Discussion paper

Structural Time Series Models for Consumer Confidence indicators

Jan van den Brakel, Sabine Krieg
and Marc Smeets

April 2024

Content

Summary

In this paper a model-based inference procedure based on a multivariate structural time series model is developed for the production of monthly figures about consumer confidence. The input for the model consists of five series of direct estimates for the indices that measure consumer confidence, which are derived from the Dutch Consumer Survey. The model improves the accuracy of the direct estimates, since it provides a better separation of measurement errors and sampling errors from estimated target parameters. The standard errors for the month-to-month changes are clearly smaller under the time series model. A second problem addressed in this paper is related to the transition to a new survey process in 2017. Structural time series models in combination with a parallel run are applied to estimate discontinuities induced by the redesign. A backcasting algorithm designed for the consumer confidence variables is developed to construct uninterrupted input series for the aforementioned structural time series model. This inference method facilitated a smooth transition to a new survey design and resulted in uninterrupted published series about consumer confidence that date back to 1986. The method is implemented for the production of official monthly figures on consumer confidence in the Netherlands.

Keywords

Small area estimation, structural time series model, discontinuities

1. Introduction

The Dutch Consumer Survey (CS) measures consumer confidence in the Netherlands with a monthly frequency. Until 2016, monthly figures were obtained with design-based inference methods applied to data obtained with probability sampling with a sample size of about 1000 respondents per month. This paper addresses two problems. First, the series of the monthly estimates are quite volatile. This is the result of the relatively small sample size and because the CS measures a relative short term emotion about the respondents opinion of the financial and economic climate, despite the fact that most questions actually refer to the last or upcoming 12 months. Increasing the sample size would therefore only partially reduce the volatility of the series. This is nevertheless not an option, since Statistics Netherlands (as many other statistical institutes) has to reduce administration costs and response burden. The second problem is the disruption of the continuity of the time series as a result of a redesign of the survey process that took place in the beginning of 2017. Changes in the questionnaire and fieldwork strategies generally have a systematic effect on the outcomes of a survey, since they affect non-sampling errors like measurement bias and selection bias. Systematic differences in time series induced by survey redesigns are further referred to as discontinuities. In a well-designed transition process, discontinuities are quantified to avoid confounding real period-to-period change from sudden changes in measurement bias and selection effects.

Since the CS is a repeated survey, a structural time series model (STM) can be developed as a solution for both problems. Seemingly Unrelated Time Series Equation (SUTSE) models are applied to the CS as a form of small area estimation and to account for discontinuities induced by the redesign of 2017. The underlying series of the five indices, which together define the consumer confidence, are the input series of the time series model. SUTSE models are multivariate structural time series models were relations between the input series are modelled with contemporaneously correlated disturbances of the state variables (Harvey, 1989, Ch. 8). Through a SUTSE model, sample information observed in previous reference periods is used to obtain more accurate estimates. Moreover, correlations between the disturbances of e.g. the trend are useful to borrow strength from the other variables used in the construct for consumer confidence. In this paper we also discuss the role of correlations between the measurement error terms of the series. Sudden real events influence all series simultaneously, which results in correlated population irregular terms.

The use of time series modelling with the aim of improving the precision of survey data has been considered by many authors and dates back to Blight and Scott (1973). It can be interpreted as a form of small area estimation by borrowing strength over time (Rao and Molina, 2015). This approach has been applied before in the context of official statistics, see e.g. Tam (1987), Binder and Dick (1989, 1990), Bell and Hillmer (1990), Tiller (1992), Rao and Yu (1994), Pfeffermann and Burck (1990), Pfeffermann and Bleuer (1993), Pfeffermann (1991), Pfeffermann and Tiller (2006), Harvey and Chung (2000), and Feder (2001), Van den Brakel and Krieg (2015) and Elliot and Zong (2019).

Discontinuities caused by a redesign of the survey process can be quantified in different ways, see Van den Brakel et al. (2020) for an overview. One approach is to collect data under both the old and the new design in parallel for some period of time, which is further referred to as a parallel run. The difference of the estimates based on both designs can be used as a direct estimate of the discontinuity. In the case of a sufficiently large parallel run this is a reliable and timely approach. The major drawback is that it requires additional data collection, which makes the approach costly. Alternatively a time series model can be applied where the discontinuities are estimated using a level intervention (Van den Brakel and Roels, 2010). In the case of a small parallel run, the information from the parallel run can be used as a-priori information in the time series model, e.g. through an informative initialization of the Kalman filter. This initial estimate is further improved with the information from the time series observed before and after the parallel run (Van den Brakel and Krieg, 2015).

Discontinuity estimates for the CS caused by the redesign are based on a parallel run of three months, with a sample size that is equal to the regular sample size. These estimates are improved with a structural time series model, as outlined above. When the estimates for discontinuities are known, it is important to communicate about them with the users of the series to avoid misinterpretation of the series. In the case of the CS the series of the past are corrected for the discontinuities. The series underlying the CS are percentages and a correction method is proposed that attempts to keep the adjusted values in the admissible range between 0% and 100%. The time series modelling approach developed in this paper has been implemented for the production of Statistics Netherlands' official monthly consumer confidence figures since April 2017.

The paper is organized as follows. Section 2 provides a description of the Dutch CS. In Section 3, a structural time series model is developed for the estimation of monthly consumer confidence figures. In this section, results are presented based on observations for the period before the redesign in 2017. In Section 4 the change-over to the new design is described. In this section a method for estimating discontinuities that combines a parallel run with a time series modelling approach is proposed. Furthermore, a correction method to adjust the series observed before the change-over to the level of the series observed under the new design is described. Finally, results for discontinuity estimates and corrected series are presented. Section 5 summarizes how the estimation method is implemented for the production of official monthly consumer confidence figures. The paper finalizes with a conclusion in Section 6.

2. The Dutch Consumer Survey

The Consumer Survey (CS) is a monthly survey. Before the redesign of 2017, each month a sample of approximately 2,500 households was drawn by stratified sampling from a sample frame of addresses that is derived from the Dutch Municipal Register. Stratification is based on the cross-classification of 12 provinces and urbanization level in five classes. All households have equal inclusion probabilities. The monthly surveys are based on independent cross-sectional samples. Sampled households are observed in one month only. Households for which a known landline telephone number was available, were contacted by an interviewer who completes the questionnaire by computer assisted telephone interviewing (CATI) during the first ten working days of the month. The questionnaire is completed by the head of the household. The head of the household is the household member that contributes the largest share to the household income. Households without a known landline telephone number were not contacted. On average a net sample of about 1,000 responding households was obtained, resulting in a response rate of about 40%. A major part of the non-response consisted of households for which no known telephone number of a landline connection is available. The response among households for which a telephone number was available was about 60%.

For the consumer confidence index, questions concerning the following five variables are relevant:

1. opinion about changes of the general economic situation of the country over the last 12 months, abbreviated as Econ. L12,
2. expectations of changes of the general economic situation of the country over the next 12 months, abbreviated as Econ. N12,
3. opinion about changes of the financial situation of the household over the last 12 months, abbreviated as Fin. L12,
4. expectations of changes of the financial situation of the household over the next 12 months, abbreviated as Fin. N12,
5. whether it is the right moment for people to make major purchases, abbreviated as Major pur.

The questions for the first four variables have two positive and two negative answer options (“a lot better”, “a little better”, “a lot worse”, “a little worse”). Furthermore, there is the neutral option “the same” as well as “do not know”. The question for the fifth variable has one positive and one negative answer option (“yes, it is the right moment now”, “no, it is not the right moment now”). Furthermore, there is a neutral option (“it is neither the right moment nor the wrong moment”) as well as “do not know”. The percentages of positive, negative and neutral answers (as percentage points of the total answers) p_i^+, p_i^-, p_i^0 with $p_i^+ + p_i^0 + p_i^- = 100$ are computed for each question $i = 1, \dots, 5$. Estimates for the five variables mentioned above are obtained by the difference of positive and negative answers, i.e. $y_i = p_i^+ - p_i^-$ for $i = 1, \dots, 5$. Furthermore, the five variables are combined by computing the following averages:

– $y_6 = (y_1 + y_2)/2$ which is the indicator for economic climate,

- $y_7 = (y_3 + y_4 + y_5)/3$ which is the indicator for willingness to buy,
- $y_8 = (y_1 + y_2 + y_3 + y_4 + y_5)/5$ which is the indicator for consumer confidence.

The indicators $y_1, \dots, y_8$ are the main target variables in the publication. Until the end of 2016 unweighted sample means were used as estimates for the target variables. Expressions for the variance of the eight series are given by Van den Brakel et al. (2017). The publication of monthly figures started in 1986. Both the original figures and seasonally adjusted figures of the indicator series are published. Furthermore, the underlying series of the percentages are also published.

Approaching households with a non-secret landline telephone will result in low coverage of the target population. In combination with the low response rate, the non-response will be not missing at random. Until the end of 2016 this problem is ignored, but it was a reason for a redesign of the survey process of the CS. In January 2017 five important changes were implemented simultaneously; 1) The sample design changed from a stratified sample of households to a stratified sample of persons. A sample of 2150 persons is drawn each month. All persons have equal inclusion probabilities. With a response rate of about 47% this results in a net sample of slightly more than 1000 respondents. 2) The data collection mode changed from CATI to a sequential mixed mode design, where the respondents are first asked to complete a questionnaire via web. After three reminders, the web non-respondents are interviewed by phone (as far as phone numbers are available, including mobile phones). 3) There are changes in the questionnaire. The most important change is the way in which the answer categories are offered. Under the old questionnaire the respondent could first choose between the options “worse”, “neutral”, or “better”. In the case “worse” or “better” was selected, the respondent had to specify whether it is “a lot” or “a little” better or worse. In the new questionnaire the two positive and two negative answer options for questions 1 to 4 are shown directly. 4) Another important change is that a conditional incentive is given to respondents to improve the response rate (a tablet is raffled among the respondents). 5) Finally, the sample estimates are based on the generalised regression (GREG) estimator (Särndal et al., 1992) to correct, at least partially, for selective non-response. The changeover to a sequential mixed-mode design will reduce the undercoverage but the response rates are still very low. It can be anticipated that the non-response is still not missing at random and that the GREG will not sufficiently correct for this. Further research in fieldwork methods that improve response rates and estimation methods that correct for selective non-response is needed, see for example Pfeffermann and Sikov (2011).

A side effect of this redesign is that it causes a sudden change in selection effects, as another part of the population is willing to respond when another mode is applied, and when an incentive is offered. Furthermore, there are sudden changes in the measurement bias due to the use of another data collection mode (partially without interviewer) and changes in the questionnaire. These cause the so-called discontinuities.

3. Inference for monthly CS figures using structural time series models

In this section a STM is developed using the series observed from January 1987 until December 2016. With a structural time series model, a series is decomposed in a trend component, a seasonal component, other cyclic components, a regression component and an irregular component for the unexplained variation. For each component a stochastic model is assumed. This allows the trend, seasonal, and cyclic component but also the regression coefficients to be time-dependent. If necessary, ARMA components can be added to capture the autocorrelation in the series beyond these structural components. See Harvey (1989) or Durbin and Koopman (2012) for an introduction to structural time series modelling.

3.1 Model description

Each month t a direct sample estimate $\hat{y}_{i,t}$ is computed for the five variables y_i $(i = 1, \dots, 5)$ measured with the questions of the CS, as explained in Section 2. Constructing time series models for series of sample estimates starts with formulating a measurement error model that decomposes the estimate into a true but unknown population parameter, say $\theta_{i,t}$, and a sampling error, say $\tilde{e}_{i,t}$:

$$\hat{y}_{i,t} = \theta_{i,t} + \tilde{e}_{i,t}, \quad (i = 1, \dots, 5). \tag{3.1}$$

For the unknown population parameter, a basic STM is assumed, i.e.

$$\theta_{i,t} = L_{i,t} + S_{i,t} + I_{i,t}, \quad (i = 1, \dots, 5), \tag{3.2}$$

with $L_{i,t}$ the level of a stochastic trend component that models the low frequency variation of the series, $S_{i,t}$ a stochastic component that models the seasonal fluctuation around the trend and $I_{i,t}$ the population irregular term. Inserting (3.2) into (3.1) gives the time series model for the observed series: $\hat{y}_{i,t} = L_{i,t} + S_{i,t} + I_{i,t} + \tilde{e}_{i,t}$, $(i = 1, \dots, 5)$. Three different models to account for the population irregular term and the sampling error are compared. The first model has a separate population irregular term and a sampling error

$$\hat{y}_{i,t} = L_{i,t} + S_{i,t} + I_{i,t} + k_{i,t}\varepsilon_{i,t}, \quad (i = 1, \dots, 5), \tag{3.3.a}$$

with $k_{i,t} = \sqrt{\mathrm{Var}(\hat{y}_{i,t})}$ the standard error of the observed series and $\varepsilon_{i,t}$ a normally distributed error term. In the second model, the population irregular term and the sampling error are combined in one measurement error, i.e., $e_{i,t} = I_{i,t} + \tilde{e}_{i,t}$ and scaled with the standard error of the input series:

$$\hat{y}_{i,t} = L_{i,t} + S_{i,t} + k_{i,t}e_{i,t}, \quad (i = 1, \dots, 5), \tag{3.3.b}$$

with $e_{i,t}$ a normally distributed error term. In the third model the measurement errors are not scaled with the standard errors of the input series, which implies that the heteroscedasticity of the sampling errors is ignored:

$$\hat{y}_{i,t} = L_{i,t} + S_{i,t} + e_{i,t}, \quad (i = 1, \dots, 5). \tag{3.3.c}$$

The standard errors of the input series $k_{i,t}$ can be derived from the survey samples and will be specified below.

The five series defined in formulas (3.3.a), (3.3.b) and (3.3.c) can be combined in a vector $\hat{\mathbf{y}}_t = (\hat{y}_{1,t}, \hat{y}_{2,t}, \hat{y}_{3,t}, \hat{y}_{4,t}, \hat{y}_{5,t})'$, which can be modelled with the so-called Seemingly Unrelated Time Series Equation (SUTSE) models (Harvey, 1989, Ch. 8). In a SUTSE model, each observed series appears in one equation and are contemporaneously correlated through the disturbance terms of their state variables. The SUTSE models for (3.3.a), (3.3.b) and (3.3.c) are defined as:

$$\hat{\mathbf{y}}_t = \mathbf{L}_t + \mathbf{S}_t + \boldsymbol{I}_t + \boldsymbol{K}_t\boldsymbol{\varepsilon}_t, \tag{3.4.a}$$

$$\hat{\mathbf{y}}_t = \mathbf{L}_t + \mathbf{S}_t + \boldsymbol{K}_t\mathbf{e}_t, \tag{3.4.b}$$

$$\hat{\mathbf{y}}_t = \mathbf{L}_t + \mathbf{S}_t + \mathbf{e}_t, \tag{3.4.c}$$

with $\mathbf{L}_t = (L_{1,t}, L_{2,t}, L_{3,t}, L_{4,t}, L_{5,t})'$, $\mathbf{S}_t = (S_{1,t}, S_{2,t}, S_{3,t}, S_{4,t}, S_{5,t})'$, $\boldsymbol{I}_t = (I_{1,t}, I_{2,t}, I_{3,t}, I_{4,t}, I_{5,t})'$, $\boldsymbol{K}_t$ a 5×5 diagonal matrix with the standard errors $k_{i,t}$ as diagonal elements, $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t}, \varepsilon_{4,t}, \varepsilon_{5,t})'$ and $\mathbf{e}_t = (e_{1,t}, e_{2,t}, e_{3,t}, e_{4,t}, e_{5,t})'$. The trends $L_{i,t}$ $(i = 1, \dots, 5)$ are modelled with the so-called smooth trend model:

$$\begin{aligned} L_{i,t} &= L_{i,t-1} + R_{i,t-1}, \\ R_{i,t} &= R_{i,t-1} + \eta_{R,i,t}, \end{aligned} \tag{3.5}$$

with $L_{i,t}$ the level of the trend and $R_{i,t}$ the slope parameter. The disturbance terms $\eta_{\mathrm{R},i,t}$ are normally distributed with

$$\begin{aligned} &\mathrm{E}(\eta_{\mathrm{R},i,t}) = 0, \\ &\mathrm{Cov}(\eta_{\mathrm{R},i,t}, \eta_{\mathrm{R},i',t'}) = \begin{cases} \sigma^2_{\mathrm{R},i} & \text{if} \quad i = i' \text{ and } t = t' \\ \varsigma_{\mathrm{R},i,i'} & \text{if} \quad i \neq i' \text{ and } t = t'. \\ 0 & \text{if} \quad t \neq t' \end{cases} \end{aligned} \tag{3.6}$$

This is a dynamic model for the low frequency variation, which has the flexibility to capture the trend as well as economic cycles. It can therefore be interpreted as the trend plus economic cycle, which is shortly referred to as trend. The smooth trend model is chosen because it results in a more stable trend estimates compared to for example the local level model (that has disturbance terms for the level parameter and does not have a slope parameter) or the local linear trend model (that has disturbance terms for both the level and the slope parameters). This is in line with the aim of the CS to evaluate economic and financial confidence over the

last and next 12 months. Note that in (3.6) the slope disturbance terms can be correlated, which makes sense since the five variables measure related phenomena.
The so-called trigonometric seasonal model is used to model the seasonal component $S_{i,t}$ ($i = 1, \dots ,5$), which is defined as:

$$S_{i,t} = \sum_{l=1}^{6} S_{i,t,l}, \tag{3.7}$$

with

$$S_{i,t,l} = S_{i,t-1,l}\cos(h_l) + S^*_{i,t-1,l}\sin(h_l) + \eta_{S,i,t,l},$$
$$S^*_{i,t,l} = S^*_{i,t-1,l}\cos(h_l) - S_{i,t-1,l}\sin(h_l) + \eta^*_{S,i,t,l},\ h_l = \frac{\pi l}{6}, l = 1, \dots ,6.$$

The disturbances $\eta_{\mathrm{S},i,t,l}$ and $\eta^*_{\mathrm{S},i,t,l}$ are normally distributed with

$$\mathrm{E}(\eta_{\mathrm{S},i,t,l}) = \mathrm{E}(\eta^*_{\mathrm{S},i,t,l}) = 0,$$

$$\begin{aligned}\mathrm{Cov}(\eta_{\mathrm{S},i,t,l}, \eta_{\mathrm{S},i',t',l'}) &= \mathrm{Cov}(\eta^*_{\mathrm{S},i,t,l}, \eta^*_{\mathrm{S},i',t',l'}) \\ &= \begin{cases} \sigma^2_{\mathrm{S},i} & \text{if} \quad i = i' \text{ and } t = t' \text{ and } l = l' \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

$\mathrm{Cov}(\eta_{\mathrm{S},i,t,l}, \eta^*_{\mathrm{S},i',t',l'}) = 0$ for all i, t, l.

