Estimating monthly labour force figures during the COVID-19 pandemic in the Netherlands

Jan van den Brakel¹,² | Martijn Souren³ | Sabine Krieg¹

¹Department of Statistical Methods, Statistics Netherlands, Heerlen, The Netherlands
²Department of Quantitative Economics, Maastricht University, Maastricht, The Netherlands
³Department of Social Statistics, Statistics Netherlands, Heerlen, The Netherlands

Correspondence
Jan van den Brakel, Department of Statistical Methods, Statistics Netherlands, CBS-weg 11, 6412 EX, Heerlen, Netherlands.
Email: ja.vandenbrakel@cbs.nl

Abstract
Official monthly statistics about the Dutch labour force are based on the Dutch Labour Force Survey (LFS). The LFS is a continuously conducted survey that is designed as a rotating panel design. Data collection among selected households is based on a mixed-mode design that uses web interviewing, telephone interviewing and face-to-face interviewing. Monthly estimates about the labour force are obtained with a structural time series model. Due to the COVID-19 pandemic, face-to-face interviewing stopped. It was anticipated that this would have a systematic effect on the outcomes of the LFS and that the lockdown at the same time affected the real monthly labour force figures. The lockdown indeed marked a sharp turning point in the evolution of the series of the monthly labour force figures and strongly increased the volatility of these series. In this paper, it is explained how Statistics Netherlands produced monthly labour force figures during the COVID-19 pandemic. It is shown how the sudden change in the mode effects, because face-to-face interviewing stopped, were separated from real period-to-period changes in the labour force figures. It is also explained how the time series model is adapted to the increased volatility in the labour force figures.

Keywords
discontinuities, Kalman filter, measurement error, mode effects, state space models
INTRODUCTION

The purpose of the Dutch Labour Force Survey (LFS) is to publish reliable monthly, quarterly and annual figures about the Dutch labour force. The LFS is based on a rotating panel design. The responding households are interviewed five times at quarterly intervals, which implies that every month five waves are interviewed. In 2010, Statistics Netherlands implemented a model-based estimation procedure based on a multivariate structural time series model (Durbin & Koopman, 2012) to produce monthly figures about the labour force (van den Brakel & Krieg, 2015; Pfeffermann, 1991). This model uses five series of general regression (GREG) estimates (Särndal et al., 1992) observed in the waves of the rotating panel as input. This time series model is used as a form of small area estimation to produce sufficiently reliable monthly estimates, despite the relatively small monthly sample sizes. The model also accounts for systematic differences between the outcomes obtained in the subsequent waves, known in the literature as rotation group bias (RGB), (Bailar, 1975). Finally, the model accounts for systematic differences or discontinuities in the outcomes due to two major survey redesigns that took place in 2010 and 2012 (van den Brakel & Roels, 2010).

Data collection of the Dutch LFS is based on a sequential mixed-mode design that starts with Web Interviewing (WI). Non-respondents are followed up with computer-assisted telephone interviewing (CATI) if a listed telephone number is available or with computer-assisted personal interviewing (CAPI) otherwise.

Due to the COVID-19 pandemic, the Netherlands went into a lockdown on 16 March 2020. Due to this lockdown, CAPI stopped. Since this resulted in a sudden change in selection bias and measurement bias, this had a systematic effect on the outcomes of the LFS. At the same time, it could be expected that the lockdown would affect the real monthly labour force figures. The Unemployed Labour Force time series had shown a steady decrease since 2014, while the evolution of the Employed Labour Force as well as the Total Labour Force showed a steady increase during these last 6 years. The lockdown and its associated policy measures indeed marked a sharp turning point in the evolution of these series and strongly increased the volatility, that is induced larger period-to-period changes, in the labour force series, which was, according to model diagnostics, not sufficiently picked up by the time series model.

In this paper, two problems related to the lockdown are addressed. The first problem is how to separate the sudden change in the mode effects because CAPI stopped from real period-to-period changes in the labour force figures. Extending the time series model with a level intervention variable is not a good solution, since the lockdown also resulted in a turning point at exactly the same time. In that case, it could be expected that a major part of the real period-to-period change would incorrectly be absorbed in the estimate for the regression coefficient of the level intervention, that models the change in mode effects due to the loss of CAPI households. In this paper, three different approaches to correct for the loss of CAPI responses are proposed and compared. The second problem is the abrupt turning point and sudden increase of the volatility in the labour force figures due to the lockdown. Three different options are considered to accommodate this in the time series model. It is discussed how the different approaches to account for both problems were evaluated during the first month of the lockdown in order to choose the most promising method for the publication of official monthly labour force figures. Consequences of these choices are discussed by comparing the different methods in an analysis that was conducted in real time until December 2020.

In Section 2, the survey design of the Dutch LFS is described. The time series model used to produce official monthly figures about the labour force is described in Section 3. Several options
that were considered to account for the sudden change in mode effects because face-to-face interviewing stopped, are described in Section 4. In Section 5, several options for how the model can accommodate the sudden increase of the volatility of the population parameters are described. Results are presented in Section 6 and the paper finishes with a discussion in Section 7. The online supplemental file of this paper contains additional details of the survey design of the Dutch LFS. It also contains more details about the different options that were considered to model the changing mode effects that arose as a result of the loss of CAPI and a simulation that was conducted to make a choice between the different options. Finally, the supplement contains additional analysis results that support the results described in Section 6 of the main paper.

2 | DUTCH LABOUR FORCE SURVEY

The objective of the Dutch LFS is to provide reliable information about the Dutch labour force. Each month a stratified two-stage cluster sample of addresses is drawn. Strata are formed by geographical regions. Municipalities are considered as primary and addresses as secondary sampling units. All households residing at an address, up to a maximum of three, are included in the sample. Until October 1999, the LFS was designed as a cross-sectional survey. Since October 1999, the LFS has been conducted as a rotating panel design. Until the redesign in 2010, data in the first wave were collected by means of CAPI. Respondents were re-interviewed four times at quarterly intervals by means of CATI. The rotation scheme of the panel is illustrated in Table 1. Monthly samples are denoted by capital letters. It can be seen how a sample that enters at the first wave is interviewed after 3, 6, 9 and 12 months. After the fifth interview, the sample leaves the panel. It can also be seen that each month data are collected from five independent samples, corresponding to the five waves of the panel.

In a redesign that took place in 2010, the data collection in the first wave changed from CAPI to a mixed data collection mode using CAPI and CATI. Households with a listed telephone number are interviewed by telephone, the remaining households are interviewed face-to-face. In a second redesign in 2012, data collection in the first wave changed to a sequential mixed-mode design that starts with WI. After three reminders, the non-responding households are contacted by telephone if they have a listed telephone number. The remaining households are interviewed face-to-face. Data collection in the re-interviews remained CATI after both redesigns.

The aforementioned redesigns have systematic effects on the sample estimates of the survey. These so-called discontinuities were quantified by collecting data under the old and new design alongside each other for some period of time. This is referred to as parallel data collection or a parallel run. In the case of both redesigns, a parallel run was conducted for the first wave for a

| Table 1 | Rotating panel design Dutch Labour Force Survey (LFS). Monthly samples are denoted by capital letters |
|---------|--------------------------------------------------------------------------------|
| Month   | Wave | t | t +1 | t +2 | t +3 | t +4 | t +5 | t +6 | t +7 | t +8 | t +9 | t +10 | t +11 | t +12 | t +13 | t +14 |
| 1       | 1    | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
| 2       | 2    |   |   |   | A | B | C | D | E | F | G | H | I | J | K | L |   |
| 3       | 3    |   |   |   |   | A | B | C | D | E | F | G | H | I |   |   |   |
| 4       | 4    |   |   |   |   |   |   | A | B | C | D | E | F |   |   |   |   |
| 5       | 5    |   |   |   |   |   |   |   |   |   |   |   | A | B | C |   |   |
period of 6 months where the sample sizes assigned to the old and new design were both equal to the regular sample size. Based on these parallel runs, direct estimates for the discontinuities were obtained and used in the time series model described in Section 3.

Key parameters of the LFS are the Employed, Unemployed and Total Labour Force, which are defined as population totals. These figures are estimated at a monthly frequency at the national level and a breakdown into six domains that is based on the cross classification of gender and age in three age classes. In the Dutch LFS, the ILO definitions for employed and unemployed are used. According to these definitions, people are classified as Employed if a person of working age has worked in paid employment or self-employment for at least 1 h a week or has a job from which he has been absent under conditions relating to the reason for absence (vacation, sick leave, maternity leave, etc.) The Unemployed comprise all persons of working age who were: (a) without work during the reference period that is were not in paid employment or self-employment; (b) currently available for work that is were available for paid employment or self-employment during the reference period; and (c) seeking work that is had taken specific steps in a specified recent period to seek paid employment or self-employment. The Labour Force is defined as the sum of the Employed and Unemployed Labour Force. People who are not classified as being Employed or Unemployed are classified as Inactive and do not belong to the Labour Force.

