Causal Inference of Non-pharmaceutical Interventions in COVID-19: An International Study

Menelaos Konstantinidis[1,2], Larry Dong[1], Rose Garrett[1,2]

[1]Dalla Lana School of Public Health, The University of Toronto
[2]Child Health Evaluative Sciences, The Hospital for Sick Children

Abstract

Background: For over a year, COVID-19 has spread across the world, infecting nearly 100 million individuals, and resulting in the death of over 2 million. With no end in sight, evaluating key interventions is of the utmost importance. Given that vaccines have only recently been approved, rigorously evaluating non-pharmaceutical interventions is essential.

Objective: To causally estimate the effect of non-pharmaceutical interventions (NPI) on the week-over-week growth rates of COVID-19 cases and deaths.

Methods: We have independely implemented two generalized estimating equations (GEE) – a baseline GEE and an inverse probably treatment-weighted GEE (marginal structural model). The former allowed us to estimate the association estimate for NPI, whereas the latter allowed for a counterfactual policy analysis, approximately estimating the causal estimate for NPI.

Results: In both the unweighted and weighted GEEs, 2 week lagged-NPI (at a level above 65.8 on the OXCGRT Overall Government Response index) was found to reduce the week-over-week increase in COVID-19 cases by 6.2-7.5% ($p < 0.05$), and 7.3-9.6% ($p < 0.05$), respectively. Moreover, while still yielding a negative effect, NPI did not yield a statistically significant effect for the growth rate in COVID-19 deaths. Furthermore, we observe that the marginal structural model assumptions (identifiability and specification) were satisfied reasonably well.

Discussion: Given the general satisfaction of the marginal structural model assumptions, we assert that the OXCGRT Overall Government Response index above 65.8 has a statistically significant negative effect on the growth rate of COVID-19 cases, and policy combinations below this NPI level have no statistically significant effect. Thus, to curtail, and ultimately eradicate the spread of COVID-19 in a given country, we recommend that policies be implemented, such that their corresponding index value falls above 65.8. Moreover, resources allocated for policy combinations falling below this level, ought to be diverted to other endeavours, as we believe them to be ineffectual.

1 Introduction

On December 31, 2019, the World Health Organization was formally made aware of a series of patients presenting with atypical pneumonia [1]. What was then referred to as the Wuhan Pneumonia [2], and later Sars-CoV-2, not only spread across the city of Wuhan, but in a few months rapidly spread across the globe. At the time of this writing, nearly a year after the disease's inception, far from being eradicated, with some 95 million cases, and over 2 million deaths, COVID-19 continues to plague the world.

While the original transmission of COVID-19 is believed to be zoonotic in nature, human-to-human transmission has resulted in exponential rates of infections world-wide. The virus is known to primarily target the respiratory system, although other organ systems are known to be affected [3]. Although the disease is known to present only mild symptoms in many individuals, the spread nevertheless poses severe public health problems on many healthcare systems, thus requiring varying degrees of public health interventions [4].

Given that vaccines have only recently been approved, non-pharmaceutical interventions (NPIs) have hitherto been one of the principal means by which to curtail the spread of the virus. NPIs include measures such as mask mandates, lockdowns, physical distancing, and closures of schools and businesses. They can be highly effective in slowing the spread of the virus; for example, a study focusing on NPIs in Canada [5] found that mask mandates were associated with a 25% reduction in the weekly new cases in July and August. However, more severe NPIs, such as lockdowns and closures of workplaces, have adverse economic consequences, and in some cases, even implications to human rights [6, 7]. It is therefore necessary to have a clear and rigorous means by which to evaluate what combination of NPIs ought to be implemented at the national level for each country.

In an endeavour to optimize public policy and their effects on the present pandemic, we present a Marginal Structural GEE to causally evaluate the effects of non-pharmaceutical interventions on COVID cases, and COVID-induced deaths in 103 countries. The remainder of the article will proceed as follows. Section 2 will be dedicated to a brief overview of the data used for the validation of our statistical model. Section 3 will describe the weighted models that were used. Section 4 will present the results from our analyses, from which we will proceed to a discussion of the results and future work in Section 5.

2 Data

The data used for the presented work consists of five principal domains, namely, (1) infections, (2) deaths, (3) non-pharmaceutical interventions, (4) social-distancing behaviour (measured using location change indices from Google Community Mobility reports), and (5) time-invariant country-level covariates. For completeness of data, and ease of analysis, countries in which data for one of the variables was missing were excluded from subsequent analyses. While this could, in principal, be remedied by more sophisticated methodologies, the scope of this work exceeds such implementations.

