
Eran Bendavid¹,², Christopher Oh¹, Jay Bhattacharya², John P.A. Ioannidis¹,³,⁴,⁵,⁶

¹ Department of Medicine, Stanford University Stanford, CA
² Center for Health Policy and the Center for Primary Care and Outcomes Research, Stanford University, Stanford, CA
³ Department of Epidemiology and Population Health, Stanford University, Stanford, CA
⁴ Department of Biomedical Data Science, Stanford University, Stanford, CA
⁵ Department of Statistics, Stanford University, Stanford, CA
⁶ Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/ECI.13484

This article is protected by copyright. All rights reserved
Abstract

Background and Aims: The most restrictive non-pharmaceutical interventions (NPIs) for controlling the spread of COVID-19 are mandatory stay-at-home and business closures. Given the consequences of these policies, it is important to assess their effects. We evaluate the effects on epidemic case growth of more restrictive NPIs (mrNPIs), above and beyond those of less restrictive NPIs (lrNPIs).

Methods: We first estimate COVID-19 case growth in relation to any NPI implementation in subnational regions of 10 countries: England, France, Germany, Iran, Italy, Netherlands, Spain, South Korea, Sweden, and the US. Using first-difference models with fixed effects, we isolate the effects of mrNPIs by subtracting the combined effects of lrNPIs and epidemic dynamics from all NPIs. We use case growth in Sweden and South Korea, two countries that did not implement mandatory stay-at-home and business closures, as comparison countries for the other 8 countries (16 total comparisons).

Results: Implementing any NPIs was associated with significant reductions in case growth in 9 out of 10 study countries, including South Korea and Sweden that implemented only lrNPIs (Spain had a non-significant effect). After subtracting the epidemic and lrNPI effects, we find no clear, significant beneficial effect of mrNPIs on case growth in any country. In France, e.g., the effect of mrNPIs was +7% (95CI -5%-19%) when compared with Sweden, and +13% (-12%-38%) when compared with South Korea (positive means pro-contagion). The 95% confidence intervals excluded 30% declines in all 16 comparisons and 15% declines in 11/16 comparisons.

Conclusions: While small benefits cannot be excluded, we do not find significant benefits on case growth of more restrictive NPIs. Similar reductions in case growth may be achievable with less restrictive interventions.
Introduction

The spread of COVID-19 has led to multiple policy responses that aim to reduce the transmission of the SARS-CoV-2. The principal goal of these so-called non-pharmaceutical interventions (NPI) is to reduce transmission in the absence of pharmaceutical options in order to reduce resultant death, disease, and health system overload. Some of the most restrictive NPI policies include mandatory stay-at-home and business closure orders (“lockdowns”). The early adoption of these more restrictive non-pharmaceutical interventions (mrNPIs) in early 2020 was justified because of the rapid spread of the disease, overwhelmed health systems in some hard-hit places, and substantial uncertainty about the virus’ morbidity and mortality.\(^1\)

Because of the potential harmful health effects of mrNPI – including hunger\(^2\), opioid-related overdoses\(^3\), missed vaccinations\(^4,5\), increase in non-COVID diseases from missed health services\(^6–9\), domestic abuse\(^10\), mental health and suicidality\(^11,12\), as well as a host of economic consequences with health implications\(^13,14\) – it is increasingly recognized that their postulated benefits deserve careful study. One approach to evaluating NPI benefits uses disease modeling approaches. One prominent modeling analysis estimated that, across Europe, mrNPIs accounted for 81% of the reduction in the effective reproduction number (\(R_e\)), a measure of disease transmission.\(^15\) However, in the absence of empirical assessment of the policies, their effects on reduced transmission are assumed rather than assessed.\(^16,17\) That analysis attributes nearly all the reduction in transmission to the last intervention, whichever intervention happened to be last, complete lockdowns in France, or banning of public events in Sweden.\(^16\)

