An Econometric Panel Data Model of the COVID-19 Pandemic

Antoine Djogbenou*, Christian Gourieroux**
Joann Jasiak*, Paul Rilstone* [1]

First Draft: June, 2020, this draft: July, 2021

New flexible-form and semi-parametric autoregressive non-linear count models for panel data are developed to analyse the spread and containment of the COVID-19 pandemic. The models are based on a discrete time form of the SIR model. These methods lead naturally to estimators of the infection process and daily reproduction numbers by jurisdiction. Two semi-parametric versions of the reproduction numbers are developed corresponding to currently popular parametric estimators. The estimators are applied to a large international data set to estimate these parameters for some 200 jurisdictions at both national and subnational levels.

1 Introduction

Anyone who has followed the reporting of the COVID-19 pandemic is familiar with the graphical presentation and comparison of the data from a variety of jurisdictions, the characterization of the growth rates as potentially "exponential" and efforts to "flatten" the curves of new infections. A common, but not uniform, initial characteristic across jurisdictions was a roughly log-quadratic trend of new infections, although this seems to have been belied by numerous resurgences of the virus. The similar, but by no means identical time-paths[2] seem to reflect a certain homogeneous underlying structure. In fact, although the time paths across jurisdictions may be eventually quite distinct, it would seem that since all spring from what is (basically) a common virus, there should be some homogeneity that can be exploited for better estimation. This paper seeks to add to this discussion by developing a number of panel and count data techniques based on an underlying epidemiological model and applying them to international data available on the pandemic. A central focus of this study are the basic or effective reproduction numbers which we show are naturally defined with respect to the parameters of the count data models we develop. By pooling the data we are able to obtain more efficient estimates, all the while allowing for heterogeneity across different jurisdictions. We allow for various assumptions regarding homogeneity of the pandemic processes across jurisdictions and time and explore the implications of these for our understanding of the pandemic. In doing so, we introduce some new statistical

[1]*York University, ** University of Toronto, Toulouse School of Economics and CREST, Correspondence to Rilstone at pril@yorku.ca .

[2]Linton (2020) categorizes international data as falling into one of five groups: Early stages, Middle Age, Over the Hill, Twin Peaks, Resurgence.

techniques, including variations of the basic exponential count model and flexible-form and semi-parametric modelling of the different deterministic trends across jurisdictions.

At writing, the pandemic has been ongoing for many months with many jurisdictions accumulating daily data. This would seem to provide a substantial amount of data for each jurisdiction, commonly with one hundred or more observations. Nevertheless, estimates based on jurisdiction data have been criticized for their unreliability, notably in terms of poor forecasts. Unfortunately, there is substantial noise in daily data[3]

We focus in this paper on the infections equation. Most of the published aggregate level forecasts are based on deterministic trend (albeit sophisticated) analysis. Linton (2020) reports a variety of country-level estimates of models using a non-parametric local polynomial approach. We refer back to this below. Many public and public health organizations have analysed the data using a variety of local trends to estimate and forecast. It is difficult to determine exactly the methods they use from any publicly available information. We model the infections equation as having stochastic and deterministic components. A non-linear long auto-regressive series can exhibit behaviour which appears to be moving along a log-quadratic deterministic trend. Simple examination of patterns in the data can lead to misleading conclusions. The notion of "flattening" the curve can be interpreted as referring to the raw moments of the reduced form projection of infections on a time trend. While this may be a pleasing from a policy perspective, it reflects the fact that the pandemic is evolving and that forecasts built on an estimated trend may be inaccurate.[4]

Count models and panel data have been used extensively in empirical health economics and epidemiology. Most of the aggregate data available on the pandemic is by definition count data and is collected at a variety of jurisdictional levels, i.e., panel data. The principal advantage of panel data is the potential it has for providing more information than a single series. This usually entails a presumption of some poolability across units. There are various ways to incorporate homogeneity, from assuming all jurisdictions are identically distributed, to allowing certain forms of heterogeneity, to modelling each jurisdiction separately. If we maintain that biologically the disease is largely time and location invariant, then it makes sense to exploit the fact that certain key parameters of the process should be alike in different jurisdictions[5]

One important consideration makes panel data especially appropriate here. The infections (or renewal) equation, at least as modelled using daily observations, is a long autoregressive process requiring the estimation of many parameters. There is scant information available with individual time series to allow for estimation of a large number of parameters with a high level of accuracy. By assuming a certain homogeneity of the disease across jurisdictions, we are able to estimate these common parameters with a high degree of accuracy.

We develop/apply a number of count data estimators to examine some of the main ag-

[3]It is common to see rolling averages reported, these smooth out the noise but do not really impart more information. In fact, this may have a substantially harmful effect on inferences in that if the time-path of infections is non-linear, a rolling average may tend to obscure this.

[4]The popular initial work used log-quadratic functional form of the pandemic in part to forecast "peaks" in the virus' spread which is immediate with a quadratic model. Subsequent waves of the virus belie any such simple trend.

[5]There is some evidence that there are various strains or mutations of the virus. While this is quite possible, small differences in the virus across outbreaks should have minimal effects on the estimation.

gregate features of the pandemic including basic regression models and maximum likelihood estimators. The count models are used to estimate the parameters of the "renewal equation" popular in epidemiology. These are particularly amenable for constructing a model which is both epidemiologically and statistically sound. Much of the focus is in a Poisson framework and its extensions. This is done largely for mathematical convenience and coherence. Its extensions allow for a wide variety of assumptions. The Poisson assumption can be relaxed substantially, retaining its main components. Poisson estimators are known to have certain robust properties even when the underlying distribution is not Poisson.[6] The panel Poisson model was perhaps first used by Hausman, Hall and Griliches (1984). There are various textbook discussions of the basic estimator including Cameron and Trevidi (2005) and Wooldridge (2010). The data is effectively an unbalanced panel with observations commencing with first reported incidence.

One of the standard methods of incorporating heterogeneity is through allowing for unobserved heterogeneity, either random or fixed effects. We estimate the fixed effects version as one way of allowing for heterogeneity across jurisdictions. This of course only allows for time-invariant tilting in the regression function. To allow for differences in the trend between jurisdictions we introduce a flexible version of the fixed-effects model and a variant of Robinson's (1988) semi-parametric estimator. If we assume that the autoregressive component of the reproduction equation is effectively homogeneous across jurisdictions, we can estimate this by a similar method.

Some consideration should be given to the issue of stationarity of the pandemic process. On the one hand, plots of individual incidents often appear explosive at the beginning of most of the outbreaks. (If they were not we clearly would not have locked down good parts of the national economies of the world.) On the other hand, many if not most outbreaks have become under control and infection levels have decreased (at writing) in most jurisdictions. Thus, it is arguable that the contagion processes are stationary, either around a trend, or conditional on other factors. From an inferential perspective non-stationarity is problematic for various reasons, not the least of which is that standard asymptotic critical values based on stationarity may be invalid. Our implicit assumption is that the processes are stationary around some trend, albeit the specific form of which is unknown. We thus make use of standard asymptotic results using George Box' caveat that any statistical model is at best an approximation. Hopefully our results can provide some useful insights. An additional argument for stationarity is that the usual tests for stationarity would not necessarily refute the hypothesis of non-stationarity if endogenous measures were taken such as to keep the spread of the pandemic in check to the side of explosiveness. Many countries have taken measures on an incremental basis, slowly ratcheting up off-setting measures. So (and this of course may sound self-serving on the side of stationarity), the "flattening of the curve" can be seen as a shifting of the trend line. Another argument for stationarity follows from the so-called herd-effect. At a certain point the percentage of the population at risk naturally diminishes which eventually puts a damper on disease propagation. The immediate extent of a herd-effect is a disputed topic. At time of writing, the introduction of various successful

[6]The textbook derivation of a Poisson distribution corresponds to situations of aggregated observations on binary events in which the individual probability of an infection is very small and the number at risk is very large. This is not a completely inaccurate characterization of the data used here on the COVID-19 pandemic.

vaccines would imply that the number of individuals at risk who are exposed to COVID-19 is decreasing.

The standard models of COVID-19 incidence, while insightful, overlook a couple of important elements which can lead to misleading results. The analyses are typically presented on a country by country basis and are based on fitting data mid-spell. They are often largely descriptive in nature, fitting what we could call a reduced form trend to the data without allowing for direct and accumulated feedback from existing infections. This can result in spurious correlations. Ignoring this and fitting a rolling trend can exaggerate the effect of a trend and potentially lead to what appears to be a flattening of the curve. While governmental policies most certainly had an impact on the spread of the disease, it is not clear to what extent this was the case, or whether individuals may have been collectively changing their behaviour regardless of government fiat.

The discussion is organized as follows. In the next section we discuss the data used in the study. Section 3 develops the simple discrete infections model linking this in a transparent way to the notion of a basic or effective reproductive number in epidemiology[7]. Section 4 develops corresponding econometric estimators for this model, provides identification and asymptotic distributional results and discusses a number of other estimation issues. Section 5 provides a summary of our empirical results. Section 6 concludes.

2 Data

The principal source of data used in this study is the Johns Hopkins githubsite. This has data on over 260 jurisdictions. Some jurisdictions have been removed from this version. We retained only those that had observations from their first reported infection (i.e. no left censoring). The data for this version is up to June 29, 2021. The earliest observation is January 22, 2020. The data is an unbalanced panel data set due to the differences between the first reported infections in different jurisdiction.

One of the issues of debate in the current COVID-19 pandemic literature is over the relation between reported and actual infection rate. This has an impact on the study in several ways.

Reported infections are most certainly an understatement of actual infections. There already has been a number of studies on this and there will certainly to be many more studies of this issue. While we would most certainly prefer to work with actual infections, if we consider the proportion of observed to actual cases to be largely constant, then we can make inferences in terms of elasticities. For this study we are using reported infections. As more information becomes available this could be incorporated. We note that changes in reported infections brought on, say, by increased testing in most jurisdictions, are allowed for in the flexible form and semi-parametric approached developed below.

A second, not-unrelated issue is the period between infection of an individual and the date when that case is reported. We may think of this an an incubation period, although incubation as referring to the time between infection and manifestation of symptoms has a

[7]Since the reproductive numbers defined in this paper vary by jurisdiction and time, they are more appropriately referred to as effective reproduction numbers or ratios.

slightly different meaning. Some individuals may stay completely asymptomatic throughout their infection or some may not find their symptoms sufficiently extreme to warrant reporting/diagnosis so will not appear as infected.[8] Through testing, some who are in this category (and this may be increasing) may now show in this category. At any rate there is most typically a lag between infection and recorded infection. The renewal equation includes all infectious individuals. In principle this can generate a very long autoregressive process, from the date of infection, through the period of incubation and to the date of recovery (or death), or quarantine, with the latter being imperfect and of variable effectiveness.

It is difficult to think of a situation in which data has been such a contentious political issue amidst claims of under or over-reporting and so on. We simply take the data as they are. It is certainly the case that there have been some administrative issues. We have, with minimal caveats used the data as collected by third parties and publicly available. Perhaps the one caveat, apart from deleting left-censored observations, is that in a few cases (probably) due to updates, cumulative numbers decreased resulting in negative counts. These were recoded as zeros when minimal; we deleted jurisdictions (including Italy and the UK) which had more than 10 negative counts. We also omitted a few obviously anomalous cases such as the Island Princess.

The cases used in the study thus correspond (caveats mentioned above) to the data freely available at the Johns Hopkins website. Most of the cases or jurisdictions are at the national level of aggregation. Some, such as Canadian (province), Australian (state) and some Chinese data[9] are at a subnational level. These could be aggregated in various ways either by country or region. We have left as is for transparency and to allow for more heterogeneity[10]. The heterogeneity in the cases in the data thus almost mirrors that of most international data sets. Thus, there are large differences between observations. This we view as a benefit as our analysis allows for a separation of the biological component of the pandemic from the environmental component. Conversely, the techniques in this paper could be applied to other data sets in which the collection is less aggregated, say at a county level. In this case, there would arguably be much more dependence between observations and this would need to be incorporated into the modelling strategy.

3 Infections, $\mathcal{R}_0$ and Count Models

Much of the modern literature on the spread of pandemics is based on the original SIR differential model of Kermack and McKendrick (1927, 1932, 1933). The SIR model and its extensions are useful for conceptualizing the characteristics of disease spread within a deterministic setting, but may be somewhat limited to characterize the observational, heterogeneous and discretely sampled data which is typically available. The simplest idea in the study of contagious diseases is that of a reproduction number or ratio. As per Heesterbeek and Dietz (1996), "$\mathcal{R}_0$ is the expected number of secondary cases produced by a typical infected individual during its entire infectious period, in a population consisting of suscepti-

[8]The Stanford (Bendavid and Bhattacharya, 2020) and other studies emphasize this.

[9]Note that since the data set begins January 22, 2020, a number of the Chinese provinces and other sub-national jurisdictions are unfortunately excluded.

[10]It could also lead to more violations of the independence assumption. See below for a discussion of this.

bles only". "$\mathcal{R}_0$" is a convenient and familiar generic term. For the most part we allow the reproduction number to vary by jurisdiction and jurisdiction and time and accordingly use the notations $\mathcal{R}_j$ and $\mathcal{R}_j(t)$.

There are numerous versions of the continuous-time SIR model (see Heesterbeek and Dietz, 1996) with correspondingly somewhat varying definitions of $\mathcal{R}_0$ which lead to somewhat different estimates (Chowell, Gerardo, Nishiura and Bettencourt, 2006 and Li, Blakeley and Smith, 2011). In its most basic form, $\mathcal{R}_0$ ignores or holds constant certain time factors (including a specific individual's infectiousness), exposure intensity factors and responses to the infection, either medical or social. Nevertheless it is a useful place to start. An infection is seen as the event of being contaminated by the virus. The probability of such an event varies across individuals and each individual's infectiousness will vary randomly. We develop a discrete version commonly (see, e.g. Allen, 1994 and Champredon et al, 2018) referred to in the epidemiology literature as the renewal equation.

The discrete version of the renewal equation as in Allen (1994) is derived from a first-order difference equation which can be seen as a time-discrete version of the corresponding first-order differential renewal equation in the original SIR model. Chowell and Nishiura (2008) and Li, Blakeley and Smith (2011) survey various methods to estimate reproduction numbers including those from survival methods and imputation from empirical growth rates.

Two alternative approaches to defining the reproduction number are reviewed by Fraser (2007). The first, the "case reproduction number" conforms with the above definition and which we develop for count models here. The second, the "instantaneous reproduction number" can be obtained from the case approach by imposing a constraint. The instantaneous number has been popularized by e.g. Cori et al. (2013) and has been widely adopted by public health authorities. It is numerically simpler to implement. We show at the end of this section how to adopt the methods here to obtain a semi-parametric version of it. We develop a discrete infections equation in an alternative manner from simple first principles and the concept of the effective reproduction number, which varies across jurisdictions and time. This leads directly into an estimable autoregressive count model. We first assume a constant (over time) exposure rate so that a random individual i who becomes infected in a jurisdiction j is considered to have the same number and kind of encounters with uninfected (susceptible or at-risk) individuals in jurisdiction j each day. This is a classic statistical experiment in a laboratory setting with exact replication each day. The assumption of a constant exposure rate is effectively plausible if a population is sufficiently large that those who become infected represent a small proportion of the at-risk set. Note that this excludes in a sense a "herd" effect which diminishes the exposure rate in terms of exposure to at risk individuals. Allowing for a herd effect is one way of thinking of subsequent draws *without* replacement. We want to connect this to an autoregressive framework. Note that in continuous time an individual may be immediately contagious and the disease's virulence within that individual would be changing continuously over time. However, this needs to be modified for discrete time. Here, everything is framed within the context of the impact of new infections and allows that their case may be "primary", "secondary" or otherwise.

Consider an individual i who becomes infected in time t. In each subsequent period, $t+s$, individual i is introduced into an environment (call it jurisdiction j) which we first assume is homogeneous over time. A person is infectious up to a maximum of m periods. Individual i will directly cause $I_j(t+s)$ infections in jurisdiction j in periods $t+s$, $s = 1, \ldots, m$. The

$I_j(t+s)$ are independent, which we assume is plausible with a fairly large population. Note that the $I_j(t+s)$ are random: they depend primarily on the amount of the dosage ingested by i and the toxicity of i. Thus the number of (secondary) infections caused by individual i in jurisdiction j is $\sum_{s=1}^{m} I_j(t+s)$.

If we decompose transmission into the product of time-varying virulence, r_s with a constant (over time) exposure factor we have $\mathbb{E}[I_j(t+s)] = r_s A_j$, where A_j is exposure rate in jurisdiction j. Normally, but not necessarily, we expect r_s to be decreasing in s. There may be a hump corresponding to initial infectiousness. After m (m finite) periods a person is no longer infecting others (they may be still sick, completely quarantined, recovered or dead, but not infecting others so that $r_s = 0$, $s > m$. Following Heesterbeek and Dietl (1990) we refer to r_s as the reproduction function which may be seen as the product of a measure of the infectivity of an individual s days from infection and the survival function of the infection. In our context these are not separately identifiable. We have that the basic (perhaps better referred to as the effective) reproduction number in jurisdiction j is

$$\mathcal{R}_j = A_j \sum_{s=1}^{m} r_s. \tag{3.1}$$

We have changed indexes to indicate that the reproduction number varies by jurisdiction. We connect this to an autoregressive process as follows. In period $t-m$ suppose there are $y_{j,t-m}$ new infections, then this contributes in conditional expectation to $A_j r_m y_{j,t-m}$ new infections in period t. In $t-m+1$ if there are $y_{j,t-m+1}$ new infections, then this contributes in conditional expectation, to $A_j r_{m-1} y_{j,t-m+1}$ new infections in t. In $t-1$ if there are $y_{j,t-1}$ new infections, then this contributes in conditional expectation to $A_j r_1 y_{j,t-1}$ new infections in period t. So, adding these together, in terms of conditional expectations, if infectiousness lasts m periods we have

$$\mathbb{E}[y_{j,t} | y_{j,t-1}, \ldots, y_{j,t-m}] = A_j \sum_{s=1}^{m} r_s y_{j,t-s}. \tag{3.2}$$

Consider now if the exposure rates, which we now indicate by $A_j(t)$, change by jurisdiction and time period. It's useful to walk through a chain of examples.

- If an individual i (in jurisdiction j) gets ill in period $t-m$,
  - in period $t-m+1$ individual i is expected to directly infect $r_1 A_j(t-m+1)$ individuals (these may go on to infect others),
  - in period $t-m+2$ individual i is expected to infect $r_2 A_j(t-m+2)$ individuals,
  - in period $t-1$ individual i is expected to infect $r_{m-1} A_j(t-1)$ individuals and so on until
  - in period t individual i is expected to infect $r_m A_j(t)$ individuals and
  - in period $t+1$ individual i is expected to infect 0 individuals.
- If an individual i gets ill in period $t-m+1$,

- in period $t-m+2$ individual i is expected to infect $r_1 A_j(t-m+2)$ individuals,
- in period $t-m+3$ individual i is expected to infect $r_2 A_j(t-m+3)$ individuals
- in period t individual i is expected to infect $r_{m-1} A_j(t)$ individuals,
- in period $t+1$ individual i is expected to infect $r_m A_j(t+1)$ individuals and
- in period $t+2$ individual i is expected to infect 0 individuals and none thereafter.

- If an individual i gets ill in period $t-1$.
  - in period t individual i is expected to infect $r_1 A_j(t)$ individuals,
  - in period $t+1$ individual i is expected to infect $r_2 A_j(t+1)$ individuals,
  - in period $t+m-1$ individual i is expected to infect $r_m A_j(t+m-1)$ individuals,
  - in period $t+m$ individual i is expected to infect 0 individuals.

Noting which previous infections cause infections in period t and by which factor, we have, in terms of conditional expectations,

$$\begin{aligned}\mathbb{E}[y_{j,t}|y_{j,t-1}, y_{j,t-2}, y_{j,t-m}, t] &= \big(r_1 y_{j,t-1} + r_2 y_{j,t-2} + \cdots + r_m y_{j,t-m}\big) A_j(t) \\ &\equiv Y'_{j,t-m} r A_j(t)\end{aligned} \tag{3.3}$$

noting that all terms are positive, where $Y'_{j,t-m} \equiv \big(y_{j,t-1}, \ldots, y_{j,t-m}\big)$ and $r = \big(r_1, \ldots, r_m\big)'$. We also have that the effective reproduction number, denoted here $\mathcal{R}_j(t)$, varies by jurisdiction and time period so that

$$\mathcal{R}_j(t) = \sum_{s=1}^{m} r_s A_j(t+s) \tag{3.4}$$

with the same parameters as from the autoregression. Note that, from a count data perspective, the reproductive number can be seen as simply a weighted sum of upcoming exposure factors, wherein the weights are corresponding measures of a typical person's infectivity at those times.

Note the formal similarity of this to the renewal equation in Cori et al. (2013) and Champredon et al. (2018)[11] Cori et al. (2013) use a Poisson distribution for the y_t with conditional mean $R_t(\theta_2) \sum_{i=1}^{t-1} w_i(\theta_1) y_{j,t-i}$ where $R_t(\theta_2)$ and the w_i's are parametrically specified (e.g, as Gamma and serial interval distribution.) Champredon et al. (2018) use a variation specified as (adapting for our notation)

$$y_{j,t} = \frac{S_{jt}}{N_j} \mathcal{R}_0 B_{jt} \sum_{i=1}^{m} \gamma(i) y_{j,t-i} \tag{3.5}$$

where S_{jt}, N_j and B_{jt} represent susceptible (at risk) individuals, population and other factors such as distancing. Champredon et al. (2018) have the same variables with two differences.

[11] This is a discrete time form of their equation as presented by Champredon (2020).

First, in our case, the effective reproduction numbers are directly derived as functions of the autoregressive parameters of the process. In their studies γ is a distribution function whose parameters are estimated. Second, although in principle at-risk population density could be introduced, there are identification and normalization issues involved when multiplying these different factors.

Note here that we are implicitly assuming independence across jurisdictions. An alternative multivariate time series count model introduced by Held and Paul (2012), also Meyer and Held (2014) allow for interdependence across jurisdictions (or regions). They allow for only one lag in their approach and their focus is more on transmission across jurisdictions rather than the dynamics of the process. A gravity model similar to this would be an interesting extension of this study, but beyond its scope. We note that the almost global ban on international travel some weeks into the pandemic effectively imposed a high degree of independence across jurisdictions. [12]

Many writers have been critical of the notion of a constant reproduction number parameter, although the concept has seen widespread use and it certainly makes sense to assign some sort of measure to the transmissibility of a disease. In a laboratory setting it is useful to consider factors by which different diseases will be transmitted by a single person over the course of their infection (this itself is unlikely to be constant for various reasons) for given rates of exposure. There are numerous secondary factors which can be held constant. Where the problems really arise is in maintaining a constant exposure rate. We use Equation 3.3 as the basis for estimation. Note that $r_s \geq 0$. In the empirical work we parametrize so that the r_s ='s are positive.

