Regensburg, den 31. Januar 2022

UR

Bachelorarbeit

zur Erlangung des akademischen Grades

Bachelor of Science

an der Fakultät für Wirtschaftswissenschaften

der Universität Regensburg

Nonlinear Unobserved Components Models for the COVID-19 Infection Rates

Eingereicht bei: Prof. Dr. Rolf Tschernig

Lehrstuhl für Ökonometrie

Betreuer: Prof. Dr. Rolf Tschernig, Tobias Hartl

Vorgelegt von: Paul Haimerl

Matrikelnummer 2099707

Schleienweg 6a

84032 Altdorf

E-Mail: paul.haimerl@gmx.de

II

Abstract.   The model developed in this paper estimates a trend signal of log COVID-19 infections. Employing an unobserved components (UC) framework with parameters following a Markov-switching (MS) process, I set up a nonlinear approach to model infections in real time and to provide medium term forecasts. Due to its regime switching characteristic, the technique reflects recurring peaks and troughs of the pandemic. This enables a long memory component to assume the underlying trend whilst still closely matching the data generating process. Inherent distributional shifts and unit root of the dependent variable can be dealt with. Forecasting performance is evaluated via a rigorous ongoing out-of-time test routine.

Contents

List of Tables

List of Figures

Symbols and abbreviations

ACF   Auto Correlation Function
ADF   Augmented Dickey Fuller test
AR   Autoregressive
ARIMA   Autoregressive integrated moving average
GNP   Gross National Product
KF   Kim filter
MAPE   Mean absolute percentage error
MedAPE   Median absolute percentage error
ML   Maximum likelihood
MS   Markov switching
NPI   Non-pharmaceutical intervention
RW   Random walk
SE   Standard error
UC   Unobserved Components

1 Introduction

As of time of writing, the coronavirus disease 2019 (COVID-19) pandemic has been tremendously influential on next to all facets of life. Ever since its onset in December 2019, a plethora of research stemming from various disciplines has accompanied the pandemic´s evolution (Telenti et al., 2021).

Robust predictions of COVID-19 dynamics have proven instrumental in the design of effective policy solutions. Moreover, underestimation of explosive short term growth and subsequent misspecified regulations pave the way to the alternation of infection waves and burdensome containment measures (Coccia, 2021). Hence, a comprehensive battery of quantitative models has been proposed to describe and forecast the ongoing pandemic.

Kraemer et al. (2020) and Yilmazkuday (2021) exploit meta mobility data provided by Big Tech to explain the spread of COVID-19 (Oliver et al., 2020). Qiu, Chen, and Shi (2020) use weather patterns as an instrument to estimate the effectiveness of public health interventions.

Furthermore, researchers have employed techniques from the econometric time series tool kit to predict COVID-19 infections. Gecili, Ziady, and Szczesniak (2021) use, among other, a traditional Autoregressive Integrated Moving Average (ARIMA) framework. Doornik, Castle, and Hendry (2020) employ a local linear trend saturation method to provide concise short term forecasts. Harvey (2021) fits logistic growth curves to daily changes of the logarithmic infection rate.

It has been widely appreciated that forecasting the pandemic´s course is littered with difficulties (Doornik, Castle, and Hendry, 2020). Two major sources of hazard are one the one hand the evolution of COVID-19 itself and on the other the respective public response (Telenti et al., 2021). Mutations are changing many characteristics of the virus on an ongoing basis. These aspects concern the duration of illness, viral load, mortality, severeness of infection, contagiousness and vaccine resistance. As mutations occur randomly and their respective dominance over other COVID-19 strains is again dependent on the interplay of various factors, accounting for such scenarios is not feasible (Moore et al., 2021). Even more crucial are the regulatory actions as well as the public´s level of

1

compliance. Containment measures like lockdowns, contact restrictions, curfews or vaccination schemes can severely dampen the spread of the disease. However, behavioural changes, inconsistent policy application, incoherent data tracking and varying vaccination rates fundamentally change the data generating process (Dolton, 2021; Hortaçsu, Liu, and Schwieg, 2021). Furthermore, single decisive events can persistently boost infection numbers and past policies may govern the effectiveness of future, seemingly unrelated, actions (Harvey and Kattuman, 2020). As pointed out by Castle, Doornik, and Hendry (2021), every aspect of the COVID-19 infections´ distribution changes over time.

Consequently, modeling the pandemic necessitates highly flexible techniques that are equipped to mirror nonstationary data, vastly changing environments and exponential growth patterns.

The contribution of this paper is to propose a nonlinear Unobserved Components (UC) model to explain log COVID-19 infections as well as to provide medium term forecasts. The preferred UC approach consists of a random walk (RW) with drift in combination with a stochastic seven day seasonal component. The RW term serves as the parameter tracking the underlying trend in COVID-19 infections. The seasonal component is to control for varying test quantities, laboratory capacity, changing contact patterns as well as measurement noise (Ricon-Becker et al., 2020).

In order to accommodate the highly variant characteristics of the pandemic, I introduce two discrete regimes. Both the innovation variance and drift of the trend component are allowed to vary based on the current state. The regimes follow an endogenous first order Markov switching (MS) process. This specification is able to cover both infection up and downturns as the distribution of estimates changes depending on the regime. Benefits over a linear approach are among other a more precise estimation of the data generating process, improved forecasting performance and inference concerning regime switches themselves. Effective policies might be able to tip the system into an infections decreasing regime, whereas misguided efforts have no effect or even cause the opposite.

Kim (1994) extended the Kalman filter in order to incorporate the

switching mechanism proposed in Hamilton (1989)[1]. The resulting Kim Filter (KF) will be employed in the following and allows for feasible estimation of switching UC models, the MS regime probabilities and potential correlation between all innovations. Unknown system parameters are estimated via Maximum Likelihood (ML).

To the best of my knowledge, this paper is the first to model endogenous MS with both regime dependent and in parallel correlated innovations using the Kim filter[2].

The forecasting performance is evaluated via an ongoing out-of-sample test routine. I estimate and project the model 30 days into the future every seven days throughout the data history. The resulting forecasts cover a wide range of scenarios and can be easily validated against actual realizations of COVID-19 infections.

Based on US data provided by the Johns Hopkins University Center for Systems Science and Engineering (JH/CSSE)[3] I find that despite the extensive flexibility inherent to the employed models, some pitfalls of the data generating process can not be overcome. Especially isolated and erratic amplifications of seasonality are absorbed by the trend component and trick the model into regime switches. As a consequence, regime dynamics are modelled as sudden increases in volatility rather than coherent episodes of increasing or decreasing infections.

Furthermore, several Monte Carlo simulation studies show the difficulty and potential treachery associated with estimating complex nonlinear UC models. Specifications that vary profoundly from the true parameters are still able to arrive at good in-sample fits. I also show that nonlinear UC models require a rich data history for robust inference.

As is to be expected, forecasting performance decreases significantly after a short period, but is still able to project a general tendency for up to 20 days.

---

[1]MS in an UC context was previously introduced by Lam (1990) and Shumway and Stoffer (1991), using other specifications. See Kim (1994) for a more exhaustive background.

[2]Klinger and Weber (2016) also employ endogenous switching with regime dependent innovations. However, the intricate endogenous dependencies and estimation technique is not laid out further.

[3]See github.com/CSSEGISandData/COVID-19 (Dong, Du, and Gardner, 2020).

The plan of the paper is as follows. Section 2 offers a brief theoretical introduction to the UC modelling framework, MS processes, Kim filtering and forecast evaluation. In section 3, I discusses the data and set up the models accordingly. Section 4 compromises a Monte-Carlo simulation study to highlight potential pitfalls of misspecification and practical application. In section 5, I present the estimation results and forecasts. The final section concludes.

2 Modeling and forecasting methodology

2.1 Linear Unobserved Components model and the Kalman filter

In its basic form, the UC model can be thought of as an additive decomposition of a time series into multiple unobserveable parts: e.g. GNP into a trend, cyclical and irregular component (see e.g. Clark (1987) or Harvey (1990)). Also called the structural time series model, this class of models is principally not unlike a multiple regression model with time varying coefficients.

UC model estimation can be performed both from a frequentist or Bayesian stance. However, both techniques are based on the Kalman recursion algorithm and give the same results in most scenarios (Durbin and Koopman, 2012, ch. 13). Due to spatial constraints, I will focus on the frequentist perspective for the remainder of the paper.

The Kalman filter was originally developed for engineering purposes in order to derive optimal inference of noisy measurements in linear systems (Kalman, 1960). However, because of its flexibility and multi-purpose application it was quickly adopted for use in other quantitative fields, like for UC estimation. The Kalman filter accepts a model in state-space form as input and iteratively adjusts estimates of an unobserveable parameter vector throughout the observational period.

A model cast into state-space form is described by a system of two equations. One, the measurement equation, provides the interaction

of an observed noisy measurement and the state vector, which houses the unobserved parameters. The second, termed transition equation, reflects a constructed model for the state vector without regard to the measurements[4]. For the conventional Kalman filter, both measurement and transition equation may be subject to Gaussian innovations.

The interplay of measurement and transition equation constitutes the foundation of the Kalman filter. Somewhat similar to Bayesian updating, one step ahead predictions for the unobserved parameters are adjusted every period according to incoming observations of the measurement. These one-step-ahead predictions are provided by the transition equation, which gives the expected value of the state vector in $t + 1$, conditional on information up to t, where $t = \{1, 2, ..., T\}$. This prediction is subsequently adjusted by an updating step: the expectation for $t + 1$, conditional on the entire information set in $t + 1$ (see Durbin and Koopman (2012, eq. 4.24), equations B.1). Naturally, updating steps contain the current noisy measurement.

This adjustment step is defined by adding a scaled one-step-ahead prediction error to the predicted state vector. The prediction error constitutes the difference between actual noisy measurements and its expected value via the measurement equation and predicted state vector. The respective scaling factor is determined by balancing variance associated with the predicted state vector against variance of shocks to the measurement equation. The higher the variance of the predicted state vector, the bigger the scaling factor. Is on the other hand the variance of the predicted state vector small, relative to the measurement equation innovations, the Kalman filter leans less on the prediction error to updated the state vector, i.e. leaving the prediction step largely unchanged. In case the measurement equation does not incorporate an irregular component, the scaling factor is solely determined by state vector variance matrix. This matrix is subject to the same prediction and updating principle as the state vector (see equations B.1). Fundamentally, it is a function of innovation variances of both measurement and transition equation (Durbin and Koopman, 2012, ch. 4 and 5). In order to start the recursions, either

---

[4]I refer to (Hamilton, 1989, ch. 13.1) for a more comprehensive introduction into state-space form.

a best guess or non-informative values associated with high variances can be employed (Durbin and Koopman, 2012, ch. 5).

Filtering only takes advantage of time periods up to and including the current one. In contrast, a similar set of backwards smoothing recursions allow for inference based on the entire observational period. Forecasting the Kalman filter is straightforward, as one can simply iterate the filter forwards, without making use of incoming measurements (Durbin and Koopman, 2012, ch. 4.11).

The concept is easily exploited for time series analysis. The structural UC model constitutes the measurement equation, where each non-irregular unobserved component gives one entry to the state vector. The transition equation explains how the state vector behaves over time as a vector autoregressive (AR) function. The structure of seasonal patterns, cyclical behaviour or trend characteristics inherent to the UC model are reflected within the transition equation and uniquely define each component in the state vector. Subsequently, every entry of the state vector also experiences individual shocks[5]. As shown by Morley, Nelson, and Zivot (2003), shocks to individual components may also be cross-correlated.

System matrices and additional parameters like innovation variances are either known, assumed or have to be estimated. In a frequentist approach, the Kalman filter gives a likelihood function which is maximized with respect to θ, where θ is a vector holding all unknown additional parameters. The likelihood reflects the joint density of each measurement in t, given all information up to $t - 1$. Prediction errors of the Kalman filter are exploited in order to arrive at these densities (Durbin and Koopman 2012, ch. 7; Harvey 1990, ch. 4).

For a deeper understanding of strucutral time series models I recommend Durbin and Koopman (2012), Hamilton (1989, ch. 13) and Harvey (1990).

2.2 Markov-switching models

In his influential article, Hamilton (1989) established the MS technique. As an original use, it is employed to model first differences of log GNP as

---

[5]Depending on the UC structure, some unobserved components may have innovations with zero mean and/or zero variance.

a nonlinear stationary series, opposing the popular specification of a linear stationary series. Building upon work by Goldfeld and Quandt (1973), nonlinearity is introduced by allowing for discrete unobservable regime shifts[6]. Those regime shifts materialize through a dummy specification of a regime dependent indicator. The Hamilton filter developed in Hamilton (1989) provides estimates regarding the individual regime realization at t[7]. Somewhat similar to the Kalman filter, it also enables inference based on the entire data history via a smoothing method (Hamilton, 1989).

The regimes themselves are assumed to follow a stationary Markov chain. The current regime realization is expressed as a latent random variable S_t, which depends on k of its lags. k denotes the order of the Markov process. S_t can take on M different values, where M defines a positive integer of possible regimes $\{1, 2, .., M\}$. A transition matrix governs the probability of each possible regime, conditional on each regime for the previous period. In most applications, k is set to 1 and M to 2. This entails a process where the current realization merely depends on the previous period, with a total of two possible regimes each (Hamilton, 1994, ch. 22.4). Introducing only a very limited number of regimes and relevant lags keeps the model easy to interpret and is sufficient for most purposes.