Of 122 potentially eligible countries from the Oxford COVID-19 Government Response Tracker (OXCGRT) data [8], nineteen were excluded – Afghanistan, Cape Verde, Georgia, Papua New Guinea, Serbia due to missing Google Mobility data, Ukraine due to missing NPI data, and Fiji and Laos due to insufficient cumulative cases (< 100) as of September 30. Furthermore, of the 122 coutries, seven countries (Benin, Ecuador, France, Jordan, Lithuania, Luxembourg and Uganda) were excluded due to an observed decrease in the cumulative number of cases – likely due to an error in recording or different criteria for case definition and, likewise, three other countries (Belgium, Kyrgyzstan and Spain) were excluded due to an observed decrease in cumulative deaths. Moreover, two countries (United Arab Emirates and Argentina) had missing mobility data on July 30, 2020 and July 2, 2020, respectively. In these two cases, we carried forward the last observed value (which also happened to be the next observed value).

Outcome: Infection/Death – Data for infections and deaths was obtained from the Johns Hopkins Centers

for Civic Impact [1]. Daily per capita data was obtained for each of 103 included countries. The final outcome, defined as a log transform of the week-over-week growth rate in cases/deaths, was formulated based on previous work [5].

Exposure: Non-pharmaceutical Interventions – The primary explanatory variable of interest (the exposure), complemented with other variables such as mobility and other covariates (see below), is non-pharmaceutical interventions (NPI). The variable was populated using the aggregation of the OXCGRT [8]. The aggregation, discussed in the working paper, provides a series of NPI indices spanning across twenty principal variables in three domains – Containment and Closure, Economic Responses, and Health systems. A fourth domain – Miscellaneous – is also included. Four policy indices are generated, namely, (1) Overall government response index, (2) Stringency index, (3) Containment and health index, and (4) Economic support index, each of which is a linear combination of a subset of the twenty variables. For the purposes of this work, we only worked with "Overall government response" which covers 15 of the 20 domains.

Mobility – Social-distancing behaviour was measured using the Google Mobility Data [9] which reports daily changes in mobility at regional, and national levels. Changes are quantified relative to baseline (median mobility for corresponding days of the week spanning Jan 3–Feb 6, 2020). The mobility data reports on changes in mobility across six categories – Groceries and Pharmacy, Parks, Transit Stations, Retail and Recreation, Residential, and Workplaces. However, in order to remediate missing data and possible collinearity between categories, following [10], we took the mean of the mobility values from "Retail", "Grocery" and "Pharmacy", and "Workplaces".

Time-invariant Covariates – In our model we also included five time-invariant covariates for each country, namely, median age, Gross Domestic Product (GDP), population density, human development index (HDI), and life expectancy.

Time-frame – In order to maintain consistency across and within the data for each country and minimize the effects of sparse data, we determined the starting point to be the day in which a 100 or more cumulative cases were recorded. Note that this calendar day will vary between countries; however, in the dataset, this the threshold date was normalized to $t = 1$. The cut-off date for this analysis was September 30, 2020.

3 Methods

In this work, we have two outcomes of interest, namely, COVID-19 cases and deaths. Each of these outcomes was modeled using a Marginal Structured Model (MSM) using an inverse probability treatment weighted generalized estimating equation (GEE). GEEs have widespread applicability to longitudinal and repeated measures data, made attractive (amongst other reasons) for their ability to provide consistent parameter estimates even with mis-specification of the underlying correlation structure [11]. However, as remarked by other authors [12, 13, 14, 15], when time-varying confounders are present, the estimate of interest may be biased. This occurs when 1) a time-varying covariate is a predictor or risk factor of the exposure, and 2) history of the exposure is a predictor of the time-varying covariate [15].

An MSM, estimated through a GEE with an independent working correlation structure and inverse probability of treatment weighting (IPTW) solves the aforementioned problems. Moreover, the use of such a model, subject to the MSM assumptions [16], allows us to causally estimate the effect of NPIs (the

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

index) on the two outcomes of interest (cases and deaths). In order to achieve this result, the IPTW weights were obtained by estimating the probability of NPI exposure based on several confounders, namely lagged mobility (1 week lag), lagged deaths (1 week lag), lagged case count (1 week lag), each of the five time-invariant covariates, and a cubic function of time (see results in section 4.1). The exposure model was estimated using a multinomial logistic regression model, and assuming a linear contribution of each covariate (except for time). While interaction and higher-order terms should be tested for correct model specification, doing so exceeded the scope of the present work.