Model (3.7) is a dynamic model that allows for seasonal patterns that gradually change over time. Like in the case of the trend component, it is possible to allow for non-zero correlations between disturbance terms of the seasonal components of the five input series. In this application, the variances of the disturbance terms tend to zero, which implies that seasonal components are nearly time invariant and their disturbance terms are almost equal to zero. Allowing for non-zero correlations between the disturbance terms does not contribute to better model fits in this application. For this reason a diagonal covariance structure for (3.7) is chosen in advance.

The population irregular terms in (3.4.a) are normally distributed with

$$\mathrm{E}(I_{i,t}) = 0,$$
$$\mathrm{Cov}(I_{i,t}, I_{i',t'}) = \begin{cases} \sigma^2_{\mathrm{I},i} & \text{if} \quad i = i' \text{ and } t = t' \\ \varsigma_{\mathrm{I},i,i'} & \text{if} \quad i \neq i' \text{ and } t = t' \\ 0 & \text{if} \quad t \neq t' \end{cases}. \tag{3.8}$$

The scaled sampling errors in (3.4.a) are normally distributed with:

$$\mathrm{E}(\varepsilon_{i,t}) = 0,$$
$$\mathrm{Cov}(\varepsilon_{i,t}, \varepsilon_{i',t'}) = \begin{cases} \sigma^2_{\varepsilon,i} & \text{if} \quad i = i' \text{ and } t = t' \\ 0 & \text{otherwise} \end{cases}. \tag{3.9}$$

The measurement errors in (3.4.b) and (3.4.c) are normally distributed with:

$$\mathrm{E}(e_{i,t}) = 0,$$

$$\mathrm{Cov}(e_{i,t}, e_{i',t'}) = \begin{cases} \sigma^2_{\mathrm{e},i} & \text{if} \quad i = i' \text{ and } t = t' \\ \varsigma_{\mathrm{e},i,i'} & \text{if} \quad i \neq i' \text{ and } t = t' \\ 0 & \text{if} \qquad t \neq t' \end{cases}. \tag{3.10}$$

Note that non-zero correlations between the population irregular terms in (3.8) and the measurement errors in (3.10) are allowed. A motivation for modelling correlations between population irregular terms or measurement errors is that some events, e.g. good or bad news about the economy, influence the answers to all questions in a similar way.

The standard errors $k_{i,t}$ of the input series are obtained from the sample surveys. The corresponding variances are approximated as:

$$\begin{aligned} \mathrm{Var}(\hat{y}_{i,t}) &= \mathrm{Var}(\hat{p}_i^+ - \hat{p}_i^-) = \mathrm{Var}(\hat{p}_i^+) + \mathrm{Var}(\hat{p}_i^-) - 2\mathrm{Cov}(\hat{p}_i^+, \hat{p}_i^-) \\ &= \frac{1}{n_t}[\hat{p}_i^+(100 - \hat{p}_i^+) + \hat{p}_i^-(100 - \hat{p}_i^-) - 2\hat{p}_i^+\hat{p}_i^-] \end{aligned}. \tag{3.11}$$

This approximation is based on the variance of proportions under simple random sampling with replacement, which is motivated by the small sampling fraction of the monthly samples. The stratification of the sample design is ignored, which implies that the variance will be slightly overestimated by formula (3.11).

The SUTSE models (3.4) can be interpreted as small area estimation models. Through the trend and seasonal components in this model, information from the past about the long-term development and seasonal fluctuations is used to improve the direct sample estimates. The covariances between the slope disturbance terms and population irregular terms of the five input series in the SUTSE model further improve the model fits.

The following versions of model (3.4) will be compared to investigate the influence of the correlations between trend and measurement error components:

- Model 1: the model as described by equations (3.4) – (3.8), i.e., with correlations between the slope disturbances of the trend and between the measurement errors of the five series. The three different ways of handling the sampling errors result in three versions of Model 1:
    a. M1a: based on (3.4.a) with correlated population irregular terms defined by (3.8) and sampling errors defined by (3.9) and (3.11).
    b. M1b: based on (3.4.b) with correlated measurement errors defined by (3.10) and (3.11).
    c. M1c: based on model (3.4.c) with correlated measurement errors defined by (3.10). The heteroscedasticity of the sampling errors is ignored under this model.
- Model 2: similar to Model 1, but without correlation between the population irregular terms or measurement errors of the five series. The three different ways of handling the sampling errors result in three versions of Model 2:
    a. M2a: based on (3.4.a) with a diagonal covariance for the population irregular terms in (3.8), i.e., $\varsigma_{\mathrm{I},i,i'} = 0$, and sampling errors defined by (3.9) and (3.11).

b. M2b: based on (3.4.b) with a diagonal covariance matrix for the measurement errors defined by (3.10), i.e., $\varsigma_{\mathrm{e},i,i'} = 0$, scaled with sampling errors defined by (3.11).
c. M2c: based on model (3.4.c) with a diagonal covariance matrix for the measurement errors defined by (3.10), i.e., $\varsigma_{\mathrm{e},i,i'} = 0$.

– Model 3: similar to Model 2, but without correlation between the slope disturbances and without correlation between the measurement errors of the five series. The three different ways of handling the sampling errors result in three versions of Model 3:
  a. M3a: based on (3.4.a) with $\varsigma_{\mathrm{R},i,i'} = 0$ in (3.6), $\varsigma_{\mathrm{I},i,i'} = 0$ in (3.8) and sampling errors defined by (3.9) and (3.11).
  b. M3b: based on (3.4.b) with $\varsigma_{\mathrm{R},i,i'} = 0$ in (3.6), $\varsigma_{\mathrm{e},i,i'} = 0$ in (3.10), scaled with sampling errors defined by (3.11).
  c. M3c: based on model (3.4.c) with $\varsigma_{\mathrm{R},i,i'} = 0$ in (3.6), $\varsigma_{\mathrm{e},i,i'} = 0$ in (3.10).

Note that the three versions of Model 3 are equivalent to fitting univariate models to each series separately, since all covariances of the state disturbance terms are equal to zero.

The general way to proceed is to put the structural time series model into state space representation. See the Appendix for state space representation of Models (3.4.a), (3.4.b) and (3.4.c). Then the Kalman filter can be applied to obtain optimal estimates for the state vector. The Kalman filter is a recursive procedure to obtain optimal estimates for the state vector at time t based on the data up to and including time period t, which 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. Let α_t denote the vector with unknown state variables for period t. Let $\alpha_{t|t'}$ denote the estimate for the state variables for period t, based on the observations obtained until (and including) period t'. If T denotes the length of the completely observed series, then $\alpha_{t|t}$ are the filtered estimates and $\alpha_{t|T}$ are the smoothed estimates.

Since revisions of published figures are no part of the normal publication strategy at Statistics Netherlands, interest in this application is mainly focused on the time series model estimates based on the complete set of information that would be available in the regular production process to produce consumer confidence figures for month t. This can be approximated with the filtered estimates. Filtered estimates for period t can be updated with the information that became available after period t. To illustrate the additional gain of applying a smoothing filter, we also compare the filtered estimates with the smoothed estimates.

The hyperparameters σ_*^2 and ς_* are estimated with maximum likelihood, using a numerical optimization procedure. The maximum likelihood estimates for the hyperparameters are inserted into the Kalman filter but treated as if they were the true values, known without error. This implies that the additional uncertainty of using the maximum likelihood estimates for the hyperparameters is ignored in the standard errors for the filtered and smoothed estimates for the trend and signal of the CS parameters. This is a standard approach in state space modelling and acceptable in this application given the long series that are available, see Bollineni-Balabay et al. (2017) for a motivation and Pfeffermann and Tiller (2005) for a bootstrap approach that accounts for the additional uncertainty in the standard errors of the estimated state variables as a result of using maximum likelihood estimates for the hyperparameters in the Kalman filter. Finally, the state variables are initialized with a diffuse initialization, unless stated differently. See Harvey (1989) or Durbin and Koopman (2012) for technical details. In this paper Ssfpack 3.0 (Koopman et al., 1999b, and Koopman et al., 2008) in combination with Ox (Doornik, 1998) is used for the computations.

After the model is estimated, model-based estimates for the five series can be computed. The trend $L_{i,t}$ and signal, which is defined as $L_{i,t} + S_{i,t}$ are used as model-based estimates for consumer confidence indicators. The model-based estimates of the combined series, i.e., economic climate, willingness to buy, and consumer confidence are computed as means of the estimates for the five series. The standard errors for the model estimates of the combined series account for the correlation between the state variables ($L_{i,t}$ and $S_{i,t}$) of the underlying series. Deriving the combined indices from the model estimates of the five input series has the advantage that the model estimates are numerically consistent. This would not be the case if three separate univariate time series models were applied to the direct estimates of the three combined indices.

With the diffuse initialization of the Kalman filter the first 12 observations are required to construct a proper prior for the Kalman filter. It is understood that in the figures in the results sections the first 14 years are omitted so that that differences between estimation results are more visible.

3.2 Results

The nine models described in Section 3.1 are applied to series from January 1987 until December 2016. The figures in this section are restricted to the period 2001 – 2016 to facilitate a better interpretation of the graphs. The model evaluation is based on the entire series.

Maximum likelihood estimates for the standard deviations of slope disturbance terms, seasonal disturbance terms, population irregular terms and the measurement error for the five baseline variables are presented in Table 3.1 for the nine models. There are only small differences between the standard deviations of the slope disturbance terms of Model M1a, M1b an M1c. This also applies to the three versions of Model 2 and the three versions of Model 3. The standard deviations of the slope disturbance terms under the three versions of Model 2 are larger compared to Model 1 and Model 3. For all models it follows that the seasonal component is time invariant for Econ. N12 and almost time invariant for Econ. L12, Fin. L12 and Fin. N12.

For model (3.4.a) it is difficult to estimate the standard deviation of both the population irregular term ($\sigma_{I,i}$) and the scaled sampling errors ($\sigma_{\varepsilon,i}$) with maximum likelihood. Therefore $\sigma_{\varepsilon,i}$ is set equal to one. As a result, the variance of the sampling error $k_{i,t}e_{i,t}$ is equal to the variance of the input series. Note that the standard deviations for the population irregular term are clearly larger than one. The standard errors $k_{i,t}$ vary between 1.2 and 1.6 with an average around 1.5. This illustrates that the population irregular term is at least as large as the sampling error. This also follows from the standard deviations of the measurement errors under models (3.4.b). In this model the measurement errors are scaled with the standard errors of the input series. If the population irregular term would be small compared to the sampling error, then the maximum likelihood estimates for $\sigma_{e,i}$ in M1b, M2b and M3b would take values close to one. The estimates for $\sigma_{e,i}$, however, are clearly larger, which indicates once again that the population irregular term has a substantial share in the measurement errors. The standard deviations of measurement errors under models (3.4.c) have the largest values since they reflect the total variance of the population irregular term and the sampling error without scaling them with the factors $k_{i,t}$.

Tables 3.2 – 3.6 show the maximum likelihood estimates of the correlations of three model specifications of Models 1 and 2. The correlations between the slope disturbances are given in Table 3.2 for Model M1a. The correlation matrices for Model M1b and M1c are almost equivalent and are therefore left out. As expected, high correlations between slope disturbance terms are observed, since opinions about financial and economic situation are related. The correlation matrix for the population irregular term under Model M1a is given in Table 3.3, for the measurement errors, scaled with the sampling errors, under Model M1b in Table 3.4, and for the measurement errors under Model M1c, which ignores the standard errors $k_{i,t}$, in Table 3.5. There are a few small differences between the correlations under these three different models but it can be concluded that the patterns are rather consistent. The correlations between the slope disturbances are given in Table 3.6 for Model M2a. The correlation matrices for Model M2b and M2c are almost equivalent and are therefore left out. The correlations of the slope disturbances are larger under the three different specifications of Model 2 than under the three different specifications of Model 1. This is because under Model 1, a part of the co-movements of the series is considered as correlations between the measurement errors. Under Models M2a, M2b and M2c this variation is interpreted as trend fluctuations. As explained in Section 3.1, there are arguments that the measurement errors are correlated. Models M1a, M1b and M1c are therefore preferred over the three different versions of Model 2, and the correlations in Table 3.6 are probably over-estimated. From Table 3.1 it follows that the standard deviations of the slope disturbance terms are consistently higher compared to the three versions of Model 1 and Model 3, which is another indication that the trends under the three versions of Model 2 tend to overfit the observed series.

Differences between filtered trends as well as their standard errors for the five input series and the three indices that are a linear combination of the five input series under M1a, M1b and M1c are very small (smaller than one percent point). This also holds for the filtered signals. In a similar way the differences between the smoothed trends as well as the smoothed signals under M1a, M1b and M1c are negligible. This also applies to the filtered and smoothed trends and signals and the standard errors under M2a, M2b and M2c and the filtered and smoothed trends and signals and the standard errors under M3a, M3b and M3c. It follows from Table 3.1 that the standard deviations of the slope disturbance terms and the disturbance terms of the seasonal component are not affected by the three different versions of modelling the population irregular term and sampling errors. Also the correlations between the slope disturbance terms presented in Tables 3.2 and 3.6 are not influenced by the three different covariance structures of the population irregular term and sampling errors. This explains why the estimates for trends and signals are hardly affected by the three different covariance structures for the population irregular terms and sampling errors.

From now on we will present estimation results for trends and signals for the models that have a separate component for the population irregular term and the sampling error, i.e. M1a, M2a and M3a, since these three models are based on model (3.4.a) which has the most intuitive interpretation of accounting for population irregular term and sampling error. It is understood that similar results are obtained for filtered and smoothed estimates for the other two versions of models 1, 2 and 3.

Comparing the filtered as well as smoothed trends for the five input series and the three indices that are a linear combination of the five input series shows that the trend under Model M3a often differs from Model M1a and Model M2a. Since the five input series show strong co-movements, Model M1a and M2a improve the accuracy of the trend estimates since they use the information from the other series via the strong positive correlations between the slope disturbance terms. Generally, the trend under Model M2a is more volatile compared to Model M1a, since a part of the measurement error under Model M1a is interpreted as trend movements under Model M2a. As an example the filtered and smoothed trends under the three models are shown for Fin. N12 in Figure 3.1.

The filtered signals under Model M2a and Model M3a are both close to the direct estimates, which indicates that these models hardly smooth the input series. The filtered signals under Model M1a are more smooth compared to the direct estimates. Also the smoothed signals under Model M2a and M3a are more volatile and closer to the direct estimates than the smoothed signals under Model M1a. As an example, Figure 3.2 compares the filtered and smoothed signal under Model M1a with the direct estimates for consumer confidence. The smoothed estimates are clearly more stable than the filtered estimates. The filtered estimates illustrate the consumer confidence figures as they are published at the end of the reference period. Publishing the more stable smoothed estimates requires a revision policy. This is, however, not in line with the publication strategy of Statistics Netherlands.