When using the the Hamilton filter, regime switching is incorporated into an overarching model. The recursions for obtaining filtered regime probabilities follow the prediction and updating principle of the Kalman filter. Thus it is important to understand, that whereas the initially predicted probability of a certain regime switch depends solely on the Markov process, the final filtered probability is further influenced by the previous realization of the model´s dependent variable.

The prediction involves the joint probability of observing a certain S_{t+1} and S_t combination, conditional on all information at t. This joint probability is defined by the transition matrix and the marginal probability of each M S_t realization. The marginal probability is provided

---

[6]Time series literature uses the terms regime and state interchangeably. To mitigate confusion with the state expression in an UC context, hereafter I will only use the term regime.

[7]The Hamilton filter referred to here is not to be confused with the filtering technique proposed in Hamilton (2018), which is also termed as Hamilton filter.

by the previous period´s updating step. The new updating step, again, incorporates the incoming observation of the dependent variable y_{t+1}. The updated joint probabilities for observing a certain S_{t+1} and S_t combination are computed by comparing the joint density of S_{t+1}, S_t and y_{t+1} to the marginal density of y_{t+1}, all conditional on the information set up to t. To close the loop and derive the marginal probability of S_{t+1}, one just has to integrate out all possible realizations for S_t (see eq. B.2). The marginal density of y_t can be further exploited to compute a sample likelihood score. Additional parameters like the entries to the transition matrix can be derived via ML (Hamilton, 1989, ch. 4.2, Steps 1 - 5).

This filtering technique is also conceptually similar to the Kalman filter in that the Markov Process represents the transition equation and the main model mirrors the measurement equation. However, instead of balancing innovation variances to obtain the updating step, the Hamilton filter evaluates the density of y_{t+1} conditional on certain S_t and predicted S_{t+1} realizations with respect to the marginal density of y_{t+1}[8]. The higher the density conditional on a particular set of S_{t+1} and S_t, the higher the updated probability value for exactly this S_{t+1} and S_t combination.

As pointed out, the predicted probability of a certain regime depends on k of its lags. As a consequence, innovations to the dependent variable merely impact the Markov process sequentially (Kang, 2014). The MS prediction step is effectively an exogenous behaviour, as regime switches only build upon the previous regime. This characteristic does not always reflect the underlying data generating process. In many situations y_t and S_t may be contemporaneously correlated. Treating endogenous switching as if it was exogenous leads to significant distortions and is furthermore not efficient (Chang, Choi, and Park, 2017).

Kim, Piger, and Startz (2008) extended the Hamilton filter to incorporate endogenous switching. Instead of assuming a Markov chain where S_{t+1} is only a function of S_t, they include dependency on a vector of other explanatory variables. This approach is executed via a Probit specification, where a latent variable is regressed on a constant, a vec-

---

[8]Becomes apparent when rewriting the joint density of S_{t+1}, S_t and y_{t+1} as the density of y_{t+1} conditional on S_t and S_{t+1}, times the joint probability of S_t and S_{t+1} (see last equation in B.2)

tor of exogenous variables and S_t. They further introduce correlation between innovations in S_{t+1} and y_{t+1} (see Kim, Piger, and Startz (2008, eq. 2.1 - 2.5)). The Hamilton filter necessitates only minor adjustment to accommodate these alterations (Kim, Piger, and Startz, 2008, eq. 2.6 - 2.10)[9]. An issue associated with contemporaneous endogenous switching is that ML becomes biased and inconsistent. However, as shown by Kim, Piger, and Startz (2008), the quasi-ML estimator still retains fairly precise performance. This aspect is further addressed theoretically in section 3.3 and in section 4 in practice.

For a substantial introduction into MS approaches, I refer to Hamilton (1989), Hamilton (1994, ch. 22) and Kim, Piger, and Startz (2008).

2.3 Nonlinear Unobserved Components Models and the Kim filter

Combining the linear UC model and MS mechanism opens up a class of highly flexible approaches. UC parameters or even system matrices are allowed to vary based on the underlying regime. Kim (1994) introduced this hybrid model which is estimated via the KF. The KF effectively inserts the Hamilton recursions in between prediction and updating step of the Kalman filter, thus creating a coherent framework for regime switching state space models.

The KF executes the Kalman filter for every possible combination of S_t and its relevant k lags at every iteration. However, even in the common case of a first order Markov chain, this introduces a $O(M^T)$ time complexity. The filter has to compute a minimum of roughly one million branches just after 20 iterations. Computational infeasibility arises since every period the filter produces M updated values of the state vector and again M predictions based on every single update. As a solution, the KF uses approximations to reduce M^2 updated state vectors into only M vectors at every iteration. This results in a manageable $O(MT)$ time complexity, at the cost of some efficiency. The approximation step employs

---

[9]Chang, Choi, and Park (2017) propose a similar specification. However, they incorporate endogeneity via sequential correlation as laid out in section 3.3.

inference regarding regime probabilities to create a single weighted average of all M updates, which are constructed from the same regime in the previous iteration. Consequently, for inference in t the dependency on the regime in $t-1$ is let go to break the exponential growth of computational demand (see Kim (1994, eq. 2.6 - 2.12; Steps 1 - 4), equations B.3). Furthermore, it is possible to again average all M estimates of the state vector to arrive at a single concise filtered value (Kim, Nelson et al., 1999, ch. 5).

Unknown parameters are derived by ML with respect to θ. As in the Kalman filter, the likelihood gives the density of y_t conditional on the information set in $t-1$. An aspect worth noting is that in the presented methodology, the KF does not reflect uncertainty of parameters estimated via ML. When using Bayesian techniques like Gibbs sampling one the other hand, inference depends on the joint distribution of input parameters as well as filtered state vectors and regime probabilities. For a discussion on competing Bayesian estimation techniques for the KF see Albert and Chib (1993) and Kim, Nelson et al. (1999, ch. 7).

For a technical representation of the Kim filter, smoother and forecasting recursions, see appendix B.1. The Kalman filter and Hamilton filter recursion are nested within the Kim filter.

2.4 Forecast methodology and evaluation

As I have already touched upon, forecasting the Kalman filter is rather trivial. Instead of balancing variances from the measurement and transition equations, one simply bases prediction steps on previous predictions (Durbin and Koopman, 2012, ch. 4.11). This mimics a situation where the state vector stops to influence the dependent variable right at the beginning of the forecasting period. Consequently, the variances associated with the forecasted state vector increase with every iteration.

Inferring predictions from the Hamilton filter is also straightforward. In the absence of realizations of the dependent variable, regime probabilities evolve according to the underlying Markov chain or Probit process. Subsequently, regime probabilities converge to their respective steady states

as the forecasting horizon increases Hamilton (1989, 1994)[10].

I evaluate the forecast performance via an ongoing out-of-sample routine. Following a brief data build up phase of 28 days after cumulative infections have surpassed 100, I run the preferred model and produce a 30 day forecast every week throughout the data history (Gecili, Ziady, and Szczesniak, 2021). Naturally, I will only include data right up to the beginning of the respective forecasted periods. Validation of forecasts is simple via the mean absolute percentage error (MAPE) and median absolute percentage error (MedAPE) (Doornik, Castle, and Hendry, 2021). As the observational period spans the entire development of the pandemic, one can easily observe and judge forecasting behaviour in a multitude of phases and distributional environments.

3 Fitting the COVID-19 infections

Data provided by JH/CSSE gives a daily count of confirmed COVID-19 infections i_t for a comprehensive list of regions. Here only the US is considered.

Choosing a sensible transformation of daily infections is not trivial. Doornik, Castle, and Hendry (2021) use the log of cumulative cases I_t. Harvey (2021) analyzes logarithmic first differences of I_t, whereas Liu, Moon, and Schorfheide (2021) examine first differences of log i_t. A specification in logs prevails in current literature. As case numbers tend to increase as well as decrease exponentially, linear trends can be fitted to a logarithmic transformation.

Cumulative infections comprise the advantages of being strictly positive and also diluting weekly seasonality. This enables easy modelling via piecewise linear trends as in Doornik, Castle, and Hendry (2021). However, as the pandemic progresses and infections begin to pile up, the cumulative series becomes fairly insensitive even to significant spikes in daily infections. Examining figure 1 makes this clearly visible. Subsequently, any analysis based on first differences of log I_t is also ruled out, except at very early

---

[10]In case of time varying transition probabilities, see Gray (1996) and Bazzi et al. (2017).

stages in the pandemic.

In contrast, logarithmic daily infections preserve sensibility to changes in the underlying level series, whilst still controlling for exponential behaviour. As the logarithm dampens the peaks and troughs of daily infections, the transformation even bears some resemblance to a stationary series. Furthermore, especially when employing the nonlinear UC approach introduced in section 2.3, strictly monotone behaviour of the dependent variable is not a necessarily a desired property. The up and downwards movement of log i_t may cater well to a two-regime scenario. In contrast, log cumulative cases stand to gain less from a regime switching specification. Subsequently, log i_t is chosen as the dependent variable for the analysis.
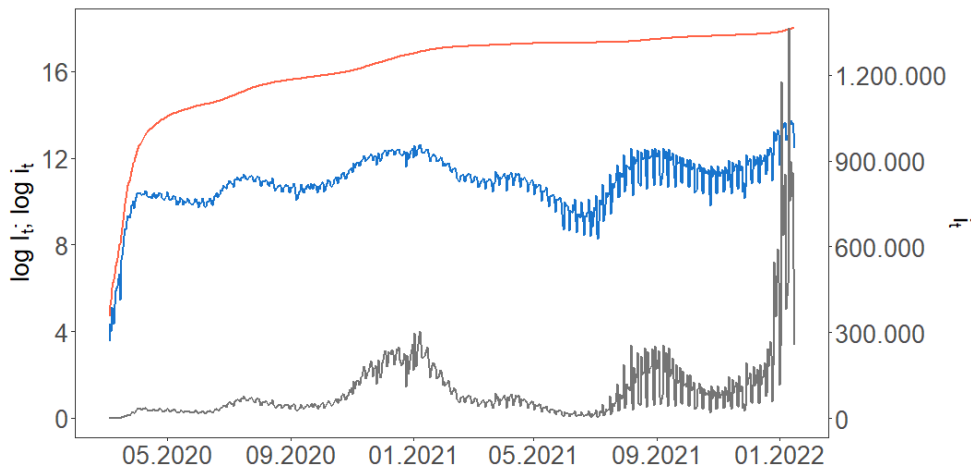
Figure 1: Log of cumulative daily COVID-19 infections log I_t (orange), log of daily infections log i_t (blue) and level COVID-19 infections i_t (grey) for the US across the observational period.

3.1 The basic model

Let $\log y_t =: \log(i_t + 1)$. Adding one reported case assures the dependent variable to be well defined for all periods and does not distort inference (Doornik, Castle, and Hendry, 2021). In the following, y_t is broken up into individual components. For each t, y_t is assumed to admit to following representation:

$$\log y_t = \log \mu_t + \gamma_t \tag{3.1}$$

where μ_t gives a trend and γ_t a seasonal component. The components themselves are assumed to satisfy:

$$\mu_t = \mu_{t-1} + \nu + \xi_t \qquad \xi_t \sim i.i.d.N(0, \sigma_\xi^2) \tag{3.2}$$

$$\gamma_t = \sum_{j=1}^{6} \gamma_{t-j} + \omega_t \qquad \omega_t \sim i.i.d.N(0, \sigma_{\omega_t}^2). \tag{3.3}$$

As previously proposed by Hartl (2021), Ricon-Becker et al. (2020) and a number of other articles, γ_t nets out seven-day seasonality. The stochastic component ω_t allows the seasonal coefficients to vary and assumes measurement errors. By their nature, ξ_t and ω_s are uncorrelated for all t and s.

The log specification of both the dependent variable and trend component in equation 3.1 forces γ_t to be estimated in proportion, rather than additively (Hartl, 2021; Doornik, Castle, and Hendry, 2021). This is a sensible choice, as both seasonality and measurement error greatly depend on y_t rather than being a constant inaccuracy. The necessity for such an approach can be easily verified in figure 1.

The trend μ_t is modelled as a RW with constant drift. Consequently, after controlling for seasonality, infections reflect accumulated unweighted innovations and a constant drift. It is fair to say that the response to the ongoing pandemic either from a medical, legislative or public standpoint matured over time. Initially, hospitals had limited experience in treating COVID-19 strains, let alone a full blown pandemic. Essential techniques like proper positioning of patients or the development of effective work routines evolved over time, whilst leaving a permanent imprint (Li et al., 2021; Jahangir et al., 2020). Other medical innovations from vaccines to anti-viral pills are poised to affect public health for the entirety of the COVID-19 life cycle (Das, Das, and Ghangrekar, 2020). The deployment of non-pharmaceutical interventions (NPI) may not only leave lasting consequences on infection events, but also greatly influence future policies

as regulatory agencies gain experience. Moreover, the effect of NPI may be complementary (Ferguson et al., 2020). A population may also initially comply with measures, grow tired of them or cycle through these states. Likewise, behaviour on the micro level can be easily imagined as a function of previous events (Hartl, 2021). As a consequence, modelling the trend as process with permanent innovation memory seems to be founded in the data generating process.