Due to the COVID-19 pandemic, the Netherlands went into a lockdown on 16 March 2020. CAPI interviewing stopped at that moment. In an attempt to obtain as many responses as possible from the households that were originally assigned to CAPI, households were sent a letter with the request to contact the interviewer assigned to them, so that they could call back to conduct the interview by telephone. In September 2020, CAPI was started up again.

An overview of the field work in the first wave of the Dutch LFS in 2019 is given in Table 2. Each month about 14,500 households were approached with the request to respond via WI. The average response after three reminders was about 24%. To increase the WI share in the total response, a large initial sample was drawn and a random sample of almost 50% among the 11,000 WI non-respondents was drawn and followed up using CATI or CAPI. As a result, the average composition of the response is 68% WI, 15% CATI and 17% CAPI. Note that the over-representation of WI respondents has been discounted in the inclusion weights of the GREG estimator. The inclusion weights account for all aspects of the sample design, including the aforementioned oversampling of the WI respondents. This avoids that this strategy introduces selection bias in the Labour Force figures.

The major part of the CAPI household sample in March was approached before the start of the lockdown and a response was obtained that was comparable with the CAPI response during 2019. It was decided to treat March as a normal month and to account for the loss of CAPI from April 2020 on. This decision was supported since the period-to-period change of the Unemployed Labour Force from February to March was in line with the claimant count figures if March was treated as a normal month with complete response for all three modes.

| Mode | Approach households | Response | Response rate (%) |
|------|---------------------|----------|-------------------|
| WI   | 14,500              | 3500     | 24                |
| CATI | 3000                | 750      | 25                |
| CAPI | 2200                | 880      | 40                |
3 | TIME SERIES MODEL FOR MONTHLY LABOUR FORCE FIGURES

Monthly figures about the labour force are based on a structural time series model. According to the rotation scheme of the panel design, households are interviewed five times at quarterly intervals. This implies that each month data are collected in five independent samples. For each sample direct estimates for the monthly Employed Labour Force, Unemployed Labour Force and Total Labour Force were obtained with the general regression (GREG) estimator (Särndal et al., 1992). Inclusion probabilities reflect the sampling design and differences in response rates between geographic regions. The weighting scheme of the GREG estimator is based on a combination of different socio-demographic categorical variables. Let \( \hat{y}_t \) denote the GREG estimate for the unknown population parameter, say \( \theta_t \), based on the \( j \)-th wave observed in month \( t \), \( j = 1, \ldots, 5 \). Since responding households are interviewed at quarterly intervals, it follows that the \( j \)-th wave interviewed in month \( t \) was sampled for the first time in month \( t - 3j + 3 \). The GREG estimates of each month can be expressed as a vector \( \hat{y}_t = (\hat{y}_t^{[1]}, \ldots, \hat{y}_t^{[5]})' \). This five-dimensional time series with GREG estimates is the input for a multivariate structural time series model, initially proposed by Pfeffermann (1991):

\[
\hat{y}_t = j[5] \theta_t + \lambda_t + \Delta_t^{[1]} \beta_t^{[1]} + \Delta_t^{[2]} \beta_t^{[2]} + e_t,
\]

with \( j[5] \) a five-dimensional column vector with each element equal to one, \( \lambda_t = (\lambda_t^{[1]}, \lambda_t^{[2]}, \lambda_t^{[3]}, \lambda_t^{[4]}, \lambda_t^{[5]})' \) a vector with time-dependent components that account for the RGB, \( \Delta_t^{[i]} = \text{Diag}(\delta_t^{[i,1]}, \delta_t^{[i,2]}, \delta_t^{[i,3]}, \delta_t^{[i,4]}, \delta_t^{[i,5]}) \) a diagonal matrix with dummy variables that change from zero to one at the moment that the survey changes from the old to the new design during redesign \( i = 1 \) in 2010 and \( i = 2 \) in 2012, \( \beta_t^{[i]} = (\beta_t^{[i,1]}, \beta_t^{[i,2]}, \beta_t^{[i,3]}, \beta_t^{[i,4]}, \beta_t^{[i,5]})' \) a five-dimensional vector with regression coefficients, \( i = 1, 2 \), and \( e_t = (e_t^{[1]}, e_t^{[2]}, e_t^{[3]}, e_t^{[4]}, e_t^{[5]})' \) the corresponding survey errors for each wave estimate. The information obtained during the parallel run in the first wave is used as a priori information in the model and the regression \( \beta_t^{[i,1]} \) is made time-varying during the period of the parallel run, \( i = 1, 2 \). See van den Brakel and Krieg (2015) for details of Model (1).

The population parameter \( \theta_t \) in Equation (1) can be decomposed into a trend component, a seasonal component, and an irregular component, that is

\[
\theta_t = L_t + S_t + \epsilon_t.
\]

In Equation (2), \( L_t \) is the so-called smooth trend model, which is defined as

\[
L_t = L_{t-1} + R_{t-1}
\]

\[
R_t = R_{t-1} + \eta_t \quad \eta_t \sim \mathcal{N}(0, \sigma^2_{\eta}).
\]

Furthermore, \( S_t \) in Equation (2) denotes a trigonometric stochastic seasonal component, see Durbin and Koopman (2012), Chapter 3 for an expression. Finally, \( \epsilon_t \) in Equation (2) denotes the irregular component, which contains the unexplained variation of the population parameter and is modelled as a white noise process, that is \( \epsilon_t \sim \mathcal{N}(0, \sigma^2_{\epsilon}) \).

The systematic differences between the subsequent waves, that is the RGB, are modelled in Equation (1) with \( \lambda_t \). The absolute bias in the monthly labour force figures cannot be estimated from the sample data only. Therefore, additional restrictions for the elements of \( \lambda_t \) are required to identify the model. Here it is assumed that an unbiased estimate for \( \theta_t \) is obtained with the
first wave, that is $\lambda_{t}^{[1]}$, see van den Brakel and Krieg (2009) for a motivation. Although the data collection strategy in the first wave changed from a uni-mode using CAPI to a mixed mode using WI, CATI and CAPI, most of the arguments mentioned in van den Brakel and Krieg (2009) are still valid. The most important reasons are that (a) selective non-response is still higher in the follow-up waves due to panel attrition, (b) the use of a strongly condensed questionnaire in the follow-up waves that focus on changes in the market position of the respondents implies that respondents tend to indicate that nothing has changed, and (c) panel effects create systematic changes in the behaviour of the respondents in the panel. Assuming that the estimates in the first wave are unbiased, implies that the first component of $\lambda_{t}$ equals zero. The other elements of $\lambda_{t}$ measure the time-dependent differences with respect to the first wave and are modelled as random walks. As a result,

$$\lambda_{t}^{[1]} = 0, \quad \lambda_{t}^{[j]} = \lambda_{t-1}^{[j]} + \eta_{\lambda,t}^{[j]}, \quad j = 2, 3, 4, 5, \quad \eta_{\lambda,t}^{[j]} \sim \mathcal{N}(0, \sigma_{\lambda}^{2}).$$

(4)

Note that the disturbance terms of the random walks for the RGB in waves 2, 3, 4 and 5 share the same variance component $\sigma_{\lambda}^{2}$.

The discontinuities induced by the redesigns in 2010 and 2012 are modelled with the third and fourth terms in Equation (1). The diagonal matrix $\Delta_{t}^{[i]}$ contains five intervention variables:

$$\delta_{t}^{[i]} = \begin{cases} 0 & \text{if } t < T_{R_{i}}^{[i]} \\ 1 & \text{if } t \geq T_{R_{i}}^{[i]} \end{cases},$$

(5)

where $T_{R_{i}}^{[i]}$ denotes the moment that wave $j$ changes from the old to the new survey design in redesign $i = 1$ in 2010 or $i = 2$ in 2012. Under the assumption that Equation (2) correctly models the evolution of the population variable, the regression coefficients in $\hat{\beta}_{t}^{[i]}$ can be interpreted as the systematic effects of the redesign on the level of the series observed in the five waves.

The last component in Equation (1) is a time series model for the survey errors. The estimates for the design variances of the sampling errors are available from the micro data and are incorporated in the time series model using the survey error model $e_{t}^{[i]} = k_{t}^{[i]} e_{t}^{[i]}$, where $k_{t}^{[i]} = \sqrt{\text{var}(\hat{y}_{t}^{[i]}))}$. Here $\text{var}(\hat{y}_{t}^{[i]})$ denotes the estimated variance of the GREG estimators allows for heterogeneous variance in the survey errors that arise, for example due to the gradually changing sample sizes over time.

The survey errors of the first wave, $e_{t}^{[1]}$, are not correlated with survey errors in the past because each new wave that enters the panel is a random sample that covers the entire target population. It is, therefore, assumed that $e_{t}^{[1]} \sim \mathcal{N}(0, \sigma_{e_{t}^{[1]}}^{2})$. As a result, the variance of the survey error equals $\text{var}(e_{t}^{[1]}) = (k_{t}^{[1]})^{2} \sigma_{e_{t}^{[1]}}^{2}$, which is approximately equal to the direct estimate of the variance of the GREG estimate for the first wave if the maximum likelihood (ML) estimate for $\sigma_{e_{t}^{[1]}}^{2}$ is close to one.