Hyperparameter	Variable	Sd								
		M1a	M1b	M1c	M2a	M2b	M2c	M3a	M3b	M3c
Slope ($\sigma_{\mathrm{R},i}$) (3.6)	Econ. L12	2.61	2.58	2.61	4.01	3.98	4.01	2.62	2.60	2.62
	Econ. N12	2.89	2.89	2.88	4.85	4.85	4.84	2.86	2.86	2.84
	Fin. L12	0.54	0.54	0.54	0.59	0.58	0.59	0.47	0.46	0.46
	Fin. N12	0.60	0.60	0.60	1.05	1.04	1.04	0.55	0.54	0.54
	Major pur.	0.94	0.93	0.94	1.13	1.12	1.13	0.88	0.87	0.87
Seasonal ($\sigma_{\mathrm{S},i}$) (3.7)	Econ. L12	0.03	0.03	0.03	0.00	0.02	0.02	0.02	0.02	0.03
	Econ. N12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Fin. L12	0.02	0.03	0.03	0.02	0.02	0.02	0.02	0.03	0.03
	Fin. N12	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.05	0.05
	Major pur.	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
Population irregular term ($\sigma_{I,i}$) (3.8)	Econ. L12	3.47			2.76			3.49		
	Econ. N12	5.94			4.87			5.87		
	Fin. L12	1.90			1.96			1.93		
	Fin. N12	2.25			2.10			2.09		
	Major pur.	2.15			2.13			2.17		
Sampling error ($\sigma_{\varepsilon,i}$)*) (3.9)	Econ. L12	1.00			1.00			1.00		
	Econ. N12	1.00			1.00			1.00		
	Fin. L12	1.00			1.00			1.00		
	Fin. N12	1.00			1.00			1.00		
	Major pur.	1.00			1.00			1.00		

Table continues on next page

Hyperparameter	Variable					Sd				
		M1a	M1b	M1c	M2a	M2b	M2c	M3a	M3b	M3c
Measurement error ($\sigma_{e,i}$) (3.10)	Econ. L12		2.47	3.83		2.05	3.18		2.47	3.83
	Econ. N12		3.85	6.14		3.21	5.12		3.81	6.08
	Fin. L12		1.55	2.49		1.58	2.53		1.57	2.51
	Fin. N12		1.77	2.73		1.70	2.61		1.69	2.62
	Major pur.		1.67	2.68		1.65	2.66		1.67	2.69

Table 3.1: maximum likelihood estimates of standard deviations slope disturbance terms, seasonal disturbance terms and the measurement error
***): the standard deviation of the sampling errors are not estimated by maximum likelihood but are set equal to 1.**

	Econ. L12	Econ. N12	Fin. L12	Fin. N12	Major pur.
Econ. L12	1				
Econ. N12	0.865	1			
Fin. L12	0.600	0.371	1		
Fin. N12	0.909	0.929	0.584	1	
Major pur.	0.528	0.384	0.771	0.614	1

Table 3.2: correlations slope disturbances, model with correlations slope disturbances and correlations measurement (Model M1a)

	Econ. L12	Econ. N12	Fin. L12	Fin. N12	Major pur.
Econ. L12	1				
Econ. N12	0.636	1			
Fin. L12	0.120	-0.110	1		
Fin. N12	0.262	0.423	0.322	1	
Major pur.	0.284	0.265	-0.127	-0.002	1

Table 3.3: correlations population irregular terms for a model with correlations between slope disturbances and correlations between population irregular terms (Model M1a)

	Econ. L12	Econ. N12	Fin. L12	Fin. N12	Major pur.
Econ. L12	1				
Econ. N12	0.564	1			
Fin. L12	0.090	-0.076	1		
Fin. N12	0.208	0.342	0.204	1	
Major pur.	0.212	0.209	-0.075	0.002	1

Table 3.4: correlations measurement errors for a model with correlations between slope disturbances and correlations between measurement errors scaled with sampling errors (Model M1b)

	Econ. L12	Econ. N12	Fin. L12	Fin. N12	Major pur.
Econ. L12	1				
Econ. N12	0.562	1			
Fin. L12	0.089	-0.079	1		
Fin. N12	0.199	0.340	0.203	1	
Major pur.	0.211	0.207	-0.078	0.002	1

Table 3.5: correlations measurement error for a model with correlations between slope disturbances and correlations between measurement errors not scaled with sampling errors (Model M1c)

	Econ. L12	Econ. N12	Fin. L12	Fin. N12	Major pur.
Econ. L12	1				
Econ. N12	0.956	1			
Fin. L12	0.712	0.577	1		
Fin. N12	0.971	0.978	0.711	1	
Major pur.	0.739	0.669	0.827	0.754	1

Table 3.6: correlations slope disturbances for a model with correlations slope between disturbances and without correlations between population irregular terms (Model M2a)

The standard errors of the filtered and smoothed signals for consumer confidence under the three models are compared with the standard errors of the direct estimates in Figure 3.3. The standard errors under Model M1a are clearly larger than under Model M2a and M3a. The standard errors of the filtered estimates under Model M1a are clearly larger than those for the direct estimates. The standard errors of the smoothed estimates under Model M1a and the standard errors of the filtered estimates under Model M2a and M3a are more or less the same as those of the direct estimates. For the other two combined series, economic climate and willingness to buy, the standard errors of the filtered and smoothed signals under Model M1a are also larger than under Models M2a and M3a. For economic climate the standard error of the direct estimates is smaller than the standard errors of the filtered signals for all three models. The standard errors of the smoothed signals for Model M2a and M3a are more or less equal to the standard errors of the direct estimates. For willingness to buy the standard errors of the filtered as well as the smoothed signals under all three models are smaller than those of the direct estimates (results not shown).

It is a remarkable result that the standard error of the filtered signals are equal or even larger than the standard errors of the direct estimates. A general finding in the literature is that state space models applied to series obtained with repeated surveys result in model estimates with standard errors that are substantially smaller compared to the standard errors of the direct survey estimates, see e.g. Pfeffermann and Bleuer (1993), Pfeffermann and Burck (1990), Pfeffermann and Tiler (2006), Krieg and Van den Brakel (2012), Van den Brakel and Krieg (2015, 2016), Boonstra and van den Brakel, (2019). The reason that this is not the case for the Dutch CS is explained as follows. Recall from the discussion of the maximum likelihood estimates in Table 3.1 that the population irregular term is at least as large as the sampling error. This additional uncertainty of the population irregular term is reflected in the standard error of the time series model estimates, which are estimates for the trend, L_t, or the signal, $L_t + S_t$. The sample estimates of the CS are direct or design-based estimates for the population parameter θ_t. Their standard errors, on the other hand, only contain the uncertainty due to the sampling error. Furthermore, the positive correlations between the measurement errors particularly increase the standard errors of the composite indices. It is nevertheless essential to model the correlation between the measurement errors since this results in a better separation of the measurement error from trend and signal.

The filtered trend estimates are volatile. Together with the relative large population irregular terms, this is an indication that the questions of the CS measure a short-term emotion and are not interpreted by the respondents as a long term evaluation over the last and next 12 months of the economy and the financial situation. This observation is supported by the fact that the time series contain a seasonal pattern, which would not be present if questions are interpreted as the situation over the last 12 and next 12 months. As a result, the largest contribution of the time series model in this application is obtained with the smoothed estimates, which are more stable and have smaller standard errors.

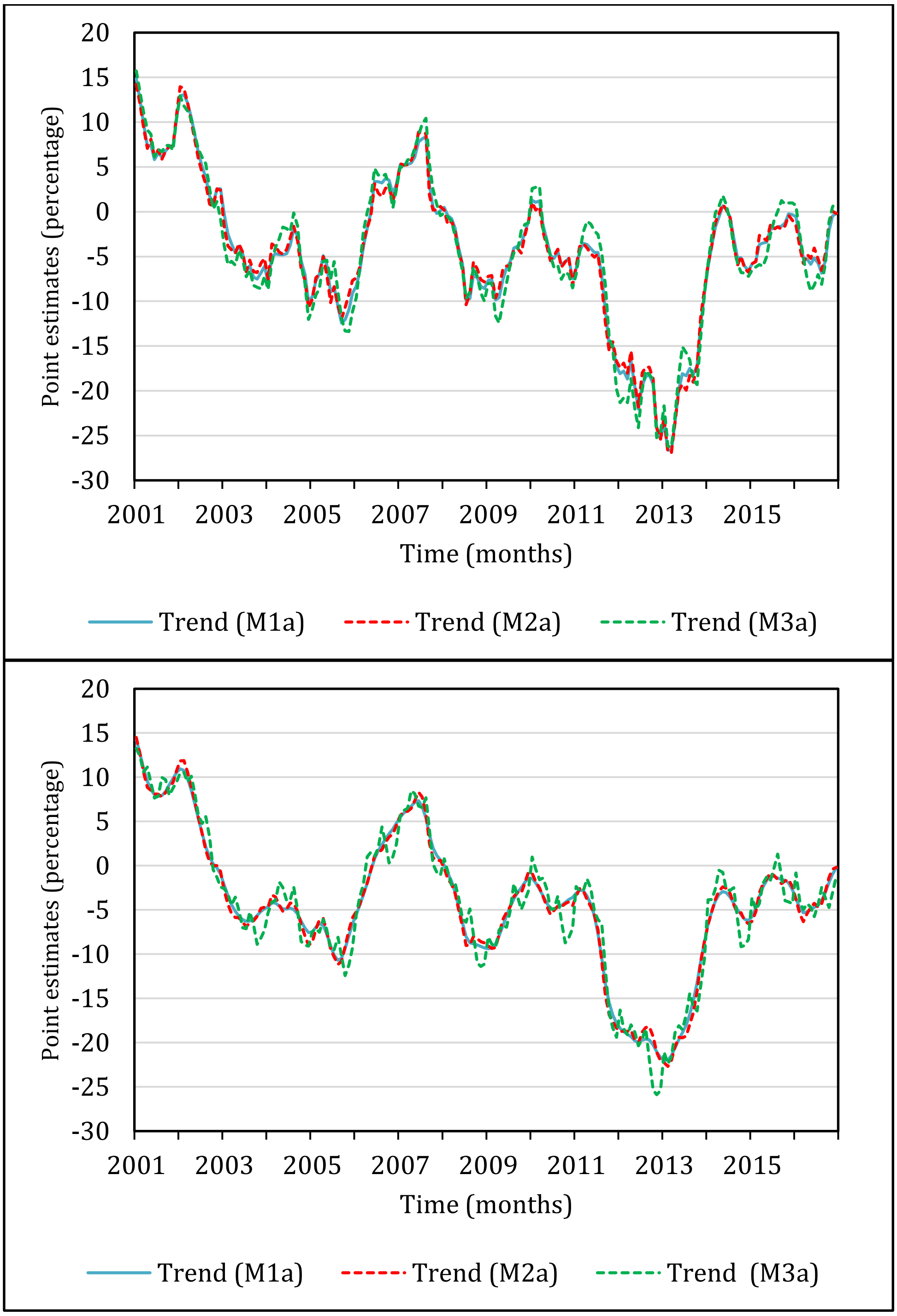

Figure 3.1: filtered trends (upper panel) and smoothed trends (lower panel) for three models for Fin. N12

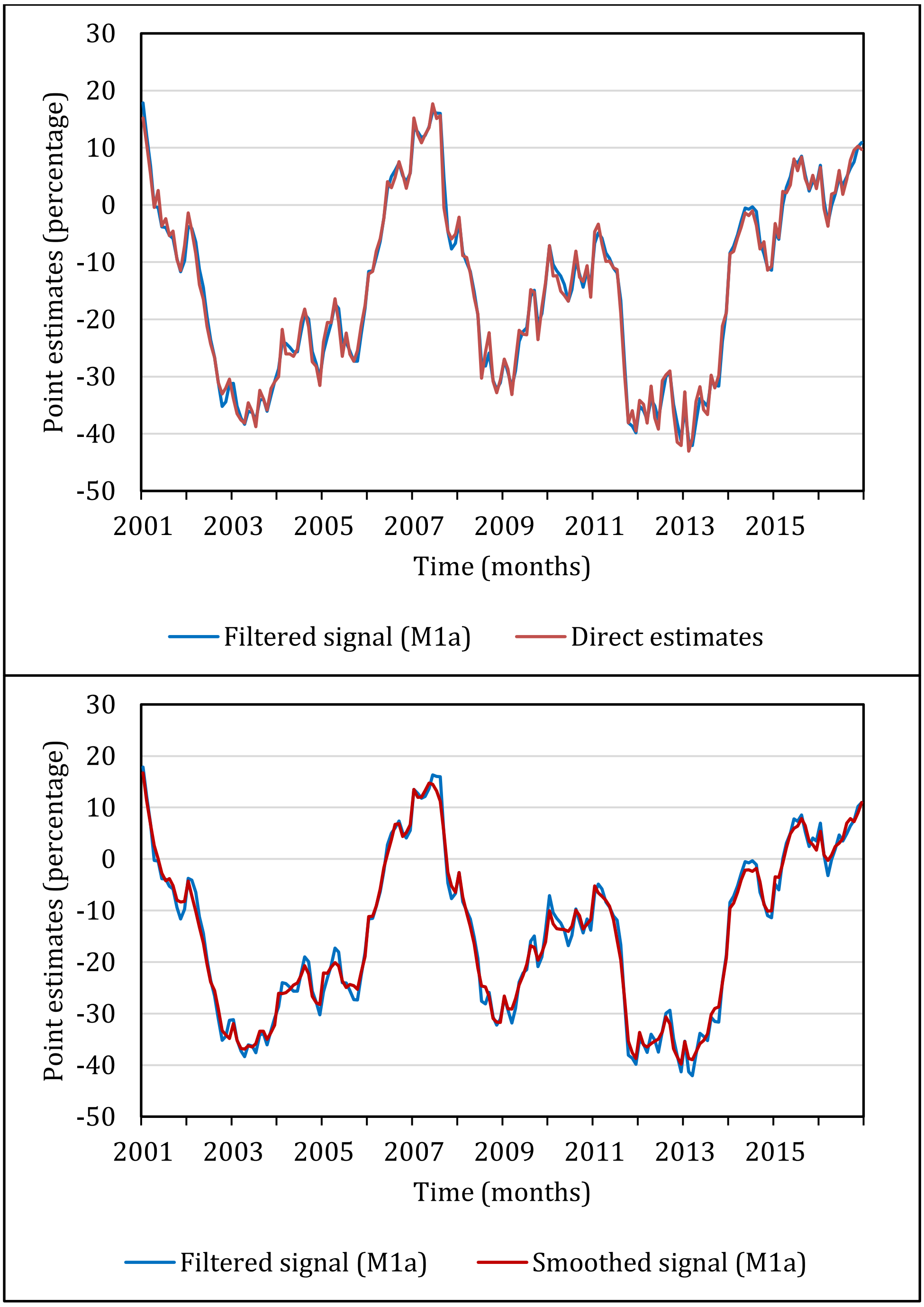

Figure 3.2: direct estimates and filtered signal (upper panel) and filtered and smoothed signal (lower panel) under model M1a for consumer confidence

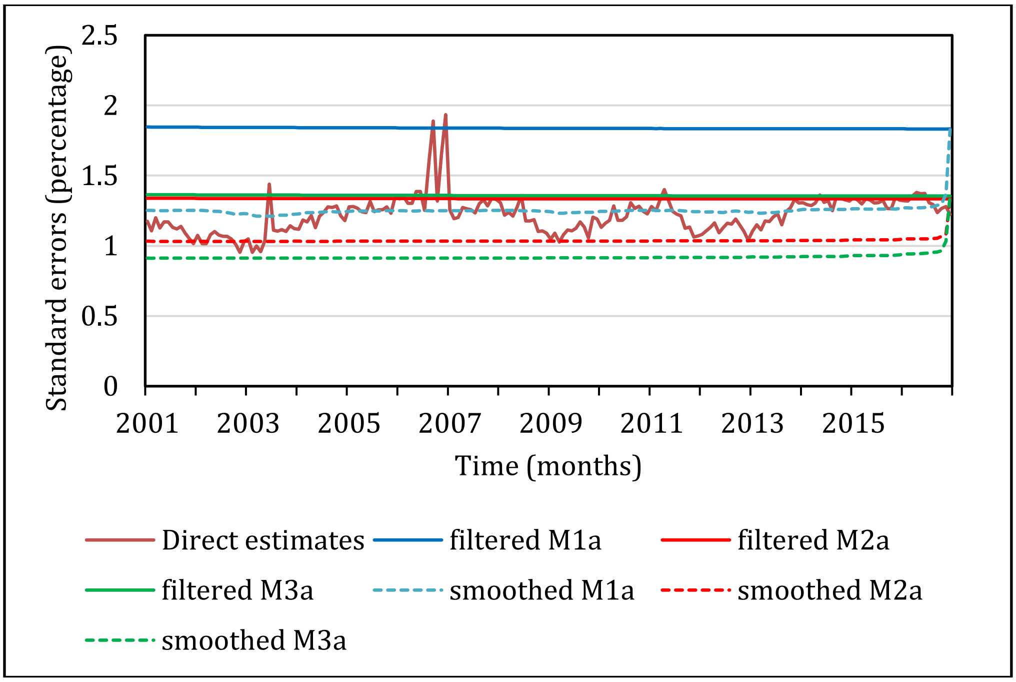

Figure 3.3: standard errors of direct estimates and of filtered and smoothed estimates of signal under three models for consumer confidence

Since the standard errors of the filtered signals are stable over time, the standard errors of the five baseline series and the three combined series are shown in Table 3.7 for December 2016 (last observation before the change-over to the new design). Standard errors for the smoothed signals are also stable over time, except for the last months of the series. Therefore, standard errors for the smoothed signals are included in Table 3.7 for December 2015. Standard errors for the trends are presented in a similar way in Table 3.8. The standard errors for the trends are very similar to the standard errors of the signals. Model M2a has the smallest standard errors for all variables. For some variables, the differences are substantial. For the five baseline series the standard error under Model M1a is smaller than those under Model M3a. For the combined series, however, the standard errors under Model M1a are larger than those under Model M3a. From the comparison between Model M2a and Model M3a, it follows that modelling cross-sectional correlations through the trend component improves the precision of the model estimates. Modelling the correlation between the population irregular terms decreases the precision, because the positive correlation inflates the variance of the population irregular terms of the combined series. As motivated above, it is necessary to account for these correlations to avoid underestimating the uncertainty of the model predictions.

