Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

Estimating monthly indicators for Consumer Confidence using Structural Time Series Models

by Jan van den Brakel, Sabine Krieg and Marc J.E. Smeets

Release date: December 23, 2025

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under “Contact us” > “Standards of service to the public.”

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

© His Majesty the King in Right of Canada, as represented by the Minister of Industry, 2025

Use of this publication is governed by the Statistics Canada Open Licence Agreement.

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, December 2025
Vol. 51, No. 2, pp. 589-627
Statistics Canada, Catalogue No. 12-001-X

Estimating monthly indicators for Consumer Confidence using Structural Time Series Models

Jan van den Brakel, Sabine Krieg and Marc J.E. Smeets[1]

Abstract

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 are 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. An 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 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.

Key Words: Break in series; Consumer confidence survey; Structural time series model.

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 1,000 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 relative short term emotions about the respondent's 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 strategy 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 a 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) is developed as a solution for both problems. Seemingly Unrelated Time Series Equation (SUTSE) models are applied to the CS as a form

1. Jan van den Brakel, Statistics Netherlands, Department of Statistical Methods and Maastricht University, Department of Quantitative Economics. E-mail: ja.vandenbrakel@cbs.nl; Sabine Krieg and Marc J.E. Smeets, Statistics Netherlands, Department of Statistical Methods.

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 where relations between the input series are modelled with contemporaneously correlated disturbances of the state variables (Harvey, 1989, Chapter 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 population irregular 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). 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), Feder (2001), Krieg and van den Brakel (2012), 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, Zhang and Tam (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 adjusted for the discontinuities. The series underlying the CS are percentages and a 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 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 adjusted 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, which 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 the response rate is about 40%, resulting in a net sample of about 1,000 responding households. 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 $i = 1,\ldots,5$. 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 $i = 1,\ldots,5$. 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 $y_1,\ldots,y_8$ 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, Söhler, Daas and Buelens (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 2,150 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 1,000 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 for p_i^+, p_i^0 and p_i^- are based on the generalised regression (GREG) estimator (Särndal, Swensson and Wretman, 1992) to correct, at least partially, for selective non-response. See also the Appendix for a detailed definition of the GREG estimator for the CS. The change-over to a sequential mixed-mode design will reduce the under coverage 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 e.g.

Sverchkov (2008) and Pfeffermann and Sikov (2011) for a review of methods to model response probabilities.

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, cyclic component and 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 $(i=1,\ldots,5)$ 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 $e_{i,t}$:

$$\hat{y}_{i,t} = \theta_{i,t} + e_{i,t},\ (i=1,\ldots,5). \tag{3.1}$$

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

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

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:

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

with $e_{i,t}$ the sampling error of the observed series with an expectation equal to zero and a variance equal to the variance of the input series, which can be derived from the survey samples.

The five series defined can be combined in a vector $\hat{\mathbf{y}}_t = \left(\hat{y}_{1,t}, \hat{y}_{2,t}, \hat{y}_{3,t}, \hat{y}_{4,t}, \hat{y}_{5,t}\right)'$, which can be modelled with a SUTSE model (Harvey, 1989, Chapter 8). In a SUTSE model, each observed series appears in one equation which are contemporaneously correlated through the disturbance terms of their state variables. The SUTSE models for (3.3) is defined as:

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

with $\mathbf{L}_t = (L_{1,t}, L_{2,t}, L_{3,t}, L_{4,t}, L_{5,t})'$, $\mathbf{S}_t = (S_{1,t}, S_{2,t}, S_{3,t}, S_{4,t}, S_{5,t})'$, $\mathbf{I}_t = (I_{1,t}, I_{2,t}, I_{3,t}, I_{4,t}, I_{5,t})'$, and $\mathbf{e}_t = (e_{1,t}, e_{2,t}, e_{3,t}, e_{4,t}, e_{5,t})'$. The trends $L_{i,t}$ $(i = 1,\ldots,5)$ 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

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

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 estimate compared to for example the local level model (that has disturbance terms for the level parameter and doesn't 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 $S_{i,t}$ $(i = 1,\ldots,5)$, which is defined as:

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

with

$$\begin{aligned} S_{i,t,l} &= S_{i,t-1,l} \cos\left(h_l\right) + S^*_{i,t-1,l} \sin\left(h_l\right) + \eta_{S,i,t,l}, \\ S^*_{i,t,l} &= S^*_{i,t-1,l} \cos\left(h_l\right) - S_{i,t-1,l} \sin\left(h_l\right) + \eta^*_{S,i,t,l}, \; h_l = \frac{\pi l}{6}, \; l = 1,\ldots,6. \end{aligned}$$

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

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

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) are normally distributed with

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

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

$$
\begin{aligned}
\mathrm{E}(e_{i,t}) &= 0, \\
\mathrm{Cov}(e_{i,t}, e_{i',t'}) &= \begin{cases} \mathrm{Var}\left(\hat{y}_{i,t}\right) & \text{if } i = i' \text{ and } t = t' \\ \mathrm{Cov}\left(\hat{y}_{i,t}, \hat{y}_{i',t}\right) & \text{if } i \neq i' \text{ and } t = t' \\ 0 & \text{if } t \neq t'. \end{cases}
\end{aligned}
\tag{3.9}
$$

Note that for $i \neq i'$ the population irregular terms $I_{i,t}$ and $I_{i',t}$ in (3.8) are allowed to be correlated, as are the sampling errors $e_{i,t}$ and $e_{i',t}$ in (3.9). A motivation for modelling correlations between population irregular terms is that some events, e.g. good or bad news about the economy, influence the answers to all questions in a similar way.