To further scrutinize the characteristics of COVID-19 infection data, table 1 provides the results of an Augmented Dickey Fuller (ADF) test on i_t, first differences Δi_t as well as log i_t.

Table 1: ADF test on US daily COVID-19 infections

Series	Lag order	Test statistic	p-value
i_t	7	2.454	0.98
Δi_t	7	-7.948	<0.01
log i_t	7	-4.849	<0.01
Alternative hypothesis: stationary			

Notes: Table 1 presents the output of an ADF test on daily COVID-19 infections i_t, first differences of COVID-19 infections Δi_t and a log transformation of daily cases log i_t in the US. The test controls for a constant. Lag order is set to reflect seasonality.

Whereas the Null hypothesis is cannot be rejected for Δi_t and log i_t, the level series is integrated of first order. Naturally, any permanent characteristic in the data generating process is still being transferred to the logarithmic transformation. Moreover, this result demands the drift term ν in equation 3.1 to be $I(0)$. When treating the coefficient of μ_{t-1} in equation 3.2 as a parameter to be estimated, the ML routine of the preferred model returns a point estimate of 0.994 with a 0.005 standard error. This, again, supports a RW specification of μ_t. Nonetheless, in light of novel mutations like the current Omikron strain and a rather endemic outlook for COVID-19, a valid argument against a RW specification can be made. Changes to the disease or to public behaviour may be severe enough to negate previous shocks. Still, much of world is yet caught in a pandemic state at time of analysis and previously adopted routines prevail. Revising equation 3.2 may therefore be a future task.

The linear UC approach defined by equations 3.1 to 3.3 is termed the L-UC (linear UC) model and can be easily executed by the Kalman filter. Parameters σ_ξ and σ_{ω_t} are estimated via ML.

It is however not hard to identify several shortcomings inherent to this approach. For once, the constant drift term ν is not conducive to reflect the underlying non-monotonous trend in $\log(i_t + 1)$. In case of an estimated drift $\hat{\nu} > 0$ drift[11], shocks in infection decreasing phases are necessarily subject to increased variance and subsequent heteroscedasticity. Moreover, even when introducing a time variant trend, the lead-ups both to infection peaks and troughs are assumed to be equivalent. As a significant reduction in cases is likely the result of certain interventions, infection downturns exist in vastly different environments than upturns. Therefore, it is intuitive that also innovation variances might differ in contrasting regimes.

3.2 The exogenous switching model

To address some flaws raised in the previous paragraph, I extend the L-UC approach to two discrete regimes. Subsequently, equation 3.2 is altered to incorporate regime switching regarding the drift term as well as innovation variance. An indicator variable S_t turns on in only one regime with $S_t = \{0, 1\}$. This new specification gives:

$$\mu_t = \mu_{t-1} + \nu_{S_t} + \xi_t \qquad \xi_t \sim i.i.d.N(0, \sigma^2_{\xi,S_t}) \tag{3.4}$$

with

$$\nu_{S_t} = \nu_0 + S_t \nu_1, \tag{3.5}$$

$$\sigma^2_{\xi,S_t} = \sigma^2_{\xi,0}(1 - S_t) + \sigma^2_{\xi,1} S_t. \tag{3.6}$$

As proposed in Hamilton (1989) and describe in section 2.2, S_t follows

---

[11]Throughout the paper I adopt the popular notation of $\hat{x}$ denoting estimates for unobserved components by the Kalman or Kim filter and $x(\hat{\theta})$ estimates derived by ML.

a first order stationary Markov chain with following fixed transition probabilities:

$$\Pr(S_t = 1 | S_{t-1} = 1) = p$$
$$\Pr(S_t = 0 | S_{t-1} = 0) = q.$$

(3.7)

The nonlinear approach substitutes equation 3.2 with 3.4 and expands the system to equations 3.4, 3.5 and 3.7. Hereafter, it will be referred to as Nl-UC (Nonlinear - UC) model. The parameters ν_1, σ_{ξ_0}, σ_{ξ_1}, σ_{ω}, p and q are estimated via ML.

The Nl-UC model allows for breaks of predetermined magnitude in an otherwise constant drift term. Hence, the RW can evolve around a drift depicting increasing as well as decreasing infections. Moreover, innovations are effectively drawn from a Gaussian mixture distribution (Hamilton, 1994, ch. 22.3). One of the two underlying distribution is switched on and innovations may occur more or less severe in size, depending on the current regime.

To keep the model identified, it is sufficient to restrict ν_1 to be negative (Sinclair, 2009). This assures regime $S_t = 1$ to represent the infections decreasing or reduced growth case and thus fixes the label switching issue often associated with MS techniques (Lu, Zeng, and Chen, 2009). Note, that it is not necessary to constrain $\sigma_{\xi,1}^2$ in relation to $\sigma_{\xi,0}^2$.

The switching structural time series model offers a large permutation of optionalities, not all of which can be covered here. Still, an additional extension worth exploring is the introduction of a cyclical component. In business cycle analysis, a cyclical component is often modelled as a stationary AR series and contrasts the non-stationary trend. The cycle in relation to the trend serves the essential tasks of identifying the output gap (see e.g. Harvey (1985), Kim, Nelson et al. (1999) or Kang (2014)).

In the case of COVID-19 infections, a cyclical component assumes short run dynamics other than seasonality. Due to the relatively high frequency data, this time span concerns roughly between three to five days. Non-permanent cyclical events could be triggered by e.g. a local outbreak in an elderly care home or hospital that is kept in check and

does not provoke a range of new infections. Still, a cyclical component did not prove sensible when applied to the data. A possible explanation is the local nature of the pandemic. Cyclical events like the proposed local outbreaks are diluted at the national level. Due to the nature of an infectious virus, significant events spur a set of new infections and thus are assumed by the permanent component. Consequently, cyclical behaviour might only exists at a relatively fine granularity or is simply insignificant.

As already laid out in section 2.3, transition probabilities solely governed by the Markov process are not realistic in many scenarios. This could very well be the case in the Nl-UC model as well. Intuitively, a sudden positive shock to the trend component may increase the probability of a high growth regime in the same period. In contrast, the population might adjust their behaviour as a reaction to contemporaneous sentiment, where sentiment gives a proxy for the regime (Hartl, 2021). In light of rapidly increasing infections caution among the public increases and subsequently dampens the spread. Furthermore, people can be expected to behave according to a set of unobservable variables. If deciding variable realizations permit themselves to be grouped as different regimes, introducing endogenous switching prevents omitted variable bias (Kim, Piger, and Startz, 2008). Also regarding regulatory action, certain thresholds might be in place which when exceeded, e.g. by excessive trend growth, trigger restrictions and thus engage the infections decreasing regime.

Even though the direction of endogeneity cannot be determined a priori, it is hard to argue that the development of infection regimes is a fully exogenous process.

3.3 The endogenous switching model

Following Kim, Piger, and Startz (2008) and Kang (2014), I further introduce endogenous switching via a Probit specification in place of the Markov chain process. For the En-Nl-UC (endogenous nonlinear UC) model, equation 3.7 is let go in favor of:

$$S_t^* = \beta_0 + \beta_1 S_{t-1} + \eta_t \qquad \eta_t \sim i.i.d. N(0,1). \tag{3.8}$$

17

The Probit model discriminates between regimes whether S_t^* is non-negative. This gives the identities:

$$S_t = 0 \quad \forall \quad \eta_t < -\beta_0 - \beta_1 S_{t-1}$$
$$S_t = 1 \quad \forall \quad \eta_t \geq -\beta_0 - \beta_1 S_{t-1}. \tag{3.9}$$

Entries to the transition matrix can be easily derived from the Probit covariates by:

$$\Pr(S_t = 0 | S_{t-1} = i) = \Phi(-\beta_0 - \beta_1 i)$$
$$\Pr(S_t = 1 | S_{t-1} = i) = 1 - \Phi(-\beta_0 - \beta_1 i), \tag{3.10}$$

where Φ denotes the standard normal cumulative distribution function (Kim, Nelson et al., 1999, ch. 4). Endogeneity arises from the joint distribution function of all innovations in the system:

$$\begin{bmatrix} \xi_{t,0} \\ \xi_{t,1} \\ \omega_t \\ \eta_t \end{bmatrix} \sim i.i.d.N(0, \Sigma), \qquad \Sigma = \begin{bmatrix} \sigma_{\xi,0}^2 & 0 & 0 & \rho_0 \\ 0 & \sigma_{\xi,1}^2 & 0 & \rho_1 \\ 0 & 0 & \sigma_\omega^2 & 0 \\ \rho_0 & \rho_1 & 0 & 1 \end{bmatrix}. \tag{3.11}$$

Note, that the covariance ρ_{S_t} between innovations to the regime process and trend innovations is itself regime dependent. This is a function of the regime induced heteroskedasticity in equation 3.6. If a certain regime innovation triggers a switch from $S_{t-1} = 0$ to $S_t = 1$, $\sigma_{\xi,0}$ follows suit and changes to $\sigma_{\xi,1}$ in t. Since η_t has a constant variance by definition, ρ_0 is forced to change in proportion to $\sigma_{\xi,1}$ and becomes ρ_1. However, the correlation between ξ_{t,S_t} and η_t stays constant and is thus uniquely defined as

$$\varrho = \frac{\rho_0}{\sigma_{\xi,0}} = \frac{\rho_1}{\sigma_{\xi,1}}, \qquad \sigma_{\xi,0}, \sigma_{\xi,1} \neq 0. \tag{3.12}$$

A non-trivial effect of the endogenous switching mechanisms is that the KF likelihood function is no longer Gaussian (Kim, Piger, and Startz,

2008). It can be shown that, in contrast to exogenous switching, the mean of ξ_t conditional on S_t and S_{t-1} is no longer zero but:

$$
\begin{aligned}
E(\xi_t|S_t = 0, S_{t-1} = i, \Psi_{t-1}) &= \rho_{S_t} \frac{-\phi(-\beta_0 - \beta_1 i)}{\Phi(-\beta_0 - \beta_1 i)} \\
E(\xi_t|S_t = 1, S_{t-1} = i, \Psi_{t-1}) &= \rho_{S_t} \frac{\phi(-\beta_0 - \beta_1 i)}{1 - \Phi(-\beta_0 - \beta_1 i)}
\end{aligned}
\tag{3.13}
$$

where ϕ denotes the standard normal probability density function and Ψ_t the information set at time t (Kim, Piger, and Startz, 2008, eq. 2.8 to 2.10). This is intuitive, since conditional on a certain set of S_t and S_{t-1} realizations the expectation of a shock to S_t^* in equation 3.8 is necessarily non-zero. This conditional expectation of ω_t is added to ξ_t.

Likewise, the conditional variance of ξ_t is augmented as:

$$
\begin{aligned}
Var(\xi_t|S_t = 0, S_{t-1} = i, \Psi_{t-1}) &= \sigma_{\xi,S_t}^2 - \\
\rho_{S_t}^2 E(\xi_t|S_t = 0, S_{t-1} = i, \Psi_{t-1}) &\left[\frac{-\phi(-\beta_0 - \beta_1 i)}{\Phi(-\beta_0 - \beta_1 i)} + \beta_0 + \beta_1 i \right] \\
Var(\xi_t|S_t = 1, S_{t-1} = i, \Psi_{t-1}) &= \sigma_{\xi,S_t}^2 - \\
\rho_{S_t}^2 E(\xi_t|S_t = 1, S_{t-1} = i, \Psi_{t-1}) &\left[\frac{\phi(-\beta_0 - \beta_1 i)}{1 - \Phi(-\beta_0 - \beta_1 i)} + \beta_0 + \beta_1 i \right]
\end{aligned}
\tag{3.14}
$$

and thus deviates from the original specification in equation 3.1 (Sinclair, 2009).

Since the quasi-ML routine now maximizes a non-Gaussian conditional density, the resulting estimator is inconsistent and due to the non-zero mean of ξ_t also biased for endogenous switching (Kim, Piger, and Startz, 2008; Campbell, 2002). Also, the irregular component of the trend process ξ_t is correlated with the regime dependent parameters ν_{S_t} and σ_{ξ,S_t}. Because the variance of ξ_t is itself regime dependent, endogenous switching might further entail serially correlated errors (Chang, Choi, and Park, 2017; Kim, Piger, and Startz, 2008). Chang, Choi, and Park (2017) overcome these correlation issues by introducing sequential endogeneity.

In contrast to the approach in Kim, Piger, and Startz (2008), they let innovations to μ_t correlate with η_{t+1} rather than η_t. This of course prohibits modelling correlations as contemporaneous. Furthermore, they relax the assumption of a stationary Markov chain by employing an AR specification to substitute the Probit model.

But as shown in Kim, Piger, and Startz (2008), the bias and inconsistency of contemporaneously correlated innovations are rather insignificant. Moreover, it is preferred to introduce endogeneity as contemporaneous.