Another, more empirically-grounded approach to assessing NPI effects uses statistical regression models and exploits variation in the location and timing of NPI implementations to identify changes in epidemic spread following various policies.\(^18\) These empirical studies find large reductions in the growth rate of new cases that are attributable to NPIs. An important challenge with these analyses is that they use pre-policy growth rates to determine the “counterfactual” trajectory of new cases – the expected case growth rate in the absence of NPIs. This is problematic because it is widely recognized that epidemic dynamics are time-varying, and brakes on disease transmission occur without any interventions (through resolution of infections), as well as from behavior changes unrelated to the NPIs.\(^19,20\) These epidemic dynamics are demonstrated by an analysis showing that...
slowing of COVID-19 epidemic growth was similar in many contexts, in a way that is more consistent with natural dynamics than policy prescriptions.\textsuperscript{21}

These challenges suggest that assessing the impact of mrNPIs is important, yet difficult. We propose an approach that balances the strengths of empirical analyses while taking into consideration underlying epidemic dynamics. We compare epidemic spread in places that implemented mrNPIs to counterfactuals that implemented only less-restrictive NPIs (lrNPIs). In this way, it may be possible to isolate the role of mrNPIs, net of lrNPIs and epidemic dynamics.

Here, we use Sweden and South Korea as the counterfactuals to isolate the effects of mrNPIs in countries that implemented mrNPIs as well as lrNPIs. Unlike most of its neighbors that implemented mandatory stay-at-home and business closures, Sweden’s approach in the early stages of the pandemic relied entirely on lrNPIs, including social distancing guidelines, discouraging of international and domestic travel, and a ban on large gatherings.\textsuperscript{22,23} South Korea also did not implement mrNPIs. Its strategy relied on intensive investments in testing, contact tracing, and isolation of infected cases and close contacts.\textsuperscript{24,25}

**Methods**

We isolate the effect of more restrictive NPIs (mrNPIs) by comparing the combined effect size of all NPIs in eight countries that implemented more restrictive policies (England, France, Germany, Iran, Italy, the Netherlands, Spain, and the United States) with the effect size of all NPIs in the two countries that only implemented less-restrictive NPIs (lrNPIs). In effect, we follow the general scheme:

\[
\text{Effects of mrNPI} = \text{Effects of (mrNPI + lrNPI + epidemic dynamics)} - \text{Effects of (lrNPI + epidemic dynamics)}
\]

We analyze only these countries because the analysis depends on sub-national data, which was only available for those countries, as explained further below.

The conceptual model underlying this approach is that, prior to meaningful population immunity, individual behavior is the primary driver of reductions in transmission rate, and that any NPI may provide a nudge towards individual behavior change, with response rates that vary between individuals and over time. lrNPIs could have large anti-contagion effects if individual behavioral
response is large, in which case additional, more restrictive NPIs may not provide much additional benefit. On the other hand, if lrNPIs provide relatively small nudges to individual behavior, then mrNPIs may result in large behavioral effects at the margin, and large reductions in the growth of new cases. However, because underlying epidemic dynamics are imprecisely characterized and are important for estimating the policy effects, our models test the extent to which mrNPIs had additional effect on reducing transmission by differencing the sum of NPI effects and epidemic dynamics in countries that did not enact mrNPIs from the sum of NPI effects and epidemic dynamics in countries that did.

We estimate the unique effects of mrNPIs on case growth rate during the northern hemispheric spring of 2020 in England, France, Germany, Iran, Italy, the Netherlands, Spain, and the United States by comparing the effect of NPIs in these countries to those in Sweden and South Korea (separately). The data we use builds on an analysis of NPI effects and consists of daily case numbers in subnational administrative regions of each country (e.g. regions in France, provinces in Iran, states in the US, and counties in Sweden), merged with the type and timing of policies in each administrative region. We use data from a COVID-19 policy databank and previous analyses of policy impacts to determine the timing and location of each NPI. Each observation in the data, then, is identified by the subnational administrative region and the date, with data on the number of cases on that date and indicators characterizing the presence of each policy. We include indicators for changes in case definitions or testing technologies to capture abrupt changes in case counts that are not the result of the underlying epidemic (these are mostly single-day indicators), as suggested in a previous analysis.