The variances and covariances for the sampling errors defined in (3.9) are ideally estimated from the sample data. However, the micro data of the samples for this project are only available for the period after the redesign. The variance approximations are derived from the point estimates as follows:

$$
\begin{aligned}
\mathrm{Var}\left(\hat{y}_{i,t}\right) &= \mathrm{Var}\left(\hat{p}^{+}_{i,t} - \hat{p}^{-}_{i,t}\right) = \mathrm{Var}\left(\hat{p}^{+}_{i,t}\right) + \mathrm{Var}\left(\hat{p}^{-}_{i,t}\right) - 2\mathrm{Cov}\left(\hat{p}^{+}_{i,t}, \hat{p}^{-}_{i,t}\right) \\
&= \frac{1}{n_t}\left[\hat{p}^{+}_{i,t}\left(100 - \hat{p}^{+}_{i,t}\right) + \hat{p}^{-}_{i,t}\left(100 - \hat{p}^{-}_{i,t}\right) - 2\hat{p}^{+}_{i,t}\hat{p}^{-}_{i,t}\right].
\end{aligned}
\tag{3.10}
$$

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. See the Appendix

for a derivation of expression (3.10). The effect of ignoring the stratification and the use of the GREG estimator, introduced with the redesign in 2017, in the variance approximations is assessed for the period January 2017 through December 2019 for which the micro data of the samples are available. Ignoring stratification leads to an overestimation of the standard errors by 0.2% on average, with values ranging from -0.1% to 0.6%. Ignoring the use of the GREG estimator leads to an underestimation of the standard errors by -0.2% on average, with values ranging from -3% to 2%. Based on these results it was decided that (3.10) provides a good approximation of the standard errors for the input series. The covariances in (3.9) are approximated as

$$\mathrm{Cov}\left(\hat{y}_{i,t}\ ,\hat{y}_{i',t}\right) = \rho_{ii'}\sqrt{\mathrm{Var}\left(\hat{y}_{i,t}\right)}\sqrt{\mathrm{Var}\left(\hat{y}_{i',t}\right)}, \tag{3.11}$$

where $\rho_{ii'}$ is the average of the monthly correlations over the period 2017 until 2019.

Through the trend and seasonal components in the SUTSE model (3.4)-(3.9), 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)-(3.9) will be compared to investigate the influence of the correlations between trend and the population irregular terms:

- Model 1: the model as described by equations (3.4)-(3.9), i.e., with correlations between the slope disturbances of the trend and between the population irregular terms of the five series.
- Model 2: similar to Model 1, but without correlation between the population irregular terms of the five series.
- Model 3: similar to Model 1, but without correlation between the slope disturbances and without correlation between the population irregular of the five series.

The general way to proceed is to put the structural time series model into state space representation. See the Appendix for the state space representation of Model (3.4)-(3.9). 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 Statistics Netherlands does not revise the figures published by the end of the month, interest in this application is mainly focussed 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, i.e. the variances and covariances of the state disturbance terms and the measurement errors, are estimated with maximum likelihood, using the numerical optimization procedure MaxBFGS (Doornik, 2007). 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, van den Brakel and Palm (2017) for a motivation and Pfeffermann and Tiller (2005) for a bootstrap approach that accounts for this additional uncertainty. 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, Shephard and Doornik, 1999; Koopman, Shephard and Doornik, 2008) in combination with Ox (Doornik, 2007) 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. In small area estimation application, the population irregular term is often included in the parameter predictions. In this application it is decided to remove the population irregular term from the parameter estimates, as the purpose of the CS is to measure the respondent's opinion of the financial and economic climate over the past and coming 12 months. This is better reflected in the parameter estimates without the population irregular term, as this term measures short-term fluctuations which is not in line with the time window of 12 months in the questions posed to the respondents.

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. With the diffuse initialization of the Kalman filter the first 13 observations are required to construct a proper prior for the Kalman filter.

3.2 Results

The three 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 and to avoid masking differences in the estimation results due to plot congestion. The model evaluation is based on the entire series.

Maximum likelihood estimates for the standard deviations of slope disturbance terms, seasonal disturbance terms and population irregular terms for the five baseline variables are presented in Table 3.1 for the three models. The standard deviations of the slope disturbance terms 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.

Table 3.1
Maximum likelihood estimates of the standard deviations of the slope disturbance terms, seasonal disturbance terms and the population irregular terms

		Standard deviation		
Hyperparameter	**Variable**	**Model 1**	**Model 2**	**Model 3**
Slope $(\sigma_{R,i})$ (3.6)	Econ. L12	2.62	3.68	2.61
	Econ. N12	2.89	4.53	2.86
	Fin. L12	0.54	0.53	0.45
	Fin. N12	0.60	0.94	0.55
	Major pur.	0.94	1.01	0.85
Seasonal $(\sigma_{S,i})$ (3.7)	Econ. L12	0.03	0.02	0.00
	Econ. N12	0.00	0.00	0.00
	Fin. L12	0.02	0.02	0.02
	Fin. N12	0.04	0.04	0.04
	Major pur.	0.12	0.11	0.12
Population irregular term $(\sigma_{I,i})$ (3.8)	Econ. L12	3.47	2.80	3.30
	Econ. N12	5.94	4.84	5.59
	Fin. L12	1.90	1.97	1.93
	Fin. N12	2.24	2.10	1.97
	Major pur.	2.15	2.21	2.18

Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months. The variable "Econ. N12" pertains to expectations of changes in the general economic situation of the country over the next 12 months. The variable "Fin. L12" refers to opinions on changes in the household financial situation over the past 12 months. The variable "Fin. N12" relates to expectations of changes in the household financial situation over the next 12 months. The variable "Major pur." addresses whether it is a good time for people to make significant purchases.

Tables 3.2 – 3.4 show the maximum likelihood estimates of the correlations of Models 1 and 2. The correlations between the slope disturbances are given in Table 3.2 for Model 1. 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 1 is given in Table 3.3. The correlations between the slope disturbances are given in Table 3.4 for Model 2. The correlations of the slope disturbances are mostly larger under Model 2 than under Model 1. This is because under Model 1, a part of the co-movements of the series is considered as correlations between the population irregular terms. Under Model 2 this variation is interpreted as trend fluctuations. As already mentioned in Section 3.1, there are arguments that the population irregular terms are correlated. Model 1 is from this point of view preferred over Model 2, and the correlations in Table 3.4 are probably over-estimated. From Table 3.1 it follows that the standard deviations of the slope disturbance terms are consistently higher for Model 2 compared to Model 1 and Model 3, in particular for Econ. L12, Econ. N12 and Fin. N12, which is another indication that the trends under Model 2 might tend to overfit the observed series.