Furthermore, we observe that the NPI score is an approximately continuous variable. While this is not strictly true given that it is the normalized mean of several ordinal and numeric variables, in the absence of more complex manipulation of the index, this serves as a reasonable assumption; though subsequent interpretations do need to be taken with a grain of salt. To partially ameliorate this problem, and that of possible positivity assumption violations, we have partitioned the observed NPI index into three categories based on the observed tritiles (54.6 and 65.8). The approximate interpretation then is low, medium, and high NPI level.

4 Results

4.1 Descriptive Statistics

Prior to presenting the modeling results, it is worth taking a moment to visualize the data, and identify any acute issues. Looking first at the histogram of the raw NPI for all countries (Figure 1), we observe that the majority of the NPIs over the time-frame are above ~ 55 (see tritiles defined above). This nevertheless implies that most countries have endeavoured to implement, at least to some degree, a medium to high set of NPI policies.

Moreover, we observe that the outcomes (log growth rate of week-over-week cases/deaths) are centred at zero, with a few tail events (Figure 2). While both distributions could be well approximated by a normal distribution, there are key differences; particularly, in that the distribution of the cases outcome is characteristically leptokurtic whereas the death outcome is characteristically platykurtic. The specific impact of this is that, in combination with the central value (0), future modeling endeavour may need to be zero-inflated, particularly in modeling the cases outcome (further substantiated by Figure 3). Lastly, we note that the mobility distribution is predominated by negative values (decrease in mobility relative to baseline) – Figure 5. This is of course expected, as many policies aim to decrease the average contact rate, perpetuated in large part by human mobility.
We also visually investigate the relationship between changes in NPI over time, and weekly case growth in the different countries in the sample (a subset shown in Figure 4). We find that in most cases, countries with a high level of NPI in place two weeks prior have relatively low case growth.

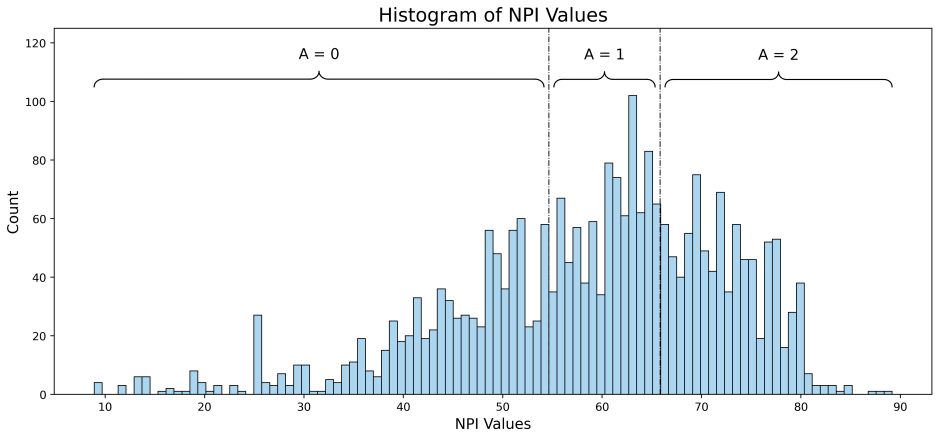

Figure 1: Histogram of NPI values across countries.

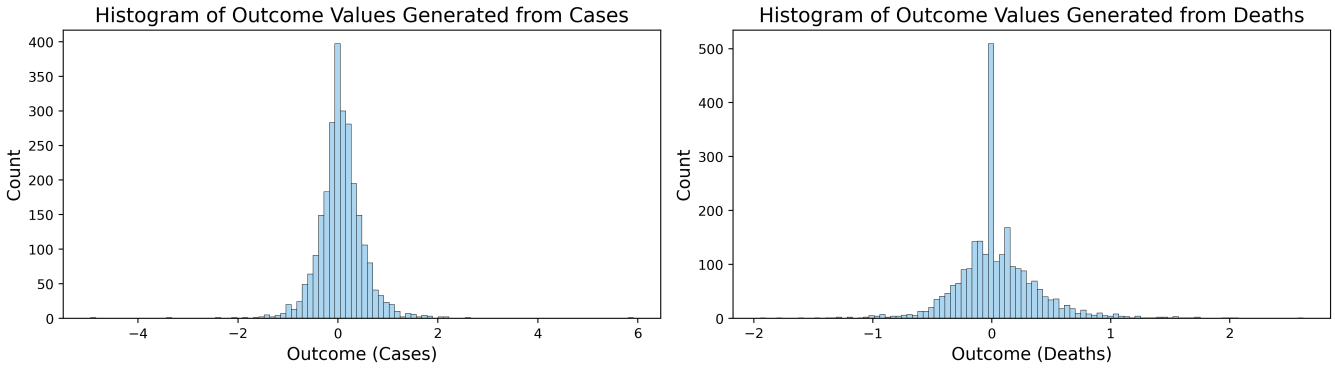

Figure 2: Left: Histogram of outcome values generated from cases. Right: Histogram of outcome values generated from deaths.