The survey errors of the second, third, fourth and fifth waves are correlated with survey errors of preceding periods. The autocorrelations between the survey errors of the subsequent waves are estimated from the survey data, using the approach proposed by Pfeffermann et al. (1998). In this application, it appears that the autocorrelation structure for the second, third, fourth and fifth waves can be modelled conveniently with an AR(1) model, van den Brakel and Krieg (2015). Therefore, it is assumed that $e_{t}^{[i]} = \rho e_{t-1}^{[i]} + v_{t}^{[i]}$, with $\rho$ the first-order autocorrelation coefficient, and $v_{t}^{[i]} \sim \mathcal{N}(0, \sigma_{e_{t}^{[i]}}^{2})$ for $j = 2, 3, 4, 5$. Since $e_{t}^{[i]}$ is an AR(1) process, $\text{var}(e_{t}^{[i]}) = (k_{t}^{[i]})^{2} \sigma_{e_{t}^{[i]}}^{2}/(1 - \rho^{2})$. 

$$\text{var}(e_{t}^{[i]}) = \frac{(k_{t}^{[i]})^{2} \sigma_{e_{t}^{[i]}}^{2}}{1 - \rho^{2}}.$$

(6)
As a result, \( \text{var}(e_t^j) \) is approximately equal to \( \hat{\text{var}}(y_t^j) \) provided that the ML estimates for \( \sigma^2_{e_j} \) are close to \((1 - \rho^2)\).

The survey redesign in 2010 and 2012 might affect the variance of the GREG estimates. Systematic differences in these variances are automatically taken into account by this model. An alternative would be to allow for different values for \( \sigma^2_{ej} \) for the periods with different survey design, which can be interpreted as interventions on the variance hyperparameters of the survey errors.

The general way to proceed is to express the model in the so-called state space representation and apply the Kalman filter to obtain optimal estimates for the state variables, see e.g. Durbin and Koopman (2012). It is assumed that the disturbances are normally distributed. Under this assumption, the Kalman filter gives optimal estimates for the state vector and the signals. Estimates for state variables for period \( t \) based on the information available up to and including period \( t \) are referred to as the filtered estimates. The filtered estimates of past state vectors can be updated if new data become available. This procedure is referred to as smoothing and results in smoothed estimates that are based on the completely observed time series. In this application, interest is mainly focused on the filtered estimates, since they are based on the complete set of information that would have been available in the regular production process to produce a model-based estimate for month \( t \).

The analysis was conducted with software developed in OxMetrics in combination with the subroutines of SsfPack 3.0, see Doornik (2009) and Koopman et al. (2008). All state variables are non-stationary with the exception of the survey errors and the population white noise. The non-stationary variables were initialised with a diffuse prior, that is the expectation of the initial states were equal to zero and the initial covariance matrix of the states was diagonal with large diagonal elements. The stationary state variables were initialised with a proper prior. The initial values for the survey errors were equal to zero and the covariance matrix was derived from its assumed process. In Ssfpack 3.0, an exact diffuse log-likelihood function is obtained with the procedure proposed by Koopman (1997).

The following estimates for the monthly labour force are derived from this model, which are published as official monthly Labour Force figures. For the Unemployed Labour Force, Employed Labour Force and the Total Labour Force, a trend-cycle \( (L_t) \) and a trend-cycle plus seasonal \( (L_t + S_t) \) is published. In this paper, the trend-cycle is briefly referred to as the trend and the trend-cycle plus seasonal is further referred to as the signal. Trends and signals are estimated for these three target variables at the national level and for a breakdown in six domains based on the cross-classification of gender and three age classes. A Lagrange function is applied to ensure that the sum over the domain totals equals the national totals, and to ensure that the sum of the Employed and Unemployed Labour Force is exactly equal to the Total Labour Force, both at the national and domain levels. See van den Brakel and Krieg (2015) for details.

4  |  ESTIMATING THE CHANGE IN MODE EFFECTS

Suddenly stopping CAPI data collection because of the lockdown had a similar effect to the monthly estimates as the redesigns of the survey process in 2010 and 2012. It resulted in a sudden change of measurement bias and selection bias in the responses of the LFS and therefore had a systematic effect on the sample estimates. In a well-planned transition process, this would be anticipated by quantifying these discontinuities to avoid confounding real developments with
systematic effects induced by the redesign (van den Brakel et al., 2020). A safe approach to quantify discontinuities is to conduct a parallel run (see Section 2). Another approach is to quantify discontinuities by fitting a structural time series model, containing intervention variables to account for discontinuities, as explained in Section 3 for the two redesigns in 2010 and 2012. This approach is cost-effective, since no additional data collection is required, but relies on the strong assumption that during the change-over the time series model correctly describes the evolution of the population parameter. All deviations from this evolution are interpreted by the model as a discontinuity due to the redesign and are absorbed into the regression coefficient of the level intervention. If the change-over exactly coincides with a turning point, it can be expected that a part of the real period-to-period change will incorrectly be absorbed into the regression coefficient of the level intervention. This will result in a biased estimate for the discontinuity and biased model predictions for the population parameter.

As explained in Section 3, the population parameter estimates in the time series model are benchmarked to the level of the series observed in the first wave, since it is assumed that the RGB in the first wave equals zero. It is therefore crucial that the first wave is measured as accurately as possible, including possible discontinuities due to a redesign. Therefore, in 2010 and 2012, it was decided to allocate the available budget for a parallel run exclusively for the first wave. In both transitions, the first wave was conducted in parallel for a period of 6 months at the size of the regular survey. This resulted in sufficiently precise estimates for the discontinuities in the first wave, which were used as a priori information for \( \beta_1 \) and \( \beta_2 \) in Model (1). The regression coefficients of the intervention variables for the follow-up waves were estimated with the Kalman filter, see van den Brakel and Krieg (2015) for details.

In an attempt to reduce the loss of CAPI households, these households were sent additional letters to motivate them to contact Statistics Netherlands by telephone to complete the questionnaire via CATI. At the same time, different strategies to account for the change in mode effects in the estimation approach were developed. A standard intervention approach was not appropriate to estimate the discontinuity due to the loss of CAPI households directly after the lockdown, because the lockdown also has a strong effect on the real period-to-period change of the labour force figures. As will be shown later, the lockdown indeed marks a sharp turning point in the series of labour force figures. Three approaches were identified to separate a sudden change in mode effects from the real period-to-period change of the target variables.

**Option 1: Using CATI and WI respondents in wave 1 only**

Under this option, only the original WI and CATI respondents in wave 1 would be used for the period from April up until August 2020. All households originally assigned to CAPI would be ignored, including the CAPI households that agreed to participate via CATI. Model 1 would be extended with an additional component that models the discontinuity due to the loss of CAPI households in the first wave, that is

\[
\hat{y}_t = j_{[5]} \theta_0 + \lambda_t + \Delta_t^{[1]} \beta_t^{[1]} + \Delta_t^{[2]} \beta_t^{[2]} + \Delta_t^{[3]} \beta_t^{[3]} + e_t.
\]

The diagonal matrix \( \Delta_t^{[3]} \) would contain five intervention variables:

\[
\delta_t^{[3,j]} = \begin{cases} 
0 & \text{if } t \not\in [T_{LD} + (j-1) \ast 3, \ldots, T_{LD} + (j-1) \ast 3 + \tau] \\
1 & \text{if } t \in [T_{LD} + (j-1) \ast 3, \ldots, T_{LD} + (j-1) \ast 3 + \tau] 
\end{cases}, \quad j = 1, \ldots, 5.
\]
where $T_{LD}$ denotes April 2020, the month that CAPI response was completely missing for the first time and $\tau$ the number of months without CAPI respondents in the first wave. In this application $\tau = 5$, since CAPI was started again in September. Note that discontinuities were also expected in the follow-up waves, since the loss of CAPI households also has a selection effect in the follow-up waves. Furthermore, the vector $\beta^{[3]} = (\beta^{[3,1]}, \beta^{[3,2]}, \beta^{[3,3]}, \beta^{[3,4]}, \beta^{[3,5]})^T$ would contain the selection effects or discontinuities due to the loss of the CAPI households in the first wave.