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 filtered trend under Model 3 often differs from Model 1 and Model 2. Differences are, however, small and the smooth trends are very similar. Since the five input series show strong co-movements, Model 1 and 2 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. As an example the filtered and smoothed trends under the three models are shown for Fin. N12 in Figure 3.1.

Table 3.2
Correlations of the slope disturbances, model with correlations between the slope disturbances and correlations between the population irregular terms (Model 1)

	Econ. L12	Econ. N12	Fin. L12	Fin. N12	Major pur.
Econ. L12	1.000				
Econ. N12	0.865	1.000			
Fin. L12	0.599	0.373	1.000		
Fin. N12	0.909	0.931	0.585	1.000	
Major pur.	0.527	0.385	0.768	0.617	1.000

Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months. The variable "Econ. N12" pertains to expectations of changes in the general economic situation of the country over the next 12 months. The variable "Fin. L12" refers to opinions on changes in the household financial situation over the past 12 months. The variable "Fin. N12" relates to expectations of changes in the household financial situation over the next 12 months. The variable "Major pur." addresses whether it is a good time for people to make significant purchases.

Table 3.3
Correlations of the population irregular terms for a model with correlations between the slope disturbances and correlations between the population irregular terms (Model 1)

	Econ. L12	Econ. N12	Fin. L12	Fin. N12	Major pur.
Econ. L12	1.000				
Econ. N12	0.581	1.000			
Fin. L12	0.002	-0.169	1.000		
Fin. N12	0.189	0.373	0.058	1.000	
Major pur.	0.175	0.214	-0.287	-0.096	1.000

Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months. The variable "Econ. N12" pertains to expectations of changes in the general economic situation of the country over the next 12 months. The variable "Fin. L12" refers to opinions on changes in the household financial situation over the past 12 months. The variable "Fin. N12" relates to expectations of changes in the household financial situation over the next 12 months. The variable "Major pur." addresses whether it is a good time for people to make significant purchases.

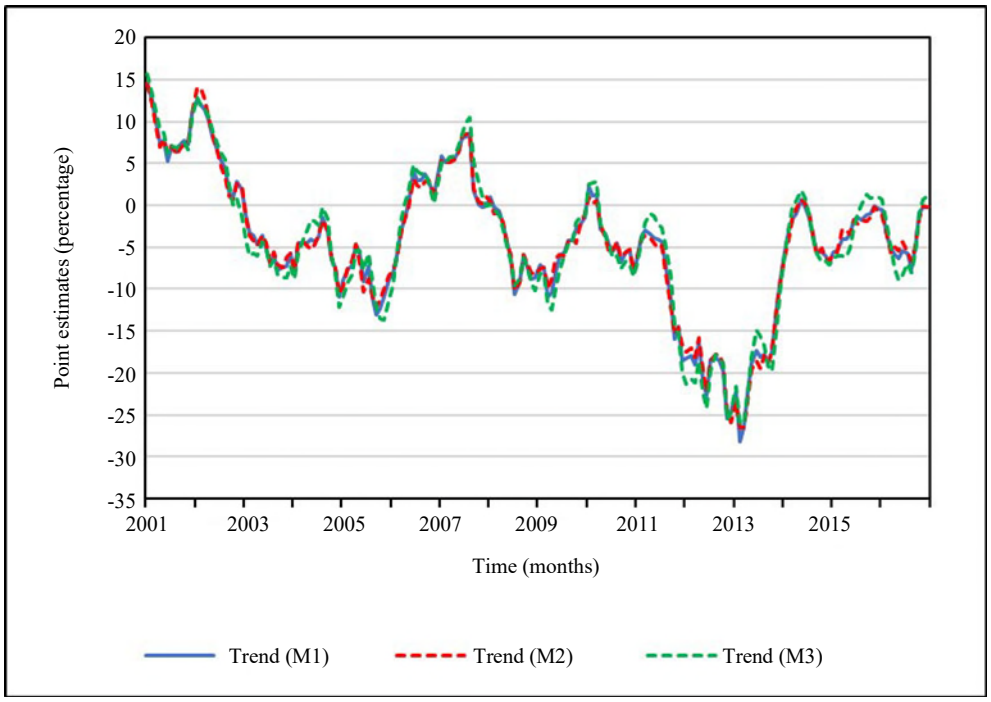
Table 3.4
Correlations of the slope disturbances for a model with correlations between the slope disturbances and without correlations between the population irregular terms (Model 2)