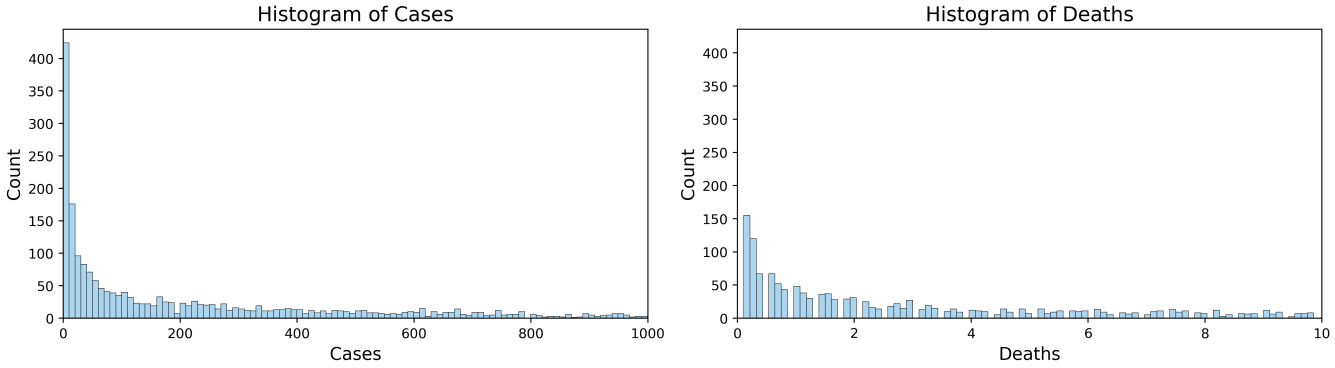

Figure 3: Left: Histogram of cases truncated from below at 1000. Right: Histogram of deaths truncated from below at 10.

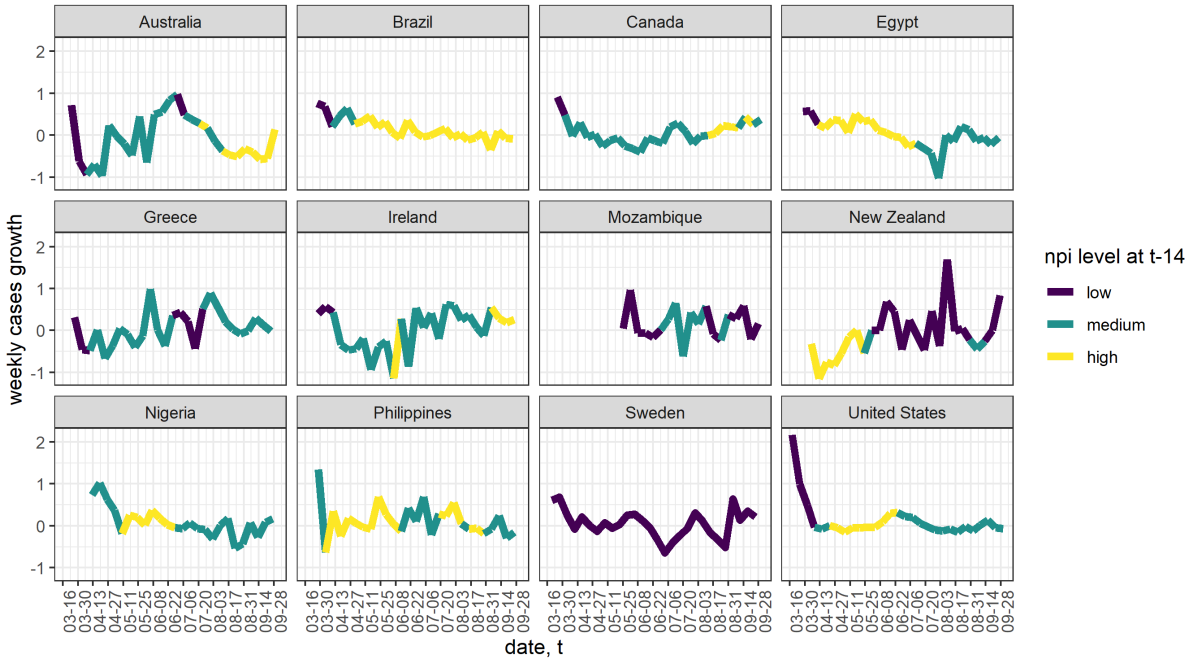

Figure 4: Average weekly case growth in twelve selected countries, with time-varying NPI.