An important consideration is the multifaceted relationship between true infections (largely unobservable) and administratively recorded infections (observable with some error). That infections show up with a lag is one issue. That increasing numbers are being tested over the sampling is another. In our autoregressive models, if true and recorded infections are proportional and with a constant factor of proportionality, then the parameters of the autoregression are the same for the observed and unobserved models.

Much is unknown about the corona virus and different mutations of it are likely to have somewhat different transmission properties and manifest themselves with somewhat different symptoms. Another issue is the incubation period (from initial transaction to manifestation in human symptoms). Indeed the extent to which the virus manifests itself is itself an important topic. From an estimation perspective we allowed for the length of infectiousness to be up to two weeks which seems quite adequate in most cases.

To close off this section we consider the interdependence of observations across jurisdictions and how that effects the analysis here, which it potentially does in two ways. First, *ipso facto*, infections across jurisdictions are clearly not independent due to migration. The above model could be modified to allow for individuals to cross from one jurisdiction to another during their infectiousness period. This kind of adjustment (not possible using the data set employed here) could have some effect on the analysis, but in our case the effects of such an adjustment would be small for a number of reasons. First, although the initial

[12]There is also an implicit assumption that a (secondary) person is only infected by one (primary) person. This assumption is made in the original SIR work, the argument being that the initial contamination swiftly dominates any subsequent contamination.

cases of the pandemic for any jurisdiction were introduced by individuals from a different jurisdiction (typically midstream of that individual's infectiousness period), the ensuing secondary cases follow the modelling above and the relative contribution of the primary cases to secondary cases is soon dwarfed by domestically infected individuals. So while allowing for inter-jurisdictional transmission of the pandemic would be important to a study focused on that, it has little relative impact here. Moreover, the enhanced controls on travel introduced by almost all countries (and many sub-national jurisdictions) by March, 2020 virtually eliminated this concern as an issue for our purposes here. It is commonly reported that foreign sources account for no more than one or two per cent of total direct infections. A second indirect way in which the independence issue impacts here is a concern in many statistical studies that there are unobservable effects which impact on all observations making them in some way dependent. We choose, at this point, to address these kind of concerns by appropriate adjustment to standard errors.

4 Econometric Models and Issues

This section deals with a number of estimation issues. We first consider a number of econometric models and corresponding estimators which are consistent with the basic epidemiological count model of Section 3 for a range of assumptions regarding the time trend component. We show how new estimators of the reproduction numbers can be computed for each of these models. We provide discussions of identification of the parameters in each model as well as basic asymptotic results for the estimators. Finally, we consider a number of other estimation considerations that arise in this context.

Since we are modelling count variables it is convenient to specify a conditional Poisson distribution for the infections, $y_{j,t}$. This framework allows for incorporation of the non-negative integer nature of $y_{j,t}$ and provides simple intuitive ways to capture heterogeneity and dependence. The approach can then simply be subsequently modified to allow for non-Poisson features while retaining the basic intuition. Poisson estimates are also robust to departures from the Poisson distribution. We first state a few well-established results regarding the standard fixed-effects Poisson model and then extend this in a couple of ways to flexible-form estimation and semi-parametric estimation.

4.1 Fixed Effects

The basic fixed-effects model can be written such that the conditional mean for an observation $y_{j,t}$ can be written as $h_{jt}(r)e^{\alpha_j}$ where $h_{jt}(r)$ is a function of conditioning variables and a parameter r. The log-likelihood function for observation j,t is written as

$$l_{jt}(Z_{jt};\alpha_j,r) = y_{j,t}\log\lambda_{jt} - \lambda_{jt} - \log y_{jt}! \tag{4.1}$$

where $\lambda_{jt} = h_{jt}(r)e^{\alpha_j}$ and Z_{jt} denotes $y_{j,t}$ and any conditioning variables. An equivalent parameterization which we also use sets $c_j = e^{\alpha_j}$.

In this paper we primarily use the specification [13] $h_{jt}(r) = Y'_{j,t-m}r$, although certain

[13] When $y_{j,t-s} = 0, s = 1,\ldots,m$ we set $l_{jt} = 0$.

other specifications, including introduction of additional covariates, can be accommodated. For now we assume that all jurisdictions are infected on the same date, i.e., a balanced panel. This is modified below to allow for different first infection dates. The pooled Poisson estimator sets $\alpha_j = \alpha$ for all $j = 1, \ldots, N$. The standard fixed-effects estimator of r can be obtained in a couple of ways. For distributional results it is simplest to follow Anderson's (1970) (adopted by Hausman, Hall and Griliches, 1984, for the Poisson case) method of deriving the likelihood of $y_{j,1}, \ldots, y_{j,T}$ conditional on $\bar{y}_j = \sum_{t=1}^{T} y_{j,t}/T$ which leads to a multinomial distribution for the $y_{j,1}, \ldots, y_{j,T}$'s. For modelling purposes we find it convenient to first maximize with respect to each α_j, for given r, and concentrate out the α_j's. The distributional results may be more difficult to derive directly in this case, but we find the interpretation of the model simpler in this form and also more readily extendable. For given r, the maximizers of $\sum_t l_{jt}(Z_{jt}; \alpha_j, r)$ with respect to α_j solve

$$e^{\alpha_j} = \frac{\bar{y}_j}{\bar{h}_j(r)} \equiv c_j, \qquad j = 1, \ldots, N \tag{4.2}$$

where $\bar{y}_j$ and $\bar{h}_j(r)$ represent averages for jurisdiction j. Summing over t, the single observation concentrated log-likelihood (up to a constant) is written

$$l_j(r, \alpha_j(r)) = \sum_t l_{jt}(r, \alpha_j(r)) = \sum_t y_{j,t} \ln \left(\frac{h_{jt}(r)}{\sum_t h_{jt}(r)} \right) + \text{const} \tag{4.3}$$

It is common in count analysis to adjust for "exposure rates" by scaling each hazard by a measure of the jurisdiction's population or population density. This is typically redundant with a fixed-effects model as multiplicative time invariant variables are absorbed into the fixed effects. Standard exposure rate adjustments do not affect the slopes of the fixed-effects estimators.

Under wide ranging conditions the Maximum Likelihood Estimator of r is known to be consistent and asymptotically normal. The score can be written as

$$s_j(r) = \sum_{t=1}^{T} \left(y_{jt} - \frac{\bar{y}_j}{\bar{h}_j(r)} h_{jt}(r) \right) \frac{\bigtriangledown h_{jt}(r)}{h_{jt}(r)} \tag{4.4}$$

where $\bigtriangledown h_{jt} \equiv \partial h_{jt}/\partial r$. Note that this holds for as few as $T = 2$ although it does fall apart if $T = 1$.

The fixed-effects estimator is useful as a benchmark to compare against other estimators and also as a segue to understanding a class of flexible-form estimators we now look at.

4.2 Time-Varying Fixed Effects

From the renewal equation we allow for the autoregressive component to be homogeneous across individuals, but the deterministic component is allowed to vary across jurisdictions. We propose a variety of flexible form and semi-parametric approaches for which the asymptotics are straightforward. Consider first a count variable whose mean over an interval is proportional to a common factor. Denote the intervals as

$$\mathcal{I}_{jl} = \begin{cases} [0, T_1), & l = 1 \\ [T_{l-1}, T_l), & 1 < l < L \\ [T_{l-1}, \infty) & l = L \end{cases} \tag{4.5}$$

where over interval $\mathcal{I}_{jl}$ we have

$$\mathbb{E}[y_{j,t} | t \in \mathcal{I}_{jl}] = h_{jt}(r) e^{\alpha_{jl}} \tag{4.6}$$

We will use somewhat of an oxymoron and refer to this as a time-varying fixed-effects model. We write the single observation log-likelihood as

$$l_j(r, \alpha_j) = \sum_{l=1}^{L} \sum_{t \in \mathcal{I}_{jl}} (y_{j,t} \ln(h_{jt} e^{\alpha_{jl}}) - h_{jt}(r) e^{\alpha_{jl}}) + \text{const}$$

where $\alpha_j(r) = \begin{pmatrix} \alpha_{j1}(r) & \cdots & \alpha_{jL}(r) \end{pmatrix}$. Let n_{jl} denote the number of observations in $\mathcal{I}_{jl}$, $\bar{y}_{jl}$, $\bar{h}_{jl}$ denote the corresponding averages of the y_{jt} and h_{jt} over those intervals. We see simply that with $c_{jl} = \bar{y}_{jl} / \bar{h}_{jl}$, maximizing the corresponding likelihood over the α_{jl}'s results in estimators: $e^{\alpha_{jl}} = c_{jl}$ and the single observation concentrated log-likelihood is written

$$\begin{aligned} l_j(r, \alpha_j(r)) &= \sum_{l=1}^{L} \sum_{t \in \mathcal{I}_{jl}} \left(y_{jt} \ln \left(h_{jt}(r) \frac{\bar{y}_{jl}(r)}{\bar{h}_{jl}(r)} \right) - h_{jt}(r) \frac{\bar{y}_{jl}}{\bar{h}_{jl}(r)} \right) + \text{const} \\ &= \sum_{l=1}^{L} \sum_{t \in \mathcal{I}_{jl}} y_{jt} \ln \left(\frac{h_{jt}(r)}{\bar{h}_{jl}(r)} \right) + \text{const} \end{aligned} \tag{4.7}$$

with the middle term on the first line being absorbed into the constant since

$$\sum_{t \in \mathcal{I}_{jl}} \left(h_{jt} \frac{\bar{y}_{jl}}{\bar{h}_{jl}} \right) = n_{jl} \bar{y}_{jl} \tag{4.8}$$

and we also note that we may interpret $\bar{y}_{jl}$ and $\bar{h}_{jl}$ as method of moments estimators of $\mathbb{E}[y_{j,t} | t \in \mathcal{I}_{jl}]$ and $\mathbb{E}[h_{jt} | t \in \mathcal{I}_{jl}]$.

We obtain the usual fixed effects estimator by setting $L = 1$ in which case the $\bar{y}_{jl}$ and $\bar{h}_{jl}$ are the averages over all the observations for jurisdiction j.

Corresponding to the standard fixed-effects case we can write down the single observation score for this extension thereof as

$$s_j(r) = \sum_{t=1}^{T} \left(y_{jt} - 1_{jlt} \frac{\bar{y}_{jr}}{\bar{h}_{jl}(r)} h_{jt}(r) \right) \frac{\bigtriangledown h_{jt}(r)}{h_{jt}(r)} \tag{4.9}$$

where

$$1_{jlt} = \begin{cases} 1, & t \in \mathcal{I}_{jl} \\ 0, & \text{otherwise} \end{cases} . \tag{4.10}$$

Note that for the estimator to work we need at least two observations to fall within each interval, although presumably more is better. Also note that we are dividing the observations up into intervals along the line, but in a more general sense we could divide the observations up into different subgroups or "clusters" so long as these are known.

We note that for given r

$$\frac{\bar{y}_{jl}}{\bar{h}_{jl}(r)} = e^{\hat{\alpha}_{jl}} \tag{4.11}$$

provides estimates of the α_{jl}'s, although without specifying the number of elements in $\mathcal{I}_{jl}$ we need to be careful when using these individually for inferences. If the number of observations in an interval is large, then we may be fairly comfortable using asymptotic theory to make inferences about the α_{jl}'s.

Note that for *fixed* L, this is a straightforward extension of the fixed-effects estimator. Its statistical properties follow from the Anderson set up. It can be seen as a multinomial estimator with the likelihood conditional on the $\bar{y}_{jl}, l = 1, \ldots, L$. This may be seen as a step-wise constant approximation to a time trend. We stay agnostic with respect to the choice of knots. Some authors have already effectively considered specific dates for these. We estimate these models for fixed values of L (actually modified to L_j, $j = 1, \ldots, N$ to allow for different initial pandemic starting dates). We could also incorporate some curvature within regions, but stay simple leaving the fixed effects constant over each interval.

It is important to link this variable fixed effects specification to the infections equation. Note here that the conditional mean of $y_{j,t}$ can be written as

$$\mathbb{E}[y_{j,t}|Y_{j,t-m}, t] = Y_{j,t-m}^{\top} r A_j(t) \tag{4.12}$$

where

$$A_j(t) = \sum_{l=1}^{L_j} e^{\alpha_{jl}} 1[t \in \mathcal{I}_{jl}] \tag{4.13}$$

where $1[]$ is the usual indicator function. Thus, in this set up we allow for a flexible approximation to the time trend via a step function. Note that the fixed-effects estimator is a special case with $A_j(t) = e^{\alpha_j}$, the deterministic trend being constant for all j.

In our empirical work we considered various choices for the length of these intervals. If these are too long, the trend effect is washed out; choosing intervals too short results in over-fitting and nonsensical values for the autoregression parameters. We found a good compromise was between 21 and 31 days. We rounded up the number of periods in the first interval for each jurisdiction (i.e. each $\mathcal{I}_{jl}$ contained L observations with the first interval containing at least L observations). We return to this estimator after the following discussion.

4.3 Semi-Parametric Modelling

The semi-parametric literature has generalizations that effectively allow for smoothing of the step-wise function. One is to allow for curvature over each interval such as having quadratic or cubic splines over the knots. The other is to allow L to increase. The latter can be problematic for technical reasons but also because of what we are willing to assume regarding the size of T: whether it is fixed or growing and at what rate. We touch briefly on

this but given our interest here is on applications we keep the discussion and assumptions at a high level. Introducing smoother functions via splines is one method of approximation which has been used by some analysts of the pandemic. We find a simpler way is through kernel and local polynomial estimators[14], which will lead to a more sophisticated estimation approach than the time-variant fixed effects estimator.

To do so we consider a variation of Robinson's (1988) semi-linear regression model. Semi-parametric estimators for other similar non-linear models have been proposed so we keep the asymptotics fairly high level. The idea is as follows. We can tautologically decompose a count (or other) random variable into its regression (conditional mean) and residual and do non-linear regression noting that the count feature restricts the regression to be positive (and puts an inequality constraint on the residual). We maintain a Poisson-type framework while noting that this can either be modified, keeping the regression component of the Poisson and/or modified with results interpreted using robust inferential techniques. To allow the trend to be jurisdiction specific we let $\tau_{j,t}$ denote periods following the first infection in jurisdiction j at time t. In our case then,

$$\begin{aligned}\mathbb{E}[y_{j,t}|Y_{j,t-m},\tau_{j,t}] &\equiv \lambda_{jt}(Y_{j,t-m},\tau_{j,t}) \\ &= h_1(Y_{j,t-m})h_2(\tau_{j,t}),\end{aligned} \tag{4.14}$$

say, so that taking the expected value of $y_{j,t}$ conditional on $\tau_{j,t}$ we have

$$\mathbb{E}[y_{j,t}|\tau_{j,t}] = \mathbb{E}[h_1(Y_{j,t-m})|\tau_{j,t}]h_2(\tau_{j,t}) \tag{4.15}$$

and rearranging we have

$$h_2(\tau_{j,t}) = \frac{\mathbb{E}[y_{j,t}|\tau_{j,t}]}{\mathbb{E}[h_1(Y_{j,t-m})|\tau_{j,t}]} \tag{4.16}$$

and under the specification for the autoregressive component of the renewal equation we have

$$\begin{aligned}\lambda_{jt} \equiv \mathbb{E}[y_{j,t}|Y_{j,t-m},\tau_{j,t}] &= \frac{h_1(Y_{j,t-m})}{\mathbb{E}[h_1(Y_{j,t-m})|\tau_{j,t}]}\mathbb{E}[y_{j,t}|\tau_{j,t}] \\ &= \Big(\frac{Y'_{j,t-m}r}{\mathbb{E}[Y'_{j,t-m}|\tau_{j,t}]r}\Big)\mathbb{E}[y_{j,t}|\tau_{j,t}] \\ &\equiv H_{jt}(r)\mathbb{E}[y_{j,t}|\tau_{j,t}],\end{aligned} \tag{4.17}$$

say, again with $r = \begin{pmatrix} r_1 & \dots & r_m \end{pmatrix}'$. A number of points arise here. First, with respect to identification, note that multiplying each r_k by a common constant leaves the conditional mean unaltered so a restriction is required. [15] We provide a more fulsome discussion of identification below.

[14] Linton (2020) uses local polynomials to estimate $\mathbb{E}[y_{j,t}|t]$.

[15] In our initial applications we normalized $\ln r_1 = \rho_1 = 1$. Subsequently we used $\sum_s r_s = 1$ to be consistent with other work in the area. A scale restriction is also required for the aforementioned fixed-effects models.

The multiplicative decomposition of $\mathbb{E}[y_{j,t}|Y_{j,t-m}, \tau_{j,t}] = \lambda_{j,t}$ has an interesting interpretation. $\mathbb{E}[y_{j,t}|\tau_{j,t}]$ is the pure trend component of the pandemic. Note that if this understates the pandemic relative to its actual spread then the coefficient $H_{jt}(r)$ will be greater than one and this will adjust the pure trend component of the pandemic upwards. Conversely if $\mathbb{E}[y_{j,t}|\tau_{j,t}]$ overstates the pandemic, $\mathbb{E}[y_{j,t}|Y_{j,t-m}, \tau_{j,t}]$ will be smaller.

4.3.1 Infeasible Semi-Parametric Modelling

Were the conditional expectations known, maximum likelihood (or non-linear regression) could be applied directly to estimate r. In that case if we substitute λ_{jt} into the single observation log-likelihood we have

$$l_{jt}(r) = -H_{jt}(r)\mathbb{E}[y_{j,t}|\tau_{j,t}] + y_{j,t} \ln H_{jt}(r) + \text{const} \tag{4.18}$$

and summing over all time periods we have

$$l_j(r) = -\sum_{t=1}^{T} H_{jt}(r)\mathbb{E}[y_{j,t}|\tau_{j,t}] + \sum_{t=1}^{T} y_{j,t} \ln H_{jt}(r) + \text{const.} \tag{4.19}$$

The maximizer, $\hat{r}^*$, of the sum of these $l_j(r)$'s we shall denote as the infeasible semi-parametric estimator. This log-likelihood is analogous to the time-varying fixed-effects case, but with an important difference. The middle term corresponds to the log-likelihood for the time-varying fixed-effects case. However, the equivalent to the lead term does not exist in the time-varying fixed-effects case (either with one or $L > 1$ effects) as we have written it. Actually, there is an equivalent term, but, in the derivation of the fixed effects estimator, the equivalent of the numerator in the first term effectively cancels with the denominator and the lead term is absorbed into the constant.

Here, differentiating and rearranging we see that the score function is of the form

$$s_j(r) \equiv \frac{\partial}{\partial r} l_j(r) = \sum_{t=1}^{T} \mathbb{E}[y_{j,t}|\tau_{j,t}] \bigtriangledown H_{jt}(r) \left(\frac{y_{jt}}{\lambda_{jt}} - 1 \right) \tag{4.20}$$

which has the familiar structure of a Poisson score function with zero mean.

4.3.2 Feasible Semi-parametric Estimation

In the absence of the conditional mean functions $\mathbb{E}[y_{j,t}|\tau_{j,t}]$ and $\mathbb{E}[Y_{j,t-1}|\tau_{j,t}]$ we can first estimate the conditional means non-parametrically and then substitute these into the likelihood function and then maximize the likelihood over the r's to obtain a feasible semi-parametric estimator, $\hat{r}$. There are numerous methods for estimating the non-parametric component, including splines, nearest neighbours, kernels or local polynomial which, under regularity conditions, result in estimators of the finite-dimensional component which are asymptotically equivalent.

Note that the time-varying fixed effects estimator is a discrete approximation to a semi-parametric estimator with the time-varying fixed effects estimates viewed as a regresso-gram or nearest neighbours estimator with a fixed number of neighbours estimated around the

mid-point of each interval. We also note that the time-varying fixed effects is analogous to Meyer's (1990) step-wise approximation to an unknown hazard rate in duration analysis. For the semi-parametric estimator we assume that the preliminary first-step non-parametric has no first-order distribution effects so that $\sqrt{NT}(\hat{r} - \hat{r}^*) = o_p(1)$. Similar adaptive results are known to hold for a wide range of models including Robinson (1988) and numerous others including Carroll (1982) and Robinson (1987) and Escanciano, Jacho-Chávez and Lewbel (2014). An often used, but not always necessary, condition for the asymptotic equivalence result is that the non-parametric first-stage estimator converge faster than the quartic root of the sample size. Since here we are using the T_j observations on jurisdiction j to obtain non-parametric estimates on $\mathbb{E}[y_{j,t}|\tau_{j,t}]$, this amounts to doing large T asymptotics or at least that T is growing at a rate faster than N. An alternative would be to do some pooling of the data across like-jurisdictions.

Feasible estimation requires non-parametric estimation of the conditional mean functions in the likelihood. We used a variety of kernel based (Nadaraya-Watson and local polynomial) estimators for a number of reasons. One is that they are simple and the ones we use have explicit representations making them straightforward to incorporate into the estimation scheme. We focused on the usual kernel estimator and first and second-order local polynomials. (The latter is used by Linton, 2020.) We used the formulation in Wand and Jones (1995) for the fixed-design case so that the trend for jurisdiction j is transformed to $\tau_{j,t}/T_j$ so that these are uniform on $U(0,1)$, eliminates a source of bias in the conditional mean estimates and simplifies the form of the estimates' approximate standard errors.

Non-parametric estimates are particularly useful as they provide preliminary unconditional estimates of the pandemic process based solely on the trend. One technical difficulty that initially concerned us with the polynomial is that it does not constrain estimates to be non-negative. In this regard we considered various methods to impose non-negativity[16]. However, after conducting some limited simulations with processes constructed similar to the observed pandemic (log-quadratic), we simply redefine the local polynomial to be the max of $\underline{\delta}$ where $\underline{\delta}$ is a small number such as 10^{-2} and $\hat{a}$ where $\hat{a}$ is the usual least squares estimate of the intercept in a local polynomial regression. We also used the standard kernel estimator for transparency and also because it leads to an intuitive representation of the coefficient in the semi-parametric regression function. (The kernel non-parametric estimator is biased although the bias may be small in the current context in that the distribution of the conditioning variable is known.)