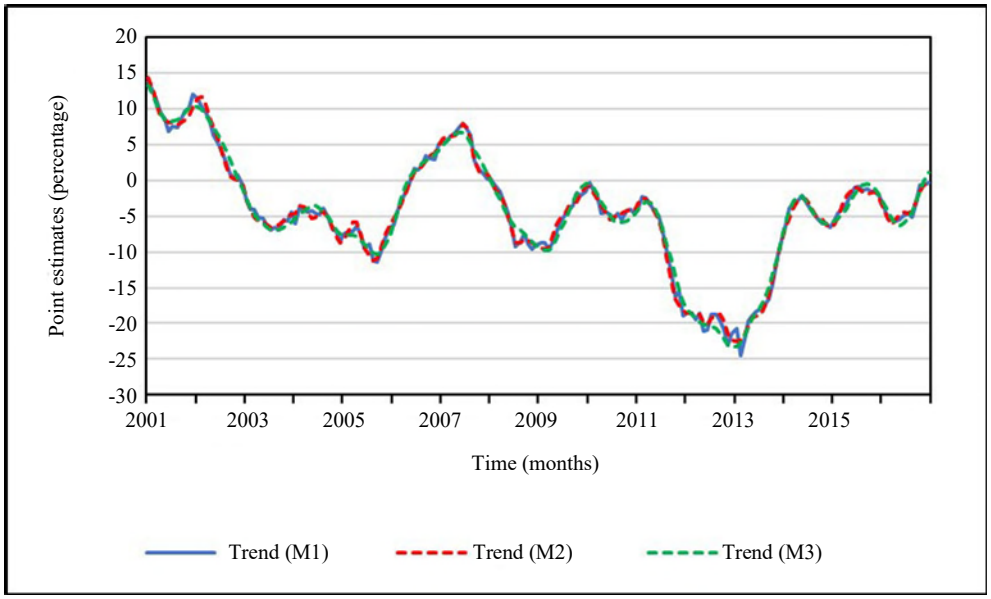
	Econ. L12	Econ. N12	Fin. L12	Fin. N12	Major pur.
Econ. L12	1.000				
Econ. N12	0.946	1.000			
Fin. L12	0.601	0.427	1.000		
Fin. N12	0.962	0.978	0.566	1.000	
Major pur.	0.668	0.591	0.730	0.691	1.000

Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months. The variable "Econ. N12" pertains to expectations of changes in the general economic situation of the country over the next 12 months. The variable "Fin. L12" refers to opinions on changes in the household financial situation over the past 12 months. The variable "Fin. N12" relates to expectations of changes in the household financial situation over the next 12 months. The variable "Major pur." addresses whether it is a good time for people to make significant purchases.

Differences between the filtered and smoothed signals under the three models are generally small. The filtered signals are close to the direct estimates. The smoothed estimates are clearly more stable than the filtered estimates. Although the differences between smoothed and filtered estimates are small, the filtered estimates contain small shocks, which are not visible in the smoothed 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. Continuously revising recently published monthly figures is, however, not in line with the publication strategy of Statistics Netherlands. As an example, Figure 3.2 compares the filtered and smoothed signal under Model 1 with the direct estimates for consumer confidence.

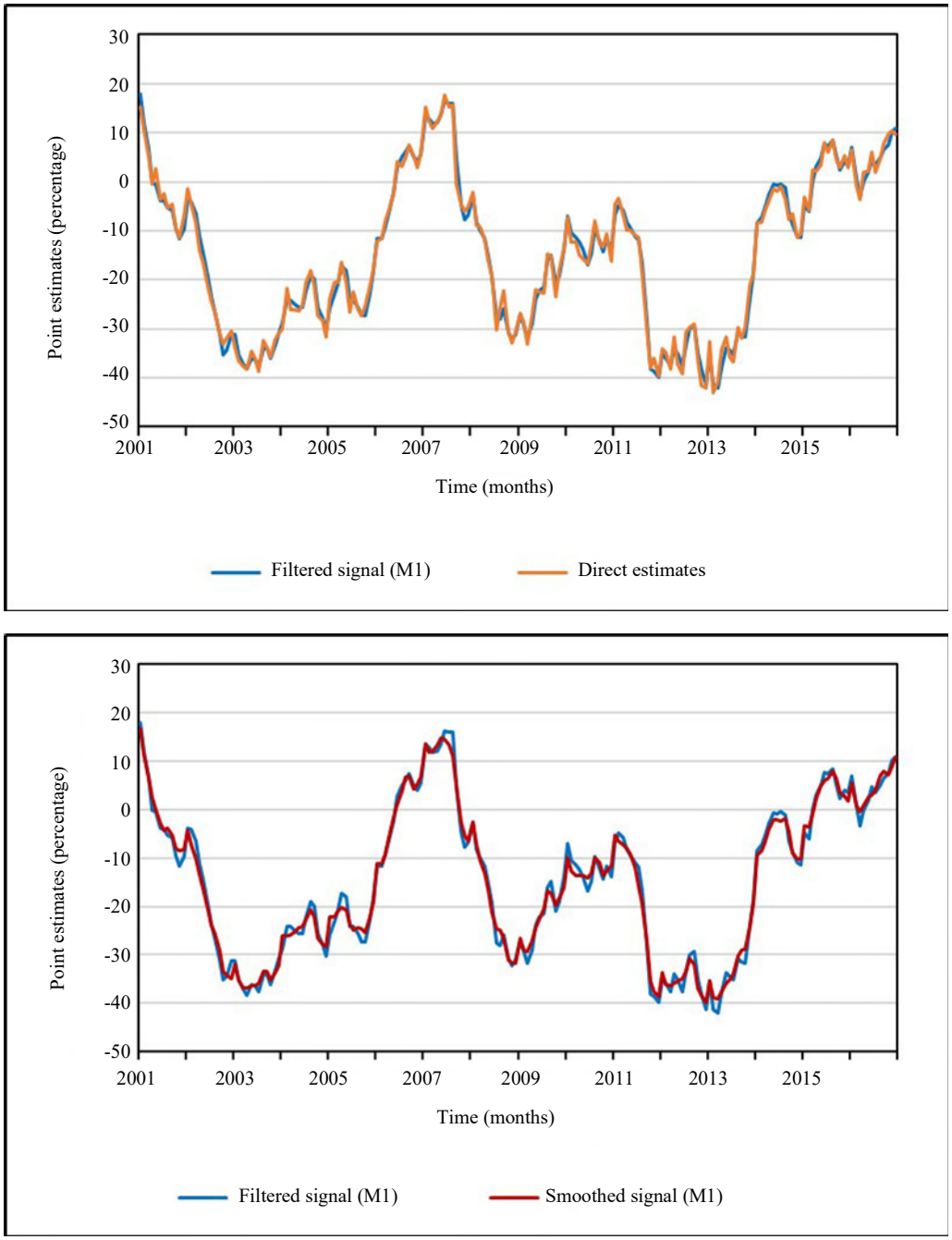
Figure 3.1 Filtered trends (upper panel) and smoothed trends (lower panel) for three models for Fin. N12





Note: The variable "Fin. N12" pertains to expectations of changes in the household financial situation over the next 12 months.