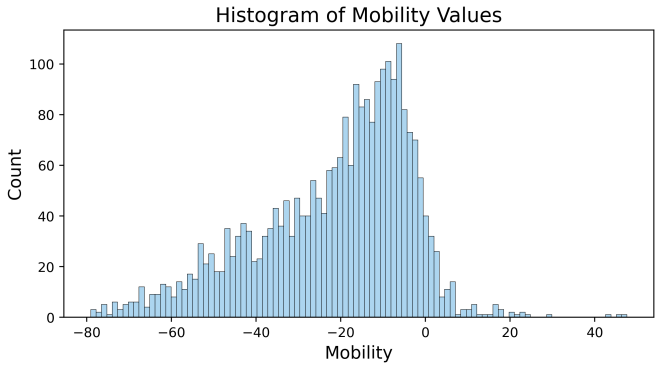

Figure 5: Histogram of Mobility Values.

4.2 Case growth modeling

We conducted case outcome modelling using three possible characterizations of the time effect: no time trend, cubic time trend in days from the beginning of the sample, and weekly dummy variable, as proposed in previous work [5]. It is plausible that time plays an important role in the evolution of case growth over time, for example through COVID fatigue. The cubic time trend is our baseline specification, and the other time schemes are investigated to perform robustness checks. For each time specification, we present both the baseline GEE and the weighted GEE approaches of estimating the outcome model (Table 1). In all of the regressions, we also control for the five time-invariant confounders mentioned previously.

The six different modelling approaches all show relatively strong evidence (p-value < 0.1) that enforcing a high NPI level within a country is associated with a statistically significant reduction in the weekly case growth two weeks later, relative to a lack of restrictions (low NPI). In other words, the results indicate that relaxing NPIs to a low level is associated with significantly higher weekly case growth. In particular,

Table 1: Summary of NPI coefficient estimates and corresponding p-values in brackets for the baseline GEEs without weighting (Models 1, 3, and 5), and the IPTW GEEs (Models 2, 4, and 6) under three possible specifications of the time trend, as proposed in previous work [5]. NPI indicates 2-week lagged NPI.

	Outcome: weekly case (log) growth rate					
	(1)	(2)	(3)	(4)	(5)	(6)
	no time trend		cubic time trend		weekly dummy variable	
medium NPI	**-0.064 ****	-0.027	**-0.056 ****	-0.013	-0.046 *	-0.001
	[0.020]	[0.503]	**[0.041]**	[0.715]	[0.079]	[0.987]
high NPI	**-0.078 *****	**-0.096 *****	**-0.064 ****	**-0.073 ****	-0.050 *	-0.058 *
	[0.008]	**[0.006]**	**[0.033]**	**[0.029]**	[0.080]	[0.069]
QIC	565.956	622.606	570.550	667.029	602.287	845.833
N. obs.	2567	2567	2567	2567	2567	2567

*** p < 0.01; ** p < 0.05; * p < 0.1. Bolded p-values < 0.05.

columns 1 and 3 (unweighted GEE) indicate that, depending on the time effect specification, high NPI is associated with a reduction in the weekly case growth of 6.4 to 7.8 log points ($p < 0.05$), which is equivalent to a 6.2-7.5 % reduction in weekly cases. Moreover, through the MSM framework, columns 2 and 4 (weighted GEE) indicate that high NPI *causally* reduced the weekly case growth by 7.3 to 9.6 log points (p-value < 0.05), which is equivalent to a 7-9.1% reduction in weekly cases. Thus after adjusting for confounding through the MSM structure, the effect of NPI is more pronounced (higher in absolute value). Indeed, the fact that there is a noticeable difference in the final result (reduction in weekly cases), suggests that there was some degree of confounding.

In addition, we note that enforcing a medium NPI level compared to low NPI, is also associated with a reduction in weekly case growth. In particular, Columns 1 and 3 (unweighted GEE) indicate that medium NPI is associated with a 5.6 to 6.4 log point reduction for case growth ($p < 0.05$), which translates to a 5.5 to 6.2% reduction in weekly cases. It is worth noting however, that in the case of the weighted GEE, while still negative, columns 1 and 3 do not provide a statistically significant NPI estimate. We therefore conclude that although medium NPI is *associated* with a reduction in weekly cases, it does not *causally* affect weekly cases.

Furthermore, we note that the NPI estimates under the weekly fixed effect time specification, while negative for both the baseline and weighted GEEs, are not significant at the 0.05 level of significance, a comparable result to previous work [17]. Lastly, these results are not included here, but neither the baseline GEE nor the weighted GEE (for any of the time schemes) resulted in statistically significant NPI estimates for the deaths outcome.