We define the dependent variable as the daily difference in the natural log of the number of confirmed cases, which approximates the daily growth rate of infections ($g$). We then estimate the following linear models:

$$g_{cit} = \theta_{0,ci} + \delta_{ct} + \sum_{p=1}^{p_c} (\gamma_{pc} policy_{pcit}) + \mu_{cit} + \varepsilon_{cit}$$

The model terms are indexed by country ($c$), sub-national unit ($i$), day ($t$), and NPI indicator ($p$). $\theta_{0,ci}$ are a series of fixed effects for the subnational unit, and $\delta_{ct}$ are country-specific day-of-week fixed effects. The parameters of interest are $\gamma_{pc}$, which identify the effect of each policy on the growth rate.
in cases. The parameters $\mu_{c,t}$ are the single-day indicators that model changes in case definitions that result in short discontinuities in case counts that are not due to underlying epidemic changes.

We estimate these models separately for each pair of countries (one with mrNPIs, one without), for a total of 16 models. We then add the coefficients of all the policies for the country with mrNPIs (yielding the combined effects of all NPIs in the mrNPI country) and subtract the combined effects of all NPIs in the comparator country without mrNPI. As noted above, the difference isolates the effect of mrNPIs on case growth rates. We estimate robust standard errors throughout, with clustering at the day-of-week level to account for serial correlation.

It is important to note that because the true number of infections is not visible in any country, it is impossible to assess the impact of national policies on transmission of new infections. Instead, we follow other studies evaluating the effects of NPIs that use case numbers, implicitly assuming that their observed dynamics may represent a consistent shadow of the underlying infection dynamics.

The code for the data preparation, analysis, and visualization is provided along with the article.

Results

The growth rate in new cases prior to implementation of any NPIs was positive in all study countries (Figure 1). The figure shows that, across all subnational units in all ten countries, the average growth rate prior to NPIs ranged from 0.23 in Spain (23% daily growth; 95CI 0.13 to 0.34) to 0.47 (95CI 0.39 to 0.55) in the Netherlands. The average across all 10 countries was 0.32, and in South Korea and Sweden, the two countries without mrNPIs, the pre-NPI growth rates were 0.25 and 0.33, respectively. The variation of pre-policy growth rates in cases may reflect epidemic intensity, testing coverage (higher growth may be a reflection of expanding testing capacity and of more people wishing to be tested), and pre-policy behavior changes that led to increased or decreased transmission.

Figure 2 and Figure 3 demonstrate the effects of individual NPIs (Figure 2) and all NPIs combined (Figure 3) on daily growth in case counts. While the effects of 3 individual NPIs were positive – that is, contributing paradoxically to case growth – and significant (one in Germany, one in Italy, and one in Spain, out of 51 individual NPIs in all 10 countries), the effects of about half of individual NPIs were negative and significant. The combined effects of all NPIs (Figure 3) were
negative and significant in 9 out of 10 countries, where their combined effects ranged from -0.10 (95CI -0.06 to -0.13) in England to -0.33 (95CI -0.09 to -0.57) in South Korea. Spain was the only country where the effect of NPIs was not distinguishable from 0 (-0.02; 95CI -0.12 to 0.07).

Figure 4 shows the effect of mrNPIs in the 8 countries where mrNPIs were implemented, after accounting for the effects of lrNPIs and underlying epidemic dynamics. In none of the 8 countries and in none out of the 16 comparisons (against Sweden or South Korea) were the effects of mrNPIs significantly negative (beneficial). The point estimates were positive (point in the direction of mrNPIs resulting in increased daily growth in cases) in 12 out of 16 comparisons (significantly positive in 3 of the 12, in Spain and in England compared with Sweden). The only country where the point estimates of the effects of mrNPIs were negative in both comparisons was Iran (-0.07 [95CI -0.21 - 0.07] compared with Sweden; -0.02 [95CI -0.28 - 0.25] compared with South Korea). The 95% confidence intervals excluded a 30% reduction in daily growth in all 16 comparisons.

**Discussion**

In the framework of this analysis, there is no evidence that more restrictive non-pharmaceutical interventions ("lockdowns") contributed substantially to bending the curve of new cases in England, France, Germany, Iran, Italy, the Netherlands, Spain, or the United States in early 2020. By comparing the effectiveness of NPIs on case growth rates in countries that implemented more restrictive measures with those that implemented less restrictive measures, the evidence points away from indicating that mrNPIs provided additional meaningful benefit above and beyond lrNPIs. While modest decreases in daily growth (under 30%) cannot be excluded in a few countries, the possibility of large decreases in daily growth due to mrNPIs is incompatible with the accumulated data.