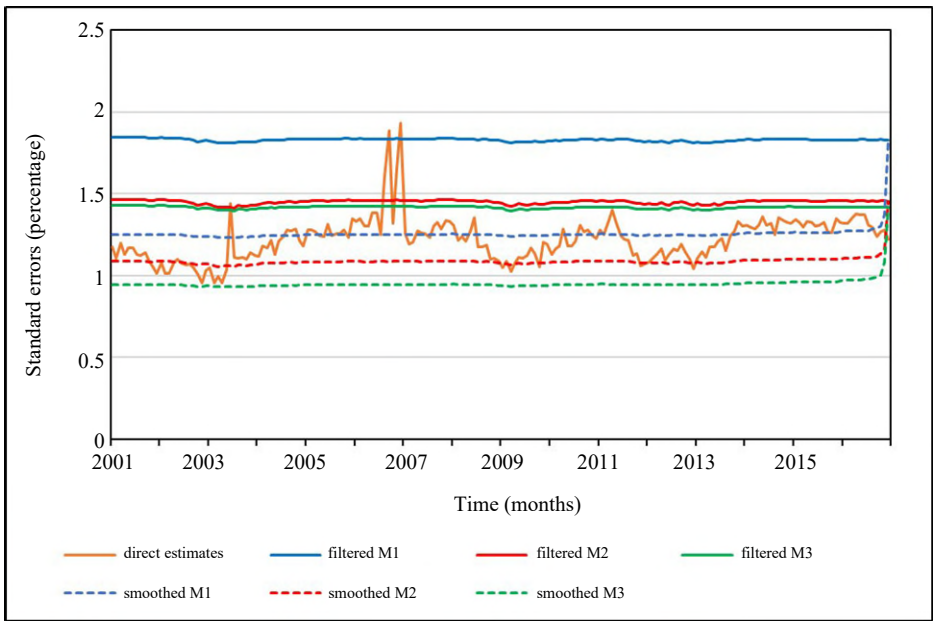
Figure 3.2 Direct estimates and filtered signal (upper panel) and filtered and smoothed signal (lower panel) under Model 1 for consumer confidence



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 1 are clearly larger than under Model 2 and 3. For Model 1, the standard errors of the filtered estimates are clearly larger than those for the direct estimates. The standard errors of the smoothed estimates under Model 1 and the standard errors of the filtered estimates under Model 2 and 3 are more or less the same as those of the direct estimates. For the other two combined series, i.e. economic climate and willingness to buy, the

standard errors of the filtered and smoothed signals under Model 1 are also larger than under Models 2 and 3. 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 2 and 3 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).

Figure 3.3 Standard errors of the direct and model estimates and for the signal under the three models for consumer confidence

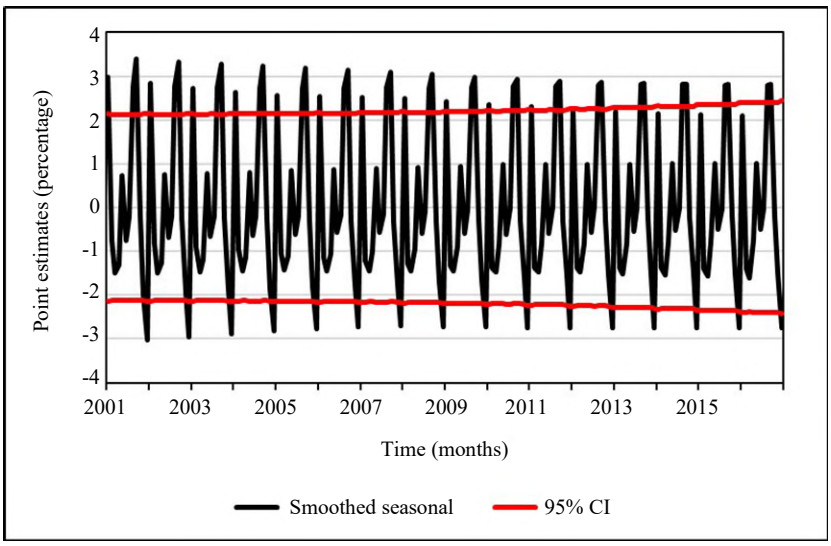


The standard error of the filtered signals are equal or even larger than those of the direct estimates. This not in line with the general finding in the literature is that state space models applied to series obtained with repeated surveys result in model estimates with standard errors that are substantially smaller compared to the standard errors of the direct survey estimates, see e.g. Pfeffermann and Bleuer (1993), Pfeffermann and Burck (1990), Pfeffermann and Tiller (2006), Krieg and van den Brakel (2012), van den Brakel and Krieg (2015, 2016), Boonstra and van den Brakel (2019). Although it is difficult to compare the standard errors of direct sample estimates with those of model-based inference methods, we can interpret our empirical finding for the Dutch CS as follows. The maximum likelihood estimates in Table 3.1 for the population irregular term is at least as large as the variance of 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 design-based estimates for the population parameter θ_t. Their standard errors, on the other hand, only contain the uncertainty of the sampling error. Furthermore, the positive correlations between the population irregular term particularly increase the standard errors of the composite indices. It is nevertheless preferred to account for the correlation between the population irregular terms since this results in a slightly better separation of short term fluctuations captured by the population irregular terms from trend and signal.

The filtered trend estimates are relatively volatile. Together with the relative large population irregular terms, this is an indication that the questions of the CS measure short-term emotions 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. See Figure 3.4 for the seasonal pattern in Econ. L12 as an example. As mentioned before, 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.