As emphasised before, estimating the discontinuities $\beta^{[3]}$ due to the loss of CAPI by applying the Kalman filter to a state-space version of model (6), results in biased estimates since the start of the lockdown also marked a sharp turning point in the real evolution of the target variable. However, the discontinuities due to the loss of CAPI households could be estimated in a reliable way using a separate time series model. Starting with the first wave, two time series of direct estimates were constructed, one with and one without CAPI respondents. This could also be done for the follow-up waves; one time series based on the full response and one time series where the CAPI respondents from the first wave were left out. This series is available from 2012 which was the moment of the change-over to the sequential mixed-mode design based on WI, CATI and CAPI. In this way, for each wave, a parallel run of about 8 years could be created. It was, however, anticipated that the composition of the modes would gradually change over this long period, resulting in time-varying differences between the two series of GREG estimates. Both series were combined in a three-dimensional time series model, which also contained a series of claimant counts as an auxiliary series:

$$
\begin{pmatrix}
y^{[1]}_t \\
L^y_t \\
S^y_t
\end{pmatrix}
= \begin{pmatrix} 1 & T_j^d - T_{B_2} - 3(j - 1) - 1 \\
0 & 0 & 0
\end{pmatrix}
\begin{pmatrix} 0 \\
S^y_t \\
\kappa^{[1]}_t \\
L^x_t
\end{pmatrix}
+ \begin{pmatrix} e^y_t \\
e^x_t \\
e^{xy}_t
\end{pmatrix},
\quad t = T_{B_2}^j \ldots T_{LD}^j + 3(j - 1) - 1.
$$

This model was applied to the series of each wave separately and it is understood that $L^y_t$, $S^y_t$, $e^y_t$, and $e^{xy}_t$ are therefore wave dependent but that the superscripts that indicate the $j$-th wave are suppressed for notational convenience. In Equation (8) $\hat{Y}^{[1]}_j$ denotes the direct estimate in month $t$ based on the complete response, $\hat{Y}^{[1]'}_j$ the direct estimate in month $t$ based on the WI and CATI respondents in the first wave only, and $\kappa_t$ the claimant counts for month $t$. The two LFS series share the same smooth trend model $L^x_t$, as defined in Equation (3), and the same trigonometric seasonal component $S^x_t$. The systematic difference between $\hat{Y}^{[1]}_j$ and $\hat{Y}^{[1]'}_j$ is modelled with a random walk, $\kappa^{[1]}_t$, i.e.

$$
\kappa^{[1]}_t = \kappa^{[1]}_{t-1} + \eta_{\kappa,t}, \quad \eta_{\kappa,t} \sim N(0, \sigma^2_{\kappa}),
$$

and has in fact the same interpretation as the RGB in Equation (4). In the case that $\kappa^{[1]}_t$ follows a relative smooth pattern, the final value for $\kappa^{[1]}_t$ could be used as an estimate for the discontinuities $\beta^{[3,j]}$ in Equation (6). In case of an erratic pattern, a smooth trend model for $\kappa^{[1]}_t$ could be considered as an alternative. In this application, the values for $\sigma^2_{\kappa}$ tended to zero, which meant that for most variables $\kappa^{[1]}_t$ was a straight horizontal line or almost a straight horizontal line. The claimant counts series has its own smooth trend $L^x_t$, which is similarly defined as (3), and its own trigonometric seasonal component $S^x_t$. The claimant count series serves as an auxiliary series to obtain more precise estimates for $\kappa^{[1]}_t$. This is achieved by modelling the correlation, say $\rho$, between the
slope disturbance terms of $L_t^y$ and $L_t^x$. If $\eta^y_t$ and $\eta^x_t$ denote the slope disturbance terms for the LFS trend and the claimant counts trend in Equation (3), respectively, then it is assumed that

$$
\begin{pmatrix}
\eta^y_t \\
\eta^x_t 
\end{pmatrix} \approx \mathcal{N}
\begin{pmatrix}
0 \\
0 
\end{pmatrix},
\begin{pmatrix}
\sigma^2_{\eta^y} & \rho \sigma_{\eta^y} \sigma_{\eta^x} \\
\rho \sigma_{\eta^y} \sigma_{\eta^x} & \sigma^2_{\eta^x}
\end{pmatrix}.
$$

Finally, $e^y_t$, $e^x_t$ and $e^x_t$ are the measurement errors of the respective series $\hat{y}^j_t$, $\hat{y}^j_t^r$, and $x_t$. It is assumed that $e^y_t \sim \mathcal{N}(0, \sigma^2_y)$ and are not correlated with $e^x_t$ and $e^x_t$. The measurement errors of the LFS series are dominated by the sampling error and are also correlated since the WI and CATI respondents in $\hat{y}^j_t$ and $\hat{y}^j_t^r$ are the same persons. For $e^y_t$ and $e^x_t$, it is assumed that $e^y_t \sim \mathcal{N}(0, \operatorname{var}(\hat{y}^j_t) \sigma^2_y)$ and $e^x_t \sim \mathcal{N}(0, \operatorname{var}(\hat{y}^j_t^r) \sigma^2_x)$. Similarly to the production model, $\operatorname{var}(\hat{y}^j_t)$ and $\operatorname{var}(\hat{y}^j_t^r)$ were estimated using the sample data and were used as a priori information in the time series model. In Section 3 of the Supplement, it is shown that the covariance between $e^y_t$ and $e^x_t$ can be approximated as

$$
\hat{\operatorname{cov}}(\hat{y}^j_t, \hat{y}^j_t^r) = \frac{\sqrt{n_i}}{\sqrt{n_r}} \sqrt{\operatorname{var}(\hat{y}^j_t)} \sqrt{\operatorname{var}(\hat{y}^j_t^r)},
$$

with $n_t$ the sample size used to estimate $\hat{y}^j_t$, that is the complete sample in month $t$ including the CAPI respondents and $n_r$ the sample size used to estimate $\hat{y}^j_t^r$, that is the sample in month $t$ with the WI and CATI respondents only. The series used in Equation (8) start in the month in 2012 that wave $j$ changed to the sequential mixed mode design based on WI, CATI and CAPI, that is $t = T^j_{R_1}$, and continue until the last period that the complete response was observed, that is $t = (T_{LD} + 3 \ast (j - 1) - 1)$. The estimate for $\kappa^j_t$ obtained for the last period with complete response with model (8) was used as a priori available information for $\beta^{3,j}$ in model (6) via an exact initialisation of the Kalman filter.

The advantage of Option 1 is that there is a long series of 8 years available that can be used to estimate the discontinuity due to the loss of the CAPI response in the first wave. The disadvantage of this option is that the responses obtained from the CAPI respondents that agreed to complete a questionnaire by CATI, are not used.

**Option 2: Using CATI and WI respondents and CAPI households in wave 1 that respond by CATI**

Under this option, all responses obtained in wave 1 are used, that is the WI and CATI respondents and the respondents that were originally assigned to the CAPI mode, but agreed to participate via CATI. In this case, compared to Option 1, a different discontinuity would be introduced. The discontinuity in wave 1 is the net result of changing measurement bias, since a part of the CAPI households participate via CATI, and changing selection bias, since not all CAPI households will contact Statistics Netherlands to participate via CATI. The discontinuities in the follow-up waves are mainly the result of increased selection bias due to the loss of response of the CAPI households in the first wave.

Under this option, production Model (1) is extended with a component that accounts for the aforementioned discontinuities in an equivalent way to Model (6) under Option 1. The
discontinuities are estimated in a separate model, which is an extension of Model (8) and is defined as:

\[
\begin{pmatrix}
\hat{y}_{jt}^{(j)}
\end{pmatrix}
= \begin{pmatrix}
L_t^y
\end{pmatrix}
+ \begin{pmatrix}
\kappa_t^{(j)}
\end{pmatrix}
+ \begin{pmatrix}
\beta^{[3,j]}\delta^{[3,j]}_t
\end{pmatrix}
+ \begin{pmatrix}
S_t^y
\end{pmatrix}
+ \begin{pmatrix}
e_t^y
\end{pmatrix}, \quad t = T_R^j \ldots T,
\]

with \( \delta^{[3,j]}_t \) defined in Equation (7). The difference from Option 1 is that this model runs until the last available observation, instead of the moment that the CAPI households from the first wave are missing in wave \( j \). Furthermore, Model (8) has been extended with a level shift \( \beta^{[3,j]}\delta^{[3,j]}_t \), that estimates the break in wave \( j \) since from period \( T_{LD} + 3 \ast (j - 1) \), the CAPI households are missing or participate via CATI in the series of \( y'_t \). The estimate for \( \beta^{[3,j]} \) obtained with model (11) is used as a priori available information for \( \beta^{[3,1]} \) in Model (6) via an exact initialisation of the Kalman filter. Similar to Model (8), Model (11) is applied to the series of each wave separately. Therefore \( L_t^y, S_t^y, e_t^y, \) and \( e_t^{e^y} \) are in fact wave dependent.

The advantage of this approach is that the CAPI households that decided to participate via CATI are also used in the estimation of the monthly figures. The major drawback of this approach is that the estimates for \( \beta^{[3,j]} \) are highly unstable. In the first month of the lockdown, there is only one observation for \( \beta^{[3,1]} \) in the first wave. These estimates will be subject to large revisions if additional observations become available in the subsequent months. A possible solution is to use Option 1 for one quarter and change to Option 2 in the fourth month of the lockdown.

**Option 3: The follow-up waves are leading**

As emphasised in Section 3, the results obtained in the first wave are considered to be more reliable than the results in the follow-up waves. Therefore, the outcomes of the follow-up waves are benchmarked to the first wave in Equation (1) by assuming that the RGB in the first wave is zero. It can be argued that, due to the sudden loss of CAPI in the first wave, the outcomes of the follow-up waves are more reliable than the first wave. An alternative approach, therefore, is to allow RGB for the first wave and make one of the follow-up waves leading. The RGBs for the follow-up waves are not significantly different during the years before the lockdown. It is therefore a natural choice to set the RGBs for all follow-up waves equal to zero. This implies that the level of the population parameter \( \theta_t \) is benchmarked to the average of the series of the follow-up waves. In this case, the RGB in Model (6) is defined as \( \lambda_t = (\lambda_t^{[1]}, 0, 0, 0, 0)' \). Estimates for the discontinuities due to the lockdown, that is \( \beta^{[3]}_t \) in Equation (6) can be obtained using either the method proposed under Option 1 or Option 2.

