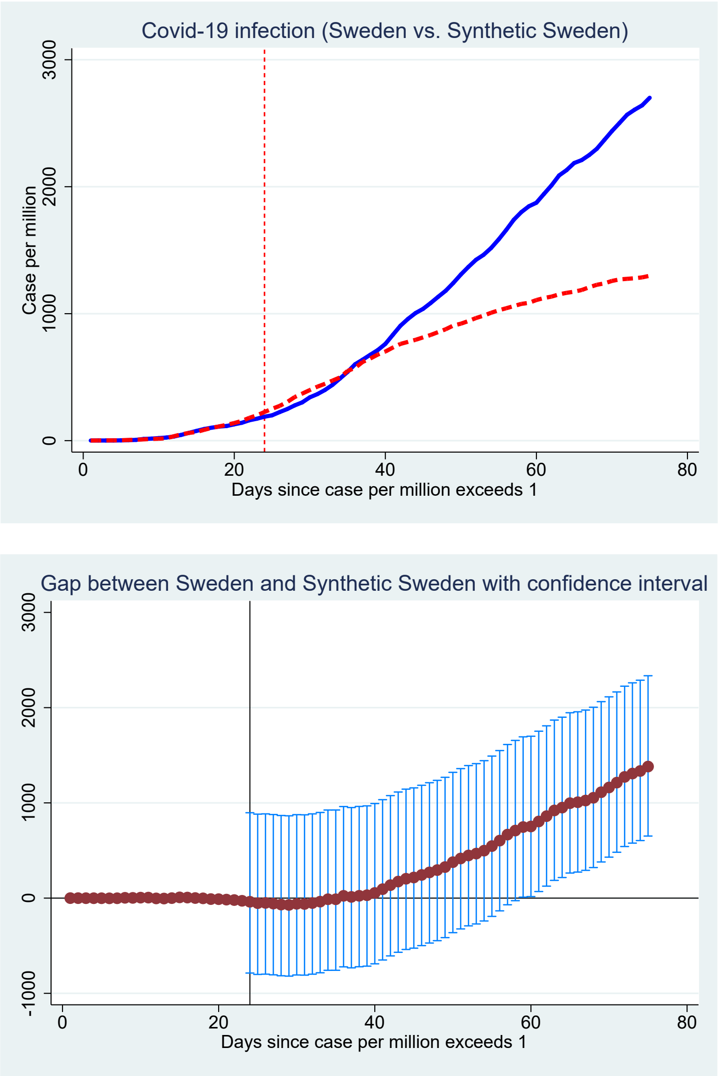
The lower panel of Figure 1 generates a 95% confidence interval for the gap between the two profiles using a methodology proposed by Firpo and Possebom (2018). The gap becomes statistically significant approximately five weeks after the implementation of the lockdown measures. On one hand, this result is consistent with that of Born, Dietrich and Müller (2020), which looks at the first five weeks of the lockdown measures and concludes that the

[6] As an alternative, I convert the outcome variable into logs and show the profile in Figure A2 in the Appendix.

[7] This day corresponds to 13 May.

Covid Economics 35, 7 July 2020: 70-95

Figure 1: Profile of Infection Rates – Sweden vs. Synthetic Sweden



Note: Top panel shows infection case per million population for Sweden (in blue) versus synthetic Sweden (in red dash). Bottom panel shows the gap between the two units with 95% confidence intervals. Vertical line indicates the date of policy intervention.

mandatory lockdown would not have made significant differences in the infection rate in Sweden.[8] However, expanding the horizon over the entire lockdown period, I show that the epidemiological impact of lockdown measures takes places with a time lag and eventually becomes more visible in the longer horizon.

3.1. Robustness Tests and Inference

To evaluate the credibility of the baseline results, I conduct placebo (or falsification) tests based on permutation techniques, as suggested in Abadie, Diamond and Hainmueller (2010). One way the design of the study may influence the outcome comes from the choice of countries in the donor pool with positive weights assigned. If dropping one country from the donor pool creates a large effect on the results without a discernible change in pre-intervention fit, this may require a reexamination if the change in the magnitude of the estimate is caused by the effects of other interventions or by particularly large idiosyncratic shocks on the outcome of the excluded country. As such, I perform a leave-one-out analysis, where I exclude from the sample one-at-a-time each of the three countries that contributes to the synthetic control in the benchmark. For each case, the new list of donors with positive weights as well as the values of the predictors are shown in Table 3. In the benchmark, Finland was the country assigned with the largest weight followed by Norway. Dropping Finland from the donor pool generates a new set of donors consisting of Bulgaria, Croatia and Norway. On the other hand, dropping out France or Norway from the donor list produced exactly the same new donor list consisting of Belgium, Finland and Iceland.

Figure 2 shows the results of a leave-one-out re-analysis. The resulting estimates for the days after the policy intervention (in dashes) are all positive and centered around the result produced under the benchmark. The main conclusion of a positive estimate of the infection rates in Sweden over its counterfactual scenario of a mandatory lockdown is robust to the exclusion of any particular country from the donor list.

Next, I run a cross-sectional placebo test (or "placebo in-space") by sequentially applying the synthetic control algorithm to each country in the pool of potential controls, which generates a distribution of placebo estimates across 29 donors. We then can compare the benchmark estimates of the truly treated economy with this distribution. The cross-sectional placebo tests are shown in Figure 3. The gray lines show the gap in the infection rates between each country in the donor pool and its respective synthetic version. The thick red line depicts the baseline results obtained for Sweden. Visual inspection shows that Sweden

[8]The construction of the synthetic control unit in Born, Dietrich and Müller (2020) differs from mine as they do not include population density as predictors. As such, their synthetic control unit has a population density that is almost ten times larger than that of Sweden.