Since the standard errors of the filtered trends and signals are stable over time, the standard errors of the five baseline series and the three combined series are shown in Table 3.5 for December 2016 (last observation before the change-over to the new design). Standard errors for the smoothed trends and 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 3.6 for December 2015. The standard errors for the trends are very similar to the standard errors of the signals. For the five baseline series the standard error under Model 1 is mostly smaller than the ones under Model 3. For the combined series, however, the standard errors under Model 1 are mostly larger than the ones under Model 2 and 3. From the comparison between Model 2 and 3, 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. This holds in particular for the indicators economic climate and consumer confidence. For the indicator willingness to buy, the population irregular terms have weak negative correlations, as shown in Table 3.3. As motivated above, it is necessary to account for these correlations to avoid underestimation the uncertainty of the model predictions.

Figure 3.4 Smoothed seasonal effect Econ. L12 with its 95% confidence interval



Notes: - The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months.
- CI = confidence interval.

Table 3.5
Standard error of the filtered estimates of the trend and signal for the last period (December 2016) (December 2015) for 8 series

	Model 1		Model 2		Model 3	
	trend	signal	trend	signal	trend	signal
Econ. L12	3.17	3.00	3.23	2.75	3.01	3.10
Econ. N12	4.77	4.67	4.79	4.00	4.25	4.68
Fin. L12	1.47	1.45	1.74	1.71	1.68	1.76
Fin. N12	1.55	1.59	1.84	1.62	1.50	1.85
Major pur.	2.07	1.84	2.34	2.22	2.30	2.22
Economic climate	3.60	3.61	2.95	2.74	3.05	2.86
Willingness to buy	1.05	1.20	1.23	1.23	1.26	1.22
Consumer confidence	1.76	1.83	1.44	1.46	1.61	1.42

Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months. The variable "Econ. N12" pertains to expectations of changes in the general economic situation of the country over the next 12 months. The variable "Fin. L12" refers to opinions on changes in the household financial situation over the past 12 months. The variable "Fin. N12" relates to expectations of changes in the household financial situation over the next 12 months. The variable "Major pur." addresses whether it is a good time for people to make significant purchases.

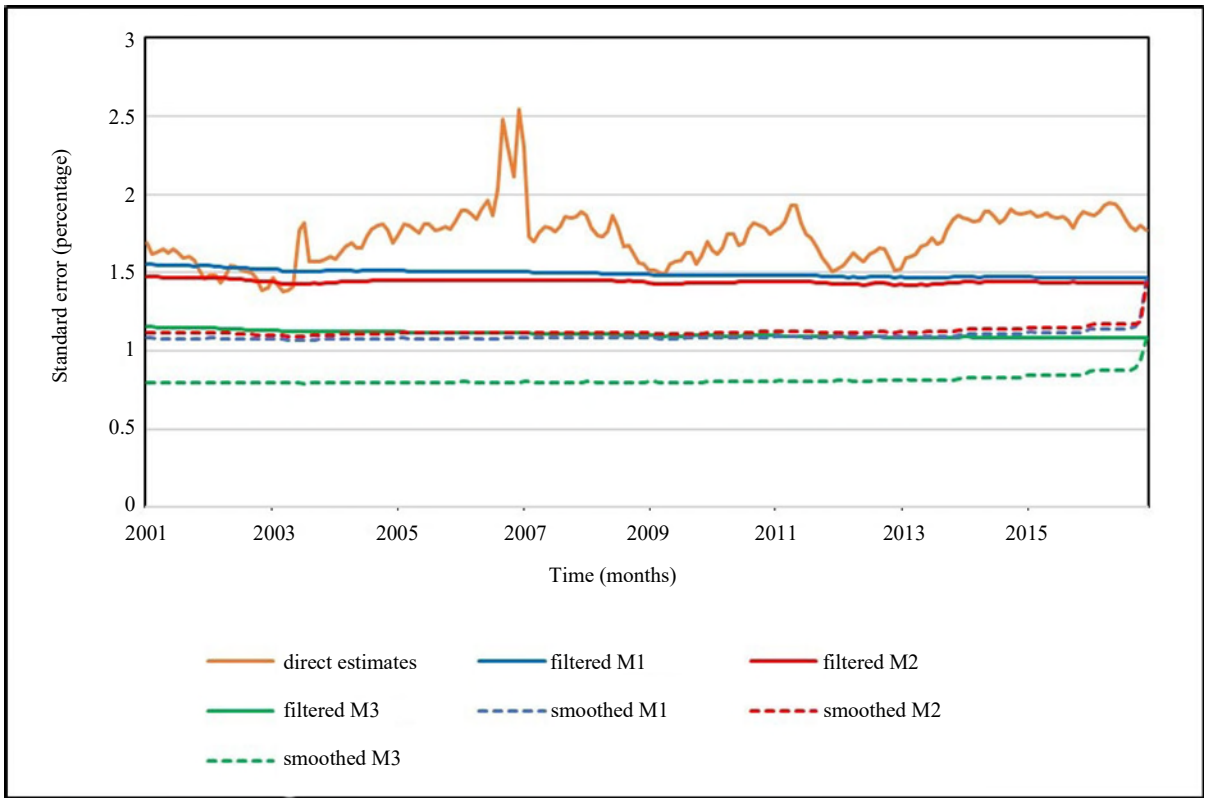
Table 3.6
Standard errors of the smoothed estimates of the trend and signal for the last period (December 2015) for 8 series

	Model 1		Model 2		Model 3	
	trend	signal	trend	signal	Trend	signal
Econ. L12	2.23	2.20	2.22	2.05	2.19	2.15
Econ. N12	3.10	3.16	2.99	2.82	3.06	3.12
Fin. L12	1.04	1.12	0.96	1.14	0.99	1.16
Fin. N12	1.11	1.29	0.90	1.22	1.05	1.34
Major pur.	1.34	1.57	1.30	1.81	1.31	1.81
Economic climate	2.44	2.46	2.33	2.08	1.91	1.93
Willingness to buy	0.76	0.90	0.76	0.92	0.69	0.89
Consumer confidence	1.22	1.27	1.19	1.11	0.90	0.97

Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months. The variable "Econ. N12" pertains to expectations of changes in the general economic situation of the country over the next 12 months. The variable "Fin. L12" refers to opinions on changes in the household financial situation over the past 12 months. The variable "Fin. N12" relates to expectations of changes in the household financial situation over the next 12 months. The variable "Major pur." addresses whether it is a good time for people to make significant purchases.