To maintain uninterrupted series, the population parameter estimates must refer to the level of the first wave. This implies that under this option the published trend estimates for target variables are obtained by \( L_t + \lambda_t^{[1]} \) and signal estimates by \( L_t + S_t + \lambda_t^{[1]} \).

**Assumptions**

All three options require strong assumptions to disentangle the change in mode effects from the real period-to-period change. More precisely, it is assumed that the difference between the
outcomes of the first wave with and without CAPI is not affected by the lockdown. This assumption is required to project the estimated difference in the first wave with and without CAPI to the period during the lockdown. It is assumed that the WI and CATI response before and after the start of the lockdown are comparable. Similar assumptions are made for the follow-up waves. Finally, it is assumed that the autocorrelation in the survey errors and the seasonal component are not affected by the lockdown. More observations after the start of the lockdown are required to establish effects on the autocorrelation and the seasonal component. Under Option 3, it is also assumed that the RGB component for the first wave is not affected by the lockdown.

**Evaluation of the options**

A simulation study, described in Section 3 of the Supplement, was conducted to choose the most appropriate option. This simulation showed that Options 1 and 2 gave comparable results (compare Table S.1 and S.2 of the Supplement). Option 3 clearly performed worse than the other two options (see Table S.3 of the Supplement). The reason is that in Option 3 a different parameterisation of the RGB is used, which introduces an additional shock in the trend estimates. Differences between the direct estimates for the first wave with and without CAPI under Option 1 were obtained by applying Model (8) to the series observed from April 2012 until March 2020 and are given in Table S.4 of the Supplement. Based on the observations obtained in April 2020, the discontinuities under Option 2 were estimated by applying Model (11) to the series observed up until and including April 2020 and are given in Table S.5 of the Supplement. The number of CAPI households that agreed to participate via CATI was very low. As a result, the discontinuity under Option 2 was predominantly a selection bias, similar to Option 1. The point estimates under Option 2 were indeed very similar to the point estimates under Option 1. The standard errors under Option 1 were much smaller, since the estimates for the discontinuity were based on series with a length of 8 years while under Option 2 there is only one month available to estimate the discontinuity. Therefore, Option 1 is further considered to accommodate the loss of CAPI. See Section 3 of the Supplement for more details.

The first cohort without CAPI respondents entered the second wave in July. Model (8) was applied to the series of the second wave to estimate the impact of losing the CAPI respondents from the first wave. See Table S.6 of the Supplement for the results. In the third quarter of 2020 it was observed that the filtered trends and signals obtained under Model (6) with and without a level intervention for the second wave, that is $\delta_t^{[3,2]}$ and $\rho_t^{[3,2]}$ are almost the same. The reason is that the RGB in the second wave absorbs the loss of CAPI households in the second wave. In order to keep the production model as parsimonious as possible, it was decided to include a level intervention in the first wave only.

## 5 MODELLING INCREASED POPULATION VOLATILITY DURING THE COVID-19 PANDEMIC

The lockdown also had a strong impact on the real development of the Employed, Unemployed and Total Labour Force. After the recovery from the Great Financial Crisis, the Unemployed Labour Force showed a steady decreasing trend from 2014 until March 2020. The Employed Labour Force and Total Labour Force showed steady increasing trends. The start of the
lockdown marked a sharp turning point in the trend of these series in April 2020. Once the lockdown regulations were relaxed in June 2020, a partial recovery was visible.

It appeared in April 2020 that the trend component (3) was not flexible enough to adapt to the suddenly increased volatility of the population parameter in Model (6). This model mis-specification became visible since standardised innovations took values around 4 in absolute terms. The most straightforward approach to increase the flexibility of the trend component is to replace the smooth trend model (3) by a local linear trend model. This implies that a level disturbance term is added to the expression for $L_t$ in the smooth trend model (3) or see Durbin and Koopman (2012), Chapter 3.

The local linear trend model only marginally increased the flexibility of the trend component and did not solve the model mis-specification after the lockdown. This option was therefore not further considered.

Three different options to adapt the time series model to accommodate the increased volatility were considered. The first option is to model a shock in the time series component for the population parameter with a level intervention. This implies that (2) is extended with a similar level intervention component as used for the discontinuities in 2010 and 2012, that is

$$\theta_t = L_t + \beta^{COVID} \delta_t^{COVID} + S_t + \epsilon_t,$$

with $\delta_t^{COVID}$ a dummy indicator that changes from zero to one in April 2020 to model the shock induced by the lockdown and $\beta^{COVID}$ a regression component. This approach is appropriate if the impact of the lockdown is concentrated in one month, which is not a realistic assumption. Alternatively, $\delta_t^{COVID}$ can gradually change from zero to one over a period of several months, but at the start of the lockdown it is hard to anticipate what a reasonable pattern for $\delta_t^{COVID}$ could be. This option was therefore not further pursued.

The second option is to temporarily increase the flexibility of the trend. The flexibility of the trend is determined by the variance of the slope disturbance terms, that is $\sigma_\eta^2$ in Equation (3). The ML estimate for $\sigma_\eta^2$ directly after the lockdown in April 2020 is based on the volatility of the population parameter before the start of the COVID-19 crisis and therefore is not sufficiently capable of picking up the sharp turning point marked by the lockdown. One way to accommodate the increased volatility is to make the variance of the slope disturbance terms time-varying. This is achieved by multiplying this variance with a time-varying factor, the value of which is assumed to be known. This implies the following covariance structure for the slope disturbance terms in Equation (3):

$$\eta_t \sim \mathcal{N}(0, k_t \sigma_\eta^2) \quad (12)$$

where $k_t$ is set in advance. Before the start of the lockdown, $k_t = 1$. Two months before the lockdown, $k_t$ is temporarily increased such that the standardised innovations in the months after the lockdown have reasonable values, say a maximum of 2.2–2.5 in absolute terms for the month after the start of the lockdown and smaller than 2.0 for the preceding months. To minimise the amount of manual and purposeful adjusting of the time series estimates, the focus in the months after the lockdown is to bring the value for $k_t$ back to one as soon as possible. Another way to tune the values of $k_t$ is to monitor the values of the ML estimates of $\sigma_\eta^2$ computed in real time. Model (6) with a time constant variance for the slope disturbance terms will accommodate the increased volatility induced by the lockdown, by increasing the value for the ML estimate for $\sigma_\eta^2$. This means that values for estimates of $\sigma_\eta^2$ obtained with the series that end before the lockdown will be smaller than
estimates obtained with series that include observations during the lockdown. One way to tune the values for \( k_t \) is to increase them until the values of the ML estimates for \( \sigma_n^2 \) are comparable with the estimates obtained with series ending before the start of the lockdown.

Increasing the variance of the slope disturbance terms through factors \( k_t \) has the following interpretation. As the variance of the slope disturbance terms increases, the influence of more distant observations on the level of the trend becomes smaller. The proposed approach implies that the filtered estimates attach less weight on the prediction based on observations from the past and more weight on the direct estimates obtained in the latest month. This seems reasonable in periods where the world suddenly changes and becomes incomparable with the past, as was the case with the COVID-19 pandemic.

Because the five series of direct estimates of the separate waves are used as inputs for the model, shocks in the innovations of the trend can be separated from shocks in the sampling errors. A shock in a separate wave will be interpreted as a sampling error. A shock in all or almost all waves will be interpreted as a shock in the trend innovation. Also the autoregressive structure of the sampling errors helps to distinguish shocks in the sampling errors from shocks in the trend innovations.

The third option is to make the variance of the slope disturbance terms time-varying by distinguishing two periods which have their own value for the slope disturbance variances, that is

\[
\eta_t \sim \mathcal{N}(0, \sigma_{\eta,t}^2), \quad \sigma_{\eta,t}^2 = \begin{cases} \sigma_{\eta,1}^2 & \text{if } t \not\in [T_{LD} - 2, \ldots, T_{EL}] \\ \sigma_{\eta,2}^2 & \text{if } t \in [T_{LD} - 2, \ldots, T_{EL}] \end{cases}
\]

(13)

where \( T_{EL} \) denotes the end of the period where an increased or different value for the variance of the slope disturbance terms is required. The slope disturbance variances \( \sigma_{\eta,1}^2 \) and \( \sigma_{\eta,2}^2 \) are estimated with ML. This option is therefore not applicable directly after the start of the lockdown, since insufficient observations are available after the start of the lockdown. This method is therefore not considered for the production of official statistics during 2020, but it is interesting to compare outcomes under option 2 with this option as a form of evaluation, once sufficient observations are available to estimate \( \sigma_{\eta,2}^2 \). Another drawback of this approach is that it is difficult to determine the month \( t = T_{EL} \) where the period with an increased variance ends. This might require a suitable parametric model that allow the variance to decay gradually to the value of the variance before the start of the corona crisis, which is hard to determine in practice.