Covid Economics 35, 7 July 2020: 70-95

CEPR PRESS

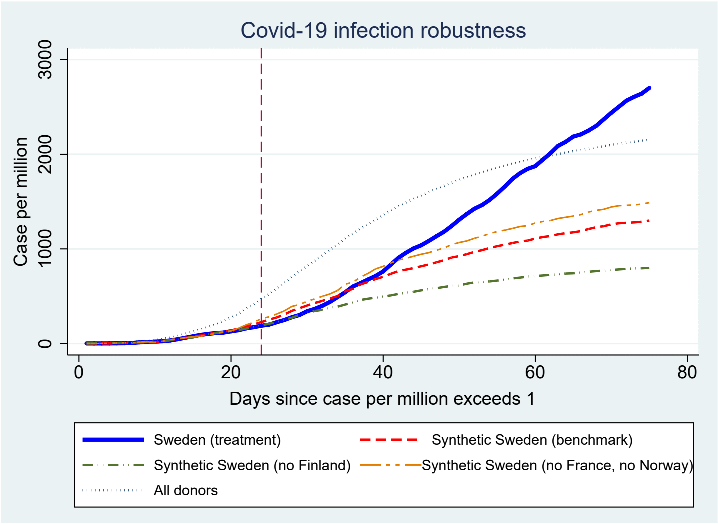
Covid Economics 35, 7 July 2020: 70-95

Table 3: Leave-one-out robustness check

	Sweden	Synthetic Sweden		
		(Benchmark)	(No Finland)	(No France/Norway)
Donors with positive weight		FIN(0.643), FRA(0.076), NOR(0.281)	BGR(0.525)*, HRV(0.088)*, NOR(0.386)	BEL(0.021)*, FIN(0.908), ISL(0.071)*
Predictors				
- Population (million)	10.099	10.047	6.102	5.299
- Urban population fraction (%)	87.431	84.126	76.138	86.246
- Population density	24.718	25.041	46.290	24.596
- Case per million (first 20 day)	42.830	42.843	43.851	42.884
RMSPE		8.605	4.304	14.117

Note: * denotes newly added countries with positive weights (shown in brackets) from each robustness check. Full list of countries in abbreviation are as follows: BEL (Belgium), BGR (Bulgaria), FIN (Finland), FRA (France), HRV (Croatia), ISL (Iceland), NOR (Norway).

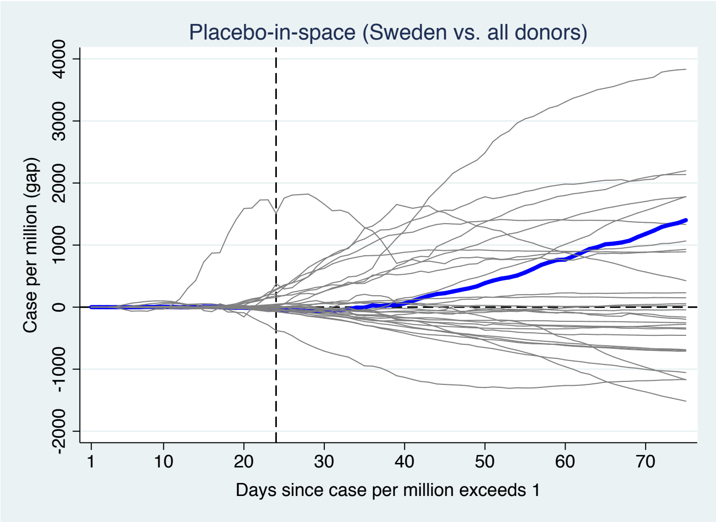
Figure 2: Leave-one-out Robustness Test



Note: New synthetic units from leave-one-out robustness check are plotted in dashes. For reference, the average profile of all donors is plotted in blue dots. Red vertical dashed line indicates the date of policy intervention.

joins in the list of countries with positive treatment effect, but not necessarily at the right tail of the distribution of treatment effects. However, towards the end of the sample period, the treatment effect for Sweden is distinctly higher than most other countries.

Figure 3: Placebo-in-space Robustness Test



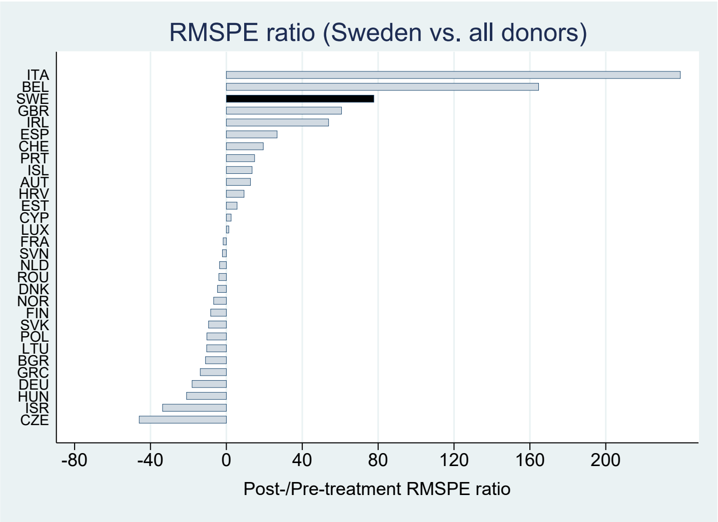
Note: Gap between the treated and synthetic control unit is plotted for Sweden (in thick blue) and each of the donors (in gray). Vertical dashed line indicates the date of policy intervention.