A major advantage of inference based on time series models is the gain in precision of period-to-period changes. 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 3.5. 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.

Figure 3.5 Standard errors for month-to-month-development of the direct estimates and filtered estimates of the signal under three models for 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 a few very weak autocorrelation for the first 12 lags for some of the series and do not indicate the presence of any residual autocorrelation. 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 three models are nested and are therefore also compared formally using a likelihood ratio test. Table 3.7 contains the log likelihoods and the results of the likelihood ratio tests. The test for Model 2 versus Model 3 indicates that modelling the correlation between the slope disturbance terms significantly improves the model fit. The test for Model 1 versus Model 2 shows that modelling the correlation between the population irregular terms further improves the model fit significantly. Finally the test for Model 1 versus

Model 3 shows that the joint test on the inclusion of a full covariance matrix for the slope disturbance terms and the population irregular terms rejects the null hypothesis that both models are equivalent. Based on the likelihood ratio tests, Model 1 is selected for the inference of the Dutch CS.

Table 3.7
Log likelihoods and results of the likelihood ratio tests

Model	Log likelihood	Comparison	LR statistic	df	p-value
M1	-5,293.09	Model 1 versus Model 2	107.96	10	0.000
M2	-5,335.70	Model 2 versus Model 3	309.20	10	0.000
M3	-5,476.66	Model 1 versus Model 3	417.16	20	0.000

Note: LR = Likelihood ratio; df = Degrees of freedom.

4. Discontinuities

An important aspect of the implementation of the new design in 2017 is to quantify and adjust 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 adjusted series are used in a second step to calculate uninterrupted series for $y_{1,t},\ldots,y_{5,t}$. These adjusted series will be used as the input for Model 1, specified by (3.4)-(3.9) that is used in the production of official monthly figures about consumer confidence. The method of adjusting the percentage series observed before the change-over is developed in this section. The proposed method ensures that the size of the adjustment 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. We contend that such an adjustment 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. 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 1,000 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 entire 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 $\hat{p}_{i,t}^{j,\mathrm{O}}$ $(j=\{+,0,-\})$ 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:

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

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

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

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

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

with

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

and n_t^{O}, n_t^{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 $i=1,\ldots,5$. 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. As in Section 3, the model has a separate component for the population irregular term and the sampling error:

$$\hat{\mathbf{p}}_t = \mathbf{L}_t + \mathbf{S}_t + \boldsymbol{\beta}'\mathbf{x}_t + \mathbf{I}_t + \mathbf{e}_t, \tag{4.4}$$

where $\hat{\mathbf{p}}_t = (\hat{p}_t^+, \hat{p}_t^0, \hat{p}_t^-)'$ 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 $\mathbf{I}_t = (I_t^+, I_t^0, I_t^-)'$ is a vector with population irregular term components and $\mathbf{e}_t = (e_t^+, e_t^0, e_t^-)'$ is a vector with sampling errors. It is understood that the regression coefficients used to model the discontinuities are modelled as time invariant parameters. It is possible to make them time varying by modelling them as random walks. This would result in an alternative way of allowing for time varying corrections for discontinuities. However, there is a good chance that the state space model can no longer identify the trend and the time varying regression coefficient of the intervention variable.

The variables L_t^j and S_t^j with $j = \{+, 0, -\}$ are smooth trend models and trigonometric seasonal models as described in Section 3.1. For the slope disturbance terms in the trend (3.5) and the disturbance terms of the seasonal component (3.7) 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:

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

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 $(\mathbf{I}_t)$ are modelled as normally and independently distributed random variables, i.e. $I_t^j \sim N(0, \sigma_{I,j}^2)$, $j = \{+, 0, -\}$. The sampling errors $\mathbf{e}_t$ are modelled as normally distributed variables $\mathbf{e}_t \sim N(0, \boldsymbol{\Sigma}_t)$ with variance and covariances derived of the input data:

$$\mathrm{Cov}\left(e_t^j, e_{t'}^{j'}\right) = \mathrm{Cov}\left(\hat{p}_t^j, \hat{p}_{t'}^{j'}\right) = \begin{cases} \hat{p}_t^j (100 - \hat{p}_t^j)/n_t & \text{if } j = j' \text{ and } t = t' \\ -\hat{p}_t^j \hat{p}_{t'}^{j'}/n_t & \text{if } j \neq j' \text{ and } t = t'. \end{cases} \tag{4.6}$$

Recall from Section 3 that ignoring the stratification and the effect of the GREG estimator in the variance approximation has a minor effect on the estimated standard errors. Therefore the above pragmatic (co)variance approximations can be used conveniently in this application.

Since Model (4.4)-(4.6) is used to estimate the discontinuities, only restriction (4.5) is applied to the coefficients that model the discontinuities. It is, however, understood that the sum over the three input series is also restricted to 100%. This translates to additional restrictions on the trend and seasonal components. For example one restriction on the sum over the signals $L_t^+ + L_t^0 + L_t^- + S_t^+ + S_t^0 + S_t^- = 100$, or a separate restriction on the trend $L_t^+ + L_t^0 + L_t^- = 100$ in combination with a zero restriction on the seasonals

$L_t^+ + S_t^+ + S_t^0 + S_t^- = 0$. Different methods are proposed in the literature to impose restrictions on time varying state variables. Doran (1992) proposed to extend the measurement equation of the state space model with additional series that impose both restrictions. Harvey and Koopman (1993) and Hungnes, Skjerpen, Hamre, Jansen, Pham and Sandvik (2024) showed how the constraint that the sum over state variables add up to zero can be imposed using a reduced-rank covariance matrix for the state disturbance terms.