The En-Nl-UC model defined by equations 3.1, 3.3 to 3.5 and 3.8 to 3.14 incorporates contemporaneous correlation between regime and trend innovations[12]. It is observationally identical to the Nl-UC model for $\rho_0 = \rho_1 = 0$. Together with all remaining additional parameters, ρ_{S_t} is estimated by quasi-ML.

The state space form of all three presented models is provided in appendix A.

4 Monte Carlo simulation study

Especially due to nonlinearity and fairly complex dynamics within the systems it is not easy to place the estimation performance of each proposed model. Depending on the data quality, environment and assumptions, the choice of model could potentially distort inference. In contrast, some approaches might prove rather robust even when certain assumptions are not met. To evaluate these characteristics, I run three Monte Carlo simulation studies. Each simulation concerns one situation where the data generating process exactly matches the respective model. I draw 500 series with a length of 100 periods from every setup and derive ML estimates via a Nelder-Mead optimizer. For the L-UC model the Kalman filter sets up the likelihood function, whereas the Kim filter assumes this task for the nonlinear approaches.

If applicable, parameters are chosen to roughly mirror the JH/CSSE data of COVID-19 infections for the US. Table 2 presents the results.

---

[12]Technically, equation 3.1 is not specified correctly for the En-Nl-UC model, since the distribution of ξ_t does not hold true as shown by equations 3.13 and 3.14.

Table 2: Results Monte Carlo simulations

	L-UC		Nl-UC		En-Nl-UC	
L-UC specification						
$\sigma_{xi_0} = 0.170$	0.170	(0)	0.192	(0.966)	16.545	(204.236)
σ_{xi_1}	-	-	0.146	(0.063)	0.165	(0.030)
$\sigma_\omega = 0.060$	0.058	(0.014)	0.056	(0.015)	0.057	(0.015)
ν_1	-	-	-0.082	(0.102)	-2.601	(4.181)
p	-	-	0.152	(0.258)	1	-
q	-	-	0.168	(0.279)	0.002	-
ϱ	-	-	-	-	0.382	(0.429)
$\log \mathscr{L}$	-48.537	(6.617)	-47.468	(6.832)	-47.868	(7.273)
Nl-UC specification						
$\sigma_{xi_0} = 0.420$	0.464	(0.04)	0.658	(5.514)	578982.081	(9639715.023)
$\sigma_{xi_1} = 0.060$	-	-	0.387	(0.154)	0.444	(0.073)
$\sigma_\omega = 0.040$	0.034	(0.024)	0.031	(0.025)	0.039	(0.024)
$\nu_1 = -0.060$	-	-	-0.245	(0.279)	-1.760	(2.824)
$p = 0.960$	-	-	0.209	(0.301)	0.999	-
$q = 0.875$	-	-	0.229	(0.308)	0.001	-
ϱ	-	-	-	-	0.590	(0.440)
$\log \mathscr{L}$	-116.785	(6.906)	-115.394	(6.967)	-116.114	(7.002)
En-Nl-UC specification						
$\sigma_{xi_0} = 0.500$	0.210	(0.067)	0.155	(0.166)	30.929	(618.075)
$\sigma_{xi_1} = 0.060$	-	-	0.177	(0.200)	0.086	(0.056)
$\sigma_\omega = 0.040$	0.038	(0.018)	0.038	(0.014)	0.041	(0.015)
$\nu_1 = -0.0004$	-	-	-0.042	(0.097)	-1.366	(2.813)
$p = 0.960$	-	-	0.263	(0.306)	0.954	-
$q = 0.829$	-	-	0.278	(0.319)	0	-
$\varrho = 0.040$	-	-	-	-	0.832	(0.220)
$\log \mathscr{L}$	-53.265	(23.138)	-42.028	(23.048)	-37.571	(28.022)

Notes: Table 2 gives summary statistics of estimates for three Monte Carlo simulations. Each simulation reflects the data generating process according to one applied model. Results are given as the mean of 500 individual estimates based on 100 periods. The values in parenthesis represent the standard deviation of the sample of estimations. No variance is given for p and q of the En-Nl-UC model, as probabilities are constructed from Probit covariates. $\mathscr{L}$ denotes likelihood.

Immediately striking is the discouragenly bad estimation performance. The linear model is at least able to reliably estimate parameters of the linear data generating process, however does not produce sensible inference in nonlinear setups. But especially the more complex nonlinear approaches disappoint. As further indicated by the second moments, the bias as well as range of estimations throughout the simulation studies is astounding across the board[13].

To uncover the reason for the underperformance of nonlinear approaches, it is elucidating to more closely examine one individual simulation. For this case study I simulate a series of 2000 realizations according to the nonlinear endogenous data generating process. I apply the En-Nl-UC model to a subset of the first 100 realizations, the first 700 and ultimately to the full series[14]. Table 3 presents the results of all three estimation runs.

As can be deduced from table 2, certain estimates, namely regime probabilities as well as $\hat{\varrho}$, are rather far off from their real values. However, consistency is clearly recognizable. This is not surprising. Regime switching is notoriously hard to estimate, even in simple specifications (Hamilton, 1994, ch. 22)[15]. Not many regime switching events take place in a subset of 100 periods and inference is necessarily constructed from very few data points. The size of the standard deviations for Probit covariates and $\hat{\varrho}$ underlines this argument. However, it may be worrying that parameters like ϱ or drift terms are not estimated precisely even when exploiting the entire significantly richer data history.

Nonetheless, the smoothed regime probabilities and trend estimates fit the simulated data staggeringly well. As shown by figure 2, using all 2000 periods the model is confidently able to discriminate between regimes and excels at uncovering the RW trend amongst seasonality and noise. The model also achieves to produce homoscedastic prediction errors, which are only serially correlated up to one lag for certain regime switches (see

---

[13]Medians are somewhat closer to real parameter values, however still fairly biased.

[14]Similar studies by Kim, Piger, and Startz (2008) or Kang (2014) employed between 200 and 1000 simulated periods.

[15]As an additional symptom of the difficult estimation process, numerical optimization of the nonlinear models did not prove very robust against different starting values. Various combinations often lead to local maxima or terminal errors.

Table 3: Results Monte Carlo simulation case study

	T = 100	T = 700	T = 2000
$\sigma_{xi,0} = 0.500$	1.127	2.622	0.578
	(0.466)	(0.327)	(0.050)
$\sigma_{xi,1} = 0.060$	0.105	0.045	0.040
	(0.132)	(0.066)	(0.034)
$\sigma_{\omega}\ \ = 0.040$	0.026	0.045	0.040
	(0.314)	(0.054)	(0.031)
$\nu_0\ \ = 0.001$	1.692	3.835	0.230
	(0.011)	(0.002)	(0.002)
$\nu_1\ \ = -0.0004$	-1.702	-3.834	-0.231
	(0.794)	(0.367)	(0.186)
$\beta_0\ \ = -0.950$	4.421	1.223	-0.645
	(10.960)	(0.210)	(0.102)
$\beta_1\ \ = 2.700$	-2.841	0.585	2.437
	(10.964)	(0.233)	(0.136)
$\varrho\ \ = 0.040$	0.861	0.629	0.350
	(10.268)	(0.463)	(0.165)
$\log\mathscr{L} = 685.192$	-21.540	210.058	693.064

Notes: Table 3 gives the estimates for three subsets of one simulated series employing the En-Nl-UC model as. The size of the respective subsets in the column names define each estimation run. The simulated data is constructed to meet all assumptions inherent to the En-Nl-UC approach. The orignial input parameters are also provided. The numbers in parenthesis represent SEs. The true log likelihood value is derived by inputting the true parameters into the En-Nl-UC model. $\mathscr{L}$ denotes likelihood.

figure 8 in appendix C).

The model is apparently able to sensibly explain observations whilst employing parameters not matching the data generating process. Consequently, the likelihood function in this example is expected to be fairly gradual around its global maximum with respect to regime switching parameters. Rather bizarrely, the log likelihood of estimates for $T = 2000$ even exceeds the one of the actual underlying parameters[16]. Naturally any optimization routine struggles immensely with identifying the true global maximum in the first place and as shown by table 3 it is further not a given that this maximum even uncovers the true underlying parameters[17].

---

[16]Difference is driven by *i.i.d.N.* stochastic variables.

[17]Naturally it would be highly insightful to run the Monte Carlo simulations again with an extensive sample size. However, due to computational budget constrains I
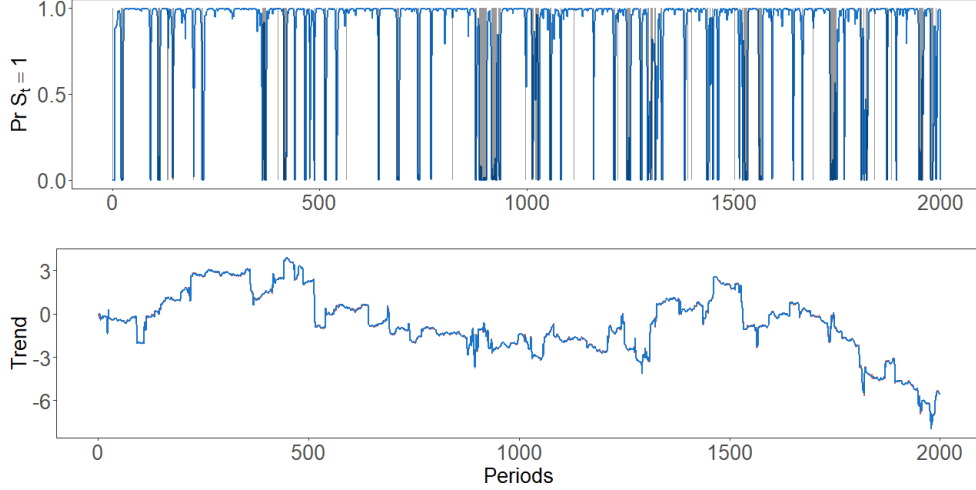
Figure 2: Top figure gives estimated regime probabilities for $S_t = 1$. Actual periods of simulated regime $S_t = 1$ are shaded. The bottom plot presents the smoothed trend $\hat{\mu}_t$ (blue) against the actual simulated switching RW trend (orange). Even though hardly visible, 90% confidence intervals of the smoothed trend are shaded.

However, as a take away for the following application, even though parameter estimates may be flawed the models can still be expected to produce reasonable estimates for unobserved components and regime probabilities. Moreover, the models improve over time as the data history becomes more and more comprehensive.

5 Empirical results and forecast

Following Liu, Moon, and Schorfheide (2021), I increase robustness by censoring all data points for which the cumulative number of infections has not yet exceeded 100. For the US, this entails an observational period spanning from 04.03.2020 until 16.01.2021 with a total of 684 daily entries. I apply all three models to the US COVID-19 infection time series. In case of the Nl-UC and En-Nl-UC approaches, p and q are initialized with their respective steady states (see equation 5.1). The trend component of the state vector is initialized with $\log i_t$ of the 03.03.2020, one day before

---

did not perform a second run. Estimating nonlinear switching models is remarkably computational demanding as further laid out in Hamilton (1994, ch. 22.3).

start of the analysis. Remaining entries to the state vector are initialized diffusely (Durbin and Koopman, 2012, ch. 5). As in the Monte Carlo study, I estimate all additional parameters with a nonlinear Nelder-Mead optimization routine. Table 4 presents the estimation results. :

Table 4: Maximum likelihood estimates

	L-UC	Nl-UC	En-Nl-UC
$\sigma_{\xi,0}$	0.171	0.491	0.482
	(0.055)	(0.080)	(0.082)
$\sigma_{\xi,1}$	-	0.061	0.063
	-	(0.098)	(0.073)
σ_ω	0.063	0.041	0.041
	(0.105)	(0.062)	(0.056)
ν_0	0.013	0.012	0.102
	(0.016)	(0.020)	(0.003)
ν_1	-	-0	-0.103
	-	(102.753)	(0.597)
p	-	0.960	0.964
	-	(0.267)	-
q	-	0.827	0.851
	-	(0.291)	-
ϱ	-	-	0.286
	-	-	(0.361)
$\log \mathscr{L}$	-52.709	198.141	199.355

Notes: Table 4 presents ML estimates for all three proposed models. Standard errors (SE) are given in parenthesis. ν_0 is part of the state vector and is thus not estimated via ML, nonetheless added for completeness. SEs for ν_0 are substituted by the estimated standard deviation. Since p and q are a function of Probit covariates for the En-Nl-UC model, no SEs are provided. The respective Probit covariates are estimated as $\beta_0(\hat{\theta}) = -1.039$ with SE = 0.176 and $\beta_1(\hat{\theta}) = 2.841$ with SE = 0.241. In case of the linear approach, parameters for $S_t = 0$ represent the general scenario. $\mathscr{L}$ denotes likelihood.

As is to be expected, ML estimates of the Nl-UC and En-Nl-UC models are far more comparable to each other than to the linear approach. Furthermore, it is also not surprising that nonlinear methods fit the data generating process better by some margin. A striking difference between the Nl-UC and En-Nl-UC models are the diverging executions of regime switches. It stands to reason that the Nl-UC model would perform almost identically without even including regime dependency of the drift term. Whereas in the Nl-UC approach different regimes only materialize through

changes in σ_{ξ_t, S_t}, the RW drift parameters also account for an additional regime induced effect in the endogenous En-Nl-UC model.