4.4 Estimation of $\mathcal{R}_0$

Riou and Althaus (2020) report a point estimate for $\mathcal{R}_0$ of 2.2. Eichenbaum, Rebelo, and Trabandt (2020) use values of 1.50 and 1.45 in their analytical macro models based on an SIR model. Biggerstaff et al. (2014) report $\mathcal{R}_0$ estimates for a wide variety of other epidemics. The count framework here allows for straightforward estimation of $\mathcal{R}_0$. We first consider the "case reproduction number". From the definition we have

[16]Non-negativity is logically required for a well-defined likelihood in our case.

$$\mathcal{R}_j(t) = \sum_{s=1}^{m} r_s A_j(t+s) \tag{4.21}$$

and we need to simply substitute estimates of r_s and $A_j(t+s)$. Note again that that in this count framework the reproduction number is simply the weighted sum of future exposure weights where the weights are measures of future toxicity of an infected individual. Note that for estimating $\mathcal{R}_j(T-s)$ for $s \leq m$, this requires estimating outside the observed data. In the fixed-effects case $A_j(t) = e^{\alpha_j}$ is constant and we have $\mathcal{R}_j(t) = e^{\alpha_j} \sum_{s=1}^{m} r_s$. In the time-varying fixed-effects case $A_j(t)$ is constant over intervals and we have

$$\mathcal{R}_j(t) = \sum_{s=1}^{m} r_s e^{\alpha_{jl}} 1[t+s \in \mathcal{I}_l] \tag{4.22}$$

Note that although the exposure factors are constant across intervals, the $\mathcal{R}_j(t)$'s in Equation 4.22 generally will vary as the infectivity factors straddle different intervals. A popular device used in the estimation of reproduction numbers as in Cori et al. (2013) is to consider variations in the length of windows of observations used to estimate $\mathcal{R}_0$. The analogous tuning parameter here is the length of the intervals used in the time-varying effects approach.

In the semi-parametric model, $A_j(t)$ is the ratio of two conditional expectations and we have

$$\mathcal{R}_j(t) = \sum_{s=1}^{m} r_s \frac{\mathbb{E}[y_{j,t+s}|\tau_{j,t+s}]}{\mathbb{E}[Y'_{j,t+s-m}|\tau_{j,t+s}]r} = \sum_{s=1}^{m} r_s \frac{\mathbb{E}[y_{j,t+s}|\tau_{j,t}]}{\mathbb{E}[Y'_{j,t+s-m}|\tau_{j,t}]r} \tag{4.23}$$

using $\mathbb{E}[y_{j,t+s}|\tau_{j,t+s}] = \mathbb{E}[y_{j,t+s}|\tau_{j,t}]$ In each of these cases note that we need to be prudent when constructing standard errors and confidence intervals. If we assume that the estimators of the homogeneous parametric component (the r_s's) converges faster than the respective estimates of the $A_j(t)$'s, then we may treat the latter as effectively fixed and consider the distribution of the $A_j(t)$'s. As is often the case, this can be a little problematic. Since T is relatively large, each estimate of $c_j = e^{\alpha_j}$ is the ratio of two averages based on a large number of observations. Standard errors based on this with approximations based on $\sqrt{T}$ asymptotics should be fairly accurate. A similar, though somewhat weaker argument can be made for the time-varying fixed effects. With regard to the pure semi-parametric estimator, the asymptotic variance can be derived using basic results for estimators such as kernel or local-polynomial based. In this situation it is well known that the convergence rate will be $\sqrt{T\gamma}$ where here $\gamma \downarrow$ is the window width. Alternatively, in each of these cases, some kind of bootstrapping could be done. In our case, our empirical discussion below provides point estimates at various stages into the pandemic accompanied by the standard deviations and quantities across jurisdictions. The latter should provide a fairly accurate measure of the variability of the point individual estimates.

In this context it is interesting to note that $\mathcal{R}_j(t)$'s are often referred to as reproduction ratios, highlighting the fact that these are often represented as a ratio whose numerator is the product of daily contacts and probability of transmission and the denominator is the product of the rate of exit from the susceptible population and the probability of contraction of the disease. In our case, note that with the basic fixed effects mode, the estimator of $\mathcal{R}_j$

is simply $\sum \hat{r}_s c_j = (\sum \hat{r}_s \bar{y}_j)/\bar{h}_j$ which can be interpreted as an estimator of the ratio of new cases to the exit rate due to past infections. The same interpretation applies (albeit a little less cleanly) for the case of the time-varying fixed coefficients and semi-parametric estimators of the reproductive numbers.

For estimation of the "instantaneous reproduction number", denoted here by $\bar{\mathcal{R}}_j(t)$ and to obtain this we simply set $A_j(t+s) = A_j(t)$ in Equation 4.23 so that $\bar{\mathcal{R}}_j(t) = A_j(t) \sum_{s=1}^m r_s$. In the instantaneous case per Fraser (2007) and Cosi et al. (2013) the r_s's are normalized to sum to unity so that term does not appear. Note that in the form written here that $\bar{\mathcal{R}}_j(t)$ is invariant with respect to scale transformation of the r_s as, for each of the estimation techniques, the r_s's appear implicitly in numerator and denominator.

Some programming remarks may be useful to practitioners. The estimation of all models is straightforward using any programmable language. We did so in two or three steps. Computation of any of the parametric models is standard. For estimating the r_s's with the semi-parametric models the first step consisted of obtaining non-parametric estimates of $\mathbb{E}[y_{j,t-s}|\tau_{j,t}]$, $s = 0, 1, \cdots, m$. (I.e. current and lagged predicted values.) In this step we also found it expedient to estimate $\mathbb{E}[y_{j,t+s}|\tau_{j,t+s}]$, $s = 0, 1, \cdots, m$. These are then inputted directly into standard optimization software for, say maximum likelihood or Poisson model estimation. This estimates the r_s's. The third step computes the exposure rates $A_j(t)$'s using the parametric or non-parametric estimates. Weighting these by the estimated r_s's produces the $\mathcal{R}_j(t)$ estimates.

A final point can be made with respect to the connection between $\mathcal{R}_0(t)$ and the statistical concept of stationarity. In a purely AR(m) process with $\mathbb{E}[y_t|Y_{t-m}] = \sum_{s=1}^m \phi_s y_{t-s}$ the usual condition for stationarity is that the roots of the equation $1 = \phi_1 z + \cdots + \phi_m z^m$ lie outside the unit circle. Sufficient conditions for these are if $\phi_s \geq 0$, $s = 1, \ldots, m$, and $\sum_{s=1}^m \phi_s < 1$. In the present context, assume that $\mathcal{R}_j(t) = \mathcal{R}_0$ is constant (or at some "steady-state" value). In our case $\mathbb{E}[y_t|Y_{t-m}, t] = \sum_{s=1}^m r_s y_{t-s} \mathcal{R}_0 = \sum_{s=1}^m \phi_s y_{t-s} \mathcal{R}_0$ with $\phi_s = r_s \mathcal{R}_0$. Since $r_s \geq 0$ and $\sum_{s=1}^m r_s = 1$, $\sum_{s=1}^m \phi_s < 1$, if $\mathcal{R}_0 < 1$. So the usual characterization of $\mathcal{R}_0(t) \geq 1$ as corresponding to an explosive epidemic can also be characterized in this way as a non-stationary process.

4.5 Identification

For each jurisdiction, (we suppress j subscripts in this subsection) the conditional means $\mathbb{E}[y_t|Y_{t-m}, t]$ and $\mathbb{E}[y_t|t]$ are identified. However, due to the multiplicative separability of the regression function, $\mathbb{E}[y_t|Y_{t-m}, t] = Y'_{t-m} r A_t$, there is an identification issue. Let r^0 and A_t^0 denote true values so that $\mathbb{E}[y_t|Y_{t-m}, t] = Y'_{t-m} r^0 A_t^0$. For any $\kappa > 0$ we also have $\mathbb{E}[y_t|Y_{t-m}, t] = Y'_{t-m} r^1 A_t^1$ where $A_t^1 = \kappa A_t^0$ and $r^1 = r^0/\kappa$ so that r^0 is not identified. We will establish conditions (scale and a form of exclusion restrictions) sufficient for identification for the three basic models considered.

Decompose $r = (r_1 \, r'_{(1)})'$ where r_1 is the first element of r and $r_{(1)}$ denotes the remaining $m - 1$ elements of r.

In each of the cases with true values of the individual parameters denoted, say, as r_q^0, $q = 1, \ldots, m$, we consider other possible values which as $r_q^1 = \kappa_q r_q^0$. We show that the κ_q's need to equal 1 to be compatible with the true values of $\mathbb{E}[y_t|t]$ and $\mathbb{E}[y_t|Y_{t-1}, t]$

For all models we impose a scale restriction. Here, for convenience, we impose $r_1 = 1$[17].

Also, for this subsection, put $Y_{t-2} = \begin{pmatrix} y_{t-2} & \dots & y_{t-m} \end{pmatrix}'$. Let Ω denote those outcomes in the population such that $y_{t-1} \neq 0$ and $Y_{t-2} = 0$. Let $Y_{t-2,m\backslash q}$ denote Y_{t-2}, excluding y_{t-q}. Let Ω_q, $q = 2, \dots, m$ denote those outcomes in the population such that $y_{t-q} \neq 0$ and $Y_{t-1,m\backslash q} = 0$. Let Ω_q^+ denote those outcomes in the population such that $y_{t-q} \neq 0$ and $Y_{t-1,m\backslash q} = 0$ and $y_{t-1} = 0$. We assume that $\Pr[\Omega] > 0$ and for $q = 2, \dots, m$, $\Pr[\Omega_q] > 0$ for the parametric models. We assume that $\Pr[\Omega] > 0$ and for $q = 2, \dots, m$, $\Pr[\Omega_q^+] > 0$ for the semi-parametric models. Since the count variables we consider have positive mass at zero this is not an issue.

4.5.1 Identification of the fixed coefficients model

It is convenient to use the parameterization $c = e^\alpha$. Note that with the restriction on r_1 we can identify $\mathbb{E}[y_t|Y_{t-m}, t] = (y_{t-1} + Y_{t-2}' r_{(1)}^0)c^0$. Suppose $(r_{(1)}^1, c^1)$ are other possible values of $(r_{(1)}, c)$ with $c^1 = \kappa c^0$. (a) On Ω we have $\mathbb{E}[y_t|Y_{t-1}, t] = y_{t-1}c^0 = y_{t-1}\kappa c^0$ so $\kappa = 1$ and c_0 is identified. (b) On each Ω_q, we have $\mathbb{E}[y_t|Y_{t-1}, t] = (y_{t-1} + r_q^0 y_{t-q})c^0 = (y_{t-1} + y_{t-q}\kappa_q r_q^0)c^0$ so each $\kappa_q = 1$ and the other r_q^0's are identified.

4.5.2 Identification of the time-varying fixed coefficients model

Identification of the time-varying fixed effects model works similarly with the identification being established on the first (or potentially other) interval.

4.5.3 Identification of the semi-parametric model

In this subsection, put $\hat{Y}_{t-1} = \mathbb{E}[Y_{t-1}|t]$ and $\hat{y}_t = \mathbb{E}[y_t|t]$. In the semi-parametric case we have $\mathbb{E}[y_t|Y_{t-1}, t] = (Y_{t-1}'r/\hat{Y}_{t-1}'r)\hat{y}_t$ and with the restriction $r_1 = 1$ we have

$$\frac{\mathbb{E}[y_t|Y_{t-1}, t]}{\hat{y}_t} = \frac{(y_{t-1} + Y_{t-2}' r_{(1)}^0)}{(\hat{y}_{t-1} + \hat{Y}_{t-2}' r_{(1)}^0)} = \frac{(y_{t-1} + Y_{t-2}' r_{(1)}^1)}{(\hat{y}_{t-1} + \hat{Y}_{t-1-m}' r_{(1)}^1)}. \tag{4.24}$$

For any $q = 2, \dots, m$, we have on Ω_q^+

$$\frac{y_{t-q} r_q^0}{\hat{Y}_{t-1}' r^0} = \frac{y_{t-q} r_q^1}{\hat{Y}_{t-1}' r^1} = \frac{y_{t-q}\kappa_q r_q^0}{\hat{Y}_{t-1}' r^1}. \tag{4.25}$$

so

$$\frac{1}{\hat{Y}_{t-1}' r^0} = \frac{\kappa_q}{\hat{Y}_{t-1}' r^1}. \tag{4.26}$$

which implies that $\kappa_q = k_p = k^*$, say, for all $p, q \geq 2$ so that

$$\frac{1}{\hat{Y}_{t-1}' r^0} = \frac{\kappa^*}{(\hat{y}_{t-1} + \hat{Y}_{t-2}' \kappa^* r_{(1)}^0)}. \tag{4.27}$$

[17] As mentioned, in the empirical section we set $\sum r_s = 1$ by parameterizing $r_s = e^{\rho_s}/(1 + \sum_{l=2}^m r_l)$, $s = 2, \dots, m$ and setting $r_1 = 1/(1 + \sum_{l=2}^m r_l)$. Other restrictions will work as well. In this case the conditional mean $\mathbb{E}[y_t|Y_{t-m}, t] = Y_{t-m}'\phi B_t$ is equivalent with $\phi_s = \rho_s/(1 + \sum_{s=2}^m r_s)$ and $B(t) = (1 + \sum_{s=2}^m r_s)A(t)$

or

$$(\hat{y}_{t-1} + \hat{Y}'_{t-2}\kappa^* r^0_{(1)}) = \kappa^*(\hat{y}_{t-1} + \hat{Y}'_{t-2} r^0_{(1)}). \tag{4.28}$$

and hence $k^* = \frac{\hat{Y}'_{t-2} r^0_{(1)} - \hat{y}_{t-1}}{\hat{Y}'_{t-2} r^0_{(1)} - \hat{y}_{t-1}} = 1$.

4.6 Asymptotics: Autoregressive Parameter Estimates

We state some high-level assumptions for the models at hand which allow us to expeditiously establish asymptotic results. Throughout we assume the log-likelihood functions are twice continuously differentiable. Throughout, we let $\hat{r}$ denote one of the four generic estimators we consider: fixed effects, time-varying fixed-effects, infeasible semi-parametric and feasible semi-parametric estimators. Where necessary we denote these respectively as $\hat{r}_{FE}$, $\hat{r}_{TFE}$, $\hat{r}_{FSP}$ and $\hat{r}_{SP}$.

For each of the fixed effects, time-varying fixed-effects and infeasible semi-parametric estimators, continue to denote by s_j their respective score functions and in each case denote

$$v_{jt}(r) \equiv \frac{\partial}{\partial r'} s_{jt}(r), \qquad v_j(r) = \sum_{t=1}^{T} v_{jt}(r) \tag{4.29}$$

$$L_{NT}(r) = \frac{1}{NT}\sum_{j=1}^{N} l_j(r), \qquad S_{NT}(r) = \frac{1}{NT}\sum_{j=1}^{N} s_j(r), \qquad V_{NT}(r) = \frac{1}{NT}\sum_{j=1}^{N} v_j(r) \tag{4.30}$$

whose exact forms will be model-specific.

Proposition 1: (Fixed effects, time-varying fixed-effects and infeasible semi-parametric estimators) Assume: 1, the true value of r, r^0, lies in the interior of a compact subset of $\mathbb{R}^{m-1}$; 2, $L_{NT}(r)$ converges uniformly in probability to a function $L_0(r)$ which has a unique maximum at r^0; 3, $V_{NT}(r)$ converges uniformly in probability to a matrix $\Phi(r)$ where $\Phi(r^0)$ is strictly positive definite. 4, $\sqrt{NT}S_{NT}(r^0) \xrightarrow{d} N(0, \Lambda)$. Then for each of the fixed effects, time-varying fixed-effects and infeasible semi-parametric estimators $\hat{r}$ is consistent (given that the model is correct) and $\sqrt{NT}(\hat{r} - r^0) \xrightarrow{d} N(0, \Phi^{-1}\Lambda\Phi^{-1})$.

Proof: Follows from standard asymptotic theory as in Newey and MacFadden (1994).

We note that the forms of Λ and Φ will depend on which model is being estimated.

For the feasible semi-parametric estimator modify the notation for the infeasible likelihood functionals with parametric components replaced by their corresponding non-parametric estimates as $\hat{l}_{jt}$, $\hat{s}_{jt}$, $\hat{v}_{jt}$, $\hat{L}_{NT}$, $\hat{S}_{NT}$ and $\hat{V}_{NT}$ with the first three potentially scaled by a trimming indicator to reduce the contribution of observations which may

Proposition 2: (Feasible semi-parametric estimators) Let the assumptions of Proposition 1 hold and assume $(L_{NT}(r)) - \hat{L}_{NT}(r)) = o_p(1)$, $(V_{NT}(r)) - \hat{V}_{NT}(r)) = o_p(1)$ uniformly in r and $\sqrt{NT}(S_{NT}(r^0)) - \hat{S}_{NT}(r^0)) = o_p(1)$. Then, $\hat{r}_{SP}$ is consistent and $\sqrt{NT}(\hat{r}_{FSP} - \hat{r}_{SP}) = o_P(1)$.

Proof: Consistency follows immediately from $(L_{NT}(r)) - \hat{L}_{NT}(r)) = o_p(1)$ and Proposition 1. Asymptotic first-order equivalence follow from a standard mean-value expansion of the estimator so that with probability approaching one.

$$0 = \sqrt{NT}\hat{S}_{NT}(\hat{r}_{SP}) = \sqrt{NT}(\hat{S}_{NT}((r^0))) + \hat{V}_{NT}(\bar{r})(\hat{r}_{SP} - r^0) \tag{4.31}$$
$$= \sqrt{NT}S_{NT}(r^0) + V_{NT}(\bar{r})(\hat{r}_{SP} - r^0) + o_P(1)$$

where $\bar{r} = r^0 + o_P(1)$ is a mean value so that

$$\sqrt{NT}(\hat{r}_{SP} - r^0) = -(V_{NT}(r^0))^{-1}\sqrt{NT}S_{NT}(r^0) + o_P(1) \tag{4.32}$$
$$= \sqrt{NT}(\hat{r}_{NSP} - r^0) + o_P(1).$$

Remark 3: Consistency of the semi-parametric estimator is straightforward given that the underlying pointwise non-parametric regression estimators are uniformly consistent for the true population regression functions so that the feasible log-likelihood converge to the same expectation as the infeasible log-likelihood. (The same applies to $\hat{V}$.) The asymptotic results (apart from considerations we have enumerated throughout the paper) for the parametric and infeasible estimators are unremarkable. To rigorously obtain the results of Proposition 2 from first principles requires considerable work, but, were the data i.i.d., the score for the semi-parametric estimator is of a form which fits into the analytical framework of Escanciano, Jacho-Chávez and Lewbel (2014) who used empirical process theory to prove results such as in Proposition 2. In that paper the authors demonstrate convergence of (scaled) averages such as our $\sqrt{NT}(S_{NT}(r^0)) - \hat{S}_{NT}(r^0))$. In their case the summands are averages of multiples of semi-parametric residuals. In our case the residual is effectively $y_{j,t} - \lambda_{jt}$. In our case the data is dependent and heterogeneous. A rigorous demonstration of Proposition 2 from primitive conditions in that case is far beyond the applied scope of this paper, although the empirical process results in Andrews (1994), Hansen (2009) and Hagemann (2014) for dependent and/or non-stationary observations suggest the high-level assumptions of Proposition 2 are quite plausible.

Remark 4: We note that each of the estimators of the auto-regressive parameters is $\sqrt{NT}$-consistent. It may be possible to analytically compare the asymptotic variance matrices. Conceptually, the estimators can be viewed as progressively less restrictive, allowing for increasing heterogeneity over jurisdictions and over time. We do not as yet have analytic results in this regard.

4.7 Asymptotics: Estimates of the time trend, $A_j(t)$ and $\mathcal{R}_j(t)$

In this subsection we approximate the distribution of the estimators the $A_j(t)$'s and $\mathcal{R}_j(t)$'s. To simplify notation, in this subsection we suppress the subscript j's and set $T_j = T$. This can be simply modified in applications. With both N and T large, we can treat the r's as fixed. For the case of the purely parametric models, the estimators of the $A(t)$'s and the $\mathcal{R}(t)$ are ratios of averages and linear combinations of these, respectively, and we approximate standard errors accordingly.

With respect to the semi-parametric estimation of the AR parameters, it was not necessary to specify the particular form of the non-parametric component so long as some generic conditions were satisfied. Various non-parametric estimators will lead to the same asymptotic results for the parametric component. However, to derive explicit results for the point-wise estimates of $A(t)$'s and $\mathcal{R}(t)$ we consider local polynomial estimators.

4.7.1 Estimation of the time trend

We adapt the following from Wand and Jones (1995) and Li and Racine (2007). To facilitate the analysis we redefine the trend variable as $X_t = t/T$ which does not change conditional expectations but which will allow us to better quantify the relationship between estimates at adjacent points. Note that $X_{t\pm j} = X_t \pm j/T$. Note that doing so allows us to approximate non-stochastic averages as $\frac{1}{T}\sum_{t=1}^{T} f(X_t) = \int f(y)dy + O(T^{-1})$ and $\gamma^{-1}\frac{1}{T}\sum_{t=1}^{T} g(X_t)f((X_t - x)/\gamma) = \int g(x+\gamma w)f(w)dw + O(T^{-1})$. For succinctness we write $g(X_t) = \mathbb{E}[Y_t|\tau_t]$ and decompose infections into the time trend and residual:

$$y_t = g(X_t) + u_t \tag{4.33}$$

where we assume

$$\mathbb{E}[u_t|X_t] = 0, \qquad \mathbb{E}[u_t u_s|X_t] = \begin{cases} v(X_t), & t = s \\ 0, & t \neq s \end{cases} \tag{4.34}$$

The assumption of uncorrelated residuals may be restrictive, but enables us to obtain tractable results. To define and manipulate the local polynomial estimators put

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix}, \quad X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_T \end{pmatrix}, \qquad g(X) = \begin{pmatrix} g(X_1) \\ g(X_2) \\ \vdots \\ g(X_T) \end{pmatrix} \tag{4.35}$$

$$X_x = X - x, \qquad K_x = \text{Diag}[K(X_x/\gamma)], \qquad V(X) = \text{Diag}[\begin{pmatrix} v(X_1) & \cdots & v(X_T) \end{pmatrix}] \tag{4.36}$$

$$Z_x = \begin{pmatrix} \iota_T & X_x^{\odot 1} & \cdots & X_x^{\odot p} \end{pmatrix}_{T\times p_1}, \qquad z(w) = \begin{pmatrix} 1 & w & \cdots & w^p \end{pmatrix}^\top \tag{4.37}$$

with $p_1 = p + 1$ where p is the order of the polynomial, $\odot$ denotes the Hadamard product and $X_x^{\odot p}$ denotes the Hadamard product of X_x with itself, p times.