While the previous figure offers a visual evidence of the treatment effects over time, it does not provide a numerical measurement that quantifies the overall significance of the results. To overcome this issue, I follow Abadie, Diamond and Hainmueller (2010), who offer an alternative approach for an inference test by constructing exact p-values based on Fisher (1935). As the root mean square prediction error (RMSPE) measures the gap between the variable of interest for the treated country and its synthetic counterpart, we can calculate a set of RMSPE values for the pre- and post-treatment period for Sweden as well as each country in the cross-sectional placebo test. Countries with negative treatment effect are assigned with a minus sign to their post-treatment RMSPE value. I then compute the country-specific ratio of the post- to pre-treatment RMSPE to quantify the post-treatment divergence in the infection rate, relative to the estimated gap pre-treatment. The distribution of this RMSPE ratio (from highest to lowest) is shown in Figure 4. For Sweden, the RMSPE ratio of around 80 is far higher than those obtained for other countries in the control group. The ranking, converted into fractions, provide the basis for a p-value for Sweden, which measures the probability of observing a ratio as high as the one obtained for Sweden if one

Covid Economics 35, 7 July 2020: 70-95

were to pick a country at random from the list of potential controls. In our case, an exact p-value for Sweden is 0.1 as Sweden ranks third out of 30 countries, which falls within the conventional range of statistical significance.

Figure 4: RMSPE ratio



Note: Countries where the post-treatment infection cases consistently fall below those of its synthetic unit are shown with a minus sign.

4. Discussion

4.1. Infection to Mortality

So far, the focus of the analysis has been the rate of infection. One caveat of our analysis using infection cases to assess the impact on the spread of COVID-19 is that Sweden conducted very little testing compared to other countries. As the infection cases depend on the number of testing, this most likely underestimates the true treatment effect. While this issue is hard to resolve, one could take a look at the rate of mortality from COVID-19, and compare how the NPIs impacted the rate of death during the COVID-19 crisis.

While national health protection agencies report daily death counts, some jurisdictions include both confirmed and probable cases and deaths while others only report confirmed cases. As such, daily reported figures for deaths are difficult to compare across countries. Instead, I use excess mortality rate—the ratio of numbers of deaths over and above the historical average between 2015 and 2019—as a more reliable source of information for comparison. The Short-Term Mortality Fluctuation data series (STMF) from Human Mortality

Covid Economics 35, 7 July 2020: 70-95

Covid Economics 35, 7 July 2020: 70-95

Database[9] offers weekly death counts by age groups and sex for 22 countries including Sweden as well as the countries assigned with positive weights in the construction of synthetic Sweden in Section 3: Finland, France and Norway. This allows me to generate weekly excess mortality for the synthetic Sweden and compare that with the profile of actual Sweden, which is shown in Figure 5. For reference, the top panel shows cumulative infection case per million population on a weekly basis, and Week 13 (22-28 March) is the week in which the policy responses began to diverge in the two groups. The bottom panel shows excess death rates for Sweden and its synthetic unit. Prior to Week 13, there is no visible difference in the excess mortality rates between the two groups. Leading to Week 13, however, the excess mortality rate rises much steeper in Sweden and remains consistently higher than its synthetic counterpart. At its peak, the mortality rate in Sweden is more than 40% above its historic average, while the corresponding peak for the synthetic unit is around 15%. On average, as summarized in Table 4, the excess death rate over the 10 weeks post-intervention period in Sweden is 28.5 percent higher than its historic average. In contrast, the corresponding rate in its synthetic version is 5.4 percent higher than the historic average. In other words, the excess death rate would have been more than 23 percentage points lower had the Swedish policymakers follow similar policies adopted by its parallel counterpart.

As the database provides mortality information by age, I apply the same analysis across different age groups as shown in Figure 6. A visual inspection shows that the gap in excess mortality after the lockdown becomes significantly more pronounced for older age cohorts. As summarized in Table 4, the average post-lockdown gap in excess mortality grows from around 13 percent among working age cohorts to more than 30 percent for the elderly cohort aged 85 and above.

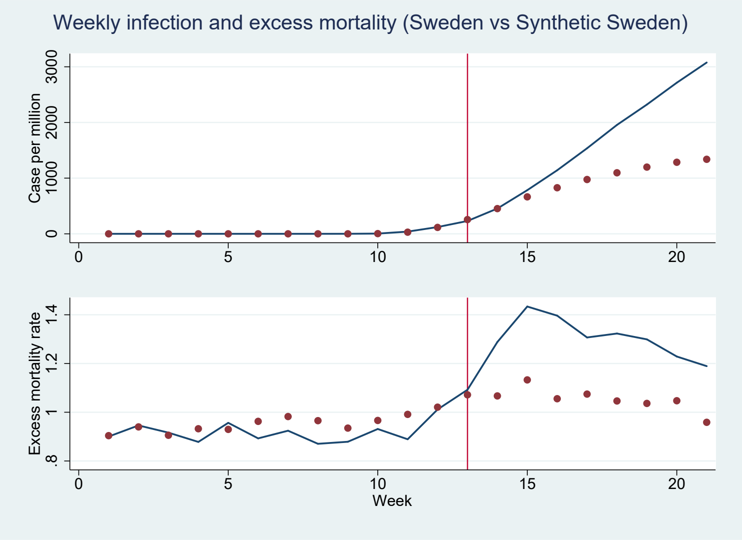
Table 4: Excess mortality pre-lockdown vs. post-lockdown

	Pre-lockdown (Week 1–12)			Post-lockdown (Week 13–22)		
	Sweden	Synthetic Sweden	**Gap**	Sweden	Synthetic Sweden	**Gap**
Total population	0.916	0.953	-0.037	1.285	1.054	0.230
Age 15-64	0.894	0.926	-0.032	1.083	0.951	0.131
Age 65-74	0.919	0.962	-0.043	1.188	1.037	0.150
Age 75-84	0.866	0.872	-0.006	1.246	0.997	0.249
Age 85 plus	0.908	0.916	-0.008	1.330	1.029	0.301

Note: Columns labeled "Gap" measure the difference in excess mortality rates between Sweden and synthetic Sweden during each sub-period. Pre-lockdown period includes the first 12 weeks of 2020 until 21 March, while post-lockdown period includes the latter 10 weeks from 22 March until 31 May.