A major advantage of inference based on time series models is that the gain in precision of period-to-period changes is large. To illustrate this, the standard errors of the month-to-month developments for filtered and smoothed signals for the consumer confidence under the three models and the direct estimates are compared in Figure 3.4. The period-to-period change and their standard errors are obtained by calculating the linear combination of $\Delta_t = L_{t|t} - L_{t-1|t} + S_{t|t} - S_{t-1|t}$ via the Kalman filter recursion. This requires that the state variables for L_{t-1} and S_{t-1} remain in the state vector. For the direct estimates these standard errors are larger, since the direct estimates of two different periods in a cross-sectional survey are independent. For the model estimates, however, these standard errors are smaller, mainly due to the strong positive correlation between the trend levels of two subsequent periods.

	Model M1a		Model M2a		Model M3a	
	filt.	smooth.	filt.	smooth.	filt.	smooth.
Econ. L12	3.22	2.27	2.68	2.01	3.24	2.27
Econ. N12	4.78	3.16	3.97	2.81	4.86	3.21
Econ. climate	3.60	2.45	2.65	2.03	2.92	1.98
Fin. L12	1.69	1.16	1.66	1.13	1.77	1.20
Fin. N12	1.73	1.26	1.58	1.21	1.93	1.38
Major pur.	2.18	1.79	2.30	1.78	2.23	1.85
Willingn. to buy	1.20	0.90	1.14	0.87	1.14	0.88
Consumer conf.	1.84	1.27	1.33	1.05	1.35	0.95

Table 3.7: standard error filtered estimates signal last period (December 2016) and smoothed estimates (December 2015) for 8 series

	Model M1a		Model M2a		Model M3a	
	filt.	smooth.	filt.	smooth.	filt.	smooth.
Econ. L12	3.34	2.24	2.83	2.24	3.38	2.28
Econ. N12	4.84	3.09	4.05	2.95	4.97	3.18
Econ. climate	3.70	2.45	2.89	2.15	3.01	1.95
Fin. L12	1.64	0.96	1.62	0.94	1.75	1.05
Fin. N12	1.60	0.92	1.45	0.89	1.91	1.12
Major pur.	2.26	1.28	2.25	1.29	2.37	1.43
Willingn. to buy	1.23	0.74	1.09	0.73	1.17	0.83
Consumer conf.	1.80	1.24	1.43	1.14	1.39	0.91

Table 3.8: standard error filtered estimates trend last period (December 2016) and smoothed estimates (December 2015) for 8 series

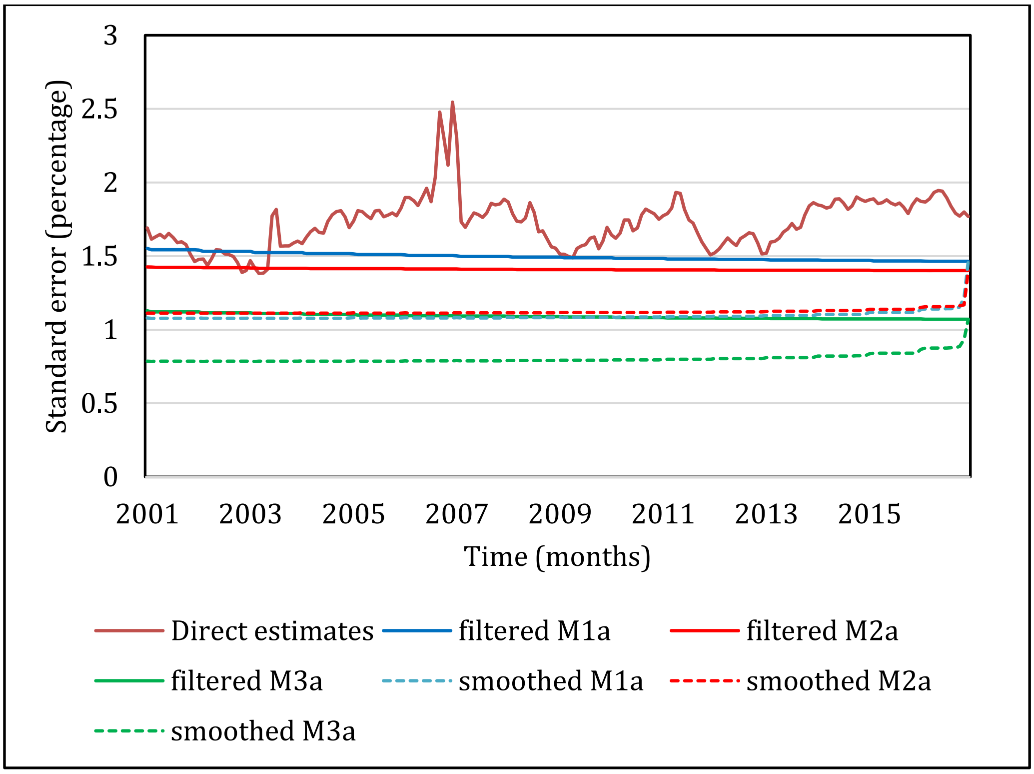

Figure 3.4: standard errors for month-to-month-development direct estimates and filtered estimates signal under three models, consumer confidence

3.3 Model evaluation

The assumptions underlying the state space model are evaluated by testing whether the standardized innovations are standard normally and independently distributed, see Durbin and Koopman (2012), Sections 2.12 and 7.5. Different tests (Bowman-Shenton normality tests, F-tests for heteroscedasticity, QQ-plots, plots of standardized innovations and sample correlograms, Durbin Watson test) indicate some small violations of these assumptions for all models. The results on normality and heteroscedasticity are comparable under all models. The correlograms show very weak autocorrelation of lag 1 for some of the series, under all versions of model 1, 2 and 3. Since the input series are very long, even small violations of the model assumptions could be significant, but the violations we found here are acceptable. Alternative models for trend, i.e., local level model or a local linear trend model, and a seasonal component with separate variance components for the harmonics in (3.7) or including AR(1), MA(1) or ARMA(1,1) components did not improve these diagnostics.

The models are also compared using AIC and BIC which are defined as $\mathrm{AIC} = -2 * \mathrm{LL} + 2(q+p)$ and $\mathrm{BIC} = -2 * \mathrm{LL} + (q+p) * \mathrm{Log}(T-p)$, with LL the loglikelihood, q the number of hyperparameters estimated with maximum likelihood, p the number state variables and T the length of the observed series. Results are presented in Table 3.9. Under both criteria Model M1a is preferred. It can be concluded that the models with a separate component for the population irregular term and the sampling error, i.e. Model (3.4.a), outperforms the models that combine the population irregular term and sampling error in one measurement error, i.e. Models (3.4.b) and (3.4.c). Model M1a, M2a and M3a are nested and can be compared with a likelihood ratio (LR) test, see Table 3.10. Similar results are obtained for the three different versions of Model 2 and the three different versions of Model 3.

Model	Loglikelihood	AIC	BIC
M1a	-5293.3	10706.52	10937.48
M2a	-5347.2	10794.48	10986.95
M3a	-5501.8	11083.68	11237.65
M1b	-5303.1	10726.24	10957.20
M2b	-5357.4	10814.88	11007.35
M3b	-5514.8	11109.62	11263.59
M1c	-5305.8	10731.60	10962.56
M2c	-5360.3	10820.68	11013.15
M3c	-5518.1	11116.20	11270.17

Table 3.9: AIC and BIC values for the three models

Comparison	LR statistic	df	p-value
Model M1a versus Model M2a	107.96	10	0.000
Model M2a versus Model M3a	309.20	10	0.000
Model M1a versus Model M3a	417.16	20	0.000

Table 3.10: Results likelihood ratio tests

The test for Model M2a versus Model M3a indicates that modelling the correlation between the slope disturbance terms significantly improves the model fit. The test for Model M1a versus Model M2a shows that modelling the correlation between the measurement errors further improves the model fit significantly. Finally the test for Model M1a versus Model M3a shows that the joint test on the inclusion of a full covariance matrix for the slope disturbance terms and the measurement errors rejects the null hypothesis that both models are equivalent. A model that allows for correlated slope disturbance terms must also allow for correlated measurement errors. Otherwise, correlated measurement errors in all input series could be incorrectly interpreted as a true development of the trend instead of measurement errors (sampling noise or noise in the population parameter). In conclusion, Model M1a is the best fitting model.

4. Discontinuities

An important aspect of the implementation of the new design in 2017 is to quantify and correct for discontinuities in the outcomes of the Dutch CS that are the result of the implementation of a new survey process. In this application discontinuities are analyzed at the level of the percentage of positive, neutral or negative answer categories of the five questions, i.e., p_i^+, p_i^-, p_i^0 for $i = 1, \ldots, 5$, as these percentages are the variables measured through the questionnaire. These discontinuity estimates are used to compute uninterrupted series for the percentages $p_{i,t}^+$, $p_{i,t}^0$, $p_{i,t}^-$ for $i = 1, \ldots, 5$, by adjusting the series observed before the change-over to the level of the series observed under the new design. These corrected series are used in a second step to calculate uninterrupted series for $y_{1,t}, \ldots, y_{5,t}$. These backcasted series will be used as the input for model (3.4) that is used in the production of official monthly figures about consumer confidence. The method to adjust the percentage series observed before the change-over will be worked out in this section. The proposed method ensures that the size of the correction at a specific time period depends on the size of the discontinuities of these percentages and the share of the percentages over the three categories at that time period. It will be shown that such a correction is more realistic than an approach that directly estimates and adjusts discontinuities at the level of the input series $y_{1,t}, \ldots, y_{5,t}$.

4.1 Quantifying discontinuities

In this application, discontinuities are estimated by means of a parallel run in combination with a time series intervention analysis, following the method proposed by Van den Brakel et al. (2020). The available budget allowed a parallel run of three months in the first quarter of 2017, where the sample sizes for both designs were equal to the net sample size of the regular survey, i.e., around 1000 persons. Initial direct estimates for the discontinuities are obtained as contrasts between direct or design-based sample estimates under the old and new survey design. The precision of the estimated discontinuities can be further improved by modelling the observed series with a STM. The population parameter of interest is modelled with a trend, seasonal component, population irregular term and the sampling error, similar to the models developed in Section 3. The discontinuity is modelled with an intervention variable that switches from zero to one at the moment that the survey is transferred from the old to the new design. Under the assumption that the other components of the time series model (trend and seasonal) describe the evolution of the population parameter correctly, the regression coefficient of the intervention variable can be interpreted as an estimate for the discontinuity. The level intervention approach with state space models was originally proposed by Harvey and Durbin (1986) to estimate the effect of seat belt legislation on British road casualties. If no information from a parallel run would be available, the regression coefficient for the intervention variable would be initialized in the Kalman filter with a diffuse prior. The direct estimate for the discontinuity and its variance obtained from the parallel run can be used as an informative initialization for the regression coefficient of the level intervention in the Kalman filter. In this way, the discontinuity estimate from the parallel run is improved with the available information from the entirely observed time series.

Based on the parallel run there are two estimates for the percentages of the five questions $p_{i,t}^{+}$, $p_{i,t}^{0}$, $p_{i,t}^{-}$, which are denoted $\hat{p}_{i,t}^{j,\mathrm{O}}$ ($j \in \{+,0,-\}$) for the estimate of $p_{i,t}^{j}$ under the old design and $\hat{p}_{i,t}^{j,\mathrm{N}}$ for the estimate of $p_{i,t}^{j}$ under the new design. Direct estimates for the discontinuities are obtained as:

$$\widehat{\Delta}(p_i^j) = \frac{1}{3}\sum_{t=2017(1)}^{2017(3)}\left[\hat{p}_{i,t}^{j,\mathrm{N}} - \hat{p}_{i,t}^{j,\mathrm{O}}\right], \text{ for } j \in \{+,0,-\}. \tag{4.1}$$

Since the percentages $p_{i,t}^{+}$, $p_{i,t}^{0}$, $p_{i,t}^{-}$ for each question i sum to 100%, it follows that

$$\sum_{j\in\{+,0,-\}} \widehat{\Delta}(p_i^j) = 0\,. \tag{4.2}$$

The variance of the estimates of the discontinuities can be estimated by

$$\widehat{\mathrm{Var}}[\widehat{\Delta}(p_i^j)] = \frac{1}{9}\sum_{t=2017(1)}^{2017(3)}[\widehat{\mathrm{Var}}(\hat{p}_{i,t}^{j,\mathrm{O}}) + \widehat{\mathrm{Var}}(\hat{p}_{i,t}^{j,\mathrm{N}})], \tag{4.3}$$

with

$$\widehat{\mathrm{Var}}\left(\hat{p}_{i,t}^{j,\mathrm{d}}\right) = \hat{p}_{i,t}^{j,\mathrm{d}}(100 - \hat{p}_{i,t}^{j,\mathrm{d}})/n_t^{\mathrm{d}}, \text{ for } d \in \{\mathrm{O},\mathrm{N}\}$$

and $n_t^{\mathrm{O}}, n_t^{\mathrm{N}}$ the sample size in month t under the old and new design.