There are other options, not further investigated, but mentioned for the sake of completeness to accommodate the sudden increased volatility in the time series. The first option is to use non-Gaussian state space models, see Durbin and Koopman (2012), Part II. In particular, non-Gaussian distributions like horseshoe, Laplace or t-distributions can be considered for the trend disturbance terms to allow for more flexibility in the trend if sudden shocks in the series occur. See Carter and Kohn (1996), Carvalho et al. (2010), Polson and Scott (2010) and Tang et al. (2018) for more details of non-Gaussian random effects. Such an approach requires the structural time series model to be expressed as a time series multilevel model that is fitted in a Bayesian framework. See Knorr-Held and Rue (2002), Chan and Jeliazkov (2009), McCausland et al. (2011), Ruiz-Cárdenas et al. (2012) and Piepho and Ogutu (2014) for a more detailed discussion of the connections between structural time series models and multilevel models. These methods are not considered in this application since it was not feasible to develop and implement such a method in the production process of the monthly labour force figures in a time period of only one month; they are left for further research.
A second option is to apply methods from financial econometrics to accommodate time-varying variance. To model time-varying volatility in financial returns, generalised autoregressive conditional heteroscedasticity (GARCH) models are proposed in the financial econometrics literature. These models are not further considered in this paper, since they are intended for time series observed at a much higher frequency that are subject to continuously changing volatility patterns. It is not likely that GARCH models are the most appropriate method to model a sudden increase in volatility that occurs at the end of a time series with about 240 observations. Also from a practical point of view it was not feasible to change the state-space model currently implemented for the production of official monthly labour force figures for such an approach within a period of one month. See Francq and Zakoian (2010) and Harvey (2013) for an overview of GARCH models to account for time-varying volatility.

A third option is the method proposed by Pfeffermann and Tiller (2006), where a multivariate state space model is proposed for domains, and the sum over the domain predictions is benchmarked to the direct estimates at the national level. This method is ideally suited to have the model adequately pick up sudden turning points, if the sample size at the national level is large enough to estimate reliable direct monthly figures. This benchmark approach is not applicable to the Dutch case, since the state space model is also applied to estimate labour force figures at the national level.

When the data for April 2020 became available, it was decided to use the second option, that is temporarily increasing the flexibility of the trend via Model (12) to accommodate the increased population volatility in combination with Option 1 from Section 4 to accommodate the loss of CAPI respondents in the model used for the production of monthly labour force figures. In the next section, the effect of both adjustments in the model is evaluated by comparing the results under this model with the unadjusted production Model (1) and Model (6) where the flexibility of the trend is temporarily increased. The chosen option is also compared with a model where the trend has separate variance components for the slope disturbance terms for two different periods, that is Model (13).

6 | RESULTS

In this section, results obtained with the following models are compared:

- Model A: unadjusted production Model (1), that is a time constant variance structure (3) for the slope disturbance terms and no correction for the loss of CAPI respondents
- Model B: production Model (1) with a time-varying variance structure (12) for the slope disturbance terms but no correction for the loss of CAPI respondents
- Model C: Option 1 to account for the loss of CAPI in the first wave, through Model (6), in combination with a time-varying variance structure (12) for the slope disturbance terms
- Model D: Option 1 to account for the loss of CAPI in the first wave, through Model (6), in combination with a time-varying variance structure (13) for the slope disturbance terms

To motivate the need for a time-varying variance for the slope disturbance terms, some plots of the standardised innovations under Model A are given in Figure S.1 of the Supplement. The start of the lockdown is marked with large innovations (in absolute terms), indicating that in the second quarter of 2020 the model is seriously mis-specified. The ML estimates of the hyperparameters of Model A, computed in real time, are given in the upper part of Table 3.
**TABLE 3** Maximum likelihood (ML) estimates hyperparameters for the Unemployed Labour Force

| St. Dev. | January | February | March | April | May | June | July | August | September | October | November | December |
|----------|---------|----------|-------|-------|-----|------|------|--------|-----------|---------|----------|----------|
| Slope ($\sigma_\eta$) | 2001.39 | 1989.41 | 1952.95 | 1884.84 | 1909.50 | 2234.10 | 2586.59 | 2520.21 | 2314.84 | 2229.64 | 2166.99 | 2212.98 |
| Seas. ($\sigma_\omega$) | 131.64 | 123.05 | 114.73 | 193.76 | 213.49 | 315.48 | 297.41 | 296.22 | 320.41 | 314.69 | 399.05 | 417.42 |
| White N. ($\sigma_\varepsilon$) | 6393.87 | 6314.45 | 6381.10 | 6731.79 | 6735.15 | 6919.48 | 6413.08 | 6490.51 | 7101.74 | 7402.67 | 7477.07 | 7191.07 |
| RGB ($\sigma_\epsilon$) | 1655.89 | 1609.37 | 1660.87 | 1744.28 | 1738.18 | 1717.45 | 1679.24 | 1627.98 | 1702.91 | 1740.79 | 1720.89 | 1808.22 |
| S.Error W1 ($\sigma_{\epsilon_1}$) | 1.14 | 1.14 | 1.14 | 1.14 | 1.14 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 | 1.14 | 1.13 |
| S.Error W2 ($\sigma_{\epsilon_2}$) | 1.16 | 1.16 | 1.17 | 1.17 | 1.18 | 1.18 | 1.17 | 1.17 | 1.17 | 1.17 | 1.17 | 1.17 |
| S.Error W3 ($\sigma_{\epsilon_3}$) | 1.11 | 1.11 | 1.11 | 1.10 | 1.10 | 1.09 | 1.10 | 1.09 | 1.09 | 1.09 | 1.09 | 1.09 |
| S.Error W4 ($\sigma_{\epsilon_4}$) | 1.11 | 1.10 | 1.10 | 1.10 | 1.11 | 1.11 | 1.11 | 1.11 | 1.11 | 1.11 | 1.11 | 1.11 |
| S.Error W5 ($\sigma_{\epsilon_5}$) | 1.12 | 1.12 | 1.12 | 1.12 | 1.12 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 |

Model B (production model with time-varying variance for the slope disturbance terms)

| St. Dev. | January | February | March | April | May | June | July | August | September | October | November | December |
|----------|---------|----------|-------|-------|-----|------|------|--------|-----------|---------|----------|----------|
| Slope ($\sigma_\eta$) | 2001.39 | 1989.41 | 1950.02 | 1991.40 | 1961.76 | 2085.15 | 2013.43 | 1974.55 | 1992.40 | 2034.68 | 2204.15 | 2282.66 |
| $k_j$ | 10 | 50 | 50 | 10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| $\sqrt{k_j \sigma_\eta}$ | 6328.94 | 14067.28 | 13788.73 | 6297.38 | 1961.76 | 2085.15 | 2013.43 | 1974.55 | 1992.40 | 2034.68 | 2204.15 | 2282.66 |
| Seas. ($\sigma_\omega$) | 131.64 | 123.05 | 119.06 | 160.31 | 153.71 | 130.46 | 160.15 | 155.65 | 140.13 | 85.25 | 281.21 | 290.35 |
| White N. ($\sigma_\varepsilon$) | 6393.87 | 6314.45 | 6346.10 | 6189.41 | 6254.74 | 6445.23 | 6406.67 | 6621.62 | 7101.74 | 7391.57 | 7024.05 | 6691.52 |
| RGB ($\sigma_\epsilon$) | 1655.89 | 1609.37 | 1660.04 | 1730.38 | 1744.28 | 1754.77 | 1730.38 | 1654.24 | 1711.45 | 1735.58 | 1708.03 | 1795.60 |
| S.Error W1 ($\sigma_{\epsilon_1}$) | 1.14 | 1.14 | 1.14 | 1.14 | 1.14 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 |
| S.Error W2 ($\sigma_{\epsilon_2}$) | 1.16 | 1.16 | 1.17 | 1.17 | 1.17 | 1.17 | 1.17 | 1.17 | 1.17 | 1.17 | 1.17 | 1.17 |
| S.Error W3 ($\sigma_{\epsilon_3}$) | 1.11 | 1.11 | 1.11 | 1.10 | 1.10 | 1.11 | 1.10 | 1.10 | 1.10 | 1.10 | 1.10 | 1.10 |
| S.Error W4 ($\sigma_{\epsilon_4}$) | 1.11 | 1.10 | 1.10 | 1.10 | 1.11 | 1.11 | 1.11 | 1.11 | 1.11 | 1.11 | 1.11 | 1.11 |
| S.Error W5 ($\sigma_{\epsilon_5}$) | 1.12 | 1.12 | 1.12 | 1.12 | 1.12 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 | 1.13 |

Rho ($\rho$) | 0.38 | 0.38 | 0.38 | 0.38 | 0.38 | 0.38 | 0.38 | 0.38 | 0.38 | 0.38 | 0.38 | 0.38 |
for the Unemployed Labour Force at the national level for the twelve months during 2020. Estimating the hyperparameters in real time means that the estimates are based on the series observed up until the particular month in the Table. See Section S.4 of the Supplement for similar tables for the Employed and Total Labour Force. It can be seen that the model reacts to the increased volatility of the population parameter with increased values for the variance estimates of the slope disturbance terms ($\sigma^2_\eta$) and with some delay also the population white noise ($\sigma^2_\epsilon$).