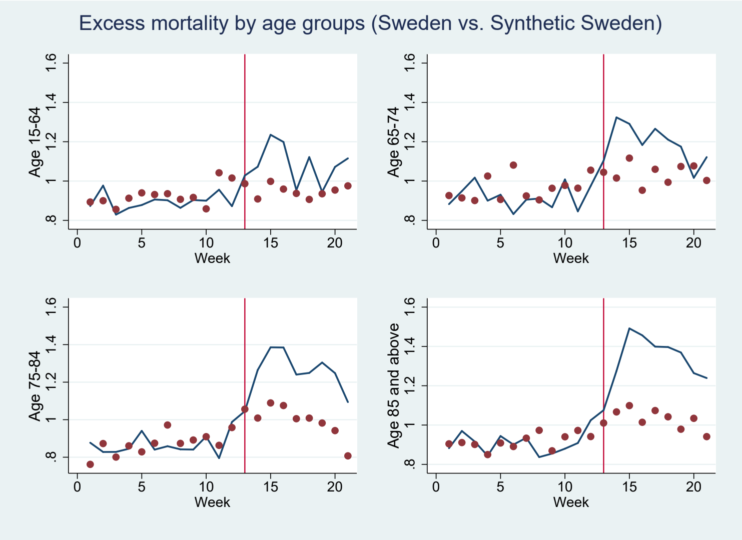
[9]https://www.mortality.org.

Figure 5: Infection to mortality



Note: Horizontal axis shows calendar weeks of 2020. Red vertical line denotes the week of policy intervention in Week 13 (22-28 March). Sweden is shown in blue line, while the synthetic Sweden is shown in red dots.

Figure 6: Excess mortality by age



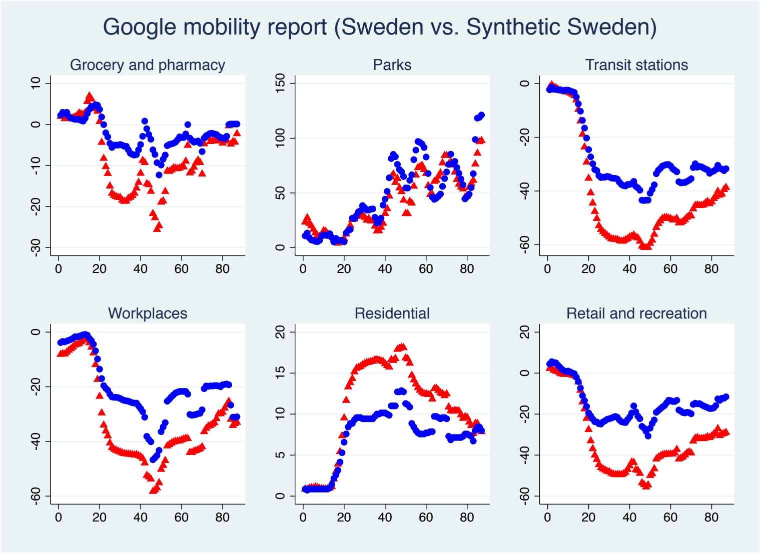
Note: Horizontal axis shows calendar weeks of 2020. Red vertical line denotes the week of policy intervention in Week 13 (22-28 March). Sweden is shown in blue line, while the synthetic Sweden is shown in red dots.

Covid Economics 35, 7 July 2020: 70-95

4.2. Voluntary Social Distancing or Involuntary Lockdown?

Naturally, infection dynamics is not only dependent on the lockdown measures. In fact, there were signs that people were already taking precautionary actions prior to various lockdown measures. For example, even before lockdown measures were announced, people made more trips to grocery stores and pharmacies to stock up on basic necessity items such as toilet papers and disinfectants. On the other hand, while the government allowed many businesses to open, most people in Sweden stayed home or followed social distancing protocols. Born, Dietrich and Müller (2020) speculate that the voluntary social distancing essentially had the same impact as a mandatory lockdown. In fact, using its location services, Google provides mobility trends by geography across different categories of places such as retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential areas.[10] The mobility trend for Sweden and the synthetic Sweden is shown in Figure 7 where the baseline—shown as zero in the vertical axis—is the median value, for the corresponding day of the week, during the 5-week period between 3 January and 6 February, 2020.

Figure 7: Google mobility report



Note: Horizontal axis measures days since 29 February. 7-day moving average for Sweden and synthetic Sweden are shown in blue and red dots, respectively.

A visual inspection of 7-day moving average of the mobility for Sweden and synthetic

[10]https://www.google.com/covid19/mobility/ for direct access of the reports.

Covid Economics 35, 7 July 2020: 70-95

Covid Economics 35, 7 July 2020: 70-95

Sweden shows that before the lockdown, there is a sudden spike in trips to grocery and pharmacy in both units before they fall dramatically. Other categories negatively impacted by the Covid-19 outbreak such as visits to transit stations, workplaces and retail and recreation facilities, there is no anticipatory effects prior to the big drop. For all these categories, post-lockdown period fall is more prominent in the synthetic Sweden than in Sweden, which reflects different behavioral patterns due to voluntary and involuntary lockdown measures. An opposite pattern is shown for home stays. Finally, visits to parks show dramatic rise reflecting the lockdown (but possibly warmer weather) with no visible difference between the two units.