5 Discussion

5.1 Model Adequacy

In order for an observational study to be conceptualized as a conditionally randomized experiment, it needs to satisfy three principal identifiability conditions: Consistency, Exchangeability, Positivity [18]. In addition, it has to be a correctly specified model [16].

Consistency – Consistency has previously been defined in terms of two conditions: (1) Having the treatment or more specifically the "counterfactuals" well-defined, and (2) Having an explicit link between the observations and the treatment (i.e., unconditional positivity). The first condition is trivially satisfied by the numerical definition of the NPI index. The second condition, is satisfied in that there are no countries where their receive the "treatment" without being treated, that is, it is very unlikely in our opinion that a country can have an observed NPI index at any point in time without having consciously implemented the corresponding policies in question.

Exchangeability – A natural way to argue that the Exchangeability condition, otherwise known as the "no unobserved confounder" assumption is satisfied is to investigate a large suite of information relevant to the exposure and outcome [19, 20]. While this is obviously true in general, it is not empirically testable. Furthermore, sources of data in the present problem make establishing "sufficient relevant information" a far from trivial task, not to mention the difficulty of arguing what constitutes "sufficient". With that in mind, we make the following note. Any unobserved covariates, by the nature of the problem at hand, would manifest themselves in the behaviour of humans in each country. That is, any unobserved covariate would directly or indirectly affect the mobility of individuals, a time-varying confounder which we have already accounted for. It therefore stands to reason that any additional unobserved confounders would have little if any effect, thereby implying that the Exchangeability assumption is reasonably satisfied.

Positivity – The third assumption of the MSM framework is that of Positivity – the assumption that the units of inference have a positive probability of treatment given confounders for each level of the treatment and confounders [16]. As touched upon in the methods section, we have partitioned the NPI index. However, several of the confounders remain continuous variables, which necessarily implies the violation of strict positivity [21]. It nevertheless stands to be argued that given sufficient density over a region of a given confounder's support (as is the case with the present data), positivity approximately holds. With that in mind, the mild violation of the assumption is likely to perpetuate a small bias in the causal estimates, a fact that should be considered in any downstream interpretations.

No Mis-specification – Over and above the identification assumptions for causal inference discussed so far, there is a fourth, and key assumption in the MSM framework, namely no mis-specification [of the model]. This a general issue in any modeling endeavour, however it is critical for MSM as mis-specification can cause biases in the causal estimates. In this regard, although we have specified a linear exposure model (with cubic time), interaction and higher-order terms need to be investigated as to confidently assert causal inference. The same holds true for the outcome model. Thus, while the three identifiability assumptions are reasonably satisfied, there is the possibility of some bias in the NPI estimate which should be addressed in future work.

Furthermore, as discussed in past works [16], there is a critical trade-off that needs to be considered in arriving at a final model. In particular, there is a trade-off between (1) an adequate model specification

as we have just discussed, and (2) the behaviour of the corresponding weights. In principal we wish the weights to have a mean of 1, and have minimal variability. A possible way to deal with variable weights is to truncate them (e.g., at the 1st and 99th percentiles). Doing so naturally reduces the weights (a desired result); however this introduces some degree of bias in the causal estimates. In the present work, we have not truncated, even though there is some variability (Figure 6 and 7).

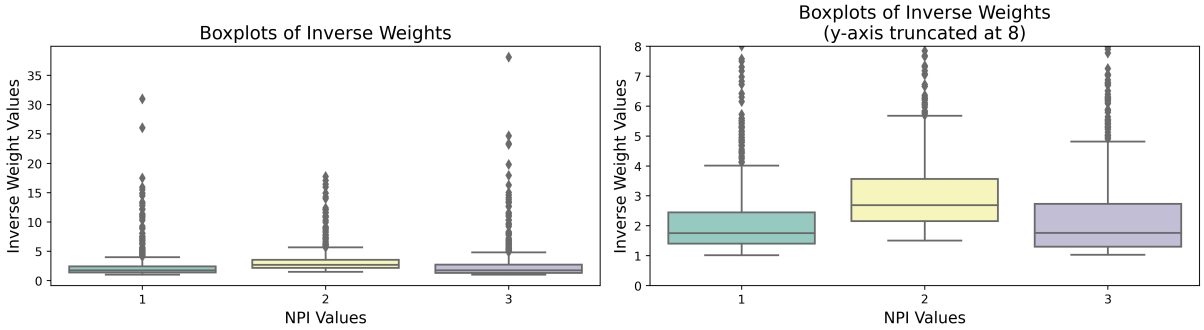

Figure 6: Left: Boxplots of inverse weights with respect to discretized NPI values. Right: Same as left, but with the y-axis truncated from above at 8.