Although the impact of regime switches transpires differently in the Nl-UC and En-NL-UC models, both approaches agree on the regime probabilities themselves. It is easily possible to derive expected regime durations from p and q, which makes the estimated persistence straightforward to interpret (Kim, Nelson et al., 1999, ch. 4). Both models project an expected duration of roughly $(1 - 0.960)^{-1} \approx 25$ days for $S_t = 1$ and ≈ 6 days for $S_t = 0$. Thus, the infections decreasing $S_t = 1$ environment is identified as predominant and alternates with short but presumably intense increases.

The correlation between regime and trend ϱ is estimated to be positive. Subsequently, a positive shock to the trend seems to favour a nudge in the direction of the infections decreasing regime $S_t = 1$. As argued in section 3.2, both a positive or negative correlation may be explained plausibly. However, in the end appropriate policy reactions or a conservative reaction to cautious sentiment appear to have the upper hand. Applying insights from the Monte Carlo simulations, the true ϱ is likely to be smaller than estimated. However, this suspicion may only be validated with a more extensive data history, given the true ϱ is constant.

Based on the estimates in table 4 alone it is not straightforward to decide between the endogenously or exogenously switching model for the preferred approach. Following Kim, Piger, and Startz (2008), I perform a t-test on $\varrho(\hat{\theta})$ as well as a likelihood ratio test in order to quantitatively support a decision. Note, that the Nl-UC model is nested within En-Nl-UC with $\varrho \overset{!}{=} 0$. A possible concern is the bias and inconsistency inherent to ML in case of endogenous switching. However, since the regime process behaves completely exogenously under the null hypothesis for both tests, asymptotic validity is preserved (Hamilton, 1994, ch. 5.8; Kim, Piger, and Startz, 2008).

Both test indicate towards the Nl-UC model. However, the likelihood ratio test if far less confident towards the endogenous approach striking out.

In spite of the fact that both tests favour the Nl-UC model, I select

Table 5: Model selection likelihood ratio and t test

	Test statistic	p-value
Likelihood ratio test	2.427	0.119
t test	0.855	0.725

Notes: Table 5 gives the result of an likelihood ratio test concerning the En-Nl-UC and nested Nl-UC models as well as a t test on $\hat{\varrho}$. The likelihood ratio test has a H_0 of $\theta \in \Theta^{Nl-UC}$ where Θ^{Nl-UC} denotes the total parameter space under the restrictions assumed by the Nl-UC model. The null hypothesis of the t test can be expressed as $H_0 : \varrho = 0$.

the endogenous En-Nl-UC for further analysis. As already touched upon, regime switches are demanding to estimate and the power of statistical tests is limited (Hamilton, 1994, ch. 22). Most importantly however, it is very hard to support an exogenous switching process qualitatively. Trend and infection regimes are most likely intertwined.

The preferred En-Nl-UC model is further scrutinized in the following section.

5.1 Model evaluation

Figure 3 presents smoothed estimates of $\log \mu_t$ together with estimated probabilities of $S_t = 1$ for the endogenous switching En-Nl-UC approach[18].

Other then expected, the system does not alternate between coherent episodes of a single regime. Besides steep inclines at the beginning and end of the observational period, the model advocates $S_t = 1$ even in phases of sustained infection increases. It seems to perceive $S_t = 1$ as the default and only transitions to $S_t = 0$ for short term episodes of increased volatility or extreme growth in logs. Switching is predominantly not driven by changes in the actual trend drift. As $\sigma_{\xi,0}(\hat{\theta})$ is roughly eight times the size of $\sigma_{\xi,1}(\hat{\theta})$, the regime $S_t = 0$ is exploited to assume sporadic breaks in the estimated trend and effectively becomes an increased volatility regime. These breaks are plucks primarily in the downwards direction. Moreover, in most instances the trend breaks seem to follow the seasonal structure. Therefore it is not far fetched, that in events of suddenly

---

[18]For sake of completeness, figures 9 and 10 in appendix C present the same output for the L-UC as well as Nl-UC model.
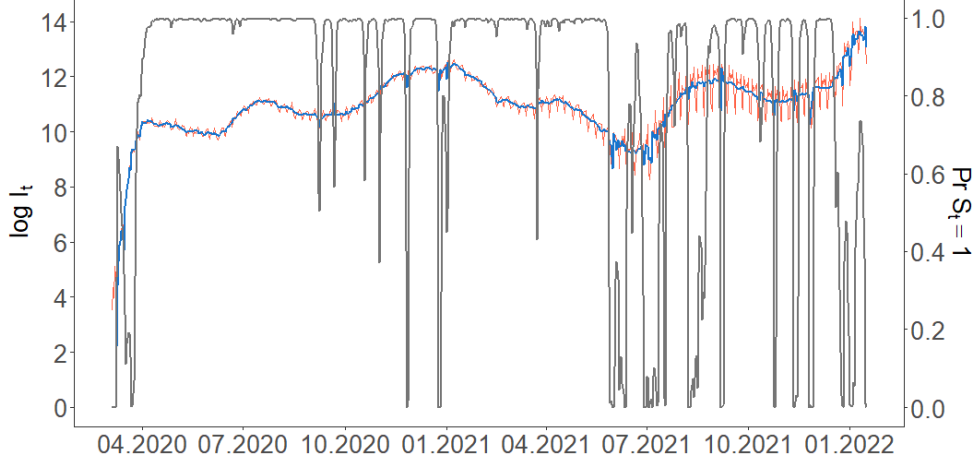
Figure 3: Smoothed trend estimates $\hat{\mu}_t$ (blue), smoothed regime probabilities $\hat{\Pr}(S_t = 1|\Psi_t)$ (grey) for the En-Nl-UC model and original $\log i_t$ series (orange). Individual regime dependent trend estimates are averaged into a single value.

inflated seasonality the model resorts to engaging the increased volatility regime $S_t = 0$ instead of relying on seasonal dummies and ω_t (see equation 3.3) to assume the pluck. Such isolated days of exacerbated seasonality could potentially occur as a consequence of reporting mishaps or e.g. policy changes that take effect at the beginning of a week. Additionally, reliable insights regarding the effectiveness of e.g. policy interventions via regime switching behaviour cannot be obtained.

Moreover, just like in the Monte Carlo simulation case study in section 4, the likelihood function does not have a distinct global maximum with respect to regime probabilities. As sketched out in figure 4, the log likelihood function forms a gradual ridge around certain sets of β_0 and β_1. The trough towards $\beta_0 \approx -1.5$ further highlights the risk of local maxima.

It is helpful to more closely examine prediction errors and especially their second moments. Abnormalities may indicate misspecifications. Figure 5 presents one-step-ahead prediction errors $\hat{v}_t^{(i,j)}$ conditinal on $S_{t-1} = i$ and $S_{t-1} = j$, together with their respective Auto Correlation Functions (ACF).

Apart from a brief initialization phase, the model is able to maintain
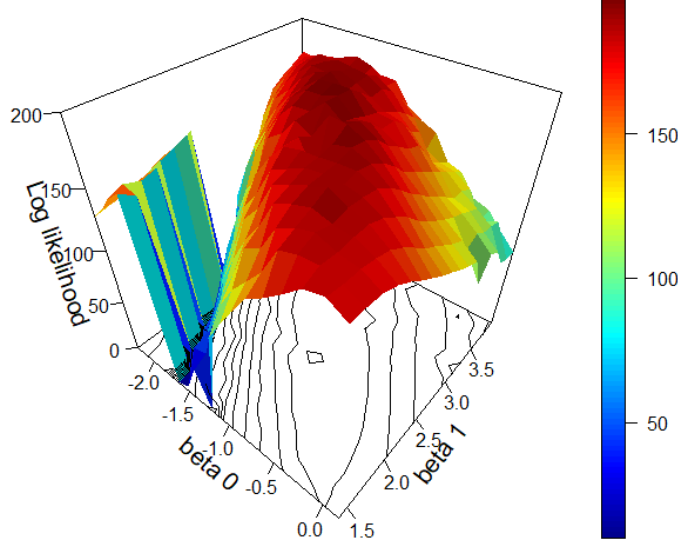
28

Figure 4: Figure outlines the log likelihood function of the En-Nl-UC model with respect to the Probit covariates in equation 3.8. Remaining additional parameters were selected individually for every grid point by ML.

homoscedasticity in the prediction errors. Some outliers and hints of heteroscedasticity emerge as a consequence of frequent breaks in the estimated trend. However, serial correlation seems to pose the main weakpoint. The model is unable to assume some serial dependencies up to ten lags, which subsequently end up in the residuals. Still, even though a structured serial correlation emerges, it only concerns few lags, certain regime transitions and vanishes quickly. Also remember that some serial correlation in the error term could be a product of endogenous switching (Chang, Choi, and Park, 2017). Issues inherent to contemporaneous endogeneity further materialize in non-zero mean prediction errors[19].

Also, even in logs the seasonality increases over time. This aids in explaining the increasing frequency of estimated regime switching at the end of the analysis. Larger seasonality of course increases the chance of the trend component catching some bumps. As another consequence, ω is also most likely subject to a time variant variance in the data generating process.

---

[19]With $\overline{v}_t^{(i,j)}$ denoting mean conditional on $S_t = j$ and $S_{t-1} = i$: $\overline{v}^{(0,0)} = -0.062$, $\overline{v}^{(1,0)} = 0.201$, $\overline{v}^{(0,1)} = -0.026$, $\overline{v}^{(1,1)} = -0.001$.
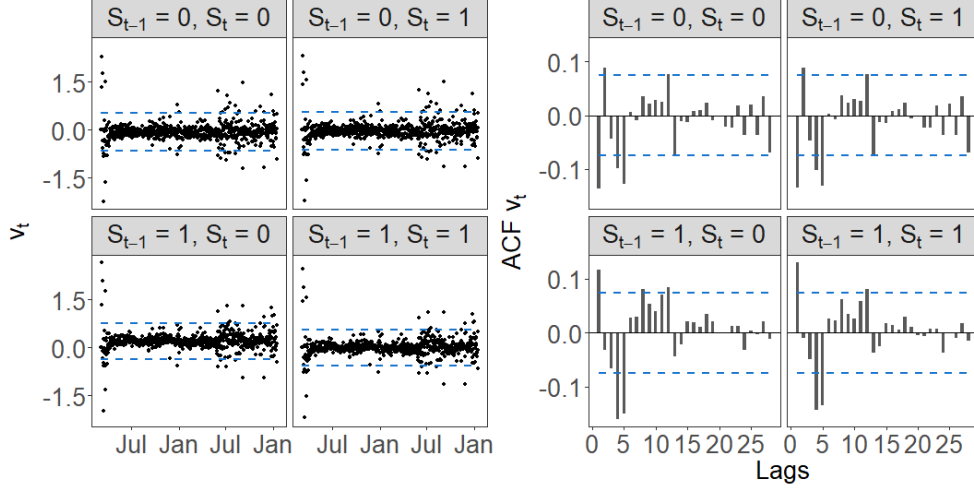
Figure 5: The figure on the left hand side presents one-step-ahead prediction errors $\hat{v}_t^{(i,j)}$ for the En-Nl-UC model together with two standard deviations (dashed, blue). The figure on the right hand side displays the ACF of the prediction errors with 95% confidence intervals (dashed, blue). Each panel gives the result conditional on a specific unobserved regime switch.

Here, again a cyclical component may pose as a potential solution. A cycle could net out non-permanent dynamics like amplified seasonality and act as a doorman for variation entering the permanent component. Still, when applied to the data, a stationary AR(2) cycle was not able to rectify these issues (see figure 11 in appendix C). Even though it is not shown here, both a cyclical component with and without correlated innovations with the trend were tested (Morley, Nelson, and Zivot, 2003). A cyclical component is also not able to net our serial correlation in $\hat{v}_t^{(i,j)}$[20] (see section 3.2). This might also point towards seasonal spikes actually adhering to the permanent component. However, this is unlikely when studying figure 3 closely.

An alternative potential fix is to manually increase σ_ω. This entails a more flexible seasonal structure and enables γ_t to assume a larger portion of problematic spikes and erratic measurements. However, an increase in σ_ω sufficient enough to net out intense seasonality also lets γ_t absorb

---

[20]Furthermore, employing a model with a cyclical component adds to the dimensionality of the likelihood function and is as a consequence not computational feasible without extensive capacity.

significant variation that is not part of seasonality or the measurement error. An inflated σ_ω would not solve the issue but simply move its symptoms to another component.

However, it is important to remember that some data driven issues were expected by nature of the data. Sporadic spikes in the data may well be caused by reporting issues or structural changes in the tracking process. Moreover, as made apparent by figure 3, the En-Nl-UC model is capable of closely tracking the dependent variable with only minor concerns regarding model misspecification.

5.2 Forecast evaluation

Forecasting the En-Nl-UC model entails projecting both the Kalman filter and Hamilton filter elements into the future. Regarding the Kalman subcomponent, the state vector is iterated onwards via the transition equation (Durbin and Koopman, 2012, ch. 4.11). Same principle applies to the Hamilton filter, where regime probabilities converge to their respective steady states (Hamilton, 1989). The steady states further impact predicted regime dependent parameters. Remember, as with filtering, the Kim filter produces M forecasts for every projected period. To form a single unconditional trend forecast, I construct a weighted average of all regime dependent forecasts for every prediction step.