Given that the mobility fell in Sweden as a result of voluntary precaution, one could impute this to the delay in the divergence of infection rate in Sweden relative to its synthetic cohort. However, in the longer horizon, infection rates diverged significantly despite voluntary social distancing. In order to better control for this behavioral change, I now turn to a difference-in-differences approach.

5. Empirical Approach—Difference-in-differences (DD)

Taking the difference-in-differences approach pioneered by Card and Krueger (1994) to quantify an unbiased estimate of the effects of lockdown measures, I run the following two-way fixed effects specification:

$$Y_{it} = \beta_0 + b_1 L_{it} + b'\mathbf{X} + \tau_t + m_i + e_{it} \tag{1}$$

where Y_{it} denotes infection case per million in country i in time t and L_{it} is an indicator dummy for lockdown status that takes a value of 1 if there is a switch to a Swedish-style recommendation and 0 otherwise. $\mathbf{X}$ captures other additional controls; τ_t is a time fixed effect dummy; m_i controls for country fixed effects, and e_{it} is the error term. Our focus is on the coefficient b_1, which essentially captures the effect of the no-lockdown policy on the infection rate dynamics.

Our treatment country is Sweden and the control group consists of Finland, France and Norway, the three countries that collectively best approximate the synthetic control unit earlier in Section 2. Table A1 in the Appendix summarizes the demographic and epidemiological characteristics of each country. The time period covers 87 days between 29 February and 25 May. The policy intervention occurs on 23 March, when the government-mandated lockdown measures are imposed in the control group countries while Sweden moves to a soft regime switch. Post-treatment period thus covers a two-month long period during which the two groups diverged in terms of voluntary social distancing vs. involuntary lockdown mandate. In light of the earlier findings, I anticipate that the estimate of b_1 to be positive, which implies that the infection cases in Sweden would be higher than the control group countries that went for a strict lockdown measure. Additional controls summarized in $\mathbf{X}$ include the Google mobility index for six different categories as shown in Figure 7. With the exception of visits to parks and residential places, the correlation between the mobility index and the infection case per million is negative.

5.1. Results

The regression DD estimates for different specifications of the model in equation (1) are summarized in Table 5. The first row shows the average treatment effect. The specification in column (1) considers the effect of lockdown without any mobility index as controls. The estimated coefficient shows that Sweden, on average, had additional infection of 482.42 cases per million when compared to the infection rates of the control group. Considering that the case in Sweden reached around 3,300 per million population on 25 May, this implies that the infection rate in Sweden would have been lower by around 15 percent had it followed a strict lockdown policy like the other countries did. Specifications from columns (2) to (7) consider the effect of lockdown while individually controlling for behavioral changes, while the specification in column (8) allows for all of the mobility categories combined. The main findings on the average treatment effects remain robust, and even stronger in magnitude, to the inclusion of additional country-specific mobility controls. For example, allowing for all behavioral changes (as shown in the specification in column (8)), the infection rate in Sweden could have been lower by around 21 percent had it followed a lockdown policy like the other countries did.

5.2. Leads and lags

The key identifying assumption of DD regression design is a parallel (or common) trend assumption, meaning that—in the absence of treatment—the average change for the treated group would have been identical to the observed average change for the control group. In our setup, this implies that infection trends would have been the same in both Sweden and its control group had Sweden followed the same policy intervention path as the control group did. A rigorous verification is necessary, especially since our data set covers a lengthy period. An alternative way to deal with this issue—referred to by Autor (2003) as a "placebo" test—is to include leads in the baseline regression:

$$Y_{it} = \beta_0 + \sum_{j=0}^{q} b_j L_{i,t+j} + \tau_t + m_i + e_{it} \tag{2}$$

Covid Economics 35, 7 July 2020: 70-95

Table 5: Regression DD estimates

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lockdown	482.42***	689.70***	681.79***	836.81***	828.90***	836.28***	801.26***	690.26***
	(76.96)	(74.98)	(74.72)	(89.35)	(88.49)	(79.17)	(87.85)	(78.86)
Groceries		-16.67***						12.94*
		(1.80)						(6.30)
Parks			-5.28***					6.38***
			(0.41)					(1.89)
Transit stations				-19.43***				18.14
				(2.05)				(10.35)
Workplaces					-24.74***			38.06***
					(2.80)			(10.61)
Residential						55.70***		261.94***
						(4.78)		(34.83)
Retail							-13.59***	8.13
							(1.47)	(7.79)
N	348	348	348	348	348	348	348	348
R^2	0.886	0.908	0.905	0.912	0.914	0.922	0.907	0.935

Note: Robust standard errors in parentheses. All specifications include country and time fixed effects. * $p < .05$, ** $p < .01$, *** $p < .001$

The basic idea behind the test is that if a variable of interest, say $L_{i,t}$, causes outcome variables, say Y_{it}, future values of $L_{i,t}$ should not have any effect on Y_{it}. This type of a falsification test allows us to check for any anticipatory effect in days prior to the policy implementation. In our specification, I include leads of up to seven days allowing for the notification time prior to the policy enforcement.