In a next step a three dimensional multivariate model is applied to the three series with percentages of positive, negative and neutral answers for each question separately, i.e. for $i = 1, \dots, 5$. For notational convenience, the index i is omitted in the formulas. It is understood that the observations of each series up until and including December 2016 are based on the old design. From January 2017 on, the observations are based on the new design. For each of the three series the basic structural time series model is extended with a level intervention. In a similar way as in Section 3, three models that account in different ways for the sampling error and the population irregular term are compared. The first model has a separate component for the population irregular term and the sampling error:

$$\widehat{\mathbf{p}}_t = \mathbf{L}_t + \mathbf{S}_t + \boldsymbol{\beta}'\,\mathbf{x}_t + \boldsymbol{I}_t + \boldsymbol{K}_t\boldsymbol{\varepsilon}_t, \tag{4.4.a}$$

where $\widehat{\mathbf{p}}_t = (\hat{p}_t^{+}, \hat{p}_t^{0}, \hat{p}_t^{-})'$ is the vector of direct estimates of the percentages, until 2016 based on the old design and from 2017 based on the new design, $\mathbf{L}_t = (L_t^{+}, L_t^{0}, L_t^{-})'$ is the vector of the trends, $\mathbf{S}_t = (S_t^{+}, S_t^{0}, S_t^{-})'$ is a vector of the seasonal patterns, $\mathbf{x}_t = (x_t, x_t, x_t)'$ the level intervention variable, i.e. x_t switches from zero to one in January 2017 when the new design is implemented and $\boldsymbol{\beta} = (\beta^{+}, \beta^{0}, \beta^{-})'$ the regression coefficients of the level interventions that can be interpreted as approximations of the discontinuities. Finally, $\boldsymbol{I}_t = (I_t^{+}, I_t^{0}, i_t^{-})'$ is a vector with population irregular term components, $\boldsymbol{\varepsilon}_t = (\varepsilon_t^{+}, \varepsilon_t^{0}, \varepsilon_t^{-})'$ is a vector containing normally distributed error terms that are pre-multiplied by a 3×3 diagonal matrix $\boldsymbol{K}_t$, with diagonal elements equal to the standard error of the input series, i.e.

$$k_t^j = \sqrt{\mathrm{Var}(\hat{p}_t^j)} = \sqrt{\hat{p}_t^j(100 - \hat{p}_t^j)/n_t}, \text{ with } j \in \{+,0,-\}.$$

The second model combines the population irregular term and the sampling error into one measurement error, say $\mathbf{e}_t = \boldsymbol{I}_t + \boldsymbol{\varepsilon}_t$, that is scaled with the standard errors of the input series:

$$\widehat{\mathbf{p}}_t = \mathbf{L}_t + \mathbf{S}_t + \boldsymbol{\beta}' \, \mathbf{x}_t + \boldsymbol{K}_t \mathbf{e}_t. \tag{4.4.b}$$

The third model combines the population irregular term and the sampling error into one measurement error and ignores the time varying behavior of the sampling errors:

$$\widehat{\mathbf{p}}_t = \mathbf{L}_t + \mathbf{S}_t + \boldsymbol{\beta}' \, \mathbf{x}_t + \mathbf{e}_t. \tag{4.4.c}$$

The variables L_t^j and S_t^j with $j \in \{+,0,-\}$ are smooth trend models and trigonometric seasonal models as described in Section 3.1. For the disturbance terms it is assumed that they are mutually independent, normally distributed with expectation zero and time-independent variance components. From (4.2) it follows that the coefficients for the discontinuities, β^+, β^0 and β^-, must obey the restriction that they add up to zero. This is enforced with the following transition equations in the state space model:

$$\begin{array}{lcl} \beta_t^+ & = & \beta_{t-1}^+ \\ \beta_t^- & = & \beta_{t-1}^- \\ \beta_t^0 & = & -\beta_{t-1}^+ - \beta_{t-1}^- \end{array} . \tag{4.5}$$

The subscript t indicates the notation of the transition equations. As there is no disturbance term, $\boldsymbol{\beta}$ is still time-independent. The population irregular terms ($\boldsymbol{I}_t$), the scaled sampling errors ($\boldsymbol{\varepsilon}_t$) and the measurement errors ($\mathbf{e}_t$) are modelled as normally and independently distributed random variables, i.e., $I_t^j \sim N(0, \sigma_{I,j}^2)$, $\varepsilon_t^j \sim N(0, \sigma_{\varepsilon,j}^2)$, $e_t^j \sim N(0, \sigma_{e,j}^2)$, $j \in \{+,0,-\}$. As in Section 3, the variances of scaled sampling errors in Model (4.4.a) are not estimated with maximum likelihood but are taken equal to one to force that the variance of $k_t^j \varepsilon_t^j$ is equal to the variance of the sampling error $\mathrm{Var}(\hat{p}_t^j)$. Finally note that since the neutral category is not used in the indices of consumer confidence, neither in the baseline nor the combined indices, a pragmatic alternative is to model the series of positive and negative percentages in a bivariate model without restrictions on the coefficients of the discontinuities. The state space representation of Models (4.4.a), (4.4.b) and (4.4.c) is given in the Appendix.

4.2 Correction methods for discontinuities

As explained at the beginning of Section 4, the series of the percentages observed before the redesign of January 2017 are corrected to the level of the percentage series observed under the new design. These so-called backcasted percentage series are used to compile backcasted input series for the time series model (3.4). The percentages can only have admissible values in the range [0,100]. Therefore correction methods are considered that result in backcasted series that have values in this admissible range. In this case the backcasted target variables $\hat{y}_1, \dots, \hat{y}_5$ will also have values in the admissible range [-100, +100]. Let $\tilde{p}_{i,t}^{j,\mathrm{N}}$ denote the backcasted series of $\hat{p}_{i,t}^{j,\mathrm{O}}$. The following correction is proposed for percentages:

$$\tilde{p}_{i,t}^{j,\mathrm{N}} = \hat{p}_{i,t}^{j,\mathrm{O}} + \hat{\beta}_i^j \frac{\hat{p}_{i,t}^{j,\mathrm{O}}(100-\hat{p}_{i,t}^{j,\mathrm{O}})}{\hat{p}_{i,\tau}^{j,\mathrm{O}}(100-\hat{p}_{i,\tau}^{j,\mathrm{O}})}, \text{ for } t = 1, \dots, T-1, \tag{4.7}$$

with T the month of the change-over to the new design, which is January 2017. Furthermore, $\hat{p}_{i,\tau}^{j,\mathrm{O}}$ denotes an estimate under the old design obtained during the entire period of the parallel run (denoted by τ). The estimated discontinuity $\hat{\beta}_i^j$ is multiplied by a factor proportional to the variance of the percentage, estimated by $\hat{p}_{i,t}^{j,\mathrm{O}}(100 - \hat{p}_{i,t}^{j,\mathrm{O}})$. The correction is zero when $\hat{p}_{i,t}^{j,\mathrm{O}} = 0$ or $\hat{p}_{i,t}^{j,\mathrm{O}} = 100$, and it is maximal when $\hat{p}_{i,t}^{j,\mathrm{O}} = 50$. The denominator of the factor is the population variance of the percentage under the old design during the three months of the parallel run and ensures that the corrected percentage estimate obtained for these three months are close to the values observed under the new design during the parallel run.

When all three percentages $\hat{p}_t^j$ for $j \in \{+,0,-\}$ are corrected with (4.7) the sum of the corrected percentages is no longer 100. Since the neutral percentage $\hat{p}_t^0$ is not used in the computation of the indicators y_i , this percentage is corrected as $\tilde{p}_{i,t}^{0,\mathrm{N}} = 100 - \tilde{p}_{i,t}^{+,\mathrm{O}} - \tilde{p}_{i,t}^{-,\mathrm{O}}$. The size of the correction in (4.7) diminishes when the percentage $\hat{p}_{i,t}^{j,\mathrm{O}}$ is close to 0 or 100. It is nevertheless not guaranteed that the values of $\tilde{p}_{i,t}^{j,\mathrm{N}}$ are in the admissible range of [0,100]. They can take values outside this range when the percentages during the parallel run $\hat{p}_{i,\tau}^{j,\mathrm{O}}$, are close to 0 or 100 and the discontinuity β_i^j is large. Note that the underlying assumption of the proposed correction method is that the discontinuity is small when the variance is small. This assumption would be violated in this case.

As an alternative approach, an additive correction after applying a logratio transformation is considered. This forces that the adjusted series obey the restriction that they do not take values outside the admissible range of [0,100] and that the sum over the three categories equals 100. The neutral percentages $\hat{p}_{i,t}^{0,d}$, for $d \in \{\mathrm{O}, \mathrm{N}\}$, are used as the reference category in the denominator of the logratio transformation, which is defined as:

$$z_{i,t}^{+,d} = \ln\left(\frac{\hat{p}_{i,t}^{+,d}}{\hat{p}_{i,t}^{0,d}}\right), \quad z_{i,t}^{-,d} = \ln\left(\frac{\hat{p}_{i,t}^{-,d}}{\hat{p}_{i,t}^{0,d}}\right), \quad \text{for } d \in \{\mathrm{O}, \mathrm{N}\}. \tag{4.8}$$

As a next step, series observed under the old design are adjusted to the level under the new design using the following additive correction to the logratio transformed series:

$$\tilde{z}_{i,t}^{+,N} = \ln\left(\frac{\hat{p}_{i,t}^{+,O}}{\hat{p}_{i,t}^{0,O}}\right) + \ln\left(\frac{\hat{\beta}_i^{+}}{\hat{\beta}_i^{0}}\right), \ \tilde{z}_{i,t}^{-,N} = \ln\left(\frac{\hat{p}_{i,t}^{-,O}}{\hat{p}_{i,t}^{0,O}}\right) + \ln\left(\frac{\hat{\beta}_i^{-}}{\hat{\beta}_i^{0}}\right), \text{for } t = 1, \dots, T-1. \tag{4.9}$$

In (4.9), $\hat{\beta}_i^{+}$, $\hat{\beta}_i^{0}$ and $\hat{\beta}_i^{-}$ denote the Kalman filter estimates for the discontinuities obtained with model (4.4) using the information from the parallel run through the informative initialization of the Kalman filter as described in Subsection 4.1. Subsequently, the anti-logratio transformation,

$$\tilde{p}_{i,t}^{+,N} = 100\frac{e^{\tilde{z}_{i,t}^{+,N}}}{e^{\tilde{z}_{i,t}^{+}} + e^{\tilde{z}_{i,t}^{-}} + 1}, \ \tilde{p}_{i,t}^{-,N} = 100\frac{e^{\tilde{z}_{i,t}^{-,N}}}{e^{\tilde{z}_{i,t}^{+}} + e^{\tilde{z}_{i,t}^{-}} + 1}, \ \tilde{p}_{i,t}^{0} = 100\frac{1}{e^{\tilde{z}_{i,t}^{+}} + e^{\tilde{z}_{i,t}^{-}} + 1}, \tag{4.10}$$

for $t = 1, \dots, T-1$, is applied to obtain three corrected series for the percentages that obey the restriction that they take values in the admissible range [0,100] and add up to 100%. A drawback of the logratio transformation is that the effect of the correction can become very large when the ratios $\hat{p}_{i,t}^{+,\mathrm{O}}/\hat{p}_{i,t}^{0,\mathrm{O}}$ or $\hat{p}_{i,t}^{-,\mathrm{O}}/\hat{p}_{i,t}^{0,\mathrm{O}}$ are smaller than 1. This will be demonstrated in Section 4.3.

4.3 Estimation results for discontinuities

Table 4.1 shows the estimates of the discontinuities obtained with the parallel run for the five variables. In this period most of the respondents were positive about the economic situation. For the variables Econ. L12, Econ. N12, Fin. L12 and Fin. N12 the percentage of “a little better” increased substantially under the new design. Also the percentage of “a little worse” increased, but to a smaller extent (not presented in Table 4.1). The percentages of the other answer options all decreased. These changes can be explained partially by the changes in the questionnaire. Under the old design the respondent could, in addition to the neutral options, only choose between “better” and “worse”. When, according to the respondent, the situation was changed only a little, the options “better” or “worse ” did probably not feel appropriate, not knowing that “a little worse” or “a little better” are also possible answers. The respondent then probably chose the neutral answer category more often. Under the new design the respondent chooses “a little worse” or “a little better” and during the parallel run mostly “a little better” more frequently. The question major purchases, the only question where the questionnaire is not changed, is the only question where the discontinuity for the positive answers is negative, and smaller than for the other questions.

	positive answers $(\widehat{\Delta}(p_i^+))$		negative answers $(\widehat{\Delta}(p_i^-))$		indicator $(\widehat{\Delta}(p_i^+) - \widehat{\Delta}(p_i^-))$
Econ. L12	11.4	(1.3)	-0.1	(0.9)	11.5
Econ. N12	16.8	(1.2)	0.1	(0.8)	16.7
Fin. L12	8.7	(1.0)	3.2	(1.1)	5.5
Fin. N12	11.1	(1.0)	3.1	(0.9)	8.0
Major pur.	-4.7	(1.2)	0.1	(0.8)	-4.8

Table 4.1: estimates discontinuities obtained with the parallel run, standard errors in brackets

Models (4.4.a), (4.4.b) and (4.4.c) are applied to estimate the discontinuities. Results are presented for the three-dimensional model applied to the series of percentages. The series start in April 1986 and run up to February 2020. Up to and including December 2016 the estimates are based on the old design and starting from January 2017 they are based on the new design. Similar to the results in Section 3, filtered and smoothed estimates for trends, seasonals and discontinuities are very similar under the three Models (4.4.a), (4.4.b) and (4.4.c). The estimates of the discontinuities are shown in Table 4.2, 4.3 and 4.4. A few minor differences can be observed between the model that ignores the time varying behaviour of the sampling errors (Table 4.4) compared to the other models that account for the time varying sampling errors (Tables 4.2 and 4.3).

The Kalman filter estimates in Table 4.2 are comparable to the direct estimates based on the parallel run in Table 4.1. It is particularly nice to see that the estimates obtained with the parallel run are in line with the Kalman filter estimates with a diffuse initialisation, which does not use the results from the parallel run. As a result the point estimates obtained with a diffuse and an informed prior match well. If the additional information from the parallel run is included in the Kalman filter through an informed initialization, the standard errors of the discontinuity estimates clearly decrease. Under the diffuse initialization, the standard errors of the discontinuity estimates are substantially larger than the standard errors of the direct estimates obtained with the parallel run. In this situation, there is no control over the precision of the discontinuity, since the precision depends on the volatility of the time series at hand. This in contrary to the parallel run, where the precision can be controlled by calculating the minimum required sample size to observe a pre-specified difference at a pre-specified power and significance level. When the results of the parallel run are combined with the time series modelling approach through an informative initialization of the Kalman filter, then the most precise estimates for the discontinuities are obtained, since all available information from the observed time series before and after the change-over and the parallel run are combined. The improvement of the precision with respect to the time series model with a diffuse initialization is substantial. With respect to the direct estimates of the parallel run there is only a slight improvement of the precision of the discontinuity estimates in this application. This might, however, be different in other applications.

The estimates of the discontinuities obtained with the time series model are instable if only a few observations after the change-over are available. They improve when more data under the new design become available. Figures 4.1 and 4.2 show the real-time estimates for the discontinuities based on the data from January 2017 until February 2020 for the percentages of respectively the positive and negative answers for Econ. L12. The figures show how the estimates of the discontinuities evolve when more data become available. The initial estimates, directly after the change-over to the new design, are in the right order of magnitude. Nevertheless, there are some visible changes during the first six months. In the case of the informative initialization the changes are much smaller. During the first six months the standard errors of the estimates decrease, where the decrease for the diffuse initialization is substantial. After this period both the point estimates and the standard errors stabilize to constant values. It is difficult to conclude in general how many observations under the new design are required before a stable estimate for the discontinuities is obtained, since this depends on the flexibility of the trend.

	Diffuse Kalman filter initialization				Informative Kalman filter initialization			
	positive answers		negative answers		positive answers		negative answers	
Econ. L12	10.7	(3.0)	0.1	(2.5)	11.2	(1.1)	-0.1	(0.8)
Econ. N12	20.1	(3.3)	-0.8	(3.1)	17.2	(1.1)	0.2	(0.7)
Fin. L12	9.7	(1.2)	2.1	(1.5)	9.1	(0.8)	3.0	(0.9)
Fin. N12	12.2	(1.2)	4.6	(1.5)	11.7	(0.7)	3.5	(0.8)
Major pur.	-6.5	(1.8)	1.3	(1.5)	-5.1	(1.0)	0.3	(0.7)

Table 4.2: estimates discontinuities based on STM, standard errors in brackets under Model (4.4.a)

	Diffuse Kalman filter initialization				Informative Kalman filter initialization			
	positive answers		negative answers		positive answers		negative answers	
Econ. L12	10.3	(3.0)	0.1	(2.4)	11.2	(1.1)	-0.1	(0.8)
Econ. N12	20.2	(3.5)	-0.8	(3.1)	17.3	(1.1)	0.2	(0.7)
Fin. L12	9.7	(1.2)	2.2	(1.5)	9.1	(0.8)	3.0	(0.9)
Fin. N12	12.2	(1.2)	4.6	(1.5)	11.7	(0.7)	3.5	(0.8)
Major pur.	-6.5	(1.8)	1.3	(1.5)	-5.1	(1.0)	0.3	(0.7)

Table 4.3: estimates discontinuities based on STM, standard errors in brackets under Model (4.4.b)

	Diffuse Kalman filter initialization				Informative Kalman filter initialization			
	positive answers		negative answers		positive answers		negative answers	
Econ. L12	9.2	(2.6)	1.0	(2.9)	11.0	(1.1)	0.0	(0.8)
Econ. N12	19.3	(3.0)	0.5	(3.7)	17.3	(1.1)	0.2	(0.7)
Fin. L12	9.8	(1.2)	2.1	(1.5)	9.1	(0.8)	2.9	(0.9)
Fin. N12	12.1	(1.2)	4.9	(1.6)	11.7	(0.7)	3.6	(0.8)
Major pur.	-6.5	(1.8)	1.3	(1.8)	-5.2	(0.9)	0.2	(0.7)

Table 4.4: estimates discontinuities based on STM, standard errors in brackets under Model (4.4.c)

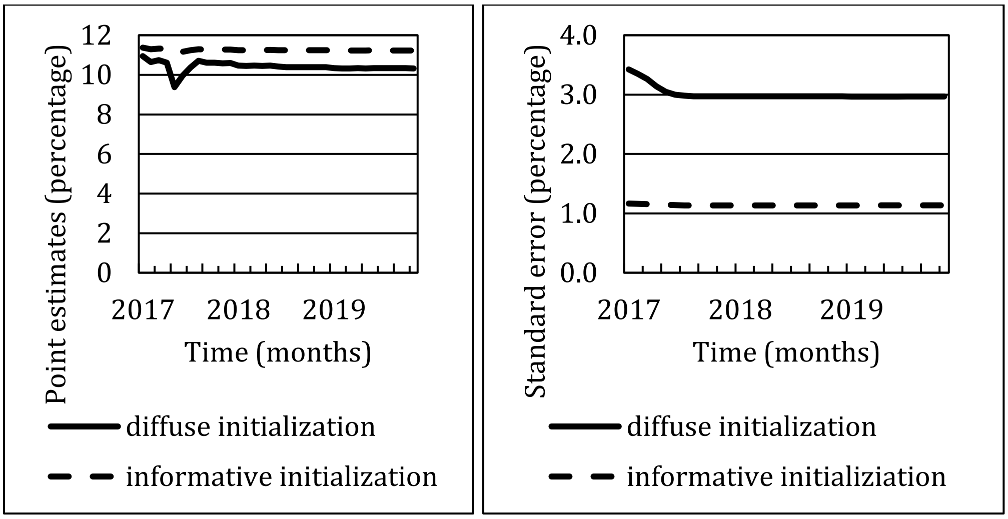

Figure 4.1: development of point estimates (left panel) and standard errors (right panel) of discontinuities percentages of positive answers, Econ. L12.