The regression and regression derivative estimators are obtained from a local regression with

$$\hat{g}(x) = e_p^\top b(x), \qquad \hat{g}^{(1)}(x) = e_p^\top b^{(1)}(x), \qquad b(x) = (Z_x^\top K_x Z_x)^{-1} Z_x^\top K_x y \tag{4.38}$$

and $e_p = (1, 0, \cdots, 0)_{1\times p_1}$. Note that $\hat{g}^{(1)}(x)$ denotes the derivative of $\hat{g}(x)$ and not necessarily the second element of $b(x)$[18]. We do not necessarily need to calculate $\hat{g}^{(1)}(x)$, but its definition and properties facilitate the derivation of the properties of $\hat{A}(t)$ and $\hat{\mathcal{R}}(t)$.

Basic properties of regression and regression derivative estimators are well-established. We state a proposition in this regard and provide some derivations at the end of this section which may not be immediately obvious.

The following moment matrices appear in the asymptotic variances of $\hat{g}(x)$ and $\hat{g}^{(1)}(x)$.

[18]With local polynomial estimation the second element of $b(x)$ is an estimate of $g^{(1)}(x)$. This does not exist for $p = 0$.

$$\hat{M}_p = \frac{1}{T\gamma} Z_x^\top K_x Z_x, \qquad M_p = \int z(\gamma w) z(\gamma w)^\top K(w) dw \tag{4.39}$$
$$\hat{M}_p^* = \frac{1}{T\gamma} Z_x^\top K_x^{\odot 2} Z_x, \qquad M_p^* = \int z(\gamma w) z(\gamma w)^\top K(w)^2 dw$$
$$\hat{M}_p^{**} = \frac{1}{T\gamma} Z_x^\top K_x^{(1)\odot 2} Z_x, \qquad M_p^{**} = \int z(\gamma w) z(\gamma w)^\top K^{(1)}(w)^2 dw$$

where $\hat{M}_p = M_p + O(1/T)$, $\hat{M}_p^* = M_p^* + O(1/T)$ and $\hat{M}_p^{**} = M_p^{**} + O(1/T)$. Put

$$\kappa_p = e_p^\top M_p^{-1} M_p^* M_p^{-1} e_p, \qquad \kappa_p' = e_p^\top M_p^{-1} M_p^{**} M_p^{-1} e_p$$

where in the cases of interest neither κ_p nor κ_p' depend on T.

Proposition 3: For some neighbourhood of x assume $g(x)$ is $(p+2)$-times continuously-differentiable. Let $K : [-1, 1] \to [0, 1]$ be a twice continuously-differentiable symmetric kernel. $\gamma \to 0$, $T\gamma \to \infty$. Then

$$\sqrt{T\gamma}(\hat{g}(x) - g(x)) \xrightarrow{d} N(0, v(x)\kappa_p) \tag{4.40}$$
$$\sqrt{T\gamma^3}(\hat{g}^{(1)}(x) - g^{(1)}(x)) \xrightarrow{d} N(0, v(x)\kappa_p').$$

The first statement is a standard result. See below for a proof of the second statement. We implicitly assume that $\hat{g}(x)$ may be bias adjusted. There are a variety of methods to reduce or remove bias such as under-smoothing, higher-order kernels, local polynomials, analytic corrections and resampling techniques. The result for $\hat{g}(x)$ can be used to obtain approximate confidence intervals for the time trend. Note also for below that $\sqrt{T\gamma^3}(\hat{g}(x) - g(x)) = o_P(1)$. Note that $v(x)$ can be estimated using a local regression of the squared residuals, $\hat{u}_t^2 = (y_t - \hat{g}_t)^2$, on the time trend.

4.7.2 Estimates of $A_j(t)$ and $\mathcal{R}_j(t)$

To derive the asymptotic distribution of $\hat{A}(t)$ and $\hat{\mathcal{R}}(t)$ it would seem straightforward to write these as functions of estimates of the time trend at various leads and lags and apply the delta method. As noted, there is a problem in doing so. Although kernel estimates at *distinct* points are asymptotically independent, the points where we are estimating are of the form x and $x_j = x + \frac{j}{T}$, with x_j converging quickly to x. This needs to be taken into account. Referring back to Proposition 3, it is straightforward to show that the asymptotic variance, of say, $\hat{g}(x_j)$ is $v(x)\kappa_p$ and the asymptotic covariance of $\hat{g}(x_j)$ and $\hat{g}(x)$ is also $v(x)\kappa_p$ hence the (usual) covariance matrix of $\hat{g}(x)$ and $\hat{g}(x_j)$ is singular (proportional to a matrix of ones). *Ergo*, the usual delta method is not applicable to derive the asymptotic distribution of non-parametric estimates of $A(t)$ and $\mathcal{R}(t)$.

We fix a point $x = x_0$ and, with some abuse of notation, respecify arguments accordingly so that we have $A(x)$ and $\mathcal{R}(x)$, $\hat{A}(x)$ and $\hat{\mathcal{R}}(x)$. Putting $g_j = g(x_j)$ and $\hat{g}_j = \hat{g}(x_j)$ we have

$$A(x_0) = \frac{g_0}{\sum_{j=1}^m r_j g_j}, \quad \mathcal{R}(x_0) = \sum_{l=1}^m r_l A(x_l), \quad \hat{A}(x_0) = \frac{\hat{g}_0}{\sum_{j=1}^m r_j \hat{g}_j}, \quad \hat{\mathcal{R}}(x_0) = \sum_{l=1}^m r_l \hat{A}(x_l) \tag{4.41}$$

Note that

$$g_j = g_0 + \frac{j}{T} g_0^{(1)} + O(1/T^2), \qquad \hat{g}_j = \hat{g}_0 + \frac{j}{T} \hat{g}_0^{(1)} + O_p(1/T^2) \tag{4.42}$$

Examining the latter we note that $\hat{A}(x_0) = 1 + O_p(1/T)$ which leads to a degeneracy if we use standard first order asymptotics to approximate the distribution of $\hat{A}(x_0)$. To circumvent this problem we use a second-order approximation to obtain approximate results. Put

$$\Sigma(x) = \frac{v(x)}{g(x)^2} \kappa_p' \Big(\sum_{j=1}^m r_j j \Big)^2 \tag{4.43}$$

Proposition 4: Let the assumptions of Proposition 3 hold with $g(x_0) > 0$. Then

$$\sqrt{(T\gamma)^3} \Sigma(x_0)^{-1/2} (\hat{A}(x_0) - A(x_0)) \xrightarrow{d} N(0,1) \tag{4.44}$$

$$\sqrt{(T\gamma)^3} \Sigma(x_0)^{-1/2} (\hat{\mathcal{R}}(x_0) - \mathcal{R}(x_0)) \xrightarrow{d} N(0,1) \tag{4.45}$$

This is proven below. The result is obtained by using the expansions in Equation 4.42 to approximate the denominators of $A(x_0)$ and $\hat{A}(x_0)$ which explains why we use the results on the derivative of $\hat{g}(x)$ and why κ_p' appears in the asymptotic variance. Note that $\hat{\mathcal{R}}(x_0)$ has the same asymptotic distribution as $\hat{A}(x_0)$, but this does not imply they are the same. They are each centred at different points and the equality of their asymptotic distributions follows from the fact that the r_l's in their definition sum to one and the points at which their respective summands are calculated are converging to x_0. Note that the instantaneous reproduction number, $\mathcal{R}(x_0)$, with $\sum r_l = 1$, corresponds to $A(x_0)$.

4.7.3 Special cases with quadratic kernel

We provide some details here for the applications in the paper. Note that the various asymptotic variances above are functions of κ_p and κ_p' which are in turn functions of M_p, M_p^* and M_p^{**}. Since these are obtained as limits of $\hat{M}_p$, $\hat{M}_p^*$ and $\hat{M}_p^{**}$, the latter can be used as estimates. Alternatively we can obtain M_p, M_p^* and M_p^{**} analytically as we do here for the application we employed.

Put $\mu_j = \int w^j K(w) dw$, $\mu_j^* = \int w^j K(w)^2 dw$ and $\mu_j^{**} = \int w^j K^{(1)}(w)^2 dw$ Note that M_p, M_p^* and M_p^{**} have individual elements $[i, j]$

$$M_p[i,j] = \int z(\gamma w) z(\gamma w)^\top K(w) dw[i,j] = \gamma^{i+j-2} \mu_{i+j-2} \tag{4.46}$$

$$M_p^*[i,j] = \int z(\gamma w) z(\gamma w)^\top K(w)^2 dw[i,j] = \gamma^{i+j-2} \mu_{i+j-2}^* \tag{4.47}$$

$$M_p^{**}[i,j] = \int z(\gamma w) z(\gamma w)^\top K(w)^{(1)2} dw[i,j] = \gamma^{i+j-2} \mu_{i+j-2}^{**} \tag{4.48}$$

Note by symmetry of the kernel that every second element of these is zero, that is $\mu_j = \mu_j^* = \mu_j{*}{*} = 0$ if j odd. We consider cases for $p \leq 2$.

$$M_0 = 1, \quad M_0^* = \mu_0^*, \quad M_0^{**} = \mu_0^{**}$$

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & \gamma^2 \mu_2 \end{pmatrix}, M_1^* = \begin{pmatrix} \mu_0^* & 0 \\ 0 & \gamma^2 \mu_2^* \end{pmatrix}, M_1^{**} = \begin{pmatrix} \mu_0^{**} & 0 \\ 0 & \gamma^2 \mu_2^{**} \end{pmatrix}$$

$$M_2 = \begin{pmatrix} 1 & 0 & \gamma^2 \mu_2 \\ 0 & \gamma^2 \mu_2 & 0 \\ \gamma^2 \mu_2 & 0 & \gamma^4 \mu_4 \end{pmatrix}, M_2^* = \begin{pmatrix} \mu_0^* & 0 & \gamma^2 \mu_2^* \\ 0 & \gamma^2 \mu_2^* & 0 \\ \gamma^2 \mu_2^* & 0 & \gamma^4 \mu_4^* \end{pmatrix}, M_2^{**} = \begin{pmatrix} \mu_0^{**} & 0 & \gamma^2 \mu_2^{**} \\ 0 & \gamma^2 \mu_2^{**} & 0 \\ \gamma^2 \mu_2^{**} & 0 & \gamma^4 \mu_4^{**} \end{pmatrix}$$

$$\kappa_j = \begin{cases} \mu_2^*, & j = 0, 1 \\ \frac{(\mu_4^2 \mu_0^* - 2\mu_2 \mu_4 \mu_2^* + \mu_2^2 \mu_4^*)}{(\mu_4 - \mu_2^2)}, & j = 2 \end{cases} \tag{4.49}$$

$$\kappa_j' = \begin{cases} \mu_2^{**}, & j = 0, 1 \\ \frac{(\mu_4^2 \mu_0^{**} - 2\mu_2 \mu_4 \mu_2^{**} + \mu_2^2 \mu_4^{**})}{(\mu_4 - \mu_2^2)}, & j = 2 \end{cases} \tag{4.50}$$

In the empirical work we used zero, first and second-order polynomials with a quadratic kernel. For the quadratic kernel: $K(w) = (3/4)(1 - w^2)1[|w| \leq 1]$, $K(w)^2 = (9/16)(1 - w^2)^2$ and $K^{(1)}(w)^2 = (9/4)w^2$. Also,

$$\begin{array}{lll} \mu_0 = 1, & \mu_2 = 1/5, & \mu_4 = 3/35 \\ \mu_0^* = 3/5, & \mu_2^* = 3/35, & \mu_4^* = 1/35 \\ \mu_0^{**} = 3/2, & \mu_2^{**} = 9/10, & \mu_4^{**} = 9/14. \end{array} \tag{4.51}$$

4.8 Derivations

Proof of Proposition 3: The basic properties of $\hat{g}(x)$ are well-established. We confirm the properties of $\hat{g}^{(1)}(x)$. To obtain the mean and variance of $\hat{g}^{(1)}$ note that $(M_p b(x))^{(1)} = M_p b^{(1)}(x)$ since $M_p^{(1)} = 0$. With $(\hat{M}_p b(x))^{(1)} = \frac{1}{T\gamma}(Z_x^\top K_x)^{(1)} y$, $\mathbb{E}[\hat{M}_p b(x))^{(1)}] = \frac{1}{T\gamma}(Z_x^\top K_x)^{(1)} g(X)$. To order $O(1/T)$,

$$
\begin{aligned}
M_p\mathbb{E}[b^{(1)}(x)] &= \frac{1}{\gamma}\int \frac{d}{dx}(Z(y-x)K(y-x))g(y)dy \qquad (4.52)\\
&= -\frac{1}{\gamma}\int \frac{d}{dy}(Z(y-x)K((y-x)/\gamma))g(y)dy\\
&= \int Z(\gamma w)K(w))g^{(1)}(x+\gamma w)dw\\
&= \int Z(\gamma w)K(w)Z(\gamma w)^\top dw\Big(\begin{array}{ccc}\big(g^{(1)}(x) & \cdots & g^{(p_1)}(x)/p_1!\big)^\top\end{array}\Big) + \gamma^{p_1}O(1)
\end{aligned}
$$

so

$$
\mathbb{E}[\hat{g}^{(1)}(x)] = \mathbb{E}[e^\top b^{(1)}(x)] = g^{(1)}(x) + \gamma^{p_1}O(1). \qquad (4.53)
$$

To order $1/T$

$$
\begin{aligned}
\mathrm{Var}[\hat{g}^{(1)}(x)] &= \mathrm{Var}\Big[e_p'\hat{M}_p^{-1}\frac{1}{T\gamma}\sum u_t(Z(X_t-x)K(X_t-x))^{(1)}\Big] \qquad (4.54)\\
&= \frac{1}{(T\gamma)^2}e_p'\hat{M}_p^{-1}\sum v(X_t)(K(X_t-x)Z(X_t-x))^{(1)}(K(X_t-x)Z(X_t-x)^\top)^{(1)}\hat{M}_p^{-1}e_p\\
&= \frac{1}{T\gamma^3}e_p'M_p^{-1}\int v(x+\gamma w)(K^{(1)}(w)^2Z(\gamma w))Z(\gamma w))'dyM_p^{-1}e_p + O((T\gamma^2)^{-1})\\
&= \frac{v(x)}{T\gamma^3}e_p'M_p^{-1}\int K^{(1)}(w)^2Z(\gamma w))Z(\gamma w))'dyM_p^{-1}e_p + o((T\gamma^3)^{-1}))\\
&\equiv \frac{v(x)}{T\gamma^3}\kappa_p' + o((T\gamma^3)^{-1})).
\end{aligned}
$$

Asymptotic normality follows since $\hat{g}^{(1)}(x)$ is proportional to a weighted average of the u_t's.
Proof of Proposition 4: We decompose

$$
\begin{aligned}
\hat{A}(x_0) - A(x_0) &= \frac{\hat{g}_0 - \sum_{j=1}^m r_j\hat{g}_{-j}}{\sum_{j=1}^m r_j\hat{g}_{-j}} - \frac{g_0 - \sum_{j=1}^m r_jg_{-j}}{\sum_{j=1}^m r_jg_{-j}} \qquad (4.55)\\
\hat{A}(x_0) - A(x_0) &= \frac{(\hat{g}_0 - \sum_{j=1}^m r_j\hat{g}_{-j}) - (g_0 - \sum_{j=1}^m r_jg_{-j})}{\sum_{j=1}^m r_jg_{-j}} + R_{1T}\\
&= \frac{(\hat{g}_0 - \sum_{j=1}^m r_j\hat{g}_{-j}) - (g_0 - \sum_{j=1}^m r_jg_{-j})}{g_0} + R_{2T} + R_{1T}\\
&= R_{3T} + R_{2T} + R_{1T}
\end{aligned}
$$

where

$$
R_{1T} = (\hat{g}_0 - \sum_{j=1}^m r_j\hat{g}_{-j})\frac{\sum_{j=1}^m r_j(g_{-j}-\hat{g}_{-j})}{(\sum_{j=1}^m r_jg_{-j})(\sum_{j=1}^m r_j\hat{g}_{-j})} = O_p\Big(\frac{1}{T}\Big)O_p\big(|\hat{g}_0 - g_0|\big) \qquad (4.56)
$$

$$
R_{2T} = \Big((\hat{g}_0 - \sum_{j=1}^m r_j\hat{g}_{-j}) - (g_0 - \sum_{j=1}^m r_jg_{-j})\Big)\Big(\frac{g_0 - \sum_{j=1}^m r_jg_{-j}}{g_0\sum_{j=1}^m r_jg_{-j}}\Big) = O_p\Big(\frac{1}{T}\Big)O\Big(\frac{1}{T}\Big) \qquad (4.57)
$$

and

$$R_{3T} = \frac{(\hat{g}_0 - \sum_{j=1}^m r_j \hat{g}_{-j}) - (g_0 - \sum_{j=1}^m r_j g_{-j})}{g_0} \tag{4.58}$$

We see that

$$\begin{aligned} R_{3T} &= \frac{\Big(\sum_{j=1}^m r_j \hat{g}_0^{(1} j/T + O_p(T^{-2})\Big) - \big(\sum_{j=1}^m r_j g_0 j/T + O_(T^{-2})\big)}{g_0} \\ &= \frac{1}{T}\frac{(\hat{g}_0^{(1)} - g_0^{(1)})}{g_0}\big(\sum_{j=1}^m r_j j\big) + O_p(T^{-2}) \end{aligned} \tag{4.59}$$

Putting these terms together, noting that $T\sqrt{T\gamma^3}R_{2T} = o_p(1) = T\sqrt{T\gamma^3}R_{1T}$, we have, from Proposition 3,

$$T\sqrt{T\gamma^3}\hat{A}(x_0) - A(x_0) = \sqrt{T\gamma^3}(\hat{g}_0^{(1)} - g^{(1)})\frac{\big(\sum_{j=1}^m r_j j\big)}{g_0} + o_p(1) \xrightarrow{d} N(0, \Sigma(x_0)).$$

To derive the distribution of $\hat{\mathcal{R}}(x_0)$, note that it is the weighted sum of $\hat{A}(x_0)$'s and we can decompose it similarly as

$$\hat{\mathcal{R}}(x_0) - \mathcal{R}(x_0) = \sum_{l=1}^m r_l R_{3T}(x_l) + O_p(T^{-1})O_p(|\hat{g}_0 - g_0|) \tag{4.60}$$

where

$$\begin{aligned} R_{3T}(x_l) &= \frac{(\hat{g}_l - \sum_{j=1}^m r_j \hat{g}_{l-j}) - (g_l - \sum_{j=1}^m r_j g_{l-j})}{g_l} \\ &= \frac{(\frac{l}{T}\hat{g}_0^{(1)} - (\sum_{j=1}^m r_j \hat{g}_0^{(1)}(l-j)/T) - (\frac{l}{T}g_0^{(1)} - (\sum_{j=1}^m r_j g_0^{(1)}(l-j)/T)}{g_0} + O_p(T^{-2}) \\ &= \frac{1}{T}\frac{(\hat{g}_0^{(1)} - g_0^{(1)})(\sum_{j=1}^m r_j j)}{g_0} + O_P(T^{-2}). \end{aligned} \tag{4.61}$$

Thus,

$$\begin{aligned} \hat{\mathcal{R}}(x_0) - \mathcal{R}(x_0) &= \sum_{l=1}^m r_l R_{3T}(x_l) + O_p(T^{-1})O_p(|\hat{g}_0 - g_0|) \\ &= \frac{1}{T}\frac{(\hat{g}_0^{(1)} - g_0^{(1)})}{g_0}\Big(\sum_{j=1}^m r_j j\Big) + O_p(T^{-1})O_p(|\hat{g}_0 - g_0|) \end{aligned} \tag{4.62}$$

which corresponds to the approximation to $\hat{A}(x_0)$.

4.9 Other Considerations

For each jurisdiction a trend is constructed commencing on the first incidence day and incremented by 1 each subsequent day. [19]

One manner in which to compare jurisdictions is with respect to their specific fixed effect. (α_j's below.) There are a variety of ways to do this. One is to simply compare the rankings. Some caution needs to be taken here. If the approximations are as $N \to \infty$ then the distribution of N of these is effectively unknown and constructing confidence intervals around them will be problematic. However, if T is relatively large then approximate intervals can be constructed based on asymptotic theory for any fixed number of these. [20] Alternatively, if we are focused simply on one jurisdiction we can introduce separate parameters for that jurisdiction and impose homogeneity on the rest. Note that the modelling does not allow separately for demographic considerations such as population density. This is a fixed effect and is incorporated into the $A_j(t)$ parameters.

To allow for day-of-the-week effects (most certainly administrative, which do show up) we defined indicators 7×1 indicators $D_t = \begin{pmatrix} D_{t1} & \cdots & D_{t7} \end{pmatrix}$. The resulting regression function can then be modified to $g(\tau, D)$ in the obvious way. The kernel used then was for mixed data of the form

$$K(\tau, D) = \frac{1}{T_j \gamma_j} K\Big(\frac{\tau_{j,t} - \tau}{\gamma}\Big) \prod_{l=1}^{7} \gamma_d^{1-D_{tl}} \tag{4.63}$$

with $\gamma_d \in [0, 1]$. The asymptotic distribution for the non-parametric components becomes amended with T_j replaced by $T_j/7$ in the standard errors. For window-width we cross-validated as follows. Putting $\gamma_j = s_j T_j^{-1/5} c$ where T_j is the number of observations on jurisdiction j and s_j the standard deviation of the $\tau_{j,t}$'s we minimized over the leave-one-out estimates over $c > 0$ and $\gamma_d > 0$. As stated, we used a quadratic kernel.

4.9.1 Outliers and trimming

Numerical difficulties can arise in semi-parametric estimation when the non-parametric component can cause an argument to be undefined or unbounded. Typically this problem occurs when the estimand is close to zero and appears in a denominator or logarithm. In these cases it is common to trim the estimators in some way. In our case this difficulty manifests itself in two ways. First, the lagged values of expected infections appear in the denominator of the semi-parametric time trend and these are often zero or close to zero. Second, this is compounded in a very few cases, but enough to create numerical difficulties by a few instances where the number of infections jumped from zero to quite large numbers. Although this is remotely possible in real life, it is more likely that some administrative clumping of reporting occurred. One possible remedy for this would be to do some preliminary smoothing. [21] However, any smoothing of this form is invariably *ad hoc* and risks removing valuable movement in the data. To alleviate the problems we trimmed the data in two ways. First,

[19] With the parametric benchmark estimates these are multiplied by 1/100 for numerical purposes. When interpreting the results and/or doing simulations, the estimates need to be rescaled.