Since the trend term μ_t is modelled as a RW with drift, the best forecast is given by the last observed trend value in addition to accumulated drift terms (Hamilton, 1994, ch. 17). Naturally, the weighted average of accumulated drift terms reflects the convergence to regime steady states as the forecasting horizons increases. As a consequence, the predicted trend approximates a linear function with slope given by $\hat{\nu}_0 \hat{\pi}_0 + (\hat{\nu}_0 + \nu_1(\hat{\theta}))\hat{\pi}_1$, where

$$
\pi_{S_t=i} = \frac{1 - \Pr(S_t = j | S_{t-1} = j)}{2 - \Pr(S_t = i | S_{t-1} = i) - \Pr(S_t = j | S_{t-1} = j)} \tag{5.1}
$$
$$
\text{with } i \neq j \text{ and } i; j \in \{0, 1\}
$$

denotes the steady state probabilities (Hamilton, 1994, ch. 22)[21]. It is further possible to add seasonality to the forecast. Since the expectation of the irregular term ω within the seasonal component is zero (see equation 3.3), the final seasonal entries to the smoothed state vector in T serve as deterministic seasonal dummies. Naturally, the variance associated with the prediction is strictly increasing in the forecasting horizon (Hamilton, 1994, ch. 17).

Because of convergence to linearity and rapidly expanding prediction intervals, the forecasting horizon needs to be kept brief. Tried and tested forecasting horizons of reduced-form models for COVID-19 range between six and ten days (Doornik, Castle, and Hendry, 2021).

I evaluate up to 30 prediction steps in en effort to possibly deduce any medium term tendencies the model might pick up. Still, robust forecasts of infections are restricted to short periods due to the turbulent data generating process.

As laid out in section 2.4, I depict forecasting performance by constructing predictions for progressively increasing subsets of the observational period. The considered subsets increase by seven days with every new forecast after a data accumulation phase of 28 days. Figure 6 plots the first 14 days of each forecasted $\hat{\mu}_{s+n|s}$ against smoothed $\hat{\mu}_t$, where n represents the forecasted step so that $s + n = t$ with $n = \{1, 2, ...\}$ and $s = \{1, 2, ..., T\}$.

It becomes quickly apparent that forecasts constructed within roughly the first 20 weeks of the exercise are next to non-informative. This mirrors experience from the Monte Carlo studies in section 4. Subsequently, I extend the initial data build-up phase by a factor of six to a total of 24 weeks up to the 18.08.2020 before further reviewing the forecasts. Examining figure 3 can give a rough idea of the point in time until the model has encountered enough regime transitions to form some sensible results.

Figure 7 displays the MAPE and MedAPE for every $n \leq 30$. MAPE and MedAPE represent the mean and median respectively of the standardized prediction error: $\frac{100}{\hat{\mu}_t}(\hat{\mu}_t - \hat{\mu}_{s+n|s})$ (Doornik, Castle, and Hendry, 2021). The smoothed trend is chosen as a reference since significant seasonality in

---

[21]Holds true for an ergodic Markov process, which is the case provide that $p < 1$, $q < 1$ and $p + q > 0$ (Hamilton, 1994, ch. 22).
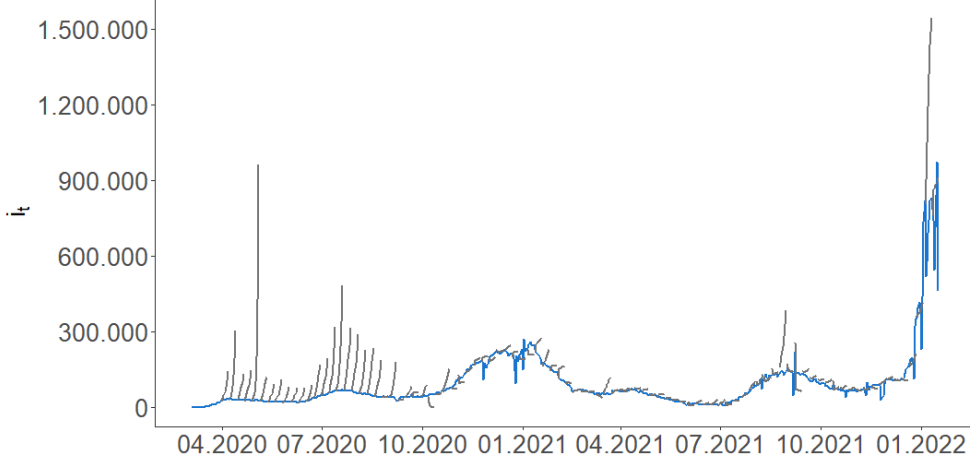
Figure 6: 14 day ahead forecasts $\hat{\mu}_{s+n|s}$ (grey) are sketched together with the smoothed trend $\hat{\mu}_t$ (blue).

the actual i_t realizations distort any results. Note, that it is also possible to add forecasted seasonality to $\hat{\mu}_{s+n|s}$, however this further complicates the evaluation.

As I have already pointed out multiple times throughout the paper, estimation of nonlinear UC models and especially the nonlinear optimization of the likelihood function pose a non-trivial challenge. This is also visibly the case for certain forecasts in figure after 6, even long after the extended initialization phase has elapsed. Individual forecasts that greatly exceed the bounds of previous and following realizations cause significant divergence between the MAPE and MedAPE. Especially when considering a forecasting period of 30 days, the MAPE increases exponentially whereas the MedAPE grows fairly linearly.

The MedAPE is roughly in line with other time series techniques employed to model COVID-19 infections (see e.g. Gecili, Ziady, and Szczesniak (2021)). However, when compared to dedicated forecasting tools like in Doornik, Castle, and Hendry (2021), the En-Nl-UC model cannot compete.

Still, it is important not to forget the core task a model is set out to perform. In the case of the proposed approaches, producing forecasts is only one feature. Moreover, the endogenous switching framework developed in this paper may be seen as a platform to construct derivative
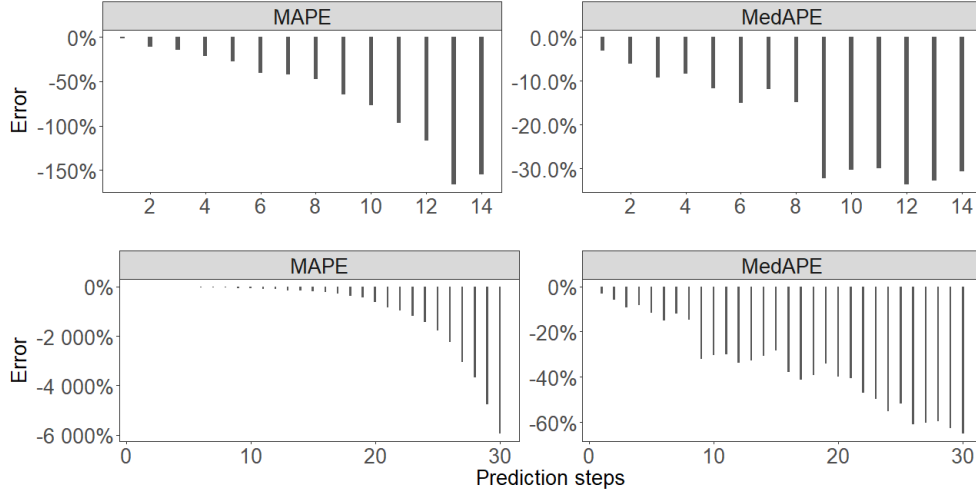
33

Figure 7: The top half gives MAPE and MedAPE of the forcasts presented in figure 6. The figures in the bottom half present the summary statistics for an extended forecasting horizon of 30 days.

models from, which are more capable at executing specific tasks.

6 Conclusion

This paper sets up an UC framework to model COVID-19 infections. Two discrete infection regimes govern the innovation variance and drift of a RW trend component. Regime switches are a function of an endogenous stationary first order Markov process.

As COVID-19 infection are a known subject to frequent distributional shifts, long memory dynamics and exogenous issues like inconsistent reporting, any reasonable econometric technique should be equipped to handle those perils. Regime switching UC models allow for such rapid and far reaching changes which are determined by unobserved regimes.

Nonetheless, even when including mechanisms to theoretically control for known trip wires, switching UC models find themselves stumbling over peculiarities in the dependent variable. Especially isolated seasonal spikes are absorbed by the permament component and prompt regime transitions.

Maybe even more extensive approaches with $M > 2$ and a higher order Markov chain might be able to assume all heterogeneity. A third regime

34

could possibly account for immediate effects of abrupt regulatory actions or simply soak up structural changes in the reporting process. In the end, the model has to explain a time series of reported positive test results, rather than actual COVID-19 infections.

Furthermore, as correctly pointed out by Liu, Moon, and Schorfheide (2021), reduced form models like the ones presented here are inherently limited to only portraying already familiar behaviour. This is clearly observable in the quickly deteriorating forecasting performance. Moreover, prediction intervals are not able to reflect the possibility of impactful future distributional shifts. Structural epidemiological approaches on the other hand already incorporate yet unobserved dynamics. Consequently, there is always a trade-off between the rigidities of a structural model and the flexibility yet performance ceiling of a reduced form approach. To counterbalance this shortcoming, one could incorporate multiple future scenarios in the forecasting behaviour of the En-Nl-UC model. Instead of allowing regime probabilities to converge to their natural steady states, it would not be hard to allocate artificial regime steady states according to planned or probable events. For example, to include the likely scenario of a highly infectious and increasingly dominant Omikron COVID-19 mutation, one allocates an inflated $S_t = 0$ steady state for a certain forecasting period. Likewise, just be fore the roll out of strict contingency measures, an analyst places more weight on the $S_t = 1$ regime for a future period. This mechanism allows for some exogenous structural influence to steer the forecasting behaviour of an otherwise reduced-form model.

A State space representations

A.1 L-UC Model

Notation is largely adopted from Durbin and Koopman (2012). The UC model defined by equations 3.1 to 3.3 in state space form is provided by the measurement equation:

$$\underset{1\times 1}{y_t} = \underset{1\times 8}{Z}\,\alpha_t \tag{A.1}$$

and transition equation:

$$\underset{8\times 1}{\alpha_{t+1}} = \begin{bmatrix} \mu_{t+1} \\ \nu \\ \gamma_{t+1} \\ \vdots \\ \gamma_{t+1-5} \end{bmatrix} = \underset{8\times 8}{T}\,\alpha_t + \underset{8\times 3}{R}\,\underset{3\times 1}{u_{t+1}} \qquad u_t \sim N(0,\underset{3\times 3}{Q}) \tag{A.2}$$

where α_t denotes the state vector and y_t the noisy measurement. The system matrices are defined as:

$$Z = \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & 0 \end{bmatrix},$$

$$T = diag\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & -1 & \cdots & -1 & -1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \right),$$

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_\omega^2 \end{bmatrix}$$

A.2 NI-UC Model

In the following, the nonlinear UC model defined by equations 3.1, 3.3 to 3.7 is cast into state form. Note, that the state space form is not uniquely identified and multiple versions provide identical results.

The measurement equation is defined as:

$$\underset{1\times1}{y_t} = \underset{1\times8}{Z} \alpha_t. \tag{A.3}$$

The transition equation is give by:

$$\underset{8\times1}{\alpha_{t+1}} = \begin{bmatrix} \mu_{t+1} \\ \nu_0 \\ \gamma_{t+1} \\ \vdots \\ \gamma_{t+1-5} \end{bmatrix} = \underset{8\times8}{T} \alpha_t + \underset{8\times1}{\lambda_{S_t}} + \underset{8\times3}{R} \underset{3\times1}{u_{t+1}} \qquad u_t \sim N(0, \underset{3\times3}{Q_{S_t}}) \tag{A.4}$$

where

$$Z = \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & 0 \end{bmatrix},$$

$$T = diag\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & -1 & \cdots & -1 & -1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \right),$$

37

$$
\lambda_{S_t} = \begin{bmatrix} S_t \nu_1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix}, \quad Q_{S_t} = \begin{bmatrix} \sigma^2_{\xi, S_t} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma^2_\omega \end{bmatrix}
$$

(Kim, Nelson et al., 1999, ch. 5).

A.3 E-Nl-UC Model

The nonlinear UC model with endogenous switching, defined by equations 3.1, 3.3 to 3.5 and 3.8 to 3.14 is cast into state space form with the measurement equation given by A.3 and transition equation as A.4. The differences between the Nl-UC and E-Nl-UC models are not reflected in the structures of measurement and transition equations (see section 3.3 and equation 3.11).