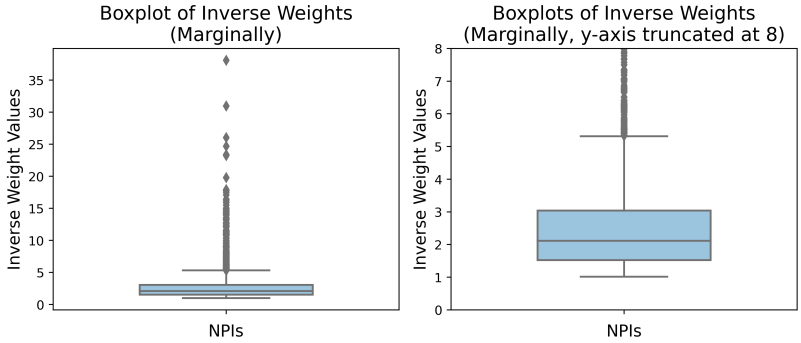

Figure 7: Left: Boxplot of all inverse weights. Right: Same as left with the y-axis truncated at 8.

The reason for doing so was that we wanted, insofar as it was possible, to arrive at an unbiased estimate of the NPI effect. Moreover, the estimated mean and standard deviation are 2.81 and 2.61, respectively. Though not at the desired mean of 1 and small variance, the observed values may be at a tolerable level for the desired unbiased estimates. With that in mind, future work should consider different truncation schemes to optimize the aforementioned trade-off.

5.2 Pandemic Implications

Even though several vaccines have been approved in many countries, a plethora of difficulties have manifested themselves, inhibiting the timely distribution of vaccines to many individuals around the world. In the meantime, the virus continues to infect and kill. It is therefore more important today than ever before to have a sound, rigorously substantiated understanding of what qualifies as effective interventions. Although in the present work we have not investigated specific non-pharmaceutical interventions, the omnibus index, imperfect though it is, provides a reasonable first step in evaluating necessary overall levels of government NPIs to curtail the spread of the pandemic. In this regard, it is of acute importance to understand precisely

what effect a set of policies will have on the number of infections and other downstream events.

The key objective throughout the present endeavour has been to estimate the causal effect of NPIs on week-over-week case growth. However, we observed several conceptual difficulties in the existing datasets such as time-varying confounding, time-invariant confounding, and variable endogeneity, all of which inhibit the ability to arrive at a causal estimate (in different ways). Nevertheless, by applying the a marginal structural model with a generalized estimating equation, subject to certain limitations (see below), we believe that this [a causal estimate] has, *en mass*, been achieved.

Of notable importance we arrived at several key inferences regarding the omnibus index. In particular, as discussed above, for any given country, policies translating to high NPI values (> 65.8) are causally related to a decrease in COVID-19 week-over-week case growth – a result in line with common intuition. A more interesting result however, and one which is seldom advocated for (heuristically speaking), is that although medium to low NPI policy schemes are associated with a decrease in COVID-19 cases, after accounting for confounding variables, such policies do no, in-fact, result in a decrease in case growth. The surprising result implies that, should a country fail to implement a policy above the aforementioned threshold (high NPI), any combination of policies would be unlikely to have any significant effect on the week-over-week case growth. That is, stringency, is unequivocally required.

Furthermore, it is worth noting that the present results did not find any causal effect of the NPI index on week-over-week mortality growth. While it naively then stands to reason that NPI has no effect on deaths, cumulative cases necessarily result in COVID-19-related deaths. Thus, NPI still has a critical role in the mitigation of deaths. Moreover, we observe that in many Countries (e.g., Canada) find themselves approaching a critical capacity in hospitals and intensive care unit facilities. Such as thing stand, an increase in cases necessarily takes a toll on these resources, which in turn compromise patient care downstream (i.e., motor vehicle accidents, heart attacks, etc.) [22].

Cumulative, we recommend, at least until such a time that herd immunity is achieved (in a given country), that combinations of policies, be consistently set above 55 points on the NPI scale.

5.3 Future Work

We appeal to the adage that "no model is perfect, but some models are useful". This certainly holds here, in that, subject to certain limitations, the analysis provides a rigorous, approximately causal inference on the effects of non-pharmaceutical interventions on COVID-19 case growth. More work has to be done for this ever-evolving crisis, but having a rigorous basis by which to inform public policy, is a large step (maybe even a leap) forward in the right direction.

Acknowledgements

We would like to thank Professor Eleanor Pullenayegum, the faculty mentor without whose support and advice this project would not have been possible. Additionally, we would like to thank Kuan Liu for her thoughts at the inception of the project. Additionally, although Times New Roman font (required by the submission) is not implementable in Latex due to proprietary reasons, we have used a close substitute [2].