The direction of the effect size in most scenarios point towards an increase in the case growth rate, though these estimates are only distinguishable from zero in Spain (consistent with non-beneficial effect of lockdowns). Only in Iran do the estimates consistently point in the direction of additional reduction in the growth rate, yet those effects are statistically indistinguishable from zero. While it is hard to draw firm conclusions from these estimates, they are consistent with a recent analysis that identified increase transmission and cases in Hunan, China during the period of stay-at-
home orders from increased intra-household density and transmission.\textsuperscript{29} In other words, it is possible that stay-at-home orders may facilitate transmission if they increase person-to-person contact where transmission is efficient such as closed spaces.

Our study builds on the findings of overall effectiveness of NPIs in reducing case growth rate. This has a plausible underlying behavioral mechanism: NPIs are motivated by the notion that they lead to anti-contagion behavior changes, either directly through personal compliance with the interventions, or by providing a signal about disease risk, as communicated by policy makers, which is used in deciding on individual behaviors. The degree to which risk communications motivate personal behaviors has been used to explain South Korea’s response to NPIs, where large personal behavior changes were observed following less restrictive NPIs.\textsuperscript{30}

This analysis ties together observations about the possible effectiveness of NPIs with COVID-19 epidemic case growth changes that appear surprisingly similar despite wide variation in national policies.\textsuperscript{31–33} Our behavioral model of NPIs – that their effectiveness depends on individual behavior for which policies provide a noisy nudge – help explain why the degree of NPI restrictiveness does not seem to explain the decline in case growth rate. Data on individual behaviors such as visits to businesses, walking, or driving show dramatic declines days to weeks prior to the implementation of business closures and mandatory stay-at-home orders in our study countries, consistent with the behavioral mechanisms noted above.\textsuperscript{34–36} These observations are consistent with a model where the severity of the risk perceived by individuals was a stronger driver of anti-contagion behaviors than the specific nature of the NPIs. In other words, reductions in social activities that led to reduction in case growth were happening prior to implementation of mrNPIs because populations in affected countries were internalizing the impact of the pandemic in China, Italy, and New York, and noting a growing set of recommendations to reduce social contacts, all of which happened before mrNPIs. This may also explain the highly variable effect sizes of the same NPI in different countries. For example the effects of international travel bans were positive (unhelpful) in Germany and negative (beneficial) in the Netherlands (Figure 2).

While this study casts doubt on any firm conclusions about the effectiveness of restrictive NPIs, it also underscores the importance of more definitive evaluations of NPI effects. NPIs can also
have harms, besides any questionable benefits, and the harms may be more prominent for some NPIs than for others. For example, school closures may have very serious harms, estimated at an equivalent of 5.5 million life years for children in the US during the spring school closures alone.\textsuperscript{37} Considerations of harms should play a prominent role in policy decisions, especially if an NPI is ineffective at reducing the spread of infections. Of note, Sweden did not close primary schools throughout 2020 as of this writing.

While we find no evidence of large anti-contagion effects from mandatory stay-at-home and business closure policies, we should acknowledge that the underlying data and methods have important limitations. First, cross-country comparisons are difficult: countries may have different rules, cultures, and relationships between the government and citizenry. For that reason, we collected information on all countries for which subnational data on case growth was obtainable. Of course, these differences may also exist across subnational units, as demonstrated in the case of different states in the US. Additional countries could provide more evidence, especially countries that had meaningful epidemic penetration and did not use mrNPIs for epidemic control. Second, confirmed case counts are a noisy measure of disease transmission. Testing availability, personal demand for or fear of getting tested, testing guidelines, changing test characteristics, and viral evolution all interfere in the relationship between the underlying infections and case counts. Because the location and timing of policies is endogenous to perceived epidemic stage, the noise in case counts is associated with the policies, making bias possible and very difficult to eradicate. The fixed effects approach provides unbiased estimates so long as the location or timing of policies is quasi-arbitrary with respect to the outcome. This may fail to hold in this assessment of NPI effects because the underlying epidemic dynamics are non-linear, and the policies respond to – and modify – the epidemic stage. This limitation also holds for all other empirical assessments of NPI effects.\textsuperscript{18}