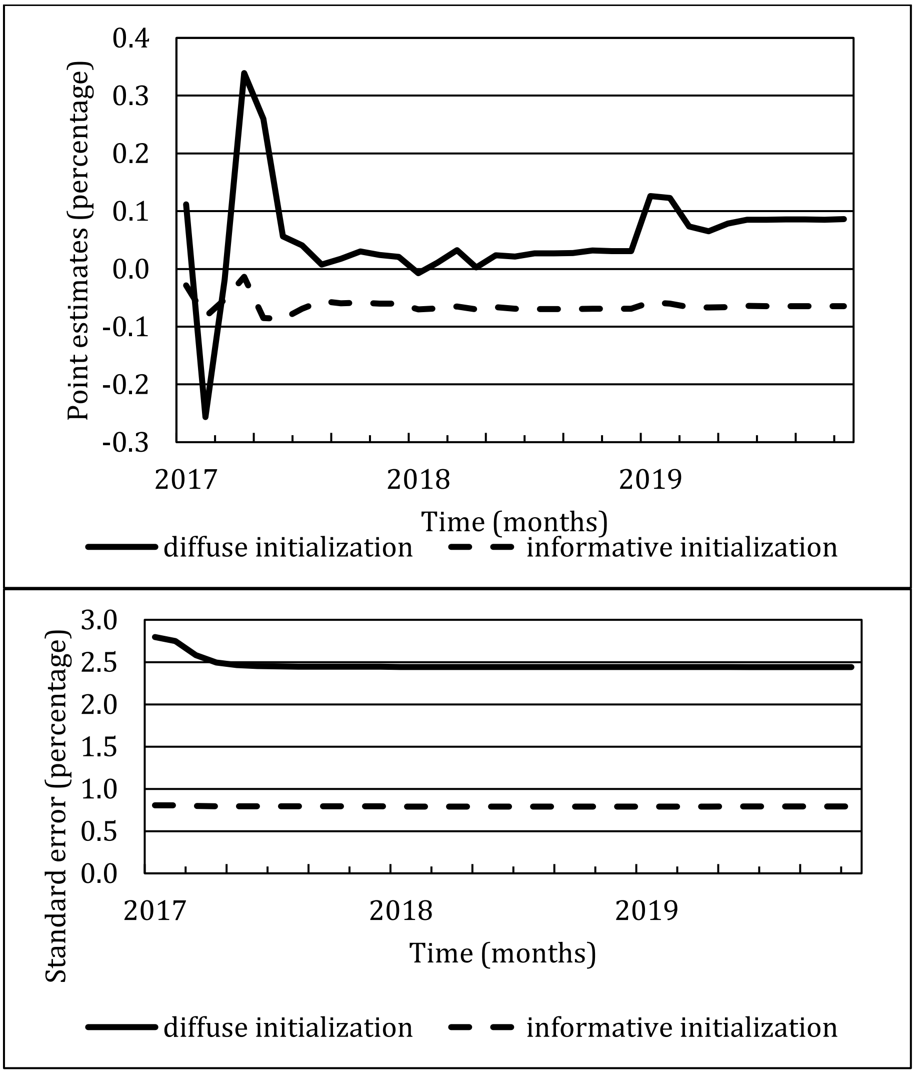

Figure 4.2: development of point estimates (upper panel) and standard errors (lower panel) of discontinuities of percentages negative answers, Econ. L12

4.4 Results for series corrected for discontinuities

In this section the effects of the correction methods described in Section 4.2 on the indicator series of the CS are investigated for the variables Econ. L12 and Major pur. The discontinuity estimates are based on the time series model (4.4.a) with an informative initialization of the Kalman filter. These discontinuity estimates are used to compute series adjusted to the level of the new design for the percentages $p_{i,t}^{+}$, $p_{i,t}^{0}$, $p_{i,t}^{-}$, using the proportional correction method defined by (4.7) and the additive correction method in combination with the logratio transformation defined by (4.9) and (4.10). These corrected series are used to calculate uninterrupted series for $y_{1,t}, \dots, y_{5,t}$, which are on their turn the input for Model (3.4.a).

Figure 4.3 shows the indicator series under the old design for Econ. L12 from January 2001 up to March 2017 together with the series corrected using the two correction methods. The discontinuities for this variable are $\widehat{\Delta}(p_1^{+}) = 11.2$ and $\widehat{\Delta}(p_1^{-}) = -0.1$ (Table 4.2) in a period where the consumer confidence is positive. Figure 4.3 shows that during periods where Econ. L12 has positive values, both methods correct in more or less the same way, i.e., the adjusted series become more positive and the corrections are equal. In periods where Econ. L12 has negative values, the logratio transformation makes the adjusted series more negative, while the proportionally corrected series stay close to the original series. This stronger correction under the logratio transformation in an opposite direction as observed during the parallel run seems less plausible compared to mild correction under the proportional correction.

Figure 4.4 shows the original and corrected series for the Major pur. For this variable the discontinuities are rather small, but compared to Econ. L12 they have opposite signs: $\widehat{\Delta}(p_5^{-}) = -5.1$ and $\widehat{\Delta}(p_5^{+}) = 0.3$. For Major pur. both correction methods give similar results: the values for the adjusted series are smaller than the original series and the effect of both corrections is the same. For this variable there is no preference for one of the two correction methods.

From these results there is a slight preference for the proportional correction, since the logratio transformation sometimes leads to larger corrections in a different direction as observed during the parallel run. A disadvantage of the proportional correction, however, is that the corrected values could fall outside the range of $[-100, +100]$. However, this only happens for unrealistic large values of the discontinuities. In Figure 4.5 it is simulated with arbitrary chosen values for the discontinuities for Econ. L12, namely $\widehat{\Delta}(p_1^{+}) = -10$ and $\widehat{\Delta}(p_1^{-}) = 20$, that the proportional correction leads to outcomes smaller than -100 in the most negative periods. In this case the logratio correction would be preferred.

Figure 4.6 shows that also with realistic values of the discontinuities the logratio transformation could result in large corrections. Here the correction methods are applied to the indicator series of Major pur. with arbitrarily, but realistic, chosen values for discontinuities of $\widehat{\Delta}(p_5^{+}) = -5$ and $\widehat{\Delta}(p_5^{-}) = -10$. In some periods the logratio correction is small (for example around 2007), but is extremely large in other periods (for example in 2003-2006 and 2009-2015). Furthermore, the logratio corrected series is always positive, even in periods where the original series is negative. This illustrates the disadvantage of the logratio correction, namely that the correction can become very large when the ratio of the original figure is much smaller than 1. This effect is also visible for the proportional correction, but to a lesser extent. From these results we conclude to use the proportional correction for backcasting the input series $\hat{y}_1, \dots, \hat{y}_5$.

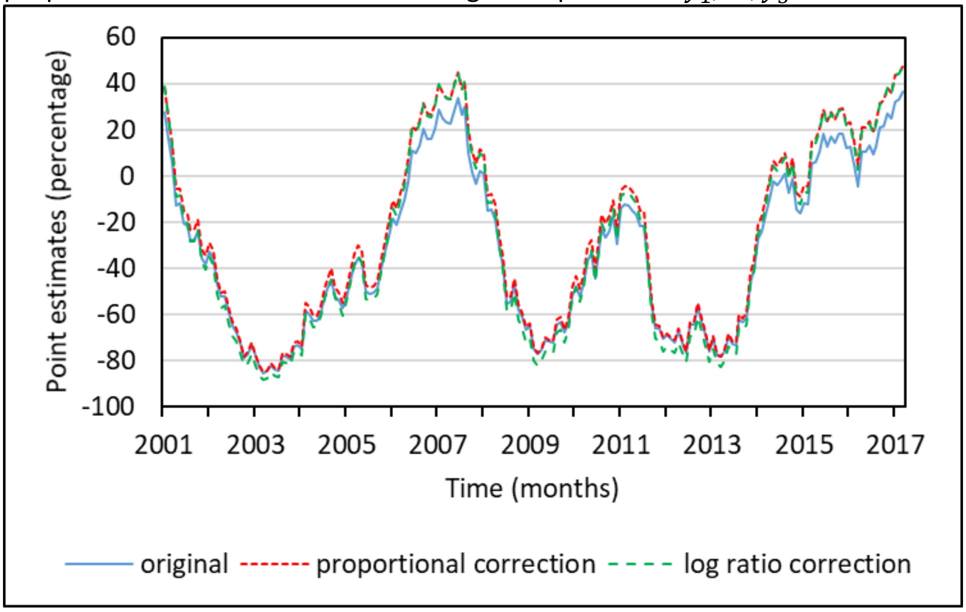

Figure 4.3: comparison of backcasting methods for indicator Econ. L12. Estimates of discontinuities are based on STM with an informative initialization

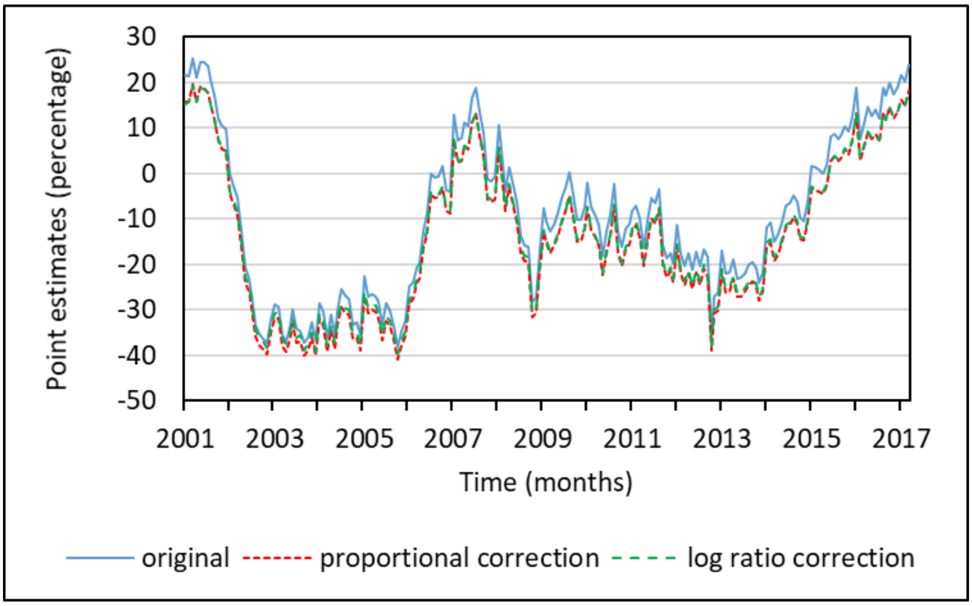

Figure 4.4: comparison of backcasting methods for indicator of Major pur. Estimates of discontinuities are based on STM with informative initialization

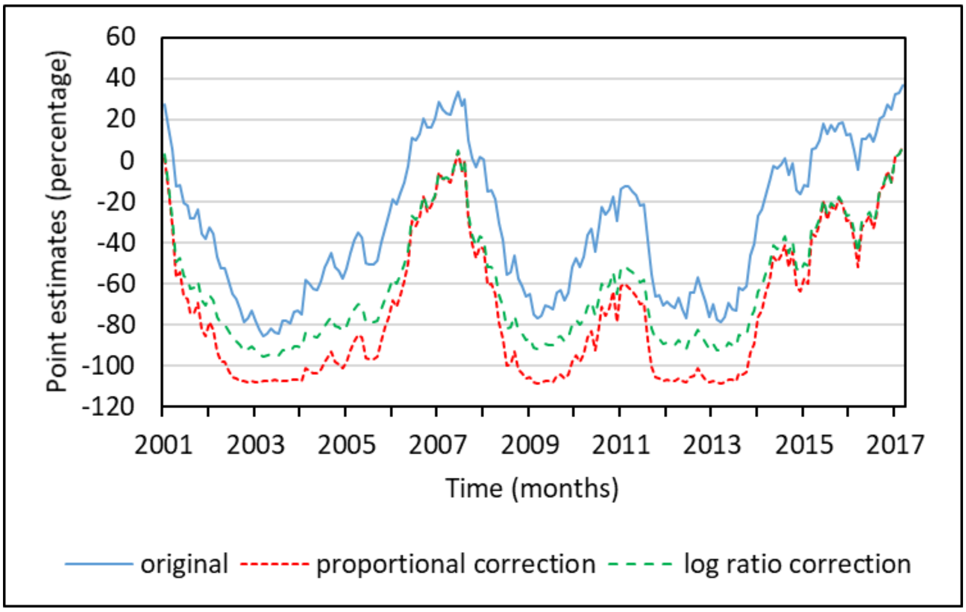

Figure 4.5: comparison of backcasting methods for indicator of Econ. L12 and arbitrarily chosen values for the discontinuities, namely: $\widehat{\Delta}(\mathbf{p}_1^+) = -10$ and $\widehat{\Delta}(\mathbf{p}_1^-) = 20$.

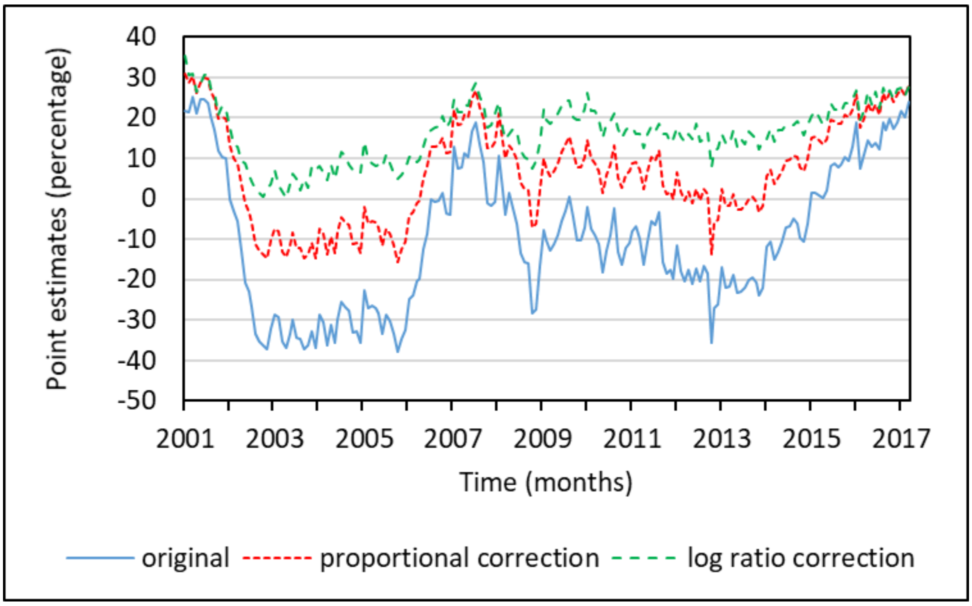

Figure 4.6: comparison of backcasting methods for indicator of Major pur. and arbitrarily chosen values for the discontinuities, namely: $\widehat{\Delta}(\mathbf{p}_5^+) = -5$ and $\widehat{\Delta}(\mathbf{p}_5^-) = -10$.

5. Official publications based on structural time series modelling

Since 2017 the time series model (3.4c) is implemented in the production process of the Dutch CS. The baseline series $\hat{y}_{1,t}, \ldots, \hat{y}_{5,t}$ are the input series for the model. This model was chosen despite model (3.4a) is preferred from a methodological point of view, as model (3.4a) is more complex and the differences between the model estimates are ignorable. These series observed for the years until December 2016 are corrected to the level of the new design following the approach described in Subsection 4.4. The filtered trend estimates are published and replace the former seasonally corrected figures. These series of filtered estimates, corrected for the discontinuity, starts in 1986.
In Figure 5.1 the unadjusted and adjusted direct estimates for consumer confidence are compared with the filtered and smoothed trend estimates obtained with time series model (3.4.a). The direct estimates observed under the old design (dashed line) are plotted until December 2016. The adjusted series of consumer confidence (solid red line) is continued with the direct estimates observed under the new design after January 2017. The adjusted series for the period until December 2016 is obtained as the average over the five backcasted baseline series $\hat{y}_{1,t}, \ldots, \hat{y}_{5,t}$. As a result an uninterrupted series of direct estimates for the monthly consumer confidence is obtained. The filtered trend (solid blue line) replaces the seasonally adjusted figures in the official publications of the Dutch CS and can be interpreted as a trend-cycle component obtained by removing seasonal fluctuations, the population irregular term and the sampling error from the series of direct estimates. For the period before January 2017, the trend is also corrected to the level observed under the new design, since time series model (3.4.a) uses the five backcasted baseline indices as input series. As a result uninterrupted trend series are obtained for the period starting in 1986. These backcasted series are published in order to provide users with uninterrupted series. The smoothed trend in the bottom panel show the more stable and optimal estimates for consumer confidence, since the trend estimates for each period is based on all available observations.