[20] For a few cases there are T_j is quite small and we would be very hesitant to read much into these.

[21] It is common to see infections reported in rolling average form.

we set the minimum for expected infections at .01. Second, we truncated expected infections to be no higher than 10 times the expected infections of the previous period.[22]

4.9.2 Goodness of Fit

There are numerous measures of goodness of fit used with Poisson models.[23] Our use of these is largely for indicative purposes rather than for rigorous model selection. We have several objectives. One is to compare models, another is to provide a rough decomposition of the relative impact of subsets of considerations, stochastic heterogeneity, deterministic trends. To maintain comparability across models we report: $\bar{R}^2$, simply based on the ratio of squared residuals or the model at hand to deviations from the jurisdictional average.

Let $T_{(j)}$ denote the date of the first reported infection in jurisdiction j, T_j the total number of observations after the first reported infection in jurisdiction j ($T_{(j)} = T - T_j + 1$) and $T(N) = \sum_{j=1}^{N} T_j$. We define

$$S = \frac{1}{T(N)} \sum_{j=1}^{N} \sum_{t=1}^{T} p_{jt}(y_{j,t} - \hat{y}_{jt})^2, \tag{4.64}$$

$$S_0 = \frac{1}{T(N)} \sum_{j=1}^{N} \sum_{t=1}^{T} p_{jt}(y_{j,t} - \bar{y}_j)^2, \tag{4.65}$$

where $\hat{y}_{j,t}$ denotes a "fitted value" from the model at hand and

$$p_{jt} = \begin{cases} 0, & t < T_{(j)} \\ 1, & \text{otherwise} \end{cases} \tag{4.66}$$

$$\bar{y}_j = \frac{1}{T_j} \sum_{t=1}^{T} p_{jt} y_{j,t}, \qquad \bar{y} = \frac{1}{N\hat{T}} \sum_{j=1}^{N} \sum_{t=T_j}^{T} y_{j,t}, \tag{4.67}$$

The measure of goodness of fit for the Poisson models was thus

$$\bar{R}^2 = 1 - \frac{S}{S_0}. \tag{4.68}$$

Note that this is not necessarily in $[0, 1]$. Given the heterogeneity between jurisdictions it makes sense to use different benchmarks measures of deviation than simply overall deviations from overall mean. The R^2's reported below for the parametric Ordinary Least Squares and Fixed Effects estimators are the usual R^2 statistics. Note that these are for goodness of fit for the logs of counts (+1).

As a practical matter the autoregression parameters were constrained to be positive and sum to one by using the estimates $r_s = e^{\rho_s}/(1 + \sum_{l=2}^{m} e^{\rho_l})$, $s = 2, \ldots, m$ and $r_1 =$

[22] Despite this, there are still a few anomalous results that crop up. Note in the tables the presence of some skewness/outliers in our estimated $\mathcal{R}_0$'s which cause the mean and median across jurisdictions to occasionally diverge (also the maximum values). Since these do not affect the overall results we have chosen to leave in all observations rather than remove in an *ad hoc* manner.

[23] Note that the non-linearity and implicit heterogeneity of count models invalidates some of the statistical properties of standard R^2 statistics.

$1/(1+\sum_{l=2}^{m} e^{\rho_l})$. (This constraint makes them comparable to the popular device of making them fit a probability distribution such as the gamma, but without the implied constraints on the coefficients including uni-modality.) The ρ_s's were then estimated without constraint with standard information matrices obtained from the Hessian of the likelihood function and gradient method. To obtain standard errors for the estimates of the r_s's the delta method was used with the $m \times m-1$ Jacobian matrix

With

$$r(\rho) = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{pmatrix} = \begin{pmatrix} 1/(1+\sum_{s=2}^{m} e^{\rho_s} \\ r_1 e^{\rho_2} \\ \vdots \\ r_1 e^{\rho_m} \end{pmatrix} \tag{4.69}$$

$$J = \frac{\partial r}{\partial \rho} = \begin{pmatrix} -r_1 r_2 & -r_1 r_3 & \cdots & -r_1 r_m \\ -r_1 r_2 e^{\rho_2} + r_1 e^{\rho_2} & -r_1 r_3 e^{\rho_2} & \cdots & -r_1 r_4 e^{\rho_2} \\ \vdots & \vdots & \vdots & \vdots \\ -r_1 r_2 e^{\rho_m} & -r_1 r_3 e^{\rho_m} & \cdots & -r_1 r_m e^{\rho_m} + r_1 e^{\rho_m} \end{pmatrix} \tag{4.70}$$

$$= \frac{\partial r}{\partial \rho} = \begin{pmatrix} -r_1 r_2 & -r_1 r_3 & \cdots & -r_1 r_m \\ -r_2^2 + r_2 & -r_2 r_3 & \cdots & -r_2 r_m \\ \vdots & \vdots & \vdots & \vdots \\ -r_m r_2 & -r_m r_3 & \cdots & -r_m^2 + r_m \end{pmatrix} \tag{4.71}$$

$$= -r r_{(1)}^{\top} + \begin{pmatrix} 0_{1\times m-1} \\ \text{diag}[r_{(1)}] \end{pmatrix} \tag{4.72}$$

5 Empirical Results

The estimators proposed in the previous section generate many different results: estimates of the autoregression parameters common across jurisdictions, as well as the daily predicted values for each of the jurisdictions as well as their daily reproduction numbers. We will summarize these as well as some benchmark estimates using the data set.

5.1 Benchmark Estimates

Many individuals and institutions have worked on various COVID-19 data sets, both formally and informally using a wide range of estimators. We initially conducted fairly extensive analysis of the data using standard regression modelling for panel data as well as purely non-parametric estimation. We report here as benchmarks estimates using a variety of specifications for the parametric regression function. These would largely agree with the literature published in the popular media. For illustrative purposes we summarize a few of these results. We first summarize some basic results using OLS and fixed effects estimators using $\ln(1+y_{j,t})$ as the dependent variable. Note again that not all jurisdictions had the same start date.

Various pooled least squares estimates are given in the first table. Here the dependent variable is $\ln(1+y_{j,t})$ with 14 lagged values of $\ln(1+y_{j,t})$ as explanatory variables as well as

a quartic polynomial in the time trend. When the pandemic first broke out, it appeared as if a quadratic trend would (perhaps hopefully) explain the spread of infections. Subsequently, although this has been the case in some jurisdictions, surges or second waves of the infection would imply that its spread is not so easily quantified. From an inspection of the coefficients in the first table we can infer that the pandemic has a fairly long lag structure, but the value of the lagged effects does dissipate over time. In our initial analysis of the data we conducted separate similar regressions on a case by case basis. What is evident is that although there is some similarity between some jurisdictions there is also substantial heterogeneity across them. We also, again more to get a sense for the data and highlight a few of its characteristics, include a variety of fixed effects estimators. These correspond to the pooled OLS results with a variety of results using 14 period lags, quartic trend and various combinations of same. We note the similarity to the pooled OLS results. We also note that (in terms of goodness of fit) there is much to be gained by including both stochastic and deterministic elements as explanatory variables.

One component of the semi-parametric results are the first stage non-parametric trends by jurisdiction. These are useful as benchmarks and correspond with some caveats to the estimates in Linton (2020). The first caveat is that there are two conditioning variables. One is the time trend which for each jurisdictions commences on the day of the first recorded infection. The other is a smoothed indicator for day-of-the-week. It became apparent that the latter has a substantial impact, probably the result of administrative recording. This was important to include: omitting it distorts the estimates of the autoregressive parameters. The plots of some of the non-parametric estimates are in the graphs. The goodness of fit for three kernel based non-parametric estimates is in the table immediately following the log-linear fixed-effects results. These correspond to using 1, local constant (aka kernel), 2, local linear and 3, local quadratic polynomials. We note that the local polynomials have a (theoretical) advantage over the local constant estimators in terms of bias at the boundary of the support which may be advantageous for forecasting.

Thus, in addition to the estimators developed in this paper, we provide, for comparison, a variety of benchmark estimators which range from entirely parametric, assuming a great deal of homogeneity (log-linear pooled and fixed effects regression estimators) to the least parametric, which assume no homogeneity across jurisdictions (kernel and local polynomial estimators).

5.2 Count Model Estimates

There are six sets of count model results containing estimates using a wide range of Poisson-like specifications. All have in common a stochastic AR(14) component. Table 5 contains three sets of parametric results: for a "pooled" Poisson model where there is one $A_j(t) = e^{\alpha}$ estimated, constant across jurisdictions and time; the second corresponds to the fixed effects where there are jurisdiction-varying, time invariant $A_j(t) = e^{\alpha_j}$'s. The third set of results correspond to time-varying fixed effects where the $A_j(t)$'s vary across jurisdictions but are constant over intervals of 26 weeks. (Estimates over 21 and 31 weeks were very similar.) Table 6 contains three sets of results corresponding to semi-parametric estimates using, respectively, kernel, local linear and local quadratic polynomial estimators of the trend. With the semi-parametric estimators the $A_j(t)$'s vary across jurisdictions on a daily basis.

We note that there is a wide agreement across all these models as to the general shape of the auto-correlation pattern.

5.2.1 Parametric Count Model Estimates

The first sets of results is when the conditional mean consists of the stochastic trend with 14 lags and a common multiplicative constant, i.e., e^{α}. This corresponds to an assumption of complete homogeneity across time and jurisdictions. Note the estimated values of r_s are relatively large for small s and then largely decreasing with s. Note that there is one implied common $\mathcal{R}_j(t)$ estimate across time and jurisdictions. The estimate is 1.0118. We note that values of $\mathcal{R}_j(t)$ greater than one generate exponential growth rates for pandemic spread. This is not an obscure point, a value of $\mathcal{R}_j(t)$ even slightly above one can lead to widespread illness in a very short time. 1.0118 may seem relatively small, but note that this is an average over jurisdictions and time so that this reflects gradual overall reductions (albeit non-monotonic) over time in infections in many if not most jurisdictions.

The second set of results allow for different rates of spread across jurisdictions, but again with no deterministic trend. This corresponds to an assumption of homogeneity across time but not jurisdictions.The pattern of the autoregressive coefficients is similar to the completely homogenous case. There are N different coefficients and $\mathcal{R}_j(t)$'s estimated implicitly and we don't report these here, but we report their quartiles mean and standard deviation across jurisdictions. Note that the median and average values are 1.0049 and 1.0170 which are quite close to the estimate with complete homogeneity. Note that with no time variation in the $A_j(t)$'s there are no separate case and instantaneous reproduction numbers.

5.2.2 Flexible-Form and Semi-Parametric Count Model Estimates

We have four sets of count model results which incorporate a deterministic trend (either piecewise constant for the time-varying fixed coefficients case or smooth, for the three semi-parametric estimators). For these cases we include tables of estimates of the autoregression parameters and tables describing the distribution of the daily, jurisdiction level estimates of the $\mathcal{R}_j(t)$'s. The latter tables provide the quartiles of these on a weekly basis [24] as well as their sample means and standard deviations. There are also two sets of illustrative graphs for five of the countries in the sample. One set plots the daily observed counts, non-parametric and semi-parametric fitted values of infections. The second set plots the estimated case and instantaneous $\mathcal{R}_j(t)$'s for these same countries.

The first set of results for models incorporating a deterministic trend, corresponds to the time-varying fixed-effects model. Here the pattern of the estimates of the autoregressive coefficients is not so precise as in the fixed-coefficients case although the coefficients for the later lags is definitely smaller. Here and in the semi-parametric cases it useful to examine the $\mathcal{R}_j(t)$'s and their distribution overtime. These diminish fairly monotonically. The quantiles' tendencies are similarly decreasing over time, though note the first quantile estimate is above 1.5 and the third quantile at 2.8 are very high from a public health perspective. There are many interesting aspects to these estimates. For example, with respect to the first quantile,

[24] That is, the $\mathcal{R}_j(t)$' as estimated on days $1, 8, 15, \ldots$. with the sample statistics over the N jurisdictions on that day.

note that it was only after the third week that barely a quarter of jurisdictions had their $\mathcal{R}_j(t)$ under one. Note that over half the jurisdictions in the sample had estimated $\mathcal{R}_j(t)$'s greater than one throughout most of the sample period.

The final three sets of tables correspond to the distribution of semi-parametric estimates of the case and instantaneous $\mathcal{R}_j(t)$'s. They correspond to estimates using local constant, linear and quadratic polynomials in the first stage. Each give similar results with some caveats. The point estimates are all consistent with gradually decreasing infectiousness but with resurgences (note in particular the third quartile). Note that with all estimators the mean and median can be quite different. We attribute this to an asymmetry in the distribution of the pandemic's severity as well as some remaining outliers which can have a large impact on some of the estimates. However, the quartiles of the $\mathcal{R}_j(t)$'s, including the median, are each quite consistent with each other and the flexible form results. In terms of goodness of fit the second-order polynomial dominates the other two somewhat, although as a caveat we found the second-order polynomial could result in some anomalous fitted values.

After the tables are graphical examples of the fits of the estimators and the $\mathcal{R}_j(t)$'s for five countries: US, Taiwan, Brazil, New Zealand and Sweden. Note that these all begin on the first day of the outbreak in each jurisdiction, which is January 22 for the US and Taiwan and January 31 for Sweden, February 26 for Brazil, February 28 for New Zealand. For each country we plot: 1, recorded infections, 2, fitted values from local regression (first-order polynomial) and 3, semi-parametric estimator using the same local regression in the first stage. From the plots we see that the semi-parametric estimator is much better at tracking some of the short-term fluctuations in the infections which are smoothed out by the pure time-trend.

Accompanying graphs plot the case and instantaneous $\mathcal{R}_j(t)$ daily estimates. These largely confirm what has been reflected in the news media regarding infections. Taiwan and New Zealand have had a lower $\mathcal{R}_j(t)$ than most countries, although with some increase at the beginning. Sweden, which was more permissive in its approach to social distancing had a fairly large initial $\mathcal{R}_j(t)$, but then diminishing with a second wave. The United States and Brazil have had a great deal of difficulty in getting the spread of COVID-19 under control, although at time of writing this seemed to be changing. Note that the case and instantaneous $\mathcal{R}_j(t)$ estimates are quite similar, the case number seems to lead the instantaneous number in most cases which may be important for prediction and policy reasons. The instantaneous reproduction number, which does not require forecasts of future infections, is simpler to estimate.

In regard to the Poisson models, there seems to be a pattern in the autoregression pattern that is quite robust across each of the specifications. In all cases there is a hump in the parameters corresponding to the sixth and seventh lags. Note that r_s can be interpreted as the product of a measure of infectiousness and the survival function at period s. Since the latter is decreasing in s this pattern in the r_s's indicates an increase in infectiousness a week after the initial reporting of the infection. After that increase there is a largely common decrease with s in the parameter estimates until very small numerically. There are also bumps in the AR process at the end of the second and third weeks which may be reflecting some accounting.

We note some limitations of the methods and/or the data. The fitted values and reproduction number estimates may be somewhat unreliable at the beginning of the processes.

This may reflect a number of issues. Recording of cases may be particularly inaccurate at the beginning of each outbreak and it may be the case that infections may have already occurred prior to the first observed case. We note that the estimates of the two different reproduction numbers track each other quite closely. The case $\mathcal{R}_j(t)$ seems to lead the instantaneous $\mathcal{R}_j(t)$ in most cases so would appear to be a better lead indicator of the direction of the pandemic.

6 Conclusion

This paper has developed new panel count data estimators for the analysis of the progress COVID-19 pandemic using an array of parametric and semi-parametric estimators which exploit the fact that the biological component of the virus' spread should be fairly homogenous across jurisdictions. The approach allows for two new direct methods of estimating of estimating the reproduction numbers associated with the disease. The estimators are applied to international panel data and produce compelling estimates of the COVID-19 process.

The autoregressive count methodology is consistent with an epidemiological model. The semi-parametric estimators in particular are capable of tracking the spread of the COVID-19 pandemic and provide simple, direct estimates of jurisdiction and time specific reproduction numbers.

Extensions we are considering at present are allowing for dependence between observations using, say, a gravity model and joint modelling of infections with deaths and recoveries.

At the time of writing new strains of the COVID-19 had manifested themselves. New strains of any virus are standard and since the beginning of the pandemic various reports had appeared of new strains, but none sufficiently different to have much impact on statistical studies. These new strains, the first appearing in the UK in mid-December, 2020, while apparently no more lethal, do appear to spread more easily. In principle, this could be modelled in a similar way with the new strain having a different set of r's.

REFERENCES

Allen, L.J.S. (1994) Some Discrete-Time SI, SIR and SIS Epidemic Models, Mathematical Biosciences, 124, 83-105.

Alvarez, F., Argente, D. and F. Lippi (2020) A Simple Planning Problem for COVID-19 Lockdown, LUISS and Einaudi Institute for Economics and Finance DP.

Atkeson, A. (2020) What will be the economic impact of COVID19 in the US? Rough estimates of Disease scenarios, NBER Working Paper 26867.

Andersen, E.B. (1970) Asymptotic Properties of Conditional Maximum Likelihood Estimators, Journal of the Royal Statistical Society, Series B, 32, 283-301.

smallskipAndrews, Donald W K. (1994) Asymptotics for semiparametric econometric models via stochastic equicontinuity, Econometrica, 62, 43-72.

Bendavid, Eran and Jay Bhattacharya (2020) Is the Coronavirus as Deadly as They Say?, Wall Street Journal, March 24.

Biggerstaff, Matthew, Simon Cauchemez, Carrie Reed, Manoj Gambhir and Lyn Finelli (2014) Estimates of the reproduction number for seasonal, pandemic, and zoonotic

influenza: a systematic review of the literature , BMC Infectious Diseases,14, Article number: 480.

Cameron, A. Colin and Pravin K. Trivedi (2005) Microeconometrics: Methods and Applications, Cambridge, New York.

Carroll, R.J. (1982) Adapting for Heteroscedasticity in Linear Models , Annals of Statistics, 10, 1223-1233.

Champredon, David, Jonathan Dushoff and David J.D. Earns (2018) Equivalence of the Erlang-distributed SEIR epidemic model and the renewal equation, SIAM Journal of Applied Mathematics, Vol. 78, No. 6, pp. 3258–3278.

Chowell G. and H. Nishiura (2008) Quantifying the transmission potential of pandemic influenza, Physics of Life Reviews, vol. 5, no. 1, pp. 50–77.

Chowell, Gerardo , Hiroshi Nishiura and Luís M.A Bettencourt (2006) Comparative estimation of the reproduction number for pandemic influenza from daily case notification data, Journal of The Royal Society Interface.

Cori, Anne, Neil M. Ferguson, Christophe Fraser and Simon Cauchemez (2013) A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics American Journal of Epidemiology, 178(9), 1505–1512.

Di Domenico, L., Pullano, G., Coletti, P., Hens, N. and V. Colizza (2020):Expected impact of school closure and telework to mitigate COVID-19 epidemic in France, www.epicx-lab.com/covid-19.html].

Eichenbaum, Martin S, Sergio Rebelo, Mathias Trabandt (2020) The Macroeconomics of Epidemics, NBER Working Paper 26882, Cambridge.

El Koufi, A., Adnani, J., Bennar, A. and N. Yousfi (2019): Analysis of a Stochastic SIR Model with Vaccination and Nonlinear Incidence Rate, International Journal of Differential Equations,Vol 2019, ID 9275051.

Escanciano, Juan Carlos, David T. Jacho-Chávez and Arthur Lewbel (2014) Uniform convergence of weighted sums of non and semiparametric residuals for estimation and testing, Journal of Econometrics, Vol 178, 426–423.

Hagemann, Andreas (2014) Stochastic equicontinuity in nonlinear time series models, Econometrics Journal, 17, pp. 188-196.

Hansen, Bruce E. (2009) Stochastic Equicontinuity for Unbounded Dependent Heterogeneous Arrays, Econometric Theory, 12, 347 - 359.

Hausman, J., B. Hall and Z. Griliches (1984) Economic Models for Count Data with an Application to the Patents–R&D Relationship, Econometrica, 52, 909-938.

Heesterbeek, J. A. P. and K. Dietz (1996) The concept of Ro in epidemic theory, Statistica Neerlandica Vol. 50(1) , pp. 89-110.

Held, L. and M. Paul (2012) Modelling seasonality in space-time infectious disease surveillance data, Biom. J. 54, 824-843.

Hens, N, Shkedy, Z., Aerts, M., Faes, C.,Van Damme, P. and P. Beutels (2012) Modelling Infectious Disease Parameters Based on Serological and Social Contact Data, Springer, New York.

Korolev, I. (2020) Estimating the Parameters of the SEIRD Model for COVID-19 using the Deaths Data, Department of Economics, Binghamton University DP.

Kermack, W. and A. McKendrick (1927) Contributions to the mathematical theory of epidemics – I, reprinted in (1991) Bulletin of Mathematical Biology. 53 (1–2): 33–55.

Kermack, W. and A. McKendrick (1932) Contributions to the mathematical theory of epidemics – II. The problem of endemicity, reprinted in (1991). Bulletin of Mathematical Biology. 53 (1–2): 57–87.

Kermack, W. and A. McKendrick (1933) Contributions to the mathematical theory of epidemics – III. Further studies of the problem of endemicity, reprinted in (1991) Bulletin of Mathematical Biology. 53 (1–2): 89–118.

Li, Jing, Daniel Blakeley and Robert J. Smith (2011) The Failure of $\mathcal{R}_0$, Computational and Mathematical Methods in Medicine Volume 2011.

Li, Qi and J.S. Racine (2007) Nonparametric Econometrics, Princeton University Press, Princeton.

Linton, Oliver (2020) Modelling the COVID-19 Pandemic, International Asssociation for Applied Econometrics Webinar, August 12, 2020.

Bruce D. Meyer (1990) Unemployment Insurance and Unemployment Spells, Econometrica, 58, 757-782.

Meyer, Sebastian and Leonhard Held (2014) Power-Law Models for Infectious Disease Spread, The Annals of Applies Statistics, 8:3, 1612-1639.

Newey, Whitney K. and Daniel McFadden (1994) Large sample estimation and hypothesis testing, in Handbook of Econometrics, Vol 4, 2113–2247.