B Kim recursions

B.1 Kim filter

Notation is identical to appendix A. Conditional on $S_{t-1} = i$ and $S_t = j$, the Kalman recursions within the Kim filter are as follows:

$$
\begin{aligned}
a_{t|t-1}^{(i,j)} &= \lambda_j + T a_{t-1|t-1}^j, \\
P_{t|t-1}^{(i,j)} &= T P_{t-1|t-1}^j T' + R Q_j R', \\
v_{t|t-1}^{(i,j)} &= y_t - Z a_{t|t-1}^{(i,j)}, \\
F_{t|t-1}^{(i,j)} &= Z P_{t|t-1}^{(i,j)} Z', \\
a_{t|t}^{(i,j)} &= a_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)} Z' F_{t|t-1}^{(i,j)\,-1} v_{t|t-1}^{(i,j)}, \\
P_{t|t}^{(i,j)} &= P_{t|t-1}^{(i,j)} - P_{t|t-1}^{(i,j)} Z' F_{t|t-1}^{(i,j)\,-1} Z P_{t|t-1}^{(i,j)}
\end{aligned}
\tag{B.1}
$$

where $a_t^{(i,j)}$ denotes the mean vector and $P_t^{(i,j)}$ the variance matrix of the state vector $\alpha_t^{(i,j)}$, $v_{t|t-1}^{(i,j)}$ the one-step-ahead prediction error and $F_{t|t-1}^{(i,j)}$ the conditional variance of the prediction error, all conditional on j and i. For the L-UC model, only equations B.1 are relevant with $\lambda_j = 0$ and irrespective of $S_{t-1} = i$ and $S_t = j$. Note, that in case of an irregular term in the measurement equation A.3, the variance of this stochastic component would be added to $F_{t|t-1}^{(i,j)}$ in B.1. In case of the En-Nl-UC model, endogeneity is already built into the B.1 recursions via the expectation of the irregular variable u_t. Subsequently the equations do not have to be adapted.

Given a transition matrix for the Markov chain or Probit covariates, regime probabilities are derived by:

$$
\Pr(S_t = j, S_{t-1} = i | \Psi_{t-1}) = \Pr(S_t = j | S_{t-1} = i) \Pr(S_{t-1} = i | \Psi_{t-1}),
$$
$$
f(y_t | \Psi_{t-1}, S_t = j, S_{t-1} = i) = (2\pi)^{-\frac{N}{2}} |F_{t|t-1}^{(i,j)}|^{-\frac{1}{2}}
$$
$$
\exp\left(-\frac{1}{2} v_{t|t-1}^{(i,j)\prime} F_{t|t-1}^{(i,j)}{}^{-1} v_{t|t-1}^{(i,j)}\right),
$$
$$
f(y_t, S_t = j, S_{t-1} = i | \Psi_{t-1}) = f(y_t | \Psi_{t-1}, S_t = j, S_{t-1} = i)
$$
$$
\Pr(S_t = j, S_{t-1} = i | \Psi_{t-1}),
$$
$$
f(y_t | \Psi_{t-1}) = \sum_{j=1}^{M} \sum_{i=1}^{M} f(y_t, S_t = j, S_{t-1} = i | \Psi_{t-1}),
$$
$$
\Pr(S_t = j | \Psi_t) = \sum_{i=1}^{M} \frac{f(y_t | \Psi_{t-1}, S_t = j, S_{t-1} = i)}{f(y_t | \Psi_{t-1})} Pr(S_t = j, S_{t-1} = i | \Psi_{t-1})
$$

$$(B.2)$$

with f denoting density and $f(y_t | \Psi_{t-1})$ also conveniently providing the likelihood score for period t. Note, that the likelihood is computed differently in the original Hamilton filter (see Hamilton (1989, ch. 4.2 step 3)). For the L-UC model, the likelihood is already given by $f(y_t | \Psi_{t-1}, S_t = j, S_{t-1} = i) = f(y_t | \Psi_{t-1})$, since regimes do not exist in this approach.

Employing the results of the adopted Hamilton filter recursions in equations B.2, it is possible to halt the exponential growth of computational demand with following approximations for the Kalman updating steps:

$$a_{t|t}^j = \frac{\sum_{i=1}^M \Pr(S_t = j, S_{t-1} = i|\Psi_t)a_{t|t}^{(i,j)}}{\Pr(S_t = j|\Psi_t)},$$

$$\tilde{\vartheta}_{t|t}^{(i,j)} = P_{t|t}^{(i,j)} + (a_{t|t}^j - a_{t|t}^{(i,j)})(a_{t|t}^j - a_{t|t}^{(i,j)})', \qquad \text{(B.3)}$$

$$P_{t|t}^j = \frac{\sum_{i=1}^M \Pr(S_t = j, S_{t-1} = i|\Psi_t)\tilde{\vartheta}_{t|t}^{(i,j)}}{\Pr(S_t = j|\Psi_t)}$$

see Kim, Nelson et al. (1999, ch. 5) for source.

B.2 Kim smoother

As already pointed out, the filter derives inference from all time periods up to the current one. The smoothing recursions draw on the entire observational period. Rather than iterating forwards, smoothing recursions run from T on backwards. With $S_t = j$ and $S_{t+1} = k$:

$$a_{t|T}^{(j,k)} = a_{t|t}^j + P_{t|t}^j F_k' P_{t+1|t}^{(j,k)^{-1}} (a_{t+1|T}^k - Ta_{t+1|t}^{(j,k)}),$$

$$P_{t|T}^{(j,k)} = P_{t|t}^j + P_{t|t}^j T' P_{t+1|t}^{(j,k)^{-1}} (P_{t+1|T}^k - P_{t+1|t}^{(j,k)}) P_{t+1|t}^{(j,k)^{-1}} {}' T P_{t|t}'^j. \qquad \text{(B.4)}$$

Smoothed regime probabilities are calculated as:

$$\Pr(S_t = j, S_{t+1} = k|\Psi_T) \approx \frac{\Pr(S_{t+1} = k|\Psi_T)\Pr(S_t = j|\Psi_t)\Pr(S_{t+1} = k|S_t = j)}{\Pr(S_{t+1} = k|\Psi_t)},$$

$$\Pr(S_t = j|\Psi_T) = \sum_{k=1}^M \Pr(S_t = j, S_{t+1} = k|\Psi_T).$$

$$\text{(B.5)}$$

Using these smoothed probabilities, one can again construct weighted averages of the smoothed state vector and according variance-covariance matrix:

$$a_{t|T}^j = \frac{\sum_{k=1}^M \Pr(S_t = j, S_{t+1} = k|\Psi_T) a_{t|T}^{j,k}}{\Pr(S_t = j|\Psi_T)},$$

$$a_{t|T} = \sum_{j=1}^M \Pr(S_t = j|\Psi_T) a_{t|T}^j,$$

$$\tilde{\vartheta}_{t|T}^{(j,k)} = P_{t|T}^{(j,k)} + (a_{t|T}^j - a_{t|T}^{(j,k)})(a_{t|T}^j - a_{t|T}^{(j,k)})', \qquad \text{(B.6)}$$

$$P_{t|T}^j = \frac{\sum_{k=1}^M \Pr(S_t = j, S_{t+1} = k|\Psi_T) \tilde{\vartheta}_{t|T}^{(j,k)}}{\Pr(S_t = j|\Psi_T)}$$

see Kim, Nelson et al. (1999, ch. 5).

B.3 Forecasting the Kim filter

To predict future values of the state vector, the filter is simply iterated forwards without updating the prediction steps. For $S_{T+n} = g$ and $S_{T+n+1} = l$ with a non-negative integer $n = \{0, 1, ...\}$:

$$a_{T+n+1|T}^{(g,l)} = \lambda_l + T a_{T+n|T}^g,$$

$$P_{T+n+1|T}^{(g,l)} = T P_{T+n|T}^g T' + R Q_g R', \qquad \text{(B.7)}$$

$$F_{T+n+1|T}^{(g,l)} = Z P_{T+n+1|T}^{(g,l)} Z'.$$

Regime probabilities are predicted to follow the underlying Markov process or in case of the E-Nl-UC model, according to their Probit specification.

Innovations are not predictable by definition. Subsequently, one might assume that endogenous switching does not play a role in forecasting. However, since the Kim filter is iterated forwards conditional on certain realizations for S_{T+n} and S_{T+n+1}, expectations for future innovations in the Probit model are non-zero (see equation 3.13). The joint distribution of system innovations in equation 3.11 still has to be considered for the endogenous model.

Regime probabilities are iterated forwards as:

$$\Pr(S_{T+n+1} = l, S_{T+n} = g|\Psi_T) = \Pr(S_{T+n+1} = l|S_{T+n} = g)\Pr(S_{T+n} = g|\Psi_T),$$

$$\Pr(S_{T+n+1} = l|\Psi_T) = \sum_{g=1}^{M} \Pr(S_{T+n+1} = l, S_{T+n} = g|\Psi_T). \tag{B.8}$$

These probabilities are exploited once again to approximate forecasted values:

$$a_{T+n+1|T}^{l} = \frac{\sum_{g=1}^{M} \Pr(S_{T+n+1} = l, S_{T+n} = g|\Psi_T)a_{T+n+1|T}^{(g,l)}}{\Pr(S_{T+n+1} = l|\Psi_T)},$$

$$a_{T+n+1|T} = \sum_{l=1}^{M} \Pr(S_{T+n+1} = l|\Psi_T)a_{T+n+1|T}^{l},$$

$$\tilde{\vartheta}_{T+n+1|T}^{(l,g)} = P_{T+n+1|T}^{(l,g)} + (a_{T+n+1|T}^{l} - a_{T+n+1|T}^{(g,l)})(a_{T+n+1|T}^{l} - a_{T+n+1|T}^{(g,l)})',$$

$$P_{T+n+1|T}^{l} = \frac{\sum_{g=1}^{M} \Pr(S_{T+n+1} = l, S_{T+n} = g|\Psi_T)\tilde{\vartheta}_{T+n+1|T}^{(l,g)}}{\Pr(S_t = l|\Psi_T)}. \tag{B.9}$$

Note, recursions B.7, B.8 and B.9 without source, but derived from filtering recursions in B.1.
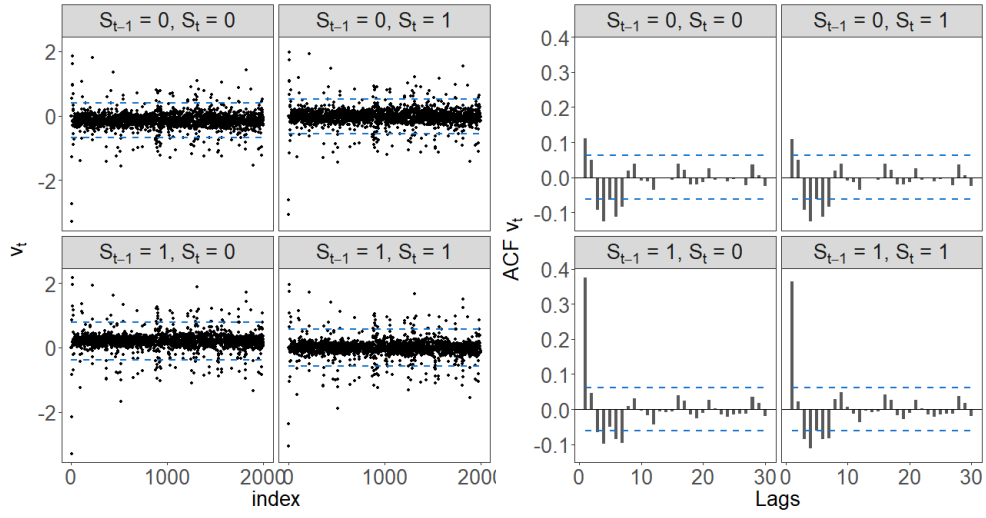
C Supplementary material

Figure 8: Figure concerns data simulated for the $T = 2000$ Monte Carlo case study in section 4. The figure on the left hand side presents one-step-ahead prediction errors $\hat{v}_t^{(i,j)}$ for the En-Nl-UC model together with two standard deviations (dashed, blue). The figure on the right hand side displays the ACF of the prediction errors with 95% confidence intervals (dashed, blue). Each panel gives the result conditional on a specific unobserved regime switch.
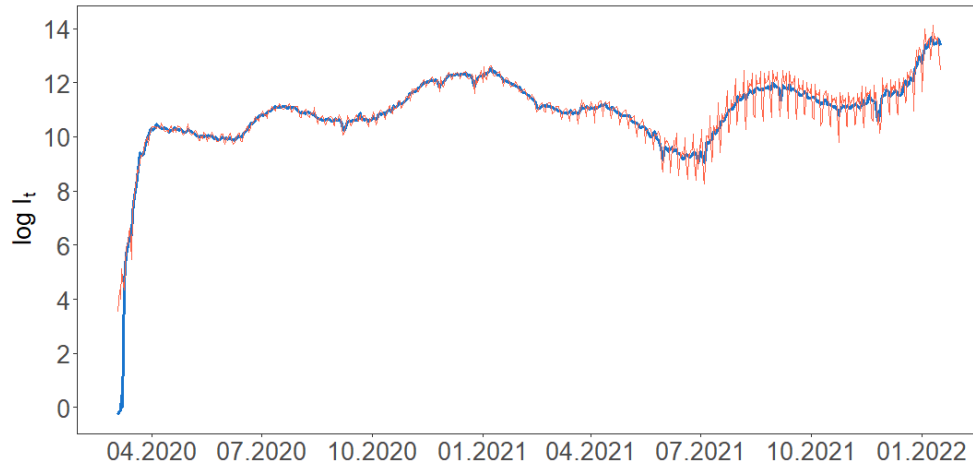