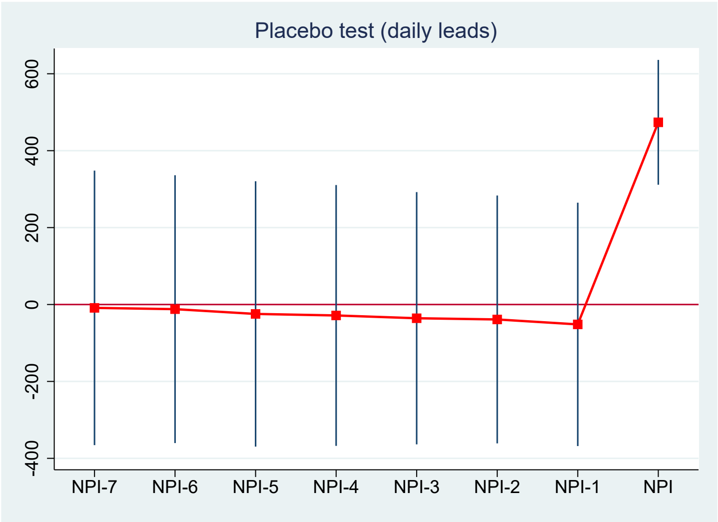
Figure 8 plots the coefficients and confidence intervals leading to the lockdown intervention.[11] As all the leads are very close to 0, I find no indication of any positive anticipatory effect for all seven days leading up to the lockdown measure. This provides some confidence that the parallel trend assumption is not violated and that the policy intervention occurs before its effect.

So far, we have implicitly assumed that the coefficient b_1 in equation (1) is constant, implying that we estimate the average treatment effects (ATE) for the whole post-treatment period. However, the impact of lockdowns could be immediate or lagged over time, and may possibly vary with time. In fact, earlier findings point out that during the first few weeks

[11]Table A2 in the Appendix presents the regression estimates for the all the specifications considered in Table 5.

Covid Economics 35, 7 July 2020: 70-95

Figure 8: Placebo tests



Note: NPI-n indicates n days prior to the policy intervention. NPI indicates all post-intervention period. Vertical lines represent 95 percent confidence intervals.

of post-lockdown, there was no discernible difference in the infection rates between Sweden and its synthetic counterpart. To explore the dynamic effects of the lockdown measures, I allow for lags in the regression specification as suggested by Autor (2003). More specifically, I add a dummy variable for each week up to the fifth week after the lockdown, as well as a dummy that captures all the weeks after week six since the lockdown is enforced. Each dummy variable takes the value of one in its relevant week. The modified specification with post-treatment dynamic effects is:

$$Y_{it} \quad = \quad \beta_0 + \sum_{j=0}^{q} b_j L_{i,t-7\times j} + \tau_t + m_i + e_{it} \tag{3}$$

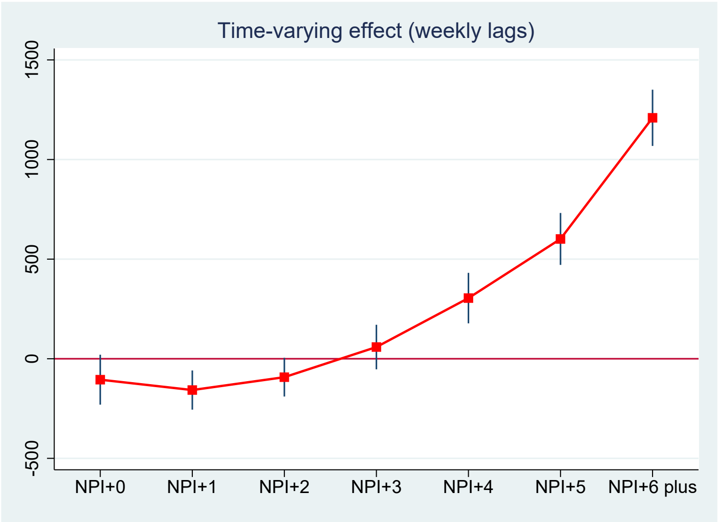
Here, b_0 captures the immediate effect of lockdown in the initial week, while the b_j $(\forall j > 0)$ coefficients pick up any subsequent weekly effects. If $b_j > b_0(> 0)$, this implies that the effect of the lockdown rises over time, while if the opposite is true then the initial impact fades with time.

Figure 9 plots the coefficients and the 95 percent confidence intervals allowing for lagged effects of the lockdown.[12] Similar to earlier findings under the synthetic control approach,

[12]Table A3 in the Appendix presents the regression estimates for the all the specifications considered in

Covid Economics 35, 7 July 2020: 70-95

Figure 9: Persistence effect of NPI



Note: NPI+n indicates n weeks after the policy intervention. NPI+6 plus indicates all post-intervention period after 6 weeks. Vertical lines represent 95 percent confidence intervals.

the estimated coefficients are not significant in the first four weeks. However, from week 5 onward, the coefficient becomes significantly positive and monotonically grows over time. This finding also coincides with the earlier outcome where the treatment effect becomes statistically significant five weeks after the implementation of the lockdown measures.

6. Conclusion

Policymakers have implemented a wide range of non-pharmaceutical interventions to fight the spread of COVID-19. Using variation in policies across countries and over time, I consider a synthetic control approach which is further complemented by a difference-in-differences (DD) research design to estimate causal effects of counter-COVID measures. I find that the lockdown measures played an important role in limiting the spread of the COVID-19 infection and that Swedish policymakers would have reduced the infection cases by more than half had they followed similar policies implemented elsewhere. I also find that as people dynamically adjust their behavior in response to information and policies, the impact of NPIs does not manifest immediately but only with a time lag of approximately 5 weeks or more.

Table 5.

Covid Economics 35, 7 July 2020: 70-95

Covid Economics 35, 7 July 2020: 70-95

One topic that the current study abstracts from is how each of the counter-COVID measures have different epidemiological impacts. A worthwhile project to pursue would be one that investigates the impact of individual measures along both epidemiological and economic aspects. Such explorations would better inform policymakers seeking to protect public health and facilitate an eventual economic recovery.

References

Abadie, A., 2020. Using synthetic controls: Feasibility, data requirements, and methodological aspects. Journal of Economic Literature (forthcoming) .

Abadie, A., Diamond, A., Hainmueller, J., 2010. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American Statistical Association 105, 493– 505.