Results in Figure 5.1 are presented until December 2019. The corona crisis that started in the first quarter of 2020 resulted in a dramatic drop of the consumer confidence. To avoid temporal model miss-specification it was necessary to increase the flexibility of the trend, by making the variance of the slope disturbance terms in (3.6) time varying, following the same approach as described in Van den Brakel et al. (2022). Details of this approach and alternative methods considered to accommodate in the model for the sudden changes in the dynamics of the underlying series, are described in a separate paper (Van den Brakel et al, 2023).

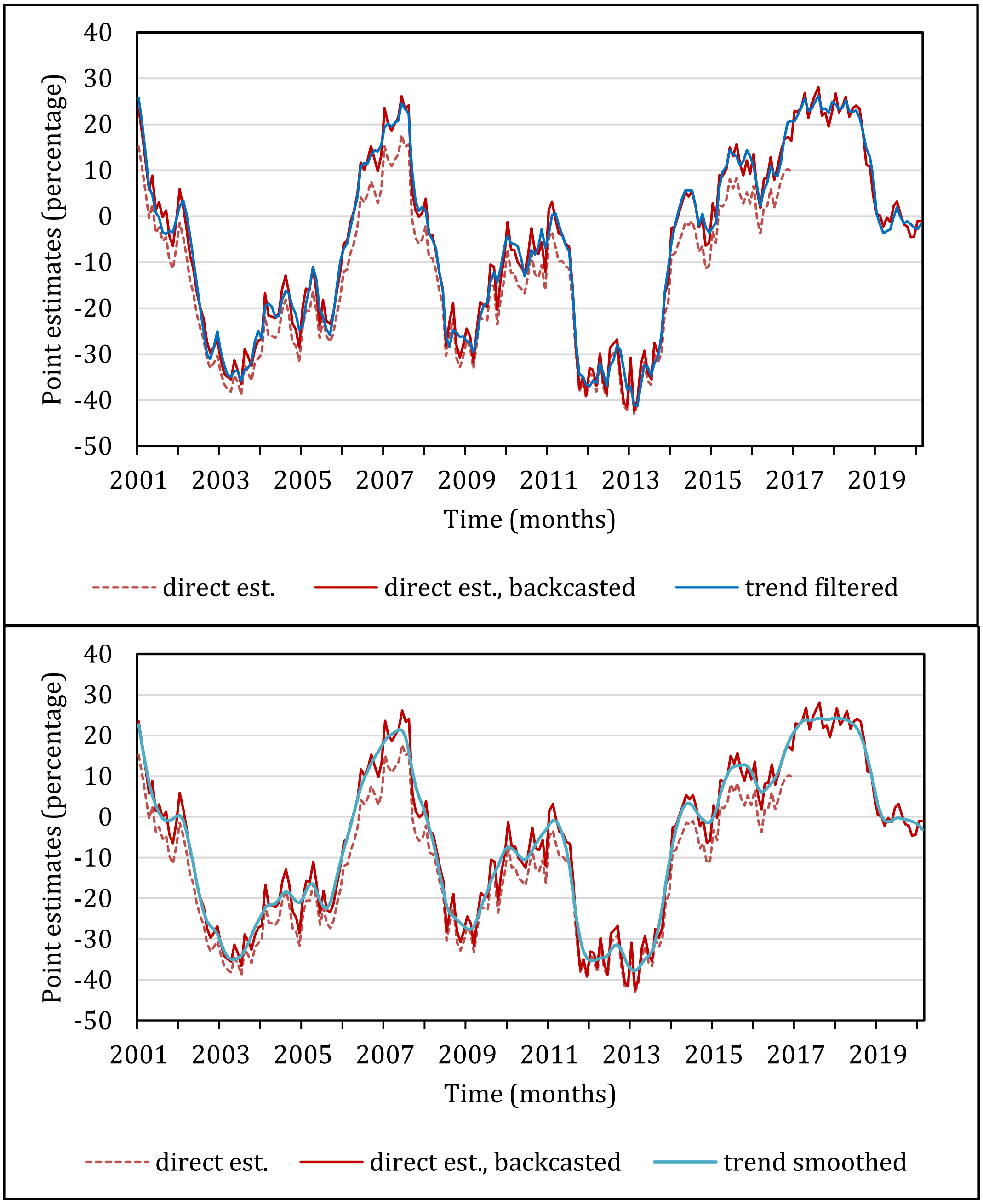

Figure 5.1: comparison direct estimates (original and backcasted) with filtered trend estimates, as published, consumer confidence (upper panel) and smooth trend estimates (bottom panel).

6. Discussion

In this paper a model-based inference method for the Dutch Consumer Survey (CS) is developed. The method is based on a Seemingly Unrelated Structural Time Series Equation (SUTSE) model where the monthly direct estimates of five baseline indices for consumer confidence are used as input. These indices are on their turn obtained from the difference between the percentage of respondents with a positive and a negative opinion on five topics on the economic and financial situation. From the five baseline indices, three composite indices are derived. One of them is the consumer confidence, which is the average over all five baseline indices. The SUTSE model is developed to produce more accurate monthly consumer confidence indicators and to estimate and correct for discontinuities induced by a change-over to a new survey design in January 2017.

The time series in the best fitting SUTSE model are contemporaneously correlated through the disturbance terms of the trend and the population irregular terms. Besides the population irregular term, the model also accounts for the sampling error. Model-based estimates including standard errors for the three composite indicators are obtained by calculating the average over the trend and signal (trend plus seasonal) of the relevant baseline indicators. Filtered trends and signals are rather volatile in this application. More precise and stable estimates are obtained with the Kalman smoother. The latter are however not published by Statistics Netherlands, since this requires a revision of earlier published figures.

The survey redesign in 2017 resulted in discontinuities in the series of the Dutch CS. To separate real month-to-month changes from sudden differences in measurement and selection bias due to the redesign, discontinuities are estimated by collecting data under both the old and the new design in parallel during the first quarter of 2017. The precision of these estimates can be improved with a time series model where the discontinuity is modelled with a level shift. An informative initialization is applied for the regression coefficient of this level shift in the Kalman filter using the direct estimates and their standard errors obtained in the parallel run. In this way the information observed with the time series before and after the change-over is used to further improve the direct estimates for the discontinuities.

Discontinuities appear in the percentage distribution across the positive, neutral, and negative response categories obtained with the questions about the economic situation, the financial situation and major purchases. The way that the discontinuities in the percentages result in discontinuities in the five baseline indices depend on the percentage distribution across these three categories. Since these distributions vary considerably over time, a time varying correction for the discontinuities is proposed. This is achieved by modelling the discontinuities in the time series of the percentages of the positive, neutral, and negative response categories. To maintain uninterrupted series, the time series of these percentages observed under the old design are backcasted to the level of the new design. An adjustment method is proposed that accounts for the fact that the values of the adjusted proportions are in the range [0,100]. Under this method the size of the adjustment in each period is made proportional to the variance of the estimated proportion. Alternatively, a logratio or logit transformation is considered. The drawback of the logratio transformation is that the adjustments can be extremely large in periods where the value of the original ratio is smaller than 1. Another alternative would be the probit link function, which is left for further research.

Backcasted series of the five baseline indices are derived from the backcasted series of the percentages of the positive, neutral, and negative response categories. The adjustments in the five baseline series are time varying, since the distribution of respondents over the positive, neutral, and negative response categories changes over time and since the adjustments in the series of these percentages is time varying. These adjusted series are extended with the direct estimates observed after April 2017 and are used as input series for the multivariate time series model to estimate monthly trends for the CS. This method has been implemented in April 2017 for the publication of uninterrupted monthly CS figures and trends that start in 1986. The trend estimates replace seasonal adjusted figures and can be interpreted as a trend-cycle derived from the direct estimates from which seasonal component, population irregular component and the sampling error are removed.

References

Bell, W.R. and S.C. Hillmer (1990). The time series approach to estimation of periodic surveys. Survey Methodology, 16, pp. 195-215.

Binder, D.A. and J.P. Dick (1989). Modeling and estimation for repeated surveys. Survey Methodology, 15, pp. 29-45.

Binder, D.A. and J.P. Dick (1990). A method for the analysis of seasonal ARIMA models. Survey Methodology, 16, pp. 239-253.

Blight, B.J.N. and A.J. Scott (1973). A Stochastic Model for Repeated Surveys. Journal of the Royal Statistical Society, B series, 35, pp. 61-66.

Bollineni-Balabay, O., J.A. van den Brakel and F. Palm (2017). State space time series modelling of the Dutch Labour Force Survey: Model selection and mean squared error estimation. Survey Methodology. 43, pp.41-67.

Boonstra, H.J. and J.A. van den Brakel (2019). Estimation of level and change for unemployment using structural time series models. Survey Methodology. 45, pp. 395-425.

Doornik, J.A. (2007). An Object-oriented Matrix Programming Language Ox 5. London: Timberlake Consultants Press.

Durbin, J. and S.J. Koopman (2012). Time series analysis by state space methods. Oxford: Oxford University Press.

Elliot, D.J. and P. Zong (2019). Improving timeliness and accuracy from the UK labour force survey. Statistical Theory and Related Fields, 3, pp. 186-198.

Feder, M. (2001). Time series analysis of repeated surveys: the state-space approach. Statistica Neerlandica, 55, pp. 182-199.

Harvey, A.C. (1989). Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge: Cambridge University Press.

Harvey, A. C. and C.H. Chung (2000). Estimating the underlying change in unemployment in the UK. Journal of the Royal Statistical Society, Series A, 163, pp. 303-339.

Koopman, S.J., N. Shephard and J.A. Doornik, (1999). Statistical Algorithms for Models in State Space using SsfPack 2.2. Econometrics Journal, 2, pp. 113-166.

Koopman, S.J., N. Shephard and J. A. Doornik (2008). Statistical Algorithms for Models in State Space using SsfPack 3.0. London: Timberlake Consultants Press.

Krieg, S. and J.A. van den Brakel (2012). Estimation of the monthly unemployment rate for six domains through structural time series modelling with cointegrated trends, Computational statistics and Data Analysis. Vol. 56, pp. 2918-2933.

Pfeffermann, D. (1991). Estimation and seasonal adjustment of population means using data from repeated surveys. Journal of Business & Economic Statistics, 9, pp. 163-175.

Pfeffermann, D. and S.R. Bleuer (1993). Robust joint modelling of labour force series of small areas. Survey Methodology, 19, pp. 149-163.

Pfeffermann, D. and L. Burck (1990). Robust small area estimation combining time series and cross-sectional data. Survey Methodology, 16, pp. 217-237.

Pfeffermann, D. and A. Sikov (2011). Imputation and Estimation under Nonignorable Nonresponse in Household Surveys with Missing Covariate Information. J. Off. Stat. 27(2), pp. 181-209.

Pfeffermann, D. and R. Tiller (2005). Bootstrap approximation to prediction MSE for state-space models with estimated parameters. Journal of Time Series Analysis, 26, pp. 893-916

Pfeffermann, D. and R. Tiller (2006). Small area estimation with state space models subject to benchmark constraints. Journal of the American Statistical Association, 101, pp. 1387-1397.

Rao, J.N.K. and I. Molina (2015). Small Area Estimation, 2nd edition. Wiley.

Rao, J.N.K. and M. Yu (1994). Small area estimation by combining time series and cross-sectional data, Canadian Journal of Statistics, 22, pp. 511-528.

Särndal, C., B. Swensson and J. Wretman (1992). Model Assisted Survey Sampling. New York: Springer Verlag.

Tam, S.M. (1987). Analysis of repeated surveys using a dynamic linear model. International Statistical Review, 55, pp. 63-73.

Tiller, R.B. (1992). Time series modelling of sample survey data from the U.S. current population survey. Journal of Official Statistics, 8, pp. 149-166.

Van den Brakel, J.A. and S. Krieg (2015). Dealing with small sample sizes, rotation group bias and discontinuities in a Rotating Panel Design. Survey Methodology, 41, pp. 267-296.

Van den Brakel, J.A. and S. Krieg (2016). Small area estimation with state-space common factor models for rotating panels. Journal of the Royal Statistical Society A series. 179, pp. 763-791

Van den Brakel, J.A., S. Krieg and M. Smeets (2023). The methodology of the Dutch consumer confidence survey during the corona crisis. Discussion paper, Statistics Netherlands, Heerlen. https://www.cbs.nl/nl-nl/achtergrond/2023/52/consumenten-vertrouwen-onderzoek-tijdens-covid-19

Van den Brakel, J.A., E. Söhler, P. Daas and B. Buelens (2017). Social media as data source for official statistics; the Dutch Consumer Confidence Index. Survey Methodology, 43, pp. 183-210.

Van den Brakel, J.A., M. Souren and S. Krieg (2022). Estimating monthly Labour Force Figures during the COVID-19 pandemic in the Netherlands. Journal of the Royal Statistical Society, Series A. 185, pp. 1560-1583.

Van den Brakel, J.A., X. Zhang and S.M. Tam (2020). Measuring discontinuities in time series obtained with repeated sample surveys. International Statistical Review, 88, pp.155-175.

Appendix: State space representations

The state space representations of the SUTSE models in Section 3 are defined by a measurement equation and a transition equation. The measurement equation defines how the observed series $\hat{\mathbf{y}}_t = (\hat{y}_{1,t}, \hat{y}_{2,t}, \hat{y}_{3,t}, \hat{y}_{4,t}, \hat{y}_{5,t})'$ depends on a p dimensional vector $\boldsymbol{\alpha}_t$ that contains the unobserved state variables:

$$\hat{\boldsymbol{y}}_t = \mathbf{Z}_t \boldsymbol{\alpha}_t, \tag{A.1}$$

with $\mathbf{Z}_t$ a design matrix. There is no measurement error in (A.1) since in this application the measurement errors are included in the state vector. The transition equation describes how the state variables evolve from period $t-1$ to t and is defined as:

$$\begin{aligned} &\boldsymbol{\alpha}_t = \boldsymbol{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t, \\ &\mathrm{E}(\boldsymbol{\eta}_t) = \mathbf{0}_p, \\ &\mathrm{Cov}(\boldsymbol{\eta}_t, \boldsymbol{\eta}_{t'}) = \begin{cases} \mathbf{Q} & \text{if} \quad t = t' \\ \mathbf{O}_{p\times p} & \text{if} \quad t \neq t', \end{cases} \end{aligned} \tag{A.2}$$

with $\mathbf{0}_p$ a p dimensional vector with each element equal to zero and $\mathbf{O}_{p\times p}$ a $p \times p$ dimensional matrix with each element equal to zero. Similarly $\mathbf{I}_p$ denotes the $p \times p$ identity matrix and $\mathbf{1}_p$ a p dimensional vector with each element equal to one. For Model (3.4.a) it follows that the state space representation is obtained with:

$$\boldsymbol{\alpha}_t = (\boldsymbol{\alpha}_t^{L1}, \boldsymbol{\alpha}_t^{L2}, \boldsymbol{\alpha}_t^{L3}, \boldsymbol{\alpha}_t^{L4}, \boldsymbol{\alpha}_t^{L5}, \boldsymbol{\alpha}_t^{S1}, \boldsymbol{\alpha}_t^{S2}, \boldsymbol{\alpha}_t^{S3}, \boldsymbol{\alpha}_t^{S4}, \boldsymbol{\alpha}_t^{S5}, \boldsymbol{\alpha}_t^{I}, \boldsymbol{\alpha}_t^{\varepsilon})', \tag{A.3}$$

$\boldsymbol{\alpha}_t^{Li} = \left(L_{i,t}, R_{i,t}\right), i = 1, \ldots, 5,$

$\boldsymbol{\alpha}_t^{Si} = (S_{i,t,1}, S_{i,t,1}^*, S_{i,t,2}, S_{i,t,2}^*, S_{i,t,3}, S_{i,t,3}^*, S_{i,t,4}, S_{i,t,4}^*, S_{i,t,5}, S_{i,t,5}^*, S_{i,t,6})\ , i = 1, \ldots, 5,$