Third, our findings rest on a conceptualization, common in the literature, of NPIs as “reduced-form” interventions: an upstream policy has expected downstream effects on transmission. This allows us to use Sweden and South Korea as comparators, since they had applied less-restrictive interventions, which then enables netting out the combined effect of lrNPIs and the underlying epidemic dynamics. While contextual factors that mediate the effects of NPIs are important –
countries implemented different variants of the same NPI, and the population responded differently – many analyses examining the effects of NPIs have a similar “reduced-form” structure. In that sense our comparison is positioned squarely within the literature on the effects of NPIs.

During the northern hemisphere fall and winter of 2020, many countries, especially in Europe and the US, experienced a large wave of COVID-19 morbidity and mortality. Those waves were met with new (or renewed) NPIs, including mrNPIs in some countries (e.g. England) and lrNPIs in others (e.g. Portugal) that had used mrNPIs in the first wave. The spread of infections in countries that were largely spared in the spring (e.g. Austria and Greece) further highlight the challenges and limited ability of NPIs to control the spread of this highly transmissible respiratory virus. Empirical data for the characteristics of fatalities in the later wave before mrNPIs were adopted as compared with the first wave (when mrNPIs had been used) shows that the proportion of COVID-19 deaths that occurred in nursing homes was often higher under mrNPIs rather than under less restrictive measures. This further suggest that restrictive measures do not clearly achieve protection of vulnerable populations. Some evidence also suggests that sometimes under more restrictive measures, infections may be more frequent in settings where vulnerable populations reside relative to the general population.

In summary, we fail to find strong evidence supporting a role for more restrictive NPIs in the control of COVID in early 2020. We do not question the role of all public health interventions, or of coordinated communications about the epidemic, but we fail to find an additional benefit of stay-at-home orders and business closures. The data cannot fully exclude the possibility of some benefits. However, even if they exist, these benefits may not match the numerous harms of these aggressive measures. More targeted public health interventions that more effectively reduce transmissions may be important for future epidemic control without the harms of highly restrictive measures.
References

doi:10.1101/2020.04.22.20076026


doi:10.1126/science.abc8637


This article is protected by copyright. All rights reserved


This article is protected by copyright. All rights reserved

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


https://www.apple.com/covid19/mobility

This article is protected by copyright. All rights reserved


Figure 1: Growth rate in cases for study countries. The black bars demonstrate the average growth rate in cases in each subnational unit (95% CI) prior to any policies implemented. The figures to the right show the daily growth rate in cases for each of the countries and demonstrate the shared decline in case growth across all countries, including the countries that did not implement mrNPIs (South Korea and Sweden).
Figure 2: Effects of individual NPIs in all study countries. The variation in the timing and location of NPI implementation allows us to identify the effects of individual NPIs on the daily growth rate of cases. Where multiple NPIs were implemented simultaneously (in the same day) across all subnational units (e.g. school closure, work from home, and no private gatherings in Spain), their overall effect cannot be identified individually and is shown combined.
Figure 3: Combined effects of all NPIs in study countries. The point estimate and 95% CI of the combined effect of NPIs on growth rate in cases, estimated from a combination of individual NPIs. The estimates show significant effects in all countries except Spain, and range from a 33% (9-57%) decline in South Korea to 10% (6%-13%) in England. The point estimate of the effect in Spain is also negative but small (2%) and not significant.
Figure 4: Effect of mrNPIs on daily growth rates after accounting for the effects of lrNPIs in South Korea and Sweden. Under no comparison is there evidence of reduction in case growth rates from mrNPIs, in any country. The point estimates are positive (point in the direction of mrNPIs resulting in increased daily growth in cases) in 12 out of 16 comparisons.
Author Contributions
EB conceived the project; EB and CO designed the analyses, prepared the data and executed the analyses; JB and JPAI were involved in discussing, interpreting the results, drafting, revising, and improving the manuscript. All authors have approved the final manuscript.

Funding/Support
The study was funded with support from the Stanford COVID-19 Seroprevalence Studies Fund.

Role of the Funding Organization or Sponsor
The funding organizations had no role in the design or execution of this analysis.