Abadie, A., Diamond, A., Hainmueller, J., 2015. Comparative politics and the synthetic control method. American Journal of Political Science 59, 495– 510.

Abadie, A., Gardeazabal, J., 2003. The economic costs of conflict: A case study of the Basque country. American Economic Review 93, 113–132.

Athey, S., Imbens, G.W., 2006. Identification and inference in nonlinear difference-in-differences models. Econometrica 74, 431–497.

Autor, D.H., 2003. Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. Journal of Labor Economics 21, 1–42.

Baldwin, R., di Mauro, B.W. (Eds.), 2020. Economics in the Time of COVID-19. CEPR Press.

Born, B., Dietrich, A.M., Müller, G.J., 2020. Do lockdowns work? A counterfactual for Sweden. Covid Economics 16, 1– 22.

Card, D., Krueger, A.B., 1994. Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. American Economic Review 84, 772–793.

Castex, G., Dechter, E., Lorca, M., 2020. COVID-19: Cross-country heterogeneity in effectiveness of non-pharmaceutical interventions. Covid Economics 14, 175–199.

Chen, X., Qiu, Z., 2020. Scenario analysis of non-pharmaceutical interventions on global Covid-19 transmissions. Covid Economics 7, 46– 67.

Chernozhukov, V., Kasahara, H., Schrimpf, P., 2020. Causal impact of masks, policies, behavior on early Covid-19 pandemic in the U.S. medRxiv .

Firpo, S., Possebom, V., 2018. Synthetic control method: Inference, sensitivity analysis and confidence sets. Journal of Causal Inference 6.

Fisher, R.A. (Ed.), 1935. The Design of Experiments. Oliver and Boyd.

Gonzalez-Eiras, M., Niepelt, D., 2020. On the optimal 'lockdown' during an epidemic. Covid Economics 7, 68– 87.

Goodman-Bacon, A., Marcus, J., 2020. Using difference-in-differences to identify causal effects of COVID-19 policies. Survey Research Methods 14, 153–158.

Hale, T., Webster, S., Petherick, A., Phillips, T., Kira, B., 2020. Variation in government responses to COVID-19. BSG Working Paper Series .

Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. Journal of Economic Literature 47, 5–86.

Rocklöv, J., Sjödin, H., 2020. High population densities catalyse the spread of COVID-19. Journal of Travel Medicine 27.

Ullah, A., Ajala, O.A., 2020. Do lockdown and testing help in curbing Covid-19 transmission? Covid Economics 13, 138–156.

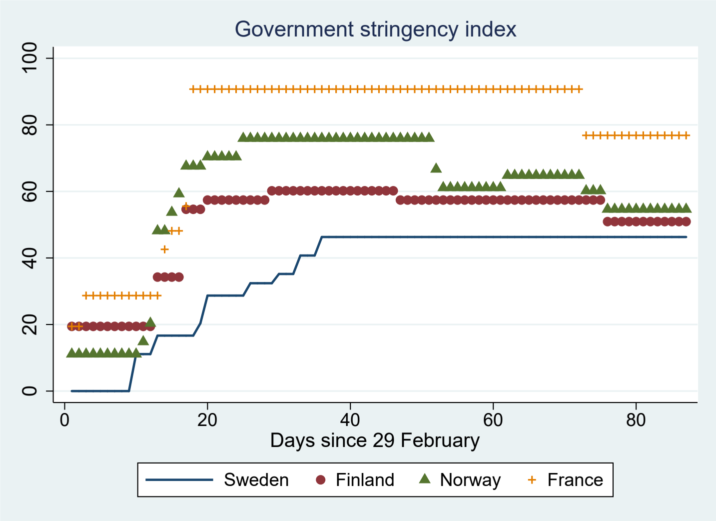
Covid Economics 35, 7 July 2020: 70-95

Appendix

Table A1: Composition of Synthetic Sweden

	Finland	France	Norway	Sweden
Weight for Synthetic Control	0.643	0.076	0.281	–
Population (million)	5.5	67.0	5.3	10.1
Urban population fraction (%)	85.4	80.4	82.2	87.4
Population density	18.1	122.3	14.5	24.7
Case per million (first 20 day)	25.9	46.1	80.7	42.8
Day 1	2 March	1 March	29 February	29 February
Lockdown Day	28 March	17 March	24 March	
Pre-lockdown duration (days)	26	16	24	
Government Stringency Index (SI) on Lockdown	68.5	90.7	75.9	

Figure A1: Government Stringency Index

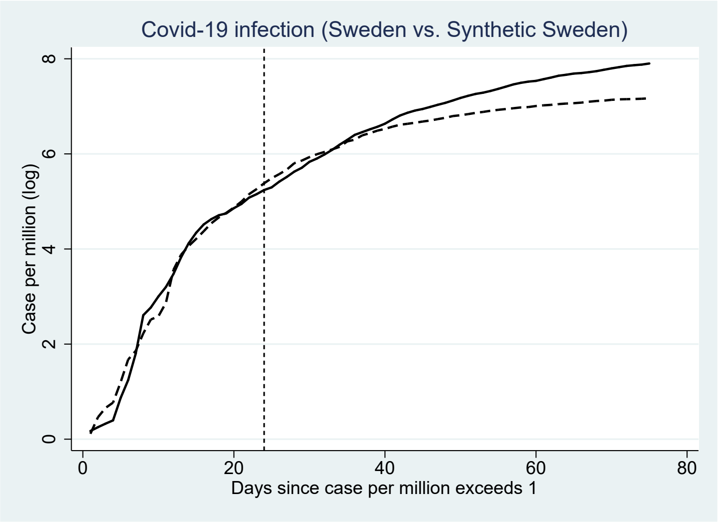


Note: Vertical axis measures the stringency index taken from Oxford COVID-19 Government Response Tracker since 29 February.