$\boldsymbol{\alpha}_t^{I} = (I_{1,t}, I_{2,t}, I_{3,t}, I_{4,t}, I_{5,t}),$

$\boldsymbol{\alpha}_t^{\varepsilon} = (\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t}, \varepsilon_{4,t}, \varepsilon_{5,t}),$

$$\mathbf{Z}_t = (\boldsymbol{Z}_t^L\ \boldsymbol{Z}_t^S\ \boldsymbol{Z}_t^I\ \boldsymbol{Z}_t^{\varepsilon}) \text{ with} \tag{A.4}$$

$\boldsymbol{Z}_t^L = \mathbf{I}_5 \otimes (1,0),$

$\boldsymbol{Z}_t^S = \mathbf{I}_5 \otimes (1,0,1,0,1,0,1,0,1,0,1),$

$\boldsymbol{Z}_t^I = \mathbf{I}_5,$

$\boldsymbol{Z}_t^{\varepsilon} = \mathrm{diag}(k_{1,t}, k_{2,t}, k_{3,t}, k_{4,t}, k_{5,t}),$

$$\mathbf{T} = \mathrm{blockdiag}(\mathbf{T}^L, \mathbf{T}^S, \mathbf{T}^I, \mathbf{T}^{\varepsilon}) \text{ with} \tag{A.5}$$

$\mathbf{T}^L = \mathbf{I}_5 \otimes \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$

$\mathbf{T}^S = \mathbf{I}_5 \otimes \mathrm{blockdiag}(\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \mathbf{C}_4, \mathbf{C}_5, 1)$ with

$\mathbf{C}_i = \begin{pmatrix} \cos(h_l) & \sin(h_l) \\ -\sin(h_l) & \cos(h_l) \end{pmatrix}, h_l = \frac{\pi l}{6}, l = 1, \ldots, 5,$

$\mathbf{T}^I = \mathbf{T}^{\varepsilon} = \mathbf{O}_{5\times 5},$

$$\boldsymbol{\eta}_t = (\boldsymbol{\eta}_t^{L1}, \boldsymbol{\eta}_t^{L2}, \boldsymbol{\eta}_t^{L3}, \boldsymbol{\eta}_t^{L4}, \boldsymbol{\eta}_t^{L5}, \boldsymbol{\eta}_t^{S1}, \boldsymbol{\eta}_t^{S2}, \boldsymbol{\eta}_t^{S3}, \boldsymbol{\eta}_t^{S4}, \boldsymbol{\eta}_t^{S5}, \boldsymbol{\eta}_t^{I}, \boldsymbol{\eta}_t^{\varepsilon}) \text{ with} \tag{A.6}$$

$\boldsymbol{\eta}_t^{Li} = \left(0, \eta_{R,i,t}\right), i = 1, \ldots, 5,$

$\boldsymbol{\eta}_t^{Si} = (\eta_{S,i,t,1}, \eta_{S,i,t,1}^*, \ldots, \eta_{S,i,t,5}, \eta_{S,i,t,5}^*, \eta_{S,i,t,6})\ , i = 1, \ldots, 5,$

$\boldsymbol{\eta}_t^{I} = \left(I_{1,t}, I_{2,t}, I_{3,t}, I_{4,t}, I_{5,t}\right),$

$\boldsymbol{\eta}_t^{\varepsilon} = \left(\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t}, \varepsilon_{4,t}, \varepsilon_{5,t}\right),$

$$\mathbf{Q} = \mathrm{blockdiag}(\mathbf{Q}^L, \mathbf{Q}^S, \mathbf{Q}^I, \mathbf{Q}^{\varepsilon}) \text{ with} \tag{A.7}$$

$\mathbf{Q}^L = 10$ x 10 matrix with diagonal elements $\left(0, \sigma_{R,1}^2, 0, \sigma_{R,2}^2, 0, \sigma_{R,3}^2, 0, \sigma_{R,4}^2, 0, \sigma_{R,5}^2\right)$ and off-diagonal elements $\varsigma_{R,2i,2i'}^2$ for $i, i' = 1, \ldots, 5$, and all other off-diagonal elements are equal to zero

$\mathbf{Q}^S = \mathrm{diag}(\mathbf{1}_{11} \otimes \sigma_{S,1}^2, \mathbf{1}_{11} \otimes \sigma_{S,2}^2, \mathbf{1}_{11} \otimes \sigma_{S,3}^2, \mathbf{1}_{11} \otimes \sigma_{S,4}^2, \mathbf{1}_{11} \otimes \sigma_{S,5}^2),$

$\mathbf{Q}^I = 5$ x 5 matrix with diagonal elements $\left(\sigma_{I,1}^2, \sigma_{I,2}^2, \sigma_{I,3}^2, \sigma_{I,4}^2, \sigma_{I,5}^2\right)$ and off-diagonal elements $\varsigma_{I,i,i'}^2$ for $i, i' = 1, \ldots, 5$,

$\mathbf{Q}^{\varepsilon} = \mathrm{diag}(\sigma_{\varepsilon,1}^2, \sigma_{\varepsilon,2}^2, \sigma_{\varepsilon,3}^2, \sigma_{\varepsilon,4}^2, \sigma_{\varepsilon,5}^2).$

The state space representation for Model (3.4.b) is obtained by:

- Replacing $\boldsymbol{\alpha}_t^I$ and $\boldsymbol{\alpha}_t^{\varepsilon}$ in (A.3) by $\boldsymbol{\alpha}_t^{e} = (e_{1,t}, e_{2,t}, e_{3,t}, e_{4,t}, e_{5,t})$,
- Skipping $\boldsymbol{Z}_t^I$ in $\boldsymbol{Z}_t$ in (A.4) and skipping $\mathbf{T}^I$ in $\mathbf{T}$ in (A.5),
- Replacing $\boldsymbol{\eta}_t^I$ and $\boldsymbol{\eta}_t^{\varepsilon}$ in (A.6) by $\boldsymbol{\eta}_t^{e} = \left(e_{1,t}, e_{2,t}, e_{3,t}, e_{4,t}, e_{5,t}\right)$,
- Replacing $\mathbf{Q}^I$ and $\mathbf{Q}^{\varepsilon}$ in $\mathbf{Q}$ in (A.7) by $\mathbf{Q}^e$, which is 5 x 5 matrix with diagonal elements $\left(\sigma_{e,1}^2, \sigma_{e,2}^2, \sigma_{e,3}^2, \sigma_{e,4}^2, \sigma_{e,5}^2\right)$ and off diagonal elements $\varsigma_{e,i,i'}^2$ for $i, i' = 1, \ldots, 5$.

The state space representation for Model (3.4.c) is obtained by:

- Replacing $\boldsymbol{\alpha}_t^I$ and $\boldsymbol{\alpha}_t^{\varepsilon}$ in (A.3) by $\boldsymbol{\alpha}_t^{e} = (e_{1,t}, e_{2,t}, e_{3,t}, e_{4,t}, e_{5,t})$,
- Skipping $\boldsymbol{Z}_t^I$ in $\boldsymbol{Z}_t$ in (A.4) and skipping $\mathbf{T}^I$ in $\mathbf{T}$ in (A.5),
- Replacing $\boldsymbol{Z}_t^{\varepsilon} = \mathrm{diag}(k_{1,t}, k_{2,t}, k_{3,t}, k_{4,t}, k_{5,t})$ in (A.4) by $\boldsymbol{Z}_t^{e} = \boldsymbol{I}_5$,
- Replacing $\boldsymbol{\eta}_t^I$ and $\boldsymbol{\eta}_t^{\varepsilon}$ in (A.6) by $\boldsymbol{\eta}_t^{e} = \left(e_{1,t}, e_{2,t}, e_{3,t}, e_{4,t}, e_{5,t}\right)$,
- Replacing $\mathbf{Q}^I$ and $\mathbf{Q}^{\varepsilon}$ in $\mathbf{Q}$ in (A.7) by $\mathbf{Q}^e$, which is 5 x 5 matrix with diagonal elements $\left(\sigma_{e,1}^2, \sigma_{e,2}^2, \sigma_{e,3}^2, \sigma_{e,4}^2, \sigma_{e,5}^2\right)$ and off diagonal elements $\varsigma_{e,i,i'}^2$ for $i, i' = 1, \ldots, 5$.

The state space representation of Model (4.4.a) in Section 4 is obtained with measurement equation (A.1) where the observed series $\hat{\boldsymbol{y}}_t$ is replaced by $\hat{\mathbf{p}}_t = (\hat{p}_t^+,\ \hat{p}_t^0,\ \hat{p}_t^-)'$ and transition equation (A.2) where:

$$\boldsymbol{\alpha}_t = (\boldsymbol{\alpha}_t^{L+}, \boldsymbol{\alpha}_t^{L0}, \boldsymbol{\alpha}_t^{L-}, \boldsymbol{\alpha}_t^{S+}, \boldsymbol{\alpha}_t^{S0}, \boldsymbol{\alpha}_t^{S-}, \boldsymbol{\alpha}_t^{\beta}, \boldsymbol{\alpha}_t^{I}, \boldsymbol{\alpha}_t^{\varepsilon})', \tag{A.8}$$
$$\boldsymbol{\alpha}_t^{Li} = (L_{i,t}, R_{i,t}), i = +,0,-,$$
$$\boldsymbol{\alpha}_t^{Si} = (S_{i,t,1}, S_{i,t,1}^*, S_{i,t,2}, S_{i,t,2}^*, S_{i,t,3}, S_{i,t,3}^*, S_{i,t,4}, S_{i,t,4}^*, S_{i,t,5}, S_{i,t,5}^*, S_{i,t,6})\ , i = +,0,-,$$
$$\boldsymbol{\alpha}_t^{\beta} = (\beta_t^+, \beta_t^0, \beta_t^-),\ \boldsymbol{\alpha}_t^{I} = (I_t^+, I_t^0, I_t^-),\quad \boldsymbol{\alpha}_t^{\varepsilon} = (\varepsilon_t^+, \varepsilon_t^0, \varepsilon_t^-),$$

$$\mathbf{Z}_t = (\boldsymbol{Z}_t^L\ \boldsymbol{Z}_t^S\ \boldsymbol{Z}_t^{\beta}\ \boldsymbol{Z}_t^I\ \boldsymbol{Z}_t^{\varepsilon}) \text{ with} \tag{A.9}$$
$$\boldsymbol{Z}_t^L = \mathbf{I}_3 \otimes (1,0),$$
$$\boldsymbol{Z}_t^S = \mathbf{I}_3 \otimes (1,0,1,0,1,0,1,0,1,0,1),$$
$$\boldsymbol{Z}_t^{\beta} = x_t \mathbf{I}_3,\quad \boldsymbol{Z}_t^I = \mathbf{I}_3,\ \boldsymbol{Z}_t^{\varepsilon} = \text{diag}(k_{1,t}, k_{2,t}, k_{3,t}),$$

$$\mathbf{T} = \text{blockdiag}(\mathbf{T}^L, \mathbf{T}^S, \mathbf{T}^{\beta}, \mathbf{T}^I, \mathbf{T}^{\varepsilon}) \text{ with} \tag{A.10}$$
$$\mathbf{T}^L = \mathbf{I}_3 \otimes \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$
$$\mathbf{T}^S = \mathbf{I}_3 \otimes \text{blockdiag}(\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \mathbf{C}_4, \mathbf{C}_5, 1) \text{ with } \mathbf{C}_i \text{ defined in (A.5)}$$
$$\boldsymbol{T}^{\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}, \mathbf{T}^I = \mathbf{T}^{\varepsilon} = \mathbf{O}_{3\times 3},$$

$$\boldsymbol{\eta}_t = (\boldsymbol{\eta}_t^{L+}, \boldsymbol{\eta}_t^{L0}, \boldsymbol{\eta}_t^{L-}, \boldsymbol{\eta}_t^{S+}, \boldsymbol{\eta}_t^{S0}, \boldsymbol{\eta}_t^{S-}, \boldsymbol{\eta}_t^{\beta}, \boldsymbol{\eta}_t^{I}, \boldsymbol{\eta}_t^{\varepsilon}) \text{ with} \tag{A.11}$$
$$\boldsymbol{\eta}_t^{Li} = (0, \eta_{R,i,t}), i = +,0,-,$$
$$\boldsymbol{\eta}_t^{Si} = (\eta_{S,i,t,1}, \eta_{S,i,t,1}^*, \dots, \eta_{S,i,t,5}, \eta_{S,i,t,5}^*, \eta_{S,i,t,6}), i = +,0,-,$$
$$\boldsymbol{\eta}_t^{\beta} = \mathbf{0}_3,\qquad \boldsymbol{\eta}_t^{I} = (I_t^+, I_t^0, I_t^-),\quad \boldsymbol{\eta}_t^{\varepsilon} = (\varepsilon_t^+, \varepsilon_t^0, \varepsilon_t^-),$$

$$\mathbf{Q} = \text{blockdiag}(\mathbf{Q}^L, \mathbf{Q}^S, \mathbf{Q}^{\beta}, \mathbf{Q}^I, \mathbf{Q}^{\varepsilon}) \text{ with} \tag{A.12}$$
$$\mathbf{Q}^L = \text{diag}(0, \sigma_{R,+}^2, 0, \sigma_{R,0}^2, 0, \sigma_{R,-}^2),$$
$$\mathbf{Q}^S = \text{diag}(\mathbf{1}_{11} \otimes \sigma_{S,+}^2, \mathbf{1}_{11} \otimes \sigma_{S,0}^2, \mathbf{1}_{11} \otimes \sigma_{S,-}^2),$$
$$\mathbf{Q}^{\beta} = \mathbf{O}_{3\times 3},$$
$$\mathbf{Q}^I = \text{diag}(\sigma_{I,+}^2, \sigma_{I,0}^2, \sigma_{I,-}^2),$$
$$\mathbf{Q}^{\varepsilon} = \text{diag}(\sigma_{\varepsilon,+}^2, \sigma_{\varepsilon,0}^2, \sigma_{\varepsilon,-}^2).$$

The state space representation for Model (4.4.b) is obtained by:
- Replacing $\boldsymbol{\alpha}_t^I$ and $\boldsymbol{\alpha}_t^{\varepsilon}$ in (A.8) by $\boldsymbol{\alpha}_t^e = (e_t^+, e, e_t^-)$,
- Skipping $\boldsymbol{Z}_t^I$ in $\boldsymbol{Z}_t$ in (A.9) and skipping $\mathbf{T}^I$ in $\mathbf{T}$ in (A.10),
- Replacing $\boldsymbol{\eta}_t^I$ and $\boldsymbol{\eta}_t^{\varepsilon}$ in (A.11) by $\boldsymbol{\eta}_t^e = (e_t^+, e_t^0, e_t^-)$,
- Replacing $\mathbf{Q}^I$ and $\mathbf{Q}^{\varepsilon}$ in $\mathbf{Q}$ in (A.12) by $\mathbf{Q}^e = \text{diag}(\sigma_{e,+}^2, \sigma_{e,0}^2, \sigma_{e,-}^2)$.

The state space representation for Model (4.4.c) is obtained by:
- Replacing $\boldsymbol{\alpha}_t^I$ and $\boldsymbol{\alpha}_t^{\varepsilon}$ in (A.8) by $\boldsymbol{\alpha}_t^e = (e_t^+, e, e_t^-)$,
- Skipping $\boldsymbol{Z}_t^I$ in $\boldsymbol{Z}_t$ in (A.9) and skipping $\mathbf{T}^I$ in $\mathbf{T}$ in (A.10),
- Replacing $\boldsymbol{Z}_t^{\varepsilon} = \text{diag}(k_{1,t}, k_{2,t}, k_{3,t})$ in (A.9) by $\boldsymbol{Z}_t^e = \boldsymbol{I}_3$,
- Replacing $\boldsymbol{\eta}_t^I$ and $\boldsymbol{\eta}_t^{\varepsilon}$ in (A.11) by $\boldsymbol{\eta}_t^e = (e_t^+, e_t^0, e_t^-)$,
- Replacing $\mathbf{Q}^I$ and $\mathbf{Q}^{\varepsilon}$ in $\mathbf{Q}$ in (A.12) by $\mathbf{Q}^e = \text{diag}(\sigma_{e,+}^2, \sigma_{e,0}^2, \sigma_{e,-}^2)$.

Explanation of symbols

Empty cell	Figure not applicable
.	Figure is unknown, insufficiently reliable or confidential
*	Provisional figure
**	Revised provisional figure
2017–2018	2017 to 2018 inclusive
2017/2018	Average for 2017 to 2018 inclusive
2017/’18	Crop year, financial year, school year, etc., beginning in 2017 and ending in 2018
2013/’14–2017/’18	Crop year, financial year, etc., 2015/’16 to 2017/’18 inclusive

Due to rounding, some totals may not correspond to the sum of the separate figures.

Colophon

Publisher
Centraal Bureau voor de Statistiek
Henri Faasdreef 312, 2492 JP Den Haag
www.cbs.nl

Prepress
Statistics Netherlands, CCN Creation and visualisation

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contactform: www.cbs.nl/information