Figure 9: Smoothed trend estimates (blue) of the L-UC model and original $\log i_t$ series (orange).
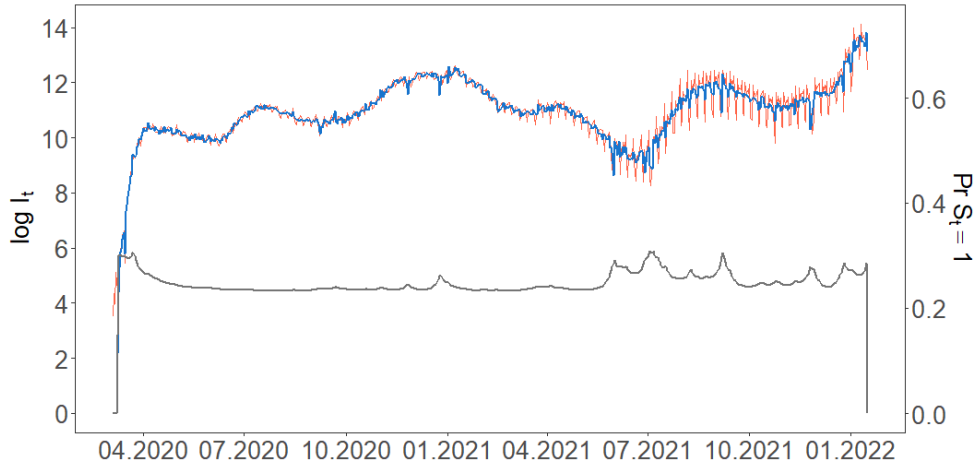
Figure 10: Smoothed trend estimates (blue) and regime probabilities for $S_t = 1$ of the Nl-UC model (grey). Figure also includes the original $\log i_t$ series (orange).
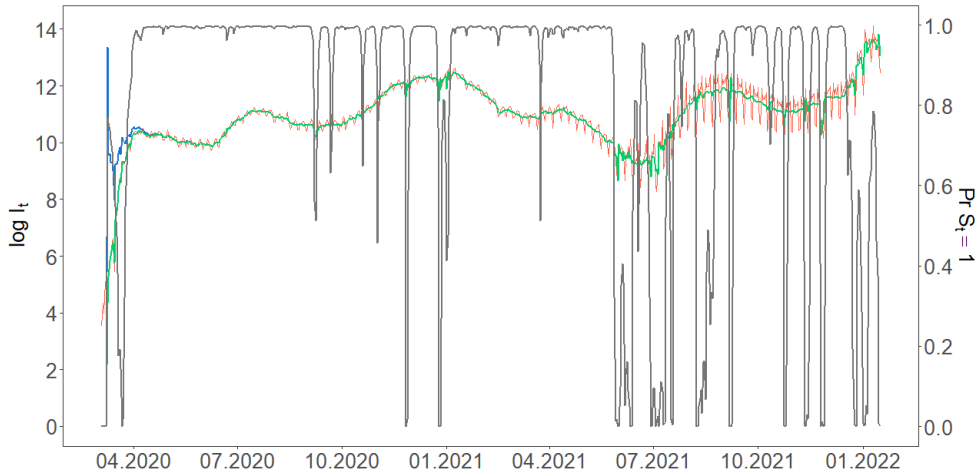

Figure 11: Smoothed trend estimates (blue), smoothed trend and cycle (green) and regime probabilities for $S_t = 1$ (grey) of the En-Nl-UC model incl. an uncorrelated stationary AR(2) cyclical component. Figure also includes the original $\log i_t$ series (orange).

44

References

Albert, J. H., and S. Chib (1993), "Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts," *Journal of Business & Economic Statistics*, 11(1), 1–15. (cited on page 10).

Bazzi, M., F. Blasques, S. J. Koopman, and A. Lucas (2017), "Time-varying transition probabilities for Markov regime switching models," *Journal of Time Series Analysis*, 38(3), 458–478. (cited on page 11).

Campbell, S. D. (2002), "Specification testing and semiparametric estimation of regime switching models: An examination of the us short term interest rate," Tech. rep., Working Paper. (cited on page 19).

Castle, J. L., J. A. Doornik, and D. F. Hendry (2021), "The value of robust statistical forecasts in the Covid-19 pandemic," *National Institute Economic Review*, 256, 19–43. (cited on page 2).

Chang, Y., Y. Choi, and J. Y. Park (2017), "A new approach to model regime switching," *Journal of Econometrics*, 196(1), 127–143. (cited on pages 8, 9, 19 and 29).

Clark, P. K. (1987), "The cyclical component of US economic activity," *The Quarterly Journal of Economics*, 102(4), 797–814. (cited on page 4).

Coccia, M. (2021), "The impact of first and second wave of the COVID-19 pandemic in society: comparative analysis to support control measures to cope with negative effects of future infectious diseases," *Environmental Research*, 197, 111099. (cited on page 1).

Das, S., S. Das, and M. Ghangrekar (2020), "The COVID-19 pandemic: biological evolution, treatment options and consequences," *Innovative Infrastructure Solutions*, 5(3), 1–12. (cited on page 13).

Dolton, P. (2021), "The statistical challenges of modelling Covid-19," *National Institute Economic Review*, 257, 46–82. (cited on page 2).

Dong, E., H. Du, and L. Gardner (2020), "An interactive web-based dashboard to track COVID-19 in real time," *The Lancet infectious diseases*, 20(5), 533–534. (cited on page 3).

Doornik, J. A., J. L. Castle, and D. F. Hendry (2020), "Short-term forecasting of the coronavirus pandemic," *International Journal of Forecasting.* (cited on page 1).

Doornik, J. A., J. L. Castle, and D. F. Hendry (2021), "Modeling and forecasting the COVID-19 pandemic time-series data," *Social Science Quarterly.* (cited on pages 11, 12, 13, 32 and 33).

Durbin, J., and S. J. Koopman (2012), *Time series analysis by state space methods*, Oxford university press. (cited on pages 4, 5, 6, 10, 25, 31, 36 and 37).

Ferguson, N., D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg, et al. (2020), "Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand," . (cited on page 14).

Gecili, E., A. Ziady, and R. D. Szczesniak (2021), "Forecasting COVID-19 confirmed cases, deaths and recoveries: revisiting established time series modeling through novel applications for the USA and Italy," *Plos one*, 16(1), e0244173. (cited on pages 1, 11 and 33).

Goldfeld, S. M., and R. E. Quandt (1973), "A Markov model for switching regressions," *Journal of econometrics*, 1(1), 3–15. (cited on page 7).

Gray, S. F. (1996), "Modeling the conditional distribution of interest rates as a regime-switching process," *Journal of financial economics*, 42(1), 27–62. (cited on page 11).

Hamilton, J. D. (1989), "A new approach to the economic analysis of non-stationary time series and the business cycle," *Econometrica: Journal of the econometric society*, 357–384. (cited on pages 3, 5, 6, 7, 8, 9, 11, 15, 31 and 39).

Hamilton, J. D. (1994), *Time series analysis*, Princeton university press. (cited on pages 7, 9, 11, 16, 22, 24, 26, 27, 31 and 32).

Hamilton, J. D. (2018), "Why you should never use the Hodrick-Prescott filter," *Review of Economics and Statistics*, 100(5), 831–843. (cited on page 7).

Hartl, T. (2021), "Monitoring the pandemic: A fractional filter for the COVID-19 contact rate," *arXiv preprint arXiv:2102.10067.* (cited on pages 13, 14 and 17).

Harvey, A., and P. Kattuman (2020), "Time series models based on growth curves with applications to forecasting coronavirus," *Harvard Data Science Review.* (cited on page 2).

Harvey, A. C. (1985), "Trends and cycles in macroeconomic time series," *Journal of Business & Economic Statistics*, 3(3), 216–227. (cited on page 16).

Harvey, A. C. (1990), "Forecasting, structural time series models and the Kalman filter," . (cited on pages 4 and 6).

Harvey, A. C. (2021), "Time series modeling of epidemics: leading indicators, control groups and policy assessment," . (cited on pages 1 and 11).

Hortaçsu, A., J. Liu, and T. Schwieg (2021), "Estimating the fraction of unreported infections in epidemics with a known epicenter: an application to covid-19," *Journal of econometrics*, 220(1), 106–129. (cited on page 2).

Jahangir, M. A., A. Muheem, M. F. Rizvi, et al. (2020), "Coronavirus (COVID-19): history, current knowledge and pipeline medications," *Int J Pharm Pharmacol 2020; 4: 140. doi: 10.31531/2581*, 3080(2). (cited on page 13).

Kalman, R. E. (1960), "A new approach to linear filtering and prediction problems," . (cited on page 4).

Kang, K. H. (2014), "Estimation of state-space models with endogenous Markov regime-switching parameters," *The Econometrics Journal*, 17(1), 56–82. (cited on pages 8, 16, 17 and 22).

Kim, C.-J. (1994), "Dynamic linear models with Markov-switching," *Journal of Econometrics*, 60(1-2), 1–22. (cited on pages 2, 3, 9 and 10).

Kim, C.-J., C. R. Nelson, et al. (1999), "State-space models with regime switching: classical and Gibbs-sampling approaches with applications," *MIT Press Books*, 1. (cited on pages 10, 16, 18, 26, 38, 40 and 41).

Kim, C.-J., J. Piger, and R. Startz (2008), "Estimation of Markov regime-switching regression models with endogenous switching," *Journal of Econometrics*, 143(2), 263–273. (cited on pages 8, 9, 17, 18, 19, 20, 22 and 26).

Klinger, S., and E. Weber (2016), "Detecting unemployment hysteresis: a simultaneous unobserved components model with Markov switching," *Economics Letters*, 144, 115–118. (cited on page 3).

Kraemer, M. U., C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, O. C.-. D. W. Group†, L. du Plessis, N. R. Faria, R. Li, et al. (2020), "The effect of human mobility and control measures on the COVID-19 epidemic in China," *Science*, 368(6490), 493–497. (cited on page 1).

Lam, P.-s. (1990), "The hamilton model with a general autoregressive component: Estimation and comparison with other models of economic time series: Estimation and comparison with other models of economic time series," *Journal of Monetary Economics*, 26(3), 409–432. (cited on page 3).

Li, H., S. Zheng, F. Liu, W. Liu, and R. Zhao (2021), "Fighting against COVID-19: innovative strategies for clinical pharmacists," *Research in Social and Administrative Pharmacy*, 17(1), 1813–1818. (cited on page 13).

Liu, L., H. R. Moon, and F. Schorfheide (2021), "Panel forecasts of country-level Covid-19 infections," *Journal of econometrics*, 220(1), 2–22. (cited on pages 11, 24 and 35).

Lu, H.-M., D. Zeng, and H. Chen (2009), "Prospective infectious disease outbreak detection using Markov switching models," *IEEE Transactions on Knowledge and Data Engineering*, 22(4), 565–577. (cited on page 16).

Moore, S., E. M. Hill, M. J. Tildesley, L. Dyson, and M. J. Keeling (2021), "Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study," *The Lancet Infectious Diseases*, 21(6), 793–802. (cited on page 1).

Morley, J. C., C. R. Nelson, and E. Zivot (2003), "Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different?" *Review of Economics and Statistics*, 85(2), 235–243. (cited on pages 6 and 30).

Oliver, N., E. Letouzé, H. Sterly, S. Delataille, M. De Nadai, B. Lepri, R. Lambiotte, R. Benjamins, C. Cattuto, V. Colizza, et al. (2020), "Mobile phone data and COVID-19: Missing an opportunity?" *arXiv preprint arXiv:2003.12347*. (cited on page 1).

Qiu, Y., X. Chen, and W. Shi (2020), "Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China," *Journal of Population Economics*, 33(4), 1127–1172. (cited on page 1).

Ricon-Becker, I., R. Tarrasch, P. Blinder, and S. Ben-Eliyahu (2020), "A seven-day cycle in COVID-19 infection, hospitalization, and mortality rates: Do weekend social interactions kill susceptible people?" *medRxiv*. (cited on pages 2 and 13).

Shumway, R. H., and D. S. Stoffer (1991), "Dynamic linear models with switching," *Journal of the American Statistical Association*, 86(415), 763–769. (cited on page 3).

Sinclair, T. M. (2009), "Asymmetry in the business cycle: Friedman's plucking model with correlated innovations," *Studies in Nonlinear Dynamics & Econometrics*, 14(1). (cited on pages 16 and 19).

Telenti, A., A. Arvin, L. Corey, D. Corti, M. S. Diamond, A. García-Sastre, R. F. Garry, E. C. Holmes, P. S. Pang, and H. W. Virgin (2021), "After the pandemic: perspectives on the future trajectory of COVID-19," *Nature*, 596(7873), 495–504. (cited on page 1).

Yilmazkuday, H. (2021), "Stay-at-home works to fight against COVID-19: International evidence from Google mobility data," *Journal of Human Behavior in the Social Environment*, 31(1-4), 210–220. (cited on page 1).

Schriftliche Versicherung

Ich versichere, dass die Arbeit von mir selbständig verfasst wurde und dass ich keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Weiterhin wurde diese Arbeit keiner anderen Prüfungsbehörde übergeben. Die elektronische Ausfertigung der Arbeit habe ich bereits beim Prüfer eingereicht.

Regensburg, den 31. Januar 2022