Riou, Julien and Christian L. Althaus (2020) Patterns of Early Humanto-human Transmission of Wuhan, bioRxivEstimating the Parameters of the SEIRD Model for COVID-19 using the Deaths Data, Department of Economics, Binghamton University DP.

Robinson, P.M. (1987) Asymptotically efficient estimation in the presence of heteroskedasticity of unknown form, Econometrica, 55, 931-954.

Robinson, P.M. (1988) Root-n consistent semiparametric regression , Econometrica, 56, 531-954.

Schmid, Volker (2006) A two-component model for counts of infectious diseases, Biostatistics, 7, 3, 422-437.

Toda, A. A. (2020): Susceptible-Infected-Recovered (SIR) Dynamics of COVID-19 and Economic Impact, ArXiv:2003.11221v2.

Wand, M.P. and M.C. Jones (1995) Kernel Smoothing, Chapman and Hall, London.

Wooldridge, Jeffrey M. (2010) Econometric Analysis of Cross Section and Panel Data, 2nd ed., MIT Press, Cambridge.

Appendix: Tables and Figures

Table 1: Pooled OLS Results with 14 lags count and/or quartic trend

	with lags and trends		with lags		with trends	
Variable	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Lag 1	0.0787	26.1281	0.0831	27.5650		
Lag 2	0.1226	40.5189	0.1264	41.7221		
Lag 3	0.1105	36.2401	0.1137	37.2370		
Lag 4	0.1042	33.9822	0.1071	34.8658		
Lag 5	0.0802	26.0394	0.0829	26.8349		
Lag 6	0.1301	42.0992	0.1325	42.7749		
Lag 7	0.3258	104.5120	0.3271	104.6505		
Lag 8	0.0590	18.8689	0.0565	18.0467		
Lag 9	-0.0476	-15.3025	-0.0510	-16.3865		
Lag 10	-0.0592	-19.0799	-0.0629	-20.2202		
Lag 11	-0.0673	-21.7182	-0.0712	-22.9469		
Lag 12	-0.0446	-14.4289	-0.0487	-15.7418		
Lag 13	0.0037	1.2174	-0.0008	-0.2574		
Lag 14	0.1995	65.0707	0.1946	63.4210		
t	-0.6877	-17.1783			2.5205	20.3739
t^2	0.4408	13.7562			-1.1683	-11.7382
t^3	-0.1119	-11.7309			0.2843	9.5784
t^4	0.0097	10.2452			-0.0264	-9.0116
Constant	0.3879	26.3332	0.0705	16.4338	1.0012	22.1171
R^2	0.9029		0.9024		0.0533	

Table 2: Fixed Effects Results with 14 lagged counts and quartic trend

Variable	Estimate	t-stat
Lag 1	0.0690	22.9090
Lag 2	0.1137	37.5886
Lag 3	0.1028	33.7685
Lag 4	0.0976	31.9191
Lag 5	0.0748	24.3532
Lag 6	0.1256	40.7730
Lag 7	0.3223	103.8212
Lag 8	0.0578	18.5794
Lag 9	-0.0485	-15.6729
Lag 10	-0.0608	-19.6780
Lag 11	-0.0697	-22.5819
Lag 12	-0.0479	-15.5476
Lag 13	-0.0003	-0.0840
Lag 14	0.1954	63.9262
t	-0.4979	-12.3293
t^2	0.3531	10.9948
t^3	-0.0917	-9.5896
t^4	0.0079	8.3196
R^2	0.6893	

Table 3: Fixed Effects Results with 14 lagged counts

Variable	Estimate	t-stat
Lag 1	0.0702	23.2886
Lag 2	0.1146	37.8863
Lag 3	0.1036	33.9985
Lag 4	0.0983	32.1142
Lag 5	0.0754	24.5273
Lag 6	0.1261	40.9123
Lag 7	0.3225	103.7877
Lag 8	0.0567	18.2211
Lag 9	-0.0499	-16.1425
Lag 10	-0.0623	-20.1797
Lag 11	-0.0713	-23.1328
Lag 12	-0.0496	-16.1307
Lag 13	-0.0021	-0.6996
Lag 14	0.1933	63.4256
R^2	0.6888	

Table 4: Fixed Effects Results with just quartic trend

Variable	Estimate	t-stat
t	2.2956	58.0509
t^2	-0.9490	-29.7866
t^3	0.2115	22.2214
t^4	-0.0190	-20.1882
R^2	0.0000	

Table 5: Point Estimates and t-stats Parametric Models

	Poisson		Fixed FX		Variable Fixed FX	
Variable	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Lag 1	0.1166	78857.8448	0.2503	165233.6504	0.12	108092.93
Lag 2	0.1195	72083.5880	0.1072	66459.5040	0.07	45882.75
Lag 3	0.0647	40275.5103	0.0363	25212.5578	0.02	16536.15
Lag 4	0.0407	27936.5396	0.0193	15255.5853	0.02	13408.67
Lag 5	0.0445	30911.3764	0.0178	15004.6973	0.01	6802.20
Lag 6	0.1380	80028.5733	0.1334	80754.3729	0.12	78015.96
Lag 7	0.4210	215258.3474	0.4026	212609.1149	0.40	266581.04
Lag 8	0.0031	2879.3392	0.0073	6145.7217	0.04	32030.05
Lag 9	0.0019	3503.4497	0.0011	3296.1438	0.00	4474.15
Lag 10	0.0013	1730.3238	0.0010	1715.1449	0.00	2144.47
Lag 11	0.0004	1231.6369	0.0005	1265.7473	0.00	1333.10
Lag 12	0.0007	2556.9607	0.0005	2532.1252	0.00	2678.55
Lag 13	0.0070	5550.1336	0.0013	1302.2248	0.01	6183.22
Lag 14	0.0405	27527.2148	0.0214	16591.8203	0.19	178613.27

Table 6: AR Point Estimates and t-stats: Semiparametric Models

	SP-Kernel		SP-Local Linear		SP-Local Quadratic	
Variable	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Lag 1	0.0654	10217.3592	0.08	11803.48	0.05	7077.67
Lag 2	0.2621	26962.0516	0.27	28567.82	0.25	23993.37
Lag 3	0.0000	402.9697	0.00	423.81	0.11	12278.51
Lag 4	0.0741	9882.0843	0.12	15805.71	0.05	5896.71
Lag 5	0.0001	469.3961	0.00	481.74	0.03	3495.55
Lag 6	0.0911	13311.4545	0.06	10823.62	0.08	11025.75
Lag 7	0.0010	4407.7208	0.00	2896.21	0.00	3157.54
Lag 8	0.1410	15895.4420	0.16	18438.75	0.09	10809.88
Lag 9	0.1601	16215.1040	0.14	15067.98	0.13	12728.97
Lag 10	0.0367	5418.9402	0.01	3694.43	0.06	6184.65
Lag 11	0.0332	4596.6403	0.05	6452.74	0.04	4761.75
Lag 12	0.0188	3535.9058	0.02	3720.07	0.04	5191.10
Lag 13	0.1113	13759.8314	0.08	12305.28	0.08	11232.15
Lag 14	0.0049	11503.9231	0.00	4942.09	0.00	11680.89

Table 7: R-squares for local polynomials

	SP-Kernel	SP-Local Linear	SP-Local Quadratic
R-square	0.9290	0.9288	0.9289

Table 8: Poisson Reproduction Number with Lagged Infections and Constant

R0	1.0118

Table 9: Fixed Effects Reproduction Number Distribution

	Q1	Median	Q3	Mean	SD
R0	1.0008	1.0049	1.0179	1.0170	0.0379
R-squared	0.8382				

Table 10: Variable Fixed Effects Reproduction Number Estimator Distribution

		Case					Instantaneous				
Week	N	Q1	Median	Q3	Mean	SD	Q1	Median	Q3	Mean	SD
1	221	1.11	1.33	1.68	1.55	0.73	1.11	1.33	1.68	1.55	0.72
2	221	1.11	1.33	1.68	1.55	0.72	1.11	1.33	1.68	1.55	0.72
3	221	1.12	1.34	1.69	1.56	0.72	1.11	1.33	1.68	1.55	0.72
4	221	1.02	1.30	1.73	1.52	0.81	1.11	1.33	1.68	1.55	0.72
5	221	0.81	1.15	1.59	1.36	1.22	1.09	1.31	1.74	1.59	0.94
6	221	0.61	1.07	1.50	1.26	1.37	0.81	1.17	1.60	1.35	1.39
7	221	0.48	1.01	1.41	1.15	1.30	0.47	1.04	1.50	1.24	1.41
8	221	0.61	1.02	1.40	1.06	1.00	0.47	1.00	1.41	1.16	1.39
9	221	0.75	1.04	1.36	1.08	0.77	0.53	1.00	1.37	1.05	1.24
10	221	0.73	1.03	1.34	1.07	0.79	0.72	1.03	1.38	1.08	0.79
11	221	0.73	1.00	1.31	1.03	0.70	0.73	1.03	1.37	1.09	0.85
12	221	0.62	0.97	1.23	0.97	0.74	0.73	1.01	1.29	1.03	0.79
13	221	0.56	0.98	1.23	0.98	0.89	0.52	0.95	1.23	0.96	0.93
14	221	0.64	0.99	1.25	1.00	0.86	0.55	0.97	1.23	0.99	0.93
15	221	0.72	1.02	1.22	1.05	0.89	0.56	0.98	1.23	0.98	0.88
16	221	0.81	1.04	1.26	1.05	0.82	0.70	1.02	1.26	1.09	1.07
17	221	0.80	1.04	1.25	1.07	0.89	0.73	1.03	1.24	1.02	0.85
18	221	0.82	1.04	1.23	1.13	1.08	0.76	1.04	1.24	1.09	1.06
19	221	0.81	1.02	1.23	1.11	1.02	0.81	1.05	1.26	1.15	1.19
20	221	0.82	1.03	1.22	1.10	0.90	0.79	1.02	1.22	1.09	1.03
21	221	0.82	1.03	1.24	1.05	0.63	0.81	1.02	1.24	1.13	1.05
22	221	0.80	1.00	1.26	1.02	0.49	0.80	1.02	1.23	1.03	0.61
23	221	0.84	1.00	1.28	1.03	0.47	0.81	1.00	1.28	1.00	0.54
24	221	0.80	0.99	1.25	1.01	0.51	0.81	1.01	1.30	1.06	0.55
25	221	0.80	0.97	1.20	0.99	0.49	0.80	0.99	1.23	1.00	0.51
26	221	0.84	0.99	1.16	0.99	0.43	0.80	0.98	1.20	0.98	0.50
27	221	0.85	1.00	1.17	1.01	0.49	0.80	1.00	1.17	0.98	0.50
28	221	0.83	0.99	1.17	1.04	0.68	0.82	0.99	1.17	1.00	0.59
29	221	0.84	1.01	1.18	1.06	0.73	0.82	0.99	1.16	1.04	0.74
30	221	0.86	1.01	1.17	1.05	0.65	0.85	1.00	1.20	1.06	0.74
31	221	0.87	1.02	1.19	1.01	0.51	0.86	1.01	1.21	1.07	0.74
32	221	0.91	1.03	1.23	1.02	0.40	0.85	1.01	1.18	1.00	0.59
33	220	0.91	1.05	1.26	1.06	0.50	0.88	1.02	1.21	1.00	0.42
34	220	0.92	1.06	1.28	1.10	0.51	0.91	1.04	1.26	1.07	0.52
35	219	0.94	1.09	1.33	1.15	0.70	0.91	1.06	1.31	1.08	0.53
36	219	0.93	1.07	1.31	1.13	0.68	0.93	1.08	1.34	1.16	0.74
37	219	0.90	1.06	1.22	1.08	0.58	0.90	1.07	1.33	1.13	0.68
38	219	0.89	1.04	1.22	1.04	0.42	0.89	1.04	1.24	1.09	0.68
39	218	0.87	1.00	1.19	1.01	0.40	0.87	1.02	1.23	1.04	0.43
40	218	0.86	0.98	1.18	1.01	0.40	0.86	0.99	1.19	1.00	0.41
41	218	0.85	1.00	1.17	0.99	0.37	0.86	0.98	1.19	1.02	0.43
42	218	0.83	0.99	1.15	0.97	0.38	0.83	0.98	1.17	0.98	0.40
43	218	0.83	0.99	1.15	0.96	0.39	0.83	0.99	1.16	0.98	0.42
44	218	0.82	0.98	1.13	0.95	0.38	0.83	0.98	1.15	0.96	0.41
45	218	0.83	0.97	1.11	0.94	0.37	0.82	0.99	1.14	0.95	0.40

Table 11: Variable Fixed Effects Reproduction Number Estimator Distribution (cont.)

		Case					Instantaneous				
Week	N	Q1	Median	Q3	Mean	SD	Q1	Median	Q3	Mean	SD
46	218	0.83	0.95	1.09	0.94	0.46	0.83	0.97	1.14	0.95	0.40
47	218	0.80	0.93	1.10	0.94	0.49	0.82	0.93	1.10	0.93	0.49
48	218	0.81	0.95	1.12	0.97	0.50	0.78	0.93	1.10	0.93	0.52
49	218	0.82	0.96	1.12	1.04	0.65	0.78	0.92	1.10	0.96	0.60
50	218	0.83	0.98	1.14	1.08	0.78	0.81	0.97	1.13	1.05	0.79
51	218	0.86	1.00	1.17	1.10	0.76	0.82	0.97	1.14	1.08	0.81
52	218	0.88	1.00	1.16	1.06	0.63	0.86	0.99	1.18	1.12	0.80
53	218	0.82	0.98	1.15	0.99	0.53	0.86	1.00	1.18	1.08	0.74
54	218	0.80	0.99	1.13	0.97	0.58	0.79	0.97	1.17	0.97	0.63
55	218	0.81	1.00	1.15	0.97	0.58	0.80	0.98	1.14	0.97	0.58
56	218	0.82	0.99	1.16	0.98	0.51	0.80	0.99	1.18	0.97	0.59
57	218	0.85	1.01	1.16	0.96	0.39	0.81	1.01	1.18	0.99	0.59
58	218	0.81	1.00	1.15	0.95	0.46	0.82	1.02	1.18	0.96	0.42
59	218	0.79	0.97	1.10	0.93	0.45	0.80	1.00	1.17	0.96	0.48
60	217	0.77	0.97	1.11	0.94	0.44	0.77	0.96	1.11	0.91	0.47
61	217	0.77	0.97	1.13	1.00	0.59	0.73	0.95	1.10	0.93	0.47
62	215	0.74	0.97	1.14	0.99	0.56	0.75	0.97	1.13	0.98	0.55
63	215	0.76	0.97	1.12	0.97	0.51	0.74	0.97	1.13	1.00	0.57
64	215	0.75	0.95	1.08	0.93	0.42	0.74	0.97	1.13	0.98	0.58
65	214	0.72	0.93	1.10	0.91	0.44	0.72	0.94	1.10	0.91	0.44
66	208	0.71	0.92	1.09	0.89	0.43	0.72	0.92	1.10	0.91	0.44
67	196	0.69	0.87	1.07	0.86	0.39	0.71	0.90	1.10	0.88	0.41
68	170	0.69	0.88	1.05	0.85	0.35	0.66	0.86	1.06	0.83	0.39
69	121	0.66	0.88	1.07	0.83	0.37	0.66	0.88	1.07	0.83	0.37
70	88	0.67	0.89	1.07	0.82	0.36	0.67	0.88	1.06	0.82	0.36
71	59	0.68	0.89	1.08	0.83	0.42	0.67	0.93	1.09	0.83	0.40
72	46	0.57	0.82	1.14	0.82	0.59	0.67	0.91	1.07	0.82	0.47
73	45	0.57	0.84	1.16	0.82	0.61	0.57	0.84	1.16	0.82	0.61
74	45	0.57	0.84	1.16	0.82	0.61	0.57	0.84	1.16	0.82	0.61
75	36	0.00	0.79	1.14	0.78	0.67	0.00	0.79	1.14	0.78	0.67
R2	0.85										

Table 12: SP estimates Zero Order (Kernel) local Polynomial Case and Instantaneous Reproduction Number Distribution

		Case					Instantaneous				
Week	N	Q1	Median	Q3	Mean	SD	Q1	Median	Q3	Mean	SD
1	221	1.37	1.98	2.68	2.17	1.28	2.38	4.31	11.10	40.25	106.83
2	221	1.16	1.65	2.37	1.86	1.18	1.43	2.55	4.75	12.15	32.48
3	221	1.05	1.49	1.95	1.69	1.39	1.07	1.66	2.93	4.69	12.23
4	221	0.86	1.34	1.79	4.65	39.15	0.68	1.34	2.09	2.42	5.97
5	221	0.67	1.17	1.70	6.54	64.15	0.55	1.18	1.70	1.69	2.95
6	221	0.67	1.10	1.61	3.99	40.33	0.38	1.01	1.52	1.31	1.50
7	221	0.77	1.14	1.67	11.23	83.74	0.21	0.94	1.43	3.20	30.12
8	221	0.75	1.18	1.62	12.79	96.32	0.18	0.88	1.43	1.13	1.55
9	221	0.78	1.20	1.52	11.20	87.67	0.41	0.89	1.49	19.44	272.68
10	221	0.71	1.11	1.42	7.66	50.37	0.40	0.92	1.49	4.44	42.81
11	221	0.70	1.07	1.42	4.32	24.24	0.30	0.94	1.43	14.22	179.19
12	221	0.65	1.07	1.38	5.50	36.16	0.20	0.89	1.28	2.39	17.05
13	221	0.75	1.09	1.39	12.93	72.43	0.10	0.87	1.23	1.27	4.72
14	221	0.78	1.10	1.45	15.99	95.56	0.01	0.89	1.26	10.80	139.17
15	221	0.83	1.10	1.60	11.54	64.21	0.03	0.89	1.27	1.98	10.52
16	221	0.87	1.10	1.54	11.33	65.79	0.06	0.88	1.26	2.10	10.51
17	221	0.81	1.08	1.48	12.34	77.32	0.14	0.90	1.31	2.06	9.45
18	221	0.83	1.06	1.46	12.06	88.63	0.26	0.91	1.31	19.90	268.17
19	221	0.82	1.09	1.46	9.61	50.18	0.15	0.94	1.29	21.89	258.87
20	221	0.83	1.10	1.40	9.01	55.67	0.24	0.93	1.31	6.44	57.24
21	221	0.82	1.06	1.49	9.26	60.02	0.30	0.96	1.32	3.07	18.91
22	221	0.85	1.10	1.49	9.86	60.29	0.33	0.95	1.28	2.27	9.90
23	221	0.83	1.08	1.50	13.35	95.93	0.18	0.95	1.30	2.10	11.95
24	221	0.78	1.05	1.46	8.86	63.02	0.28	0.94	1.44	14.68	188.94
25	221	0.84	1.07	1.40	2.83	14.10	0.29	0.93	1.32	1.81	11.44
26	221	0.86	1.06	1.37	3.10	17.44	0.28	0.93	1.32	1.66	8.23
27	221	0.87	1.04	1.38	10.91	71.50	0.33	0.97	1.32	2.26	14.11
28	221	0.88	1.04	1.37	10.95	85.52	0.49	0.99	1.33	1.55	3.58
29	221	0.84	1.03	1.40	6.99	28.86	0.53	0.99	1.36	18.52	253.81
30	221	0.83	1.04	1.33	7.53	49.50	0.52	1.00	1.40	5.09	54.21
31	221	0.83	1.07	1.38	4.11	26.50	0.48	0.97	1.34	2.71	17.34
32	221	0.86	1.08	1.42	5.83	48.12	0.48	0.97	1.34	1.93	8.57
33	220	0.88	1.11	1.42	9.91	93.39	0.53	0.98	1.34	1.51	3.93
34	220	0.88	1.09	1.45	3.90	28.84	0.50	0.97	1.33	1.26	1.88
35	219	0.91	1.12	1.45	3.40	17.14	0.46	1.02	1.30	1.25	1.99
36	219	0.89	1.12	1.44	2.13	6.24	0.45	1.02	1.30	2.20	15.84
37	219	0.90	1.12	1.38	3.50	17.83	0.59	1.03	1.35	3.48	30.82
38	219	0.89	1.10	1.35	8.71	62.82	0.67	1.04	1.31	1.30	2.38
39	218	0.90	1.06	1.37	8.95	71.54	0.72	1.07	1.38	4.50	46.46
40	218	0.89	1.05	1.34	3.87	34.19	0.55	0.99	1.29	1.17	1.96
41	218	0.90	1.04	1.37	4.25	32.35	0.42	0.97	1.29	1.15	2.09
42	218	0.88	1.07	1.36	4.07	32.35	0.39	0.95	1.36	18.45	255.56
43	218	0.87	1.05	1.29	2.54	8.88	0.35	0.91	1.28	1.30	4.37
44	218	0.81	1.03	1.28	2.34	9.98	0.39	0.91	1.23	1.11	2.08
45	218	0.82	1.03	1.29	3.83	32.79	0.38	0.88	1.18	18.20	252.75

Table 13: SP estimates Zero Order (Kernel) local Polynomial Case and Instantaneous Reproduction Number Distribution (cont.)