Covid Economics 35, 7 July 2020: 70-95

Covid Economics 35, 7 July 2020: 70-95

Figure A2: Profile of Infection Rates in Logs – Treatment vs. Synthetic Control



Note: Infection case per million population in logs is shown for Sweden versus synthetic Sweden (in dashed line).

Covid Economics 35, 7 July 2020: 70-95

Table A2: Placebo test

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NPI-7	-8.72	-4.17	8.30	65.05	5.16	14.92	10.90	-13.52
	(181.22)	(104.06)	(70.45)	(153.87)	(138.47)	(130.68)	(152.61)	(82.99)
NPI-6	-12.05	-28.31	7.24	107.97	66.34	52.72	33.46	43.89
	(176.80)	(103.93)	(65.93)	(165.85)	(158.12)	(147.74)	(158.23)	(105.72)
NPI-5	-24.47	-16.71	-0.90	151.43	143.24	111.57	66.17	134.20
	(175.18)	(109.72)	(63.97)	(167.27)	(171.14)	(146.64)	(163.75)	(102.96)
NPI-4	-28.45	28.17	-3.88	209.11	233.54	192.57	117.94	260.74*
	(172.18)	(116.06)	(74.54)	(170.29)	(181.40)	(144.90)	(171.86)	(103.88)
NPI-3	-35.67	94.64	29.67	283.79	324.32	289.52	187.71	365.27**
	(166.53)	(136.16)	(86.91)	(171.65)	(187.19)	(147.30)	(179.42)	(128.04)
NPI-2	-38.81	143.55	37.85	338.45	400.57*	376.84*	238.32	528.06*
	(163.62)	(130.20)	(104.06)	(176.68)	(192.60)	(159.41)	(185.90)	(206.78)
NPI-1	-51.65	153.94	30.05	373.79*	455.97*	402.39*	269.95	501.83*
	(160.66)	(134.97)	(122.95)	(189.25)	(220.07)	(181.22)	(196.99)	(220.04)
NPI	473.74***	707.71***	686.70***	918.05***	918.04***	910.81***	849.71***	796.83***
	(82.39)	(79.66)	(78.97)	(94.20)	(89.43)	(79.61)	(95.06)	(81.62)
Grocery_and_pharmacy		-16.82***						15.04*
		(1.84)						(6.15)
Parks			-5.28***					6.94***
			(0.42)					(1.91)
Transit_stations				-20.23***				15.39
				(2.15)				(10.65)
Workplaces					-26.05***			38.69***
					(2.93)			(10.87)
Residential						57.57***		268.51***
						(4.94)		(34.18)
Retail_and_recreation							-13.94***	8.48
							(1.56)	(7.70)
N	348	348	348	348	348	348	348	348
R^2	0.886	0.908	0.905	0.913	0.915	0.923	0.907	0.937

Note: NPI-n indicates n days prior to the policy intervention. NPI refers to the all post-treatment. Robust standard errors in parentheses. All specifications include country and time fixed effects. * $p < .05$, ** $p < .01$, *** $p < .001$

Covid Economics 35, 7 July 2020: 70-95

Table A3: Time-varying effects of lockdown

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NPI+0	-105.10	70.54	45.40	181.42*	212.33*	212.74**	118.84	77.95
	(63.65)	(72.26)	(63.18)	(80.93)	(87.72)	(81.10)	(74.63)	(72.30)
NPI+1	-157.16**	16.84	7.83	141.45*	159.44*	178.55**	95.43	70.17
	(49.91)	(60.16)	(46.00)	(64.10)	(69.62)	(65.20)	(61.94)	(62.48)
NPI+2	-92.50	51.98	110.46*	166.91**	168.81**	201.83***	149.49*	178.69**
	(49.27)	(58.31)	(55.55)	(58.80)	(57.57)	(56.19)	(62.96)	(54.69)
NPI+3	58.71	236.18***	292.40***	287.56***	228.25***	322.44***	290.21***	248.17***
	(56.76)	(55.10)	(54.56)	(60.41)	(61.93)	(56.46)	(59.15)	(66.82)
NPI+4	304.43***	418.25***	553.84***	556.74***	556.80***	581.13***	523.43***	617.89***
	(64.33)	(58.41)	(63.91)	(67.76)	(69.99)	(62.11)	(64.98)	(64.21)
NPI+5	601.36***	721.69***	764.95***	830.28***	836.36***	832.04***	829.47***	776.95***
	(66.14)	(58.15)	(41.91)	(61.18)	(62.72)	(55.31)	(61.85)	(57.24)
NPI+6 plus	1209.40***	1257.31***	1361.02***	1343.93***	1330.91***	1327.60***	1340.16***	1338.46***
	(71.70)	(64.50)	(69.74)	(64.67)	(61.99)	(59.89)	(64.53)	(74.89)
Grocery_and_pharmacy		-9.29***						14.87**
		(1.60)						(5.23)
Parks			-4.75***					-3.97**
			(0.37)					(1.53)
Transit_stations				-11.86***				33.93***
				(1.52)				(8.11)
Workplaces					-15.11***			-9.33
					(2.13)			(9.17)
Residential						35.99***		92.30**
						(3.63)		(30.52)
Retail_and_recreation							-8.43***	-5.54
							(1.15)	(6.47)
N	348	348	348	348	348	348	348	348
R^2	0.952	0.958	0.966	0.960	0.961	0.965	0.959	0.974

Note: NPI+n indicates n weeks after the policy intervention. NPI+6 plus indicates all post-intervention periods after 6 weeks. Robust standard errors in parentheses. All specifications include country and time fixed effects. * $p < .05$, ** $p < .01$, *** $p < .001$