[2]https://ctan.org/pkg/newtx

References

[1] S. Chaplin, "COVID-19: a brief history and treatments in development," Prescriber, vol. 31, pp. 23–28, may 2020.

[2] Y.-C. Liu, R.-L. Kuo, and S.-R. Shih, "COVID-19: The first documented coronavirus pandemic in history," Biomedical Journal, vol. 43, pp. 328–333, aug 2020.

[3] K. Yuki, M. Fujiogi, and S. Koutsogiannaki, "Covid-19 pathophysiology: A review," Clinical Immunology, vol. 215, p. 108427, 2020.

[4] C. I. Paules, H. D. Marston, and A. S. Fauci, "Coronavirus Infections—More Than Just the Common Cold," JAMA, vol. 323, pp. 707–708, 02 2020.

[5] A. Karaivanov, S. E. Lu, H. Shigeoka, C. Chen, and S. Pamplona, "Face masks, public policies and slowing the spread of covid-19: Evidence from canada," tech. rep., National Bureau of Economic Research, 2020.

[6] M. D'cruz and D. Banerjee, "'an invisible human rights crisis': The marginalization of older adults during the covid-19 pandemic – an advocacy review," Psychiatry Research, vol. 292, p. 113369, 2020.

[7] A. Lebret, "COVID-19 pandemic and derogation to human rights," Journal of Law and the Biosciences, vol. 7, 05 2020. lsaa015.

[8] T. Hale, N. Angrist, E. Cameron-Blake, L. Hallas, B. Kira, S. Majumdar, A. Petherick, T. Phillips, H. Tatlow, T. Baby, and W. Samuel, "Variation in Government Responses to COVID- 19." 2020.

[9] G. LLC, "Google COVID-19 Community Mobility Reports."

[10] A. Karaivanov, S. E. Lu, H. Shigeoka, C. Chen, and S. Pamplona, "Face masks, public policies and slowing the spread of covid-19: Evidence from canada," medRxiv, 2020.

[11] G. M. Fitzmaurice, N. M. Laird, and J. H. Ware, Applied longitudinal analysis, vol. 998. John Wiley & Sons, 2012.

[12] M. Sullivan Pepe and G. L. Anderson, "A cautionary note on inference for marginal regression models with longitudinal data and general correlated response data," Communications in Statistics - Simulation and Computation, vol. 23, pp. 939–951, jan 1994.

[13] J. M. Robins, "Marginal structural models versus structural nested models as tools for causal inference," in Statistical models in epidemiology, the environment, and clinical trials, pp. 95–133, Springer, 2000.

[14] R. Crouchley and R. B. Davies, "A comparison of population average and random-effect models for the analysis of longitudinal count data with base-line information," Journal of the Royal Statistical Society. Series A (Statistics in Society), vol. 162, no. 3, pp. 331–347, 1999.

[15] J. M. Robins, M. Á. Hernán, and B. Brumback, "Marginal Structural Models and Causal Inference in Epidemiology," Epidemiology, vol. 11, no. 5, 2000.

[16] S. R. Cole and M. A. Hernan, "Constructing Inverse Probability Weights for Marginal Structural Models," American Journal of Epidemiology, vol. 168, pp. 656–664, jul 2008.

[17] V. Chernozhukov, H. Kasaha, and P. Schrimpf, "Causal impact of masks, policies, behavior on early covid-19 pandemic in the us," arXiv preprint arXiv:2005.14168, 2020.

[18] M. A. Hernan and J. M. Robins, Causal Inference: What If. Boca Raton: Chapman & Hall/CRC, 2020.

[19] O. Saarela, D. A. Stephens, and E. E. Moodie, "The role of exchangeability in causal inference," arXiv preprint arXiv:2006.01799, 2020.

[20] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, Bayesian data analysis. CRC press, 2013.

[21] D. Westreich and S. R. Cole, "Invited Commentary: Positivity in Practice," American Journal of Epidemiology, vol. 171, pp. 674–677, 02 2010.

[22] S. J. Lange, M. D. Ritchey, A. B. Goodman, T. Dias, E. Twentyman, J. Fuld, L. A. Schieve, G. Imperatore, S. R. Benoit, A. Kite-Powell, Z. Stein, G. Peacock, N. F. Dowling, P. A. Briss, K. Hacker, A. V. Gundlapalli, and Q. Yang, "Potential indirect effects of the covid-19 pandemic on use of emergency departments for acute life-threatening conditions — united states, january–may 2020," American Journal of Transplantation, vol. 20, no. 9, pp. 2612–2617, 2020.