		Case					Instantaneous				
Week	N	Q1	Median	Q3	Mean	SD	Q1	Median	Q3	Mean	SD
46	218	0.79	0.99	1.30	4.49	34.83	0.35	0.86	1.14	1.11	2.46
47	218	0.76	0.97	1.29	3.32	18.25	0.29	0.87	1.17	0.98	1.39
48	218	0.78	0.99	1.23	3.30	15.26	0.39	0.85	1.17	0.94	1.34
49	218	0.81	1.02	1.27	7.88	52.58	0.44	0.90	1.15	1.14	2.07
50	218	0.83	1.03	1.27	12.73	78.89	0.47	0.94	1.19	4.95	49.37
51	218	0.85	1.06	1.26	12.44	83.43	0.49	0.96	1.23	4.64	46.10
52	218	0.86	1.04	1.29	4.71	36.15	0.51	0.99	1.24	1.40	3.11
53	218	0.84	1.04	1.33	2.15	7.93	0.47	0.98	1.27	1.22	1.78
54	218	0.83	1.06	1.33	2.02	8.17	0.42	1.01	1.30	1.16	1.51
55	218	0.85	1.09	1.34	4.07	25.24	0.49	1.00	1.36	22.07	306.47
56	218	0.82	1.09	1.31	4.54	31.25	0.42	0.96	1.36	6.59	56.89
57	218	0.81	1.01	1.32	7.75	66.16	0.49	0.97	1.36	2.25	14.74
58	218	0.79	1.01	1.23	9.40	70.39	0.43	0.97	1.28	1.87	7.21
59	218	0.78	1.00	1.23	9.74	81.00	0.47	0.95	1.26	4.55	46.05
60	217	0.74	0.99	1.24	8.21	80.98	0.42	0.92	1.24	1.30	3.30
61	217	0.77	0.98	1.27	4.12	24.40	0.45	0.91	1.23	1.30	3.04
62	215	0.76	0.97	1.29	1.82	6.95	0.44	0.90	1.22	18.99	257.33
63	215	0.74	1.00	1.29	1.56	3.74	0.38	0.91	1.15	5.13	55.02
64	215	0.74	0.99	1.24	1.99	9.69	0.42	0.89	1.16	2.89	19.72
65	214	0.74	0.99	1.28	3.63	21.74	0.33	0.88	1.19	2.18	12.22
66	208	0.72	0.97	1.32	5.94	40.01	0.48	0.90	1.19	1.80	9.47
67	196	0.68	0.95	1.30	6.32	39.63	0.42	0.84	1.16	1.78	9.31
68	170	0.70	0.97	1.28	2.49	9.85	0.33	0.81	1.14	1.57	8.99
69	121	0.71	0.94	1.13	1.16	1.50	0.39	0.76	1.15	0.84	0.83
70	88	0.73	0.90	1.13	0.95	0.62	0.38	0.80	1.13	0.83	0.83
71	59	0.61	0.88	1.10	0.85	0.54	0.17	0.79	1.14	0.79	0.75
72	46	0.20	0.77	1.05	0.76	0.60	0.01	0.69	1.15	0.72	0.70
73	45	0.53	0.90	1.24	1.13	1.63	0.00	0.76	1.16	0.93	1.17
74	45	0.63	0.92	1.30	4.46	23.39	0.01	0.78	1.26	2.93	10.98
75	36	0.00	0.85	1.32	5.29	25.94	0.00	0.67	1.29	39.88	221.83
R^2	0.9383										

Table 14: SP estimates First Order local Polynomial Case and Instantaneous Reproduction Number Distribution

		Case					Instantaneous				
Week	N	Q1	Median	Q3	Mean	SD	Q1	Median	Q3	Mean	SD
1	221	1.31	2.00	2.70	2.12	1.19	2.90	5.44	20.89	227.95	735.32
2	221	1.20	1.64	2.25	1.80	1.08	1.45	2.54	5.39	37.60	133.25
3	221	1.03	1.44	1.88	1.67	1.37	1.03	1.63	2.83	8.88	34.37
4	221	0.89	1.33	1.77	2.76	13.20	0.68	1.32	2.02	2.47	7.14
5	221	0.73	1.18	1.66	3.82	25.89	0.60	1.11	1.60	1.66	2.69
6	221	0.74	1.11	1.58	1.94	9.51	0.44	1.00	1.56	1.31	1.54
7	221	0.83	1.14	1.65	4.42	22.18	0.29	0.96	1.47	1.20	1.54
8	221	0.85	1.21	1.63	7.13	42.03	0.49	1.00	1.43	1.17	1.46
9	221	0.93	1.19	1.57	11.61	74.17	0.56	0.99	1.44	5.53	65.16
10	221	0.89	1.11	1.43	11.63	71.10	0.61	1.00	1.49	9.53	116.56
11	221	0.82	1.06	1.39	6.78	35.47	0.60	1.00	1.43	99.93	1454.43
12	221	0.84	1.05	1.37	6.34	28.72	0.60	1.00	1.34	5.32	58.49
13	221	0.87	1.10	1.41	9.24	39.51	0.55	0.99	1.21	1.52	6.40
14	221	0.93	1.12	1.50	12.60	73.14	0.55	0.99	1.25	41.54	592.39
15	221	0.94	1.10	1.60	12.33	77.84	0.68	1.00	1.32	2.69	22.13
16	221	0.97	1.10	1.50	11.17	55.50	0.69	1.00	1.32	4.17	40.13
17	221	0.94	1.08	1.45	12.79	77.70	0.71	1.00	1.34	4.08	35.55
18	221	0.93	1.06	1.45	12.39	62.92	0.70	1.00	1.38	8.53	75.45
19	221	0.95	1.07	1.48	9.01	49.77	0.72	1.00	1.33	13.86	136.28
20	221	0.95	1.08	1.39	7.59	34.01	0.72	1.00	1.36	13.62	131.64
21	221	0.93	1.07	1.46	8.57	37.13	0.71	1.00	1.33	7.11	64.39
22	221	0.91	1.09	1.49	7.79	37.75	0.70	1.00	1.34	4.07	30.65
23	221	0.89	1.07	1.51	9.00	48.00	0.77	1.00	1.47	2.14	10.97
24	221	0.89	1.04	1.45	11.05	80.99	0.74	1.00	1.41	28.88	401.56
25	221	0.91	1.05	1.38	6.02	46.82	0.71	1.00	1.36	2.67	22.84
26	221	0.92	1.05	1.35	3.03	12.34	0.71	1.00	1.37	3.43	32.42
27	221	0.93	1.05	1.37	6.24	27.91	0.78	1.00	1.37	5.85	57.15
28	221	0.92	1.03	1.38	10.56	57.91	0.81	1.00	1.34	1.65	4.44
29	221	0.87	1.03	1.35	13.24	67.23	0.75	1.00	1.31	15.10	200.49
30	221	0.91	1.04	1.29	9.82	47.62	0.70	1.00	1.36	16.69	222.41
31	221	0.90	1.06	1.39	7.25	46.90	0.70	1.00	1.38	6.53	71.71
32	221	0.92	1.09	1.39	5.07	27.73	0.74	1.00	1.39	3.91	32.82
33	220	0.92	1.10	1.40	6.62	49.63	0.71	1.00	1.36	2.35	12.58
34	220	0.94	1.10	1.42	9.11	83.87	0.71	1.00	1.35	1.36	2.35
35	219	0.97	1.11	1.48	6.40	54.15	0.60	1.00	1.29	1.37	2.52
36	219	0.97	1.13	1.45	3.77	19.13	0.68	1.03	1.29	5.08	57.47
37	219	0.94	1.11	1.35	4.50	28.94	0.79	1.03	1.32	6.05	62.95
38	219	0.94	1.09	1.35	9.27	75.28	0.84	1.04	1.33	1.60	5.48
39	218	0.93	1.05	1.34	12.74	132.16	0.83	1.04	1.33	1.27	1.67
40	218	0.92	1.04	1.36	2.99	15.95	0.77	1.00	1.33	1.26	2.65
41	218	0.92	1.04	1.34	4.08	20.45	0.63	0.98	1.30	1.25	3.00
42	218	0.93	1.06	1.34	5.18	32.99	0.66	1.00	1.41	10.36	134.56
43	218	0.92	1.05	1.31	4.38	23.02	0.54	0.98	1.27	1.25	3.03
44	218	0.87	1.03	1.31	2.57	8.91	0.53	0.98	1.21	1.12	1.68
45	218	0.86	1.05	1.30	15.66	197.34	0.56	0.94	1.20	10.26	134.70

Table 15: SP estimates First Order local Polynomial Case and Instantaneous Reproduction Number Distribution (cont.)

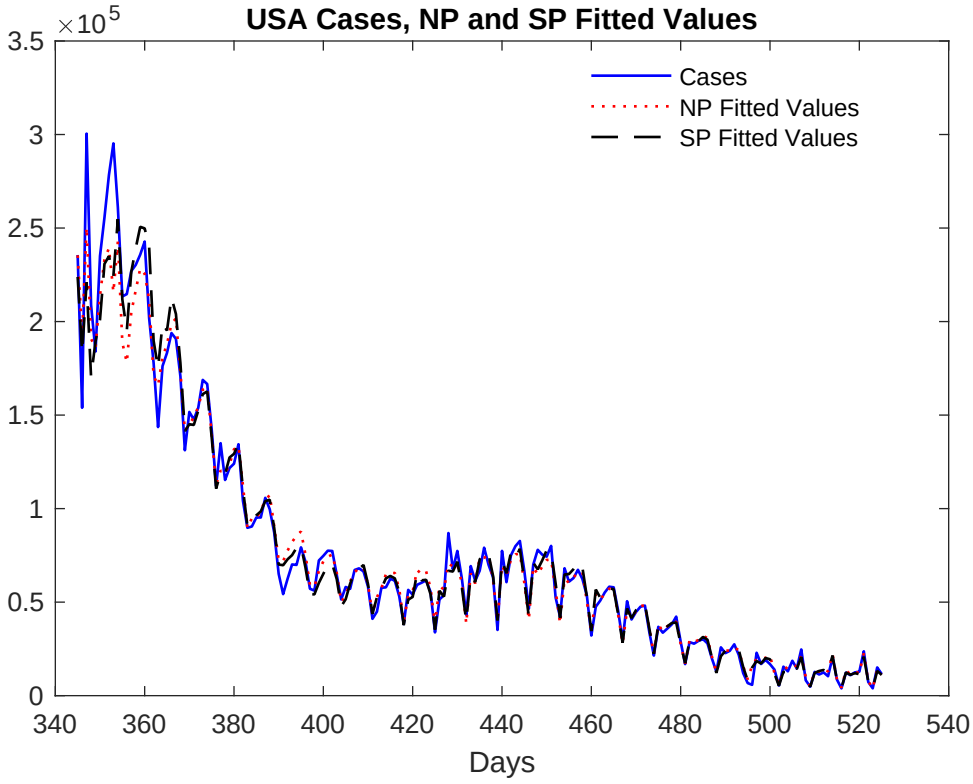
		Case					Instantaneous				
Week	N	Q1	Median	Q3	Mean	SD	Q1	Median	Q3	Mean	SD
46	218	0.87	1.02	1.32	15.79	193.53	0.56	0.94	1.14	1.15	2.04
47	218	0.86	1.00	1.28	5.18	35.00	0.52	0.96	1.17	1.05	1.23
48	218	0.87	1.00	1.23	3.28	13.63	0.60	0.97	1.17	1.02	1.18
49	218	0.86	1.01	1.29	14.28	117.45	0.62	0.97	1.16	1.23	2.27
50	218	0.89	1.04	1.27	24.03	203.40	0.62	0.99	1.17	1.47	4.50
51	218	0.90	1.04	1.26	11.75	76.47	0.73	1.00	1.26	1.53	4.14
52	218	0.91	1.05	1.30	7.84	50.98	0.68	1.00	1.24	1.44	2.96
53	218	0.89	1.04	1.38	12.38	125.02	0.67	1.00	1.28	1.30	2.06
54	218	0.90	1.06	1.35	5.95	42.16	0.57	1.02	1.29	1.24	1.77
55	218	0.91	1.09	1.36	3.67	20.41	0.56	1.00	1.33	5851.92	86382.18
56	218	0.88	1.06	1.31	3.71	16.06	0.62	1.00	1.33	15.00	199.20
57	218	0.89	1.03	1.30	18.43	223.74	0.69	1.00	1.36	4.19	41.29
58	218	0.86	1.01	1.26	73.25	1017.58	0.70	1.00	1.29	2.55	15.11
59	218	0.84	1.01	1.29	159.65	2248.58	0.65	1.00	1.25	1.68	4.79
60	217	0.83	1.00	1.25	243.98	3481.90	0.63	1.00	1.24	1.47	4.26
61	217	0.84	1.00	1.33	297.94	4293.91	0.67	1.00	1.27	1.44	3.90
62	215	0.84	1.00	1.30	1.78	4.06	0.65	1.00	1.22	6.18	67.78
63	215	0.84	1.00	1.26	36.51	512.54	0.65	0.97	1.20	9.63	119.66
64	215	0.84	1.01	1.23	151.86	2199.22	0.62	0.97	1.18	7.96	71.72
65	214	0.81	1.00	1.31	250.36	3633.53	0.64	0.99	1.19	5.68	48.09
66	208	0.84	1.00	1.36	337.95	4812.59	0.64	0.97	1.17	4.41	39.29
67	196	0.82	1.00	1.39	4.82	20.55	0.66	0.95	1.16	3.69	37.01
68	170	0.81	1.00	1.39	4.49	27.31	0.62	0.91	1.13	3.95	39.08
69	121	0.81	1.00	1.25	1.65	4.95	0.57	0.87	1.16	0.94	0.93
70	88	0.83	1.00	1.18	1.15	1.07	0.61	0.89	1.13	0.92	0.76
71	59	0.79	0.96	1.12	1.01	0.52	0.43	0.99	1.14	0.90	0.75
72	46	0.74	0.97	1.03	0.92	0.48	0.61	1.00	1.24	0.87	0.62
73	45	0.82	1.00	1.26	1.92	5.07	0.76	1.00	1.31	1.07	0.92
74	45	0.89	1.00	1.22	6.23	32.64	0.80	1.00	1.44	3.61	11.93
75	36	1.00	1.00	1.37	10.40	50.70	1.00	1.00	1.41	167.37	968.49
R^2	0.94										

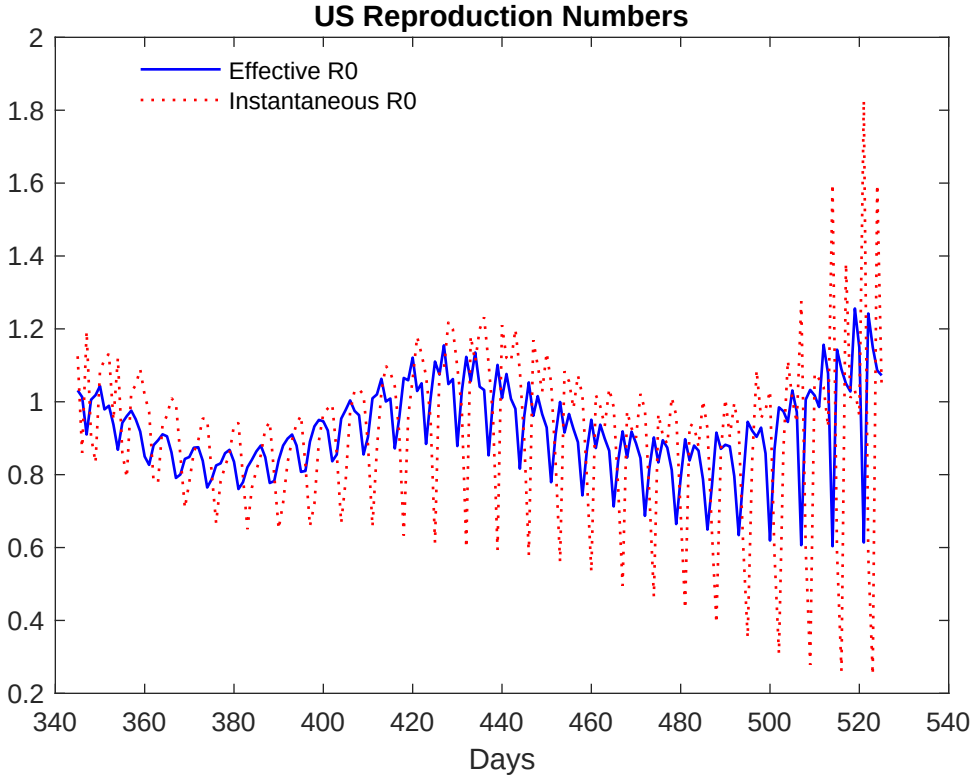
Table 16: SP estimates Second Order local Polynomial Case and Instantaneous Reproduction Number Distribution

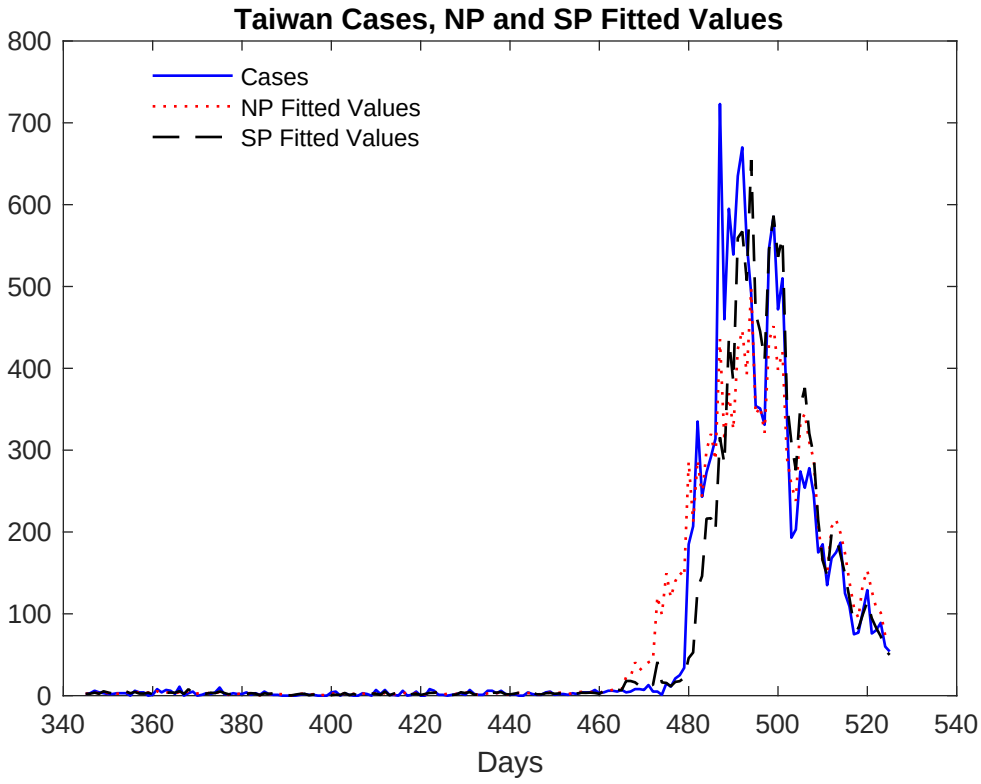
		Case					Instantaneous				
Week	N	Q1	Median	Q3	Mean	SD	Q1	Median	Q3	Mean	SD
1	221	1.50	2.14	2.72	2.14	0.99	3.06	5.36	17.20	55.48	143.26
2	221	1.28	1.64	2.25	1.81	0.99	1.52	2.46	5.02	12.57	36.76
3	221	1.00	1.46	1.87	1.63	1.21	1.02	1.55	2.78	4.07	10.41
4	221	0.88	1.29	1.69	1.99	8.18	0.72	1.34	1.93	2.13	5.36
5	221	0.71	1.14	1.60	2.38	15.83	0.58	1.15	1.63	1.59	2.56
6	221	0.73	1.07	1.56	1.78	8.28	0.45	1.03	1.45	1.27	1.33
7	221	0.81	1.09	1.58	3.67	19.12	0.29	0.97	1.45	1.16	1.33
8	221	0.83	1.14	1.52	4.72	27.39	0.49	0.97	1.43	1.11	1.10
9	221	0.87	1.15	1.42	5.25	37.47	0.58	1.00	1.47	5.42	63.59
10	221	0.89	1.12	1.38	3.91	20.64	0.61	1.00	1.50	1.37	2.30
11	221	0.86	1.09	1.38	2.75	9.65	0.64	1.00	1.43	2.07	8.40
12	221	0.85	1.05	1.35	3.46	13.19	0.61	1.00	1.33	2.37	17.93
13	221	0.87	1.07	1.35	5.94	23.50	0.53	1.00	1.24	1.16	2.11
14	221	0.91	1.07	1.36	7.91	40.61	0.51	1.00	1.28	8.07	102.11
15	221	0.94	1.12	1.47	6.42	26.15	0.65	1.00	1.35	1.62	7.68
16	221	0.96	1.10	1.54	6.04	22.31	0.67	1.00	1.33	1.88	10.29
17	221	0.95	1.10	1.45	6.10	34.99	0.71	1.00	1.35	1.83	9.05
18	221	0.97	1.11	1.44	5.73	27.50	0.68	1.00	1.34	6.11	64.54
19	221	0.98	1.09	1.44	5.11	20.34	0.74	1.00	1.36	9.17	84.18
20	221	0.93	1.10	1.37	3.74	12.29	0.75	1.00	1.36	6.06	56.39
21	221	0.92	1.10	1.42	3.65	12.65	0.69	1.00	1.30	3.06	20.14
22	221	0.91	1.08	1.46	3.61	12.55	0.72	1.00	1.36	2.21	9.97
23	221	0.89	1.07	1.45	4.97	27.41	0.75	1.00	1.41	2.29	15.83
24	221	0.91	1.06	1.34	5.89	36.61	0.71	1.00	1.47	9.95	127.71
25	221	0.92	1.06	1.34	3.64	19.96	0.70	1.00	1.41	1.62	8.51
26	221	0.94	1.05	1.29	2.43	8.49	0.72	1.00	1.36	1.71	8.63
27	221	0.94	1.05	1.33	3.64	12.92	0.77	1.00	1.33	1.31	2.41
28	221	0.93	1.05	1.33	3.81	18.30	0.77	1.00	1.34	1.27	1.62
29	221	0.91	1.06	1.33	3.73	14.60	0.72	1.00	1.38	10.47	136.21
30	221	0.93	1.06	1.33	4.76	25.00	0.69	1.00	1.36	5.12	56.73
31	221	0.93	1.09	1.33	2.94	14.26	0.65	1.00	1.35	2.52	18.25
32	221	0.96	1.10	1.40	3.55	20.38	0.69	1.00	1.38	1.85	8.62
33	220	0.94	1.11	1.45	3.83	26.71	0.72	1.00	1.37	1.47	3.48
34	220	0.96	1.10	1.41	3.81	23.97	0.68	1.00	1.33	1.25	1.44
35	219	0.98	1.11	1.42	2.69	8.51	0.66	1.03	1.33	1.33	2.34
36	219	0.97	1.12	1.41	2.86	13.38	0.66	1.05	1.32	1.30	1.98
37	219	0.96	1.13	1.36	2.89	12.54	0.78	1.06	1.35	1.32	1.80
38	219	0.94	1.11	1.34	4.67	26.43	0.83	1.07	1.35	1.26	1.30
39	218	0.94	1.06	1.33	5.31	44.33	0.79	1.06	1.36	1.30	1.64
40	218	0.93	1.06	1.36	2.12	9.47	0.70	1.00	1.30	1.23	1.99
41	218	0.94	1.06	1.34	2.62	12.76	0.66	1.00	1.28	1.15	1.72
42	218	0.95	1.06	1.32	2.78	15.74	0.65	1.00	1.39	8.59	110.32
43	218	0.93	1.05	1.31	2.37	8.36	0.55	0.97	1.27	1.47	6.40
44	218	0.88	1.04	1.28	2.34	8.42	0.55	0.97	1.27	1.19	2.60
45	218	0.86	1.04	1.29	3.99	25.58	0.56	0.95	1.21	8.36	107.23

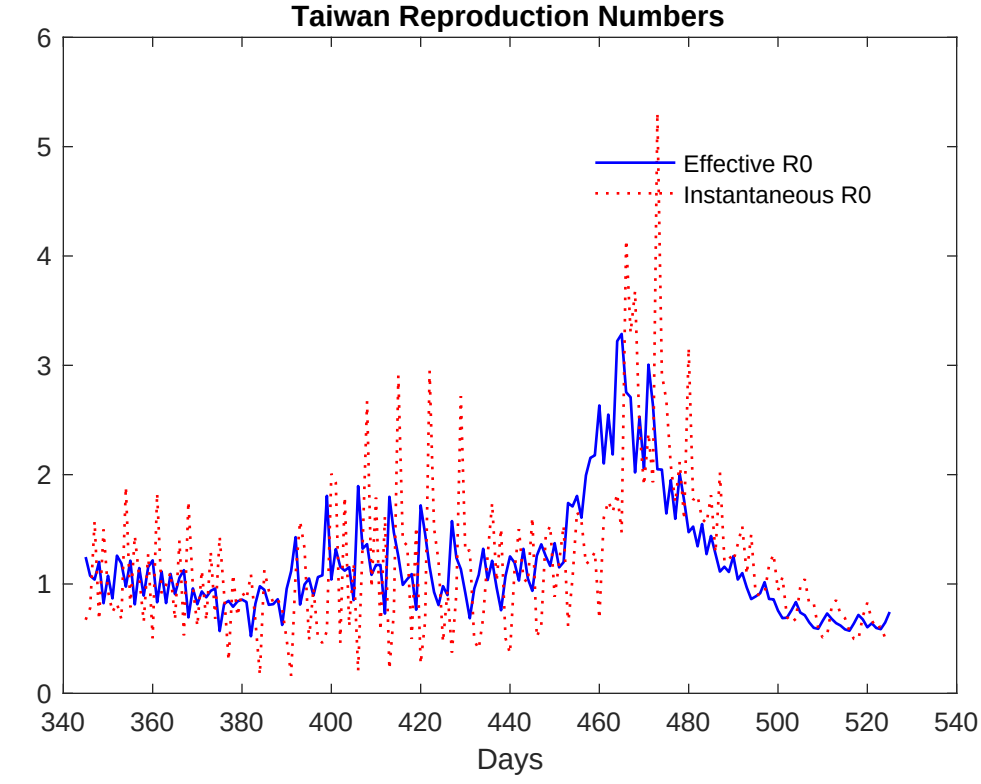
Table 17: SP estimates Second Order local Polynomial Case and Instantaneous Reproduction Number Distribution (cont.)