Based on these findings, the variance of the slope disturbance term is made time-varying. The factors $k_t$ used in Models B and C are specified in the lower part of Table 3 for the Unemployed Labour Force. These values are minimally increased such that the standardised innovations are reduced until they have values just outside their admissible range of 1.96 in absolute terms. Examples of the standardised innovations for Model B are given in Figure S.1 of the Supplement. Note that the structure of the smooth trend model (3) implies that the innovations react with a lag of two months on adjustments of the flexibility factor $k_t$. Therefore, the value of $k_t$ is increased in February to avoid large standardised innovations in April. To avoid a large shock in the variance of the slope disturbance terms from January to February, the value for $k_t$ was also slightly increased in January. The values for $k_t$ were brought back to one as soon as possible, again by looking at the behaviour of the standardised innovations.

The ML estimates of the hyperparameters of Model B, computed in real time, are given in the lower part of Table 3 for the 12 months during 2020. The estimates for the variance of the slope disturbance terms are now more stable over time. There is still a slight increase that can be avoided by using larger values for $k_t$ for longer periods, but from the perspective of minimally increasing the $k_t$ values to have admissible values for the standardised innovations, this was not necessary and therefore not considered.

The ML estimates for the other hyperparameters are very similar under Model A and Model B. The ML estimates for the hyperparameters obtained with Model C, which also account for the loss of CAPI, are similar to the hyperparameter estimates for Model B. Also, the standardised innovations of both models are comparable with each other. These results are therefore omitted to save space.

In Figure 1, filtered trend estimates for the Unemployed, Employed and Total Labour Force under Models A, B and C are compared together with the GREG input series for the period from January 2018 until December 2020. It is understood that the few CAPI households that responded via CATI during the lockdown, were not used in the computations of the direct estimates of the input series. The figure for the Unemployed Labour Force also shows the claimant counts. As expected, the filtered trends under the three models are almost equal up until March 2020. Deviations start after the start of the lockdown in April 2020. For all variables, the effect of making the trend flexible is clearly larger (Model A vs. Model B and C) compared to the effect of compensating for the loss of CAPI (Model B vs. Model C). CAPI interviewing stopped in April and was restarted in September. As a result, the filtered trends under Models B and C are almost equal again in the last three months of 2020.

The need to accommodate the increased volatility and the loss of CAPI becomes clear if the filtered trend for the Unemployed Labour Force under Model A is compared with the claimant counts. In April there was a lot of uncertainty on the impact of COVID-19 on the Unemployed Labour Force. On the one hand, a sharp increase was to be expected due to the loss of jobs. On the other hand, people may not have been available for a job or were temporarily not searching for a job due to the restrictions. From previous economic downturns, we expect a strong
Figure 1  Trend estimates under Models A, B and C for the Unemployed, Employed and Total Labour Force with their input series from January 2018 until December 2020. CC in the top panel is the abbreviation for claimant counts. [Colour figure can be viewed at wileyonlinelibrary.com]
correlation between the claimant counts and the Unemployed Labour Force. Therefore, the claimant counts were used to assess the unemployment estimates. The increase in the claimant counts was actually also not representing all job losses because many of the people that lost their jobs were not entitled to benefits because they had worked only for a short period. This was especially the case this time because at once a lot of temporary jobs and on-call jobs were terminated. The claimant counts show a much sharper turn in the slope of the series than the estimates of the Employed and Unemployed Labour Force under model A. This is, in addition to the high values of the standardised innovations, an indication that Model A is mis-specified in the second quarter of 2020 since it cannot pick up the sharp turning point.

For the interpretation of the evolution of these figures, filtered estimates for the month-to-month change of the trends are calculated. The month-to-month change of the trend is defined as \((L_t - L_{t-1})\). The filtered estimates for \(L_t\) and \(L_{t-1}\) have a strong positive correlation that cannot be ignored by calculating the standard error of this month-to-month change. An estimate for the standard error that accounts for this correlation is obtained by retaining \(L_{t-1}\) in the state vector of period \(t\) and calculate the standard error of the linear combination \((L_t - L_{t-1})\) from the covariance matrix of the filtered state vector obtained with the Kalman filter recursions See van den Brakel and Krieg (2016) for more details. Now, the evolution of the figures during the lockdown can be interpreted as follows. At the start of the lockdown the Employed Labour Force dropped by about 150,000 (s.e. = 33,500) people in April. The Unemployed Labour Force increased by only 35,000 (s.e. = 14,000) people. A major part of the people who lost their jobs, left the Labour Force. The Total Labour Force indeed decreased by about 115,000 (s.e. = 32,500) people. People who lost their jobs in the second quarter of 2020 were not able to apply for or accept a job, due to the strict lockdown regulations in this period. This explains the relatively mild increase of the Unemployed Labour Force and the strong decrease of the Total Labour Force in April. In June the lockdown regulations were relaxed and people were able to apply for and accept jobs. This explains the delayed strong increase in the Unemployed Labour Force and the recovery of the Total Labour Force in July. Particularly in the Employed and the Total Labour Force, these period-to-period changes are not visible in the filtered trends under Model A where the flexibility of the trend was not adapted. In the third and fourth quarter of 2020 the Unemployed Labour Force gradually decreases, and the Employed Labour Force gradually increases, which indicates that people start finding paid work again.

The movements from June onwards are not visible in the period-to-period change of the claimant counts. This is because it mainly concerns part-time jobs in branches like retail, hotels, bars, and restaurants. Particularly young people with part-time jobs are not qualified to receive claimant counts after losing this job.

As explained in the previous sections, Model C is used for the production of official monthly figures about the Labour Force. The filtered trend for the Unemployed Labour Force from January 2003 until December 2020 is shown in Figure 2 together with GREG input series of the five waves. See Figure S.2 in Section 4 of the Supplement for similar figures for the Employed and Total Labour Force. The filtered trends are benchmarked to the level of the first wave through the assumption that the RGB of the first wave is zero, at the level of the survey design that is used for data collection from 2012. The filtered trends under the designs used before 2012 are corrected for the discontinuities through \(\beta^{[1,1]}\) and \(\beta^{[1,2]}\) in Equation (6), which explains the deviation of the filtered trend from the input series of the first wave in the first part of the series. The lockdown in April 2020 marks a sharp turning point in the trends of all three variables.
The standard errors of the filtered trends under Model A and Model C are compared in Figure 3. See Figure S.3 in Section 4 of the Supplement for similar figures for the Employed and Total Labour Force. Until the start of the lockdown, the standard errors are almost equal under both models. Modelling the discontinuities due to the redesigns in 2010 and 2012 resulted in a substantial increase of the uncertainty in these periods. After the last redesign in 2012, the standard errors gradually decreased again as more and more information under the new design became available. During the lockdown the standard errors under both models clearly deviate. Making the variance of the slope disturbance terms larger increases the uncertainty of the filtered trend under Model C. Artificially increasing the variance of the slope disturbance terms implies that the time series model relies less on information collected in the past and gives more weight to
the GREG estimates observed in the current month. The standard errors reflect this temporarily increased uncertainty, as intended.

After having observed eight months after the start of the lockdown, it is possible to compare the chosen approach with Model D, which has two separate variance components for the slope disturbance terms, where $\sigma^2_{\eta|1}$ refers to the period before January 2020 and $\sigma^2_{\eta|2}$ refers to the period from January 2020 until December 2020. Results for the ML estimates for the standard deviations of the slope disturbance terms can be found in Table S.9 in Section 4 of the Supplement. A comparison of the filtered trends under Model C and D can be found in Figure S.4 in Section 4 of the Supplement.

Model D has several drawbacks which makes this approach inappropriate for this problem. First, it takes a relatively long time before sufficient observations are available to obtain stable ML estimates for the variance of the slope disturbance term for the second period. Second, if the period for which a more flexible trend is required is short, then there will never be enough observations to have a reliable ML estimate. On top of that, it is difficult to choose the end of the period for which a more flexible trend is required. In this application it appears that the period where the trend requires increased flexibility is short. The advantage of Model C is that the size of the variance of the slope disturbance terms can be increased temporarily and brought back gradually. Under Model D, the variance of the slope disturbance terms already appears to be unnecessarily large in Q4 of 2020. This leads to too volatile trends, while the standard errors of the filtered trends are also unnecessarily large.