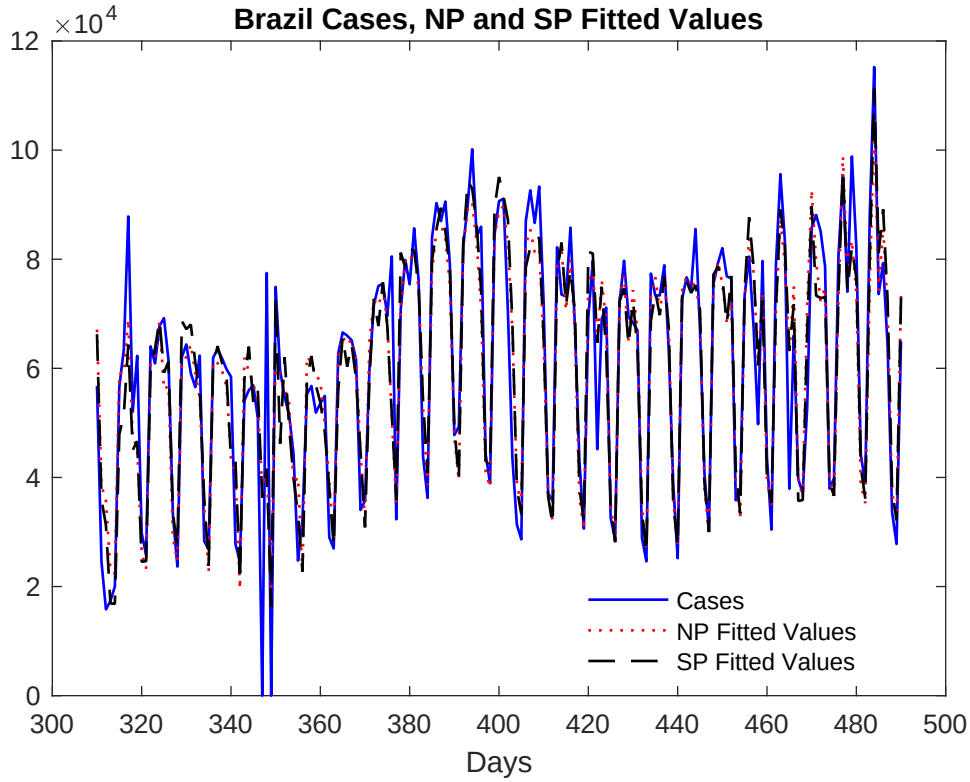
		Case					Instantaneous				
Week	N	Q1	Median	Q3	Mean	SD	Q1	Median	Q3	Mean	SD
46	218	0.87	1.01	1.26	3.58	26.19	0.53	0.95	1.17	1.06	1.54
47	218	0.87	1.00	1.33	2.83	12.83	0.52	0.95	1.19	1.04	1.45
48	218	0.86	1.00	1.23	2.91	11.19	0.58	0.98	1.17	1.03	1.54
49	218	0.88	1.00	1.27	4.74	25.93	0.65	0.98	1.16	1.21	2.04
50	218	0.89	1.03	1.29	6.22	35.31	0.68	1.00	1.25	1.66	5.91
51	218	0.90	1.04	1.27	4.72	22.83	0.75	1.00	1.28	1.56	4.14
52	218	0.91	1.06	1.28	2.86	12.00	0.66	1.00	1.24	1.53	4.03
53	218	0.90	1.06	1.33	2.23	7.46	0.69	1.00	1.32	1.27	1.82
54	218	0.90	1.06	1.30	1.95	5.88	0.61	1.01	1.33	1.19	1.43
55	218	0.94	1.09	1.32	2.42	8.86	0.57	1.01	1.37	22.27	308.50
56	218	0.92	1.08	1.31	3.64	17.60	0.62	1.00	1.34	4.88	52.68
57	218	0.90	1.03	1.29	5.05	35.82	0.64	1.00	1.35	2.27	15.77
58	218	0.87	1.02	1.25	5.62	39.33	0.67	1.00	1.32	1.75	6.88
59	218	0.84	1.01	1.19	5.74	39.46	0.63	1.00	1.30	1.41	2.87
60	217	0.83	1.00	1.24	4.74	42.02	0.60	1.00	1.24	1.24	2.06
61	217	0.85	1.00	1.26	3.77	19.54	0.61	0.98	1.20	1.28	2.10
62	215	0.86	1.00	1.24	3.16	20.42	0.62	0.95	1.22	6.03	65.90
63	215	0.83	1.00	1.25	2.35	9.30	0.64	0.97	1.21	5.19	54.33
64	215	0.83	1.00	1.28	2.09	7.09	0.59	0.98	1.20	3.05	21.10
65	214	0.83	1.00	1.29	2.05	6.48	0.64	0.99	1.20	2.29	12.91
66	208	0.86	1.00	1.30	3.44	15.07	0.63	1.00	1.20	1.91	10.01
67	196	0.82	1.00	1.29	3.62	12.07	0.63	0.96	1.16	1.63	9.06
68	170	0.80	1.00	1.23	3.30	14.61	0.60	0.92	1.13	1.61	9.14
69	121	0.81	1.00	1.15	2.53	10.76	0.59	0.89	1.14	0.93	0.78
70	88	0.85	1.00	1.22	1.54	4.00	0.60	0.91	1.14	0.91	0.59
71	59	0.79	0.98	1.08	1.32	2.20	0.45	0.98	1.15	0.89	0.68
72	46	0.77	0.99	1.11	1.10	0.82	0.38	1.00	1.15	0.87	0.63
73	45	0.87	1.00	1.17	1.65	3.62	0.80	1.00	1.29	1.13	1.17
74	45	0.94	1.00	1.36	6.65	36.57	0.83	1.00	1.39	3.17	12.82
75	36	0.98	1.00	1.58	12.72	68.93	1.00	1.00	1.47	27.63	154.54
R^2	0.94										

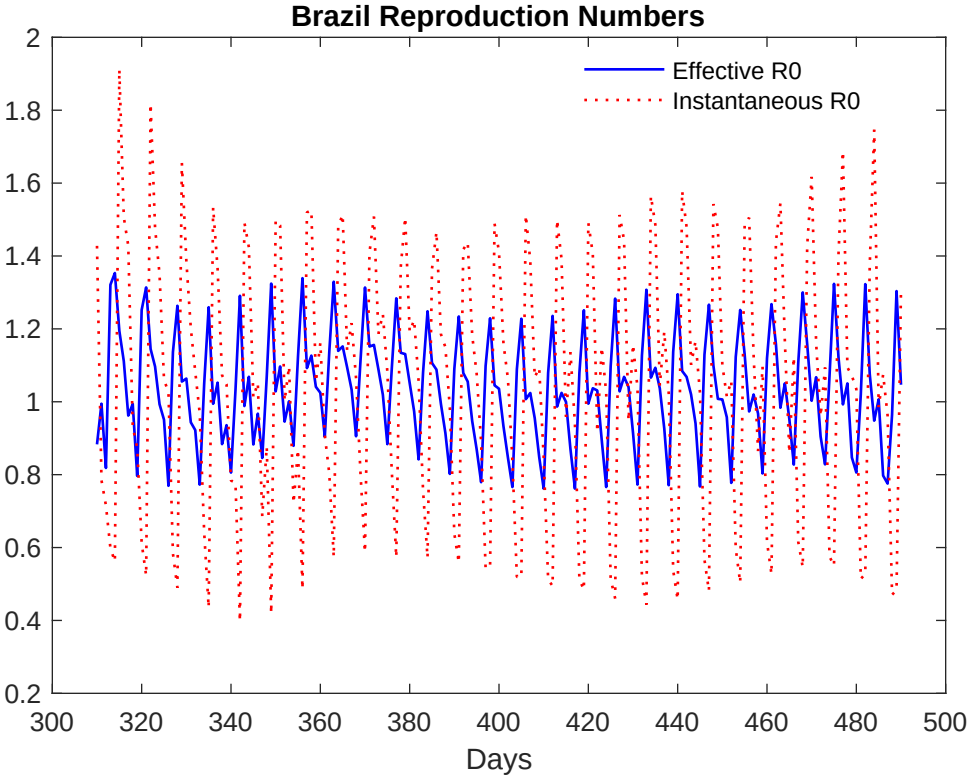

USA Cases, NP and SP Fitted Values
×10^5
3.5
3
2.5
2
1.5
1
0.5
0
Cases
NP Fitted Values
SP Fitted Values
340
360
380
400
420
440
460
480
500
520
540
Days

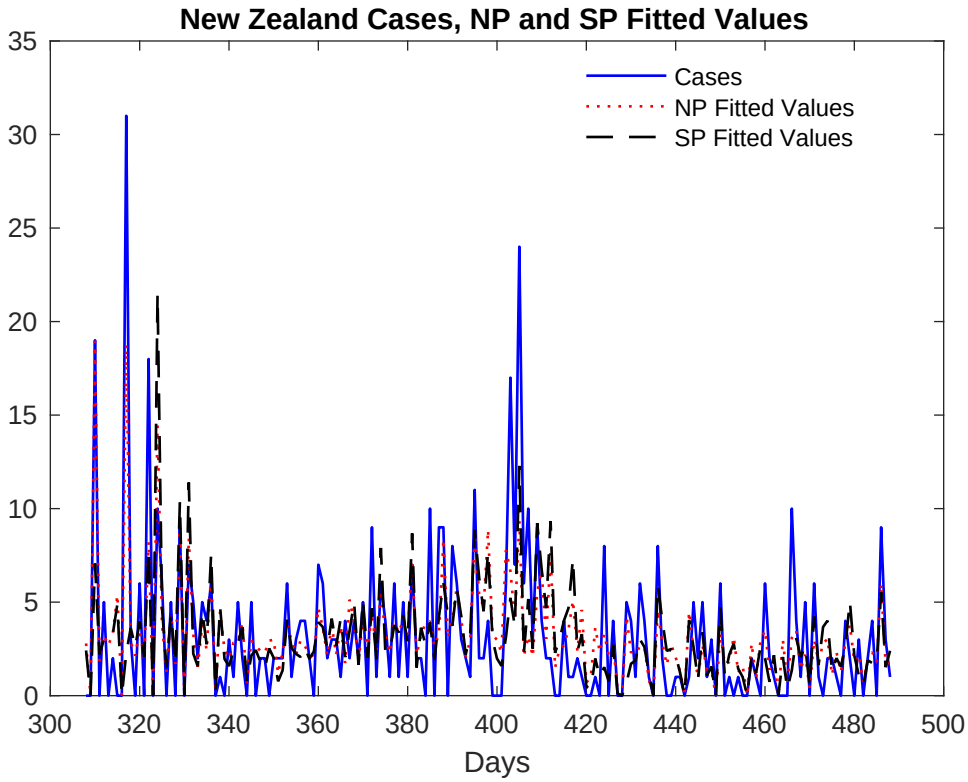

US Reproduction Numbers
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
Effective R0
Instantaneous R0
340
360
380
400
420
440
460
480
500
520
540
Days

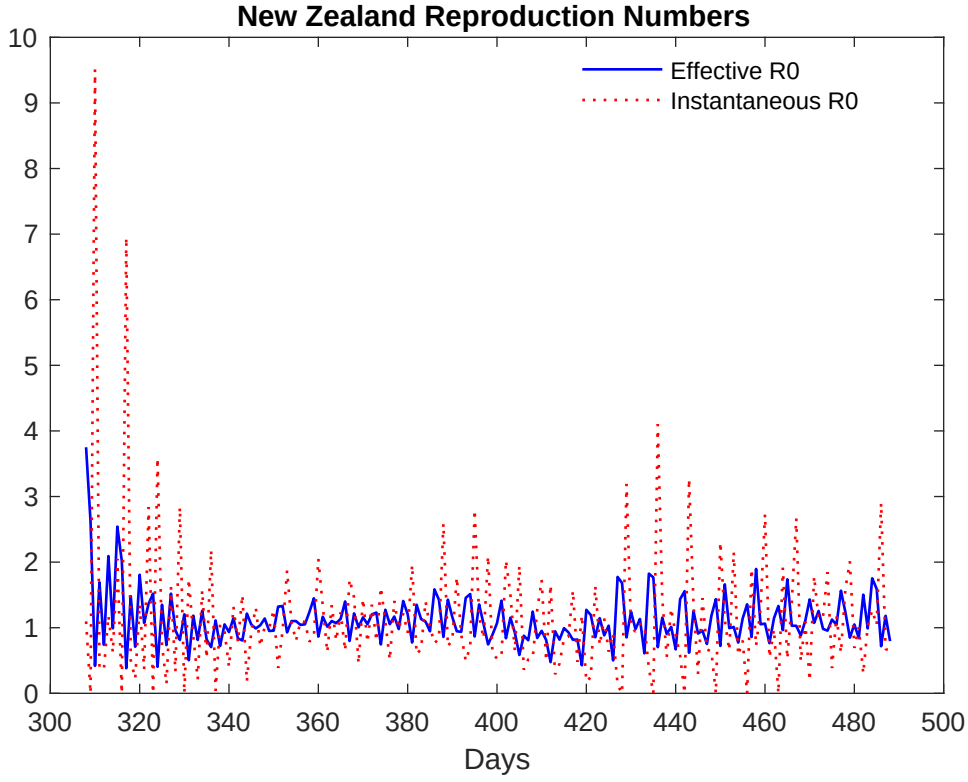

Taiwan Cases, NP and SP Fitted Values
Cases
NP Fitted Values
SP Fitted Values
800
700
600
500
400
300
200
100
0
340
360
380
400
420
440
460
480
500
520
540
Days

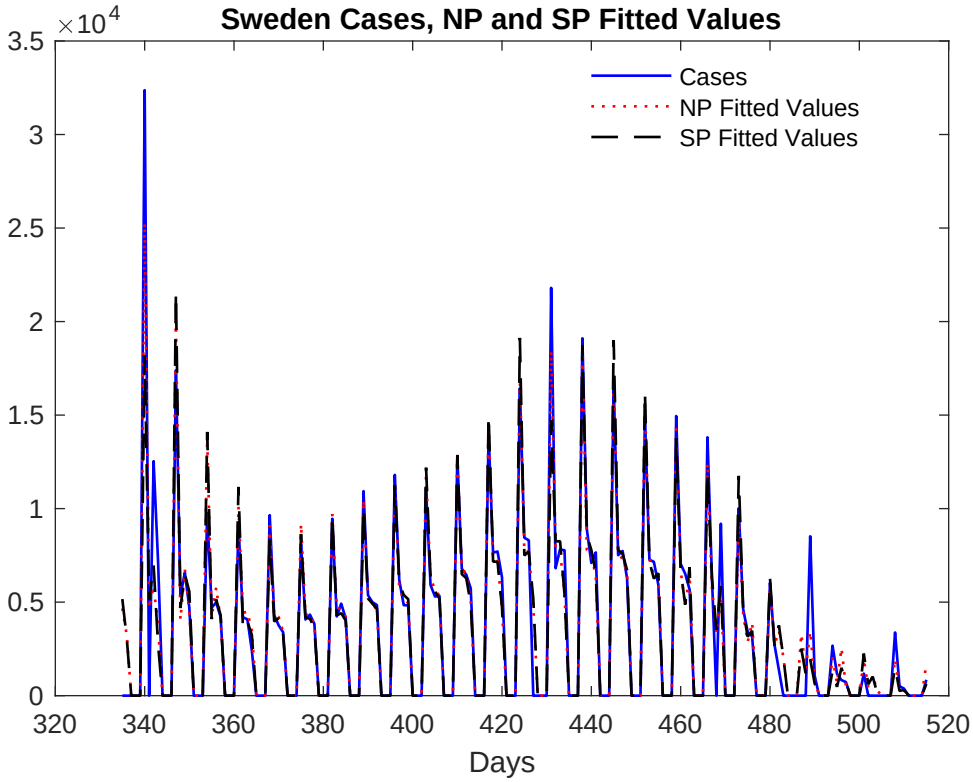

Taiwan Reproduction Numbers
Effective R0
Instantaneous R0
6
5
4
3
2
1
0
340
360
380
400
420
440
460
480
500
520
540
Days

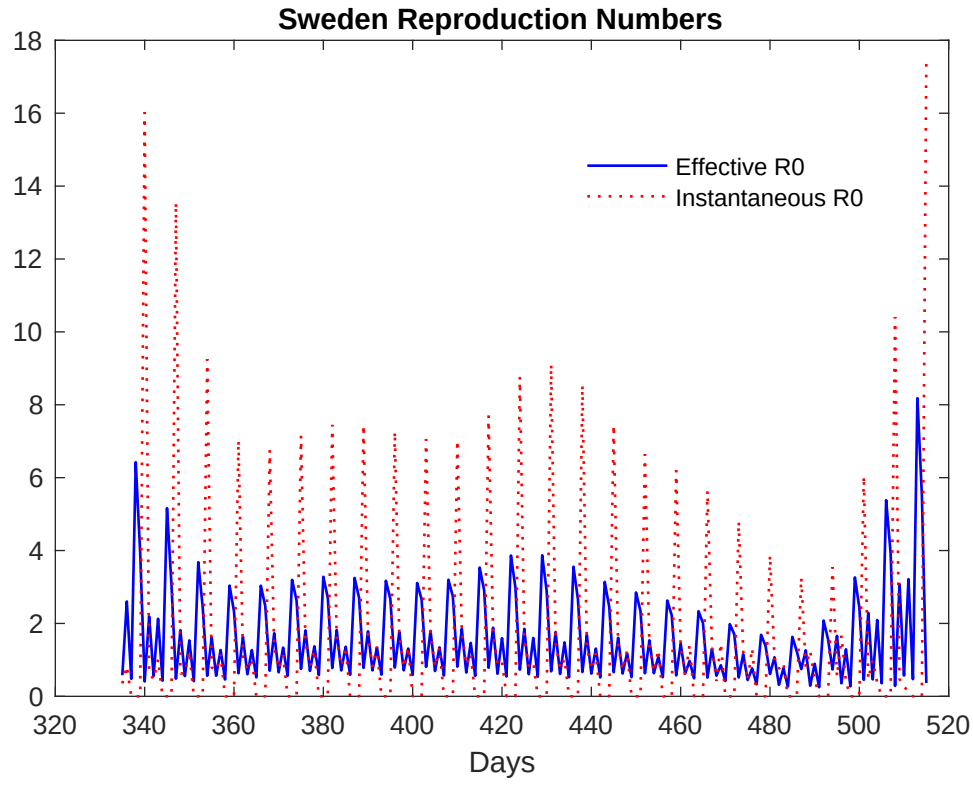

Brazil Cases, NP and SP Fitted Values
×10^4
12
10
8
6
4
2
0
300
320
340
360
380
400
420
440
460
480
500
Days
Cases
NP Fitted Values
SP Fitted Values


Brazil Reproduction Numbers
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
300
320
340
360
380
400
420
440
460
480
500
Days
Effective R0
Instantaneous R0


New Zealand Cases, NP and SP Fitted Values
Cases
NP Fitted Values
SP Fitted Values
35
30
25
20
15
10
5
0
300
320
340
360
380
400
420
440
460
480
500
Days


New Zealand Reproduction Numbers
Effective R0
Instantaneous R0
10
9
8
7
6
5
4
3
2
1
0
300
320
340
360
380
400
420
440
460
480
500
Days


Sweden Cases, NP and SP Fitted Values
×10^4
3.5
3
2.5
2
1.5
1
0.5
0
Cases
NP Fitted Values
SP Fitted Values
320
340
360
380
400
420
440
460
480
500
520
Days


Sweden Reproduction Numbers
18
16
14
12
10
8
6
4
2
0
Effective R0
Instantaneous R0
320
340
360
380
400
420
440
460
480
500
520
Days