Finally, the extent to which two important model assumptions, mentioned in Section 4, are met is evaluated. The first assumption is that differences between the outcomes in the first wave with and without CAPI are not affected by the lockdown. To evaluate the validity of this assumption, Model (9) is also applied to the time series available until December 2020, where observations for $\hat{y}_{ij}^{[j]}$ are missing from April until August and are available for the last 4 months of 2020. The estimates for $\kappa_{ij}^{[j]}$ remain very close to the values based on the time series observed until March 2020 (see Table S.4 in the Supplement). This is a strong indication that the differences did not change after the lockdown. It is also observed that the response rates of the WI and CATI households did not change after the start of the lockdown, which indicates that the assumption that the WI and CATI responses before and after the start of the lockdown are comparable is reasonable.

7 | DISCUSSION

Monthly figures about the Dutch labour force are based on a structural time series model. This estimation approach solves problems with small monthly sample sizes, rotation group bias, and discontinuities induced by survey redesigns. The lockdown caused by the COVID-19 pandemic has two effects that need to be taken into account in the production of monthly labour force figures. It changed the data measurement process because CAPI interviewing stopped, which affects selection and measurement bias in the direct estimates of the monthly labour force figures. At the same time, it changed the data generating process, since the volatility of the population parameters were increased due to the lockdown. The crisis induced by COVID-19 marked a sharp turning point in the evolution of the monthly labour force figures followed by a partial recovery. Since both processes changed at exactly the same time, a method is proposed in this paper to avoid confounding effects that are the result of differences in the measurement process and the data generating process. As a first step, the effect in the outcomes
due to the loss of CAPI in the first wave is estimated with a separate time series model applied to series of direct monthly estimates with and without CAPI responses and a series of claimant counts as an auxiliary series. In a next step, the time series model used for the production of official monthly labour force figures is extended with a level intervention component that compensates for the loss of CAPI in the input series. The estimate for this effect, obtained in the first step, is used as an approximation for the regression coefficient of the level intervention. This estimate is treated in the model as if its value is available a priori without uncertainty. The increased volatility due to the lockdown results in a temporal mis-specification of the time series model. The model accommodates this by making the trend temporarily more flexible. This is achieved by increasing the variance of the slope disturbance terms of the trend.

The empirical results show that the effect of the lockdown on the evolution of the population parameters is much larger than the effect due to the loss of the CAPI responses. Alternative approaches to account for the loss of CAPI responses and the increased volatility of the population parameters are considered. The advantage of the chosen method is that it uses all the available information from the past to estimate the effect of the loss of CAPI. Temporarily increasing the flexibility of the trend proved to be a pragmatic solution to accommodate the sudden mis-specification in the time series model. The model temporarily gives more weight to the direct estimates of the current period and less weight to predictions based on past observations. The increased uncertainty is reflected in a temporary increase of the standard errors of the filtered trend. Finally, the chosen approach proved to be manageable to accommodate unexpected effects in the time series model used for the production of timely monthly labour force figures.

The COVID-19 pandemic induced unprecedented changes into the economy and at the same time affected the data measurement process of the surveys that are in place to measure the economy. Although an all out effort was made to adapt the structural time series models that are used for the production of official monthly labour force figures, it is emphasised that any modelling strategy will not be adequate to accommodate these changes. In this paper, it is proposed to increase the flexibility of the model, so that the model estimates are less strongly based on past observations. This comes at the cost of a temporary increase in the uncertainty of the model estimates. Moreover, the method relies on the strong assumption that the response behaviour of the WI and CATI respondents is not affected by the lockdown. Response analysis of the WI and CATI modes for the months before and after the lockdown supports the validity of this assumption. The estimates for the first wave based on the WI, CATI, and CAPI responses in the last four months of 2020, also support the assumption that differences between the first wave with and without CAPI did not change after the start of the lockdown.

The proposed model adjustments only affect the estimates during the lockdown but will remain in the model that is used for the production of official monthly figures permanently, to avoid model misspecification. At this point, the effect of the COVID-19 pandemic on the seasonal effects and the autocorrelation of the sampling errors is unknown. An advantage of the structural time series model is that seasonal effects are modelled dynamically, which implies that the model can pick up small changes in the seasonal effects. If the model, however, is not capable of picking up a large change in the seasonal effects, it might be necessary to adjust the seasonal component. This can be done by modelling seasonal breaks or by making the variance of the seasonal disturbance terms more flexible in a similar way to that proposed for the trend. This issue is left as further research since several years of observations after the start of the crisis are required.
to evaluate possible effects on the seasonal component and the autocorrelation of the sampling error.

ACKNOWLEDGEMENTS
The views expressed in this paper are those of the authors and do not necessarily reflect the policies of Statistics Netherlands. The authors are grateful to the Editor, Guest Editor, Associate Editor and two anonymous reviewers for careful reading of a former draft of the manuscript and providing constructive comments. They proved to be very helpful to improve the quality of the paper.

ORCID
Jan van den Brakel https://orcid.org/0000-0002-4238-196X

REFERENCES
Bailar, B. (1975) The effects of rotation group bias on estimates from panel surveys. Journal of the American Statistical Association, 70(349), 23–30.
van den Brakel, J. & Krieg, S. (2009) Structural time series modelling of the monthly unemployment rate in a rotating panel. Survey Methodology, 35(2), 177–190.
van den Brakel, J. & Krieg, S. (2015) Dealing with small sample sizes, rotation group bias and discontinuities in a rotating panel design. Survey Methodology, 41(2), 267–296.
van den Brakel, J. & Krieg, S. (2016) Small area estimation with state-space common factor models for rotating panels. Journal of the Royal Statistical Society, 179, 763–791.
van den Brakel, J. & Roels, R. (2010) Intervention analysis with state-space models to estimate discontinuities due to a survey redesign. Annals of Applied Statistics, 4, 1105–1138.
van den Brakel, J., Zhang, X.M. & Tam, S.-M. (2020) Measuring discontinuities in time series obtained with repeated sample surveys. International Statistical Review, 88, 155–175.
Carter, C.K. & Kohn, R. (1996) Markov chain Monte Carlo in conditionally gaussian state space models. Biometrika, 83(3), 589–601.
Carvalho, C.M., Polson, N.G. & Scott, J.G. (2010) The horseshoe estimator for sparse signals. Biometrika, 97(2), 465–480.
Chan, J. & Jeliazkov, I. (2009) Efficient simulation and integrated likelihood estimation in state space models. International Journal of Mathematical Modelling and Numerical Optimisation, 1, 101–120.
Doornik, J. (2009) An object-oriented matrix programming language Ox 6. London: Timberlake Consultants Press.
Durbin, J. & Koopman, S. (2012) Time series analysis by state space methods. Oxford: Oxford University Press.
Francq, C. & Zakoian, J.-M. (2010) GARCH models. Hoboken: Wiley and Sons.
Harvey, A. (2013) Dynamic models for volatility and heavy tails. Cambridge: Cambridge University Press.
Knorr-Held, L. & Rue, H. (2002) On block updating in Markov random field models for disease mapping. Scandinavian Journal of Statistics, 29(4), 597–614.
Koopman, S. (1997) Exact initial Kalman filtering and smoothing for non-stationary time series models. Journal of the American Statistical Association, 92, 1630–1638.
Koopman, S., Shephard, A. & Doornik, J. (2008) Ssfpack 3.0: statistical algorithms for models in state-space form. London: Timberlake Consultants, Press London.
McCausland, W., Miller, S. & Pelletier, D. (2011) Simulation smoothing for statespace models: A computational efficiency analysis. Computational Statistics and Data Analysis, 55, 199–212.
Pfeffermann, D. (1991) Estimation and seasonal adjustment of population means using data from repeated surveys. Journal of Business & Economic Statistics, 9, 163–175.
Pfeffermann, D. & Tiller, R. (2006) Small area estimation with state-space models subject to benchmark constraints. Journal of the American Statistical Association, 101, 1387–1397.
Pfeffermann, D., Feder, M. & Signorelli, D. (1998) Estimation of autocorrelations of survey errors with application to trend estimation in small areas. Journal of Business & Economic Statistics, 16, 339–348.
Piepho, H.-P. & Ogutu, J. (2014) Simple state-space models in a mixed model framework. *The American Statistician*, 61(3), 224–232.
Polson, N.G. & Scott, J.G. (2010) Shrink globally, act locally: Sparse bayesian regularization and prediction. *Bayesian Statistics*, 9, 501–538.
Ruiz-Cárdenas, R., Krainski, E. & Rue, H. (2012) Direct fitting of dynamic models using integrated nested laplace approximations – inla. *Computational Statistics and Data Analysis*, 56, 1808–1828.
Särndal, C.-E., Swensson, B. & Wretman, J. (1992) *Model assisted survey sampling*. Berlin: Springer.
Tang, X., Ghosh, M., Ha, N.S. & Sedransk, J. (2018) Modeling random effects using global–local shrinkage priors in small area estimation. *Journal of the American Statistical Association*, 113, 1–14.

**SUPPORTING INFORMATION**

Additional supporting information may be found in the online version of the article at the publisher’s website.

**How to cite this article:** van den Brakel, J., Souren, M. & Krieg, S. (2022) Estimating monthly labour force figures during the COVID-19 pandemic in the Netherlands. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4), 1560–1583. Available from: [https://doi.org/10.1111/rssa.12869](https://doi.org/10.1111/rssa.12869)