Note that since the neutral category is not used in the baseline series $\hat{y}_{1,t},\ldots,\hat{y}_{5,t}$ and also not in the combined indices $\hat{y}_{6,t}$, $\hat{y}_{7,t}$ and $\hat{y}_{8,t}$, 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. Differences between discontinuity estimates based on this approach and Model (4.4)-(4.6) where smaller than 1 percent point and also smaller than the standard error of the discontinuity estimates. The state space representation of Model (4.4)-(4.6) is given in the Appendix.

4.2 Adjustment methods for discontinuities

As explained at the beginning of Section 4, the series of the percentages observed before the redesign of January 2017 are adjusted to the level of the percentage series observed under the new design. These adjusted percentage series are used to compile adjusted input series for the time series model (3.4)-(3.9). The percentages can only have admissible values in the range [0, 100]. Therefore methods are considered that result in adjusted series that have values in this admissible range. In this case the adjusted target variables $\hat{y}_1,\ldots,\hat{y}_5$ will also have values in the admissible range $[-100,+100]$. Let $\tilde{p}_{i,t}^{j,\mathrm{N}}$ denote the adjusted series of $\hat{p}_{i,t}^{j,\mathrm{O}}$. The following adjustment is proposed for percentages:

$$\tilde{p}_{i,t}^{j,\mathrm{N}} = \hat{p}_{i,t}^{j,\mathrm{O}} + \hat{\beta}_i^j \frac{\hat{p}_{i,t}^{j,\mathrm{O}}\left(100-\hat{p}_{i,t}^{j,\mathrm{O}}\right)}{\hat{p}_{i,\tau}^{j,\mathrm{O}}\left(100-\hat{p}_{i,\tau}^{j,\mathrm{O}}\right)} \quad . \quad t=1,\ldots,T-1, \tag{4.7}$$

with T the month of the change-over to the new design, i.e. 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 $\hat{p}_{i,t}^{j,\mathrm{O}}\left(100-\hat{p}_{i,t}^{j,\mathrm{O}}\right)$. In this way, the size of the adjustment for the discontinuity is time varying and depends on the value of $\hat{p}_{i,t}^{j,\mathrm{O}}$. The adjustment 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 adjusted 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 $j=\{+,0,-\}$ are adjusted with (4.7) the sum of the adjusted 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 adjusted as $\tilde{p}_{i,t}^{0,\mathrm{N}} = 100 - \tilde{p}_{i,t}^{+,\mathrm{N}} - \tilde{p}_{i,t}^{-,\mathrm{N}}$. The size of the adjustment 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 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 adjustment 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 $d = \{\mathrm{O}, \mathrm{N}\}$, are used as the reference category in the denominator of the logratio transformation, which is defined as (recall that "O" denotes "old design" and "N" denotes "new design"):

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

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

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

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)-(4.6) using the information from the parallel run through the informative initialization of the Kalman filter as described in Section 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 $t = 1, \ldots, T-1$, is applied to obtain three adjusted 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 adjustment 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 4.4.

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.

Table 4.1
Estimates of the discontinuities obtained with the parallel run, standard errors in brackets

	positive answers ($\hat{\Delta}(p_i^+)$)		negative answers ($\hat{\Delta}(p_i^-)$)		indicator ($\hat{\Delta}(p_i^+) - \hat{\Delta}(p_i^-)$)
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

Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months. The variable "Econ. N12" pertains to expectations of changes in the general economic situation of the country over the next 12 months. The variable "Fin. L12" refers to opinions on changes in the household financial situation over the past 12 months. The variable "Fin. N12" relates to expectations of changes in the household financial situation over the next 12 months. The variable "Major pur." addresses whether it is a good time for people to make significant purchases.

Model (4.4)-(4.6) is applied to estimate the discontinuities. 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. The estimates of the discontinuities are shown in Table 4.2.

Table 4.2
Smoothed estimates of the discontinuities based on the STM under Model (4.4)-(4.6), standard errors in brackets

	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)

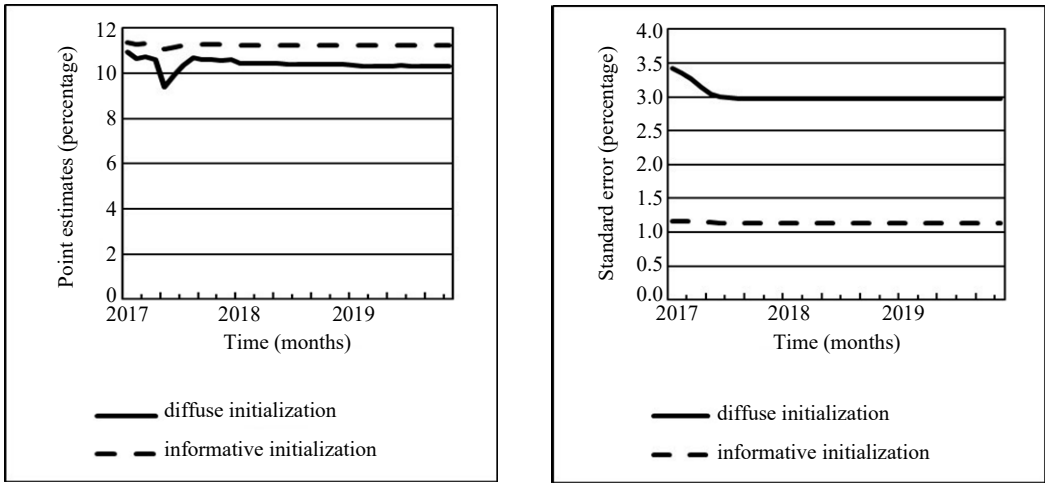
Notes:
- The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months. The variable "Econ. N12" pertains to expectations of changes in the general economic situation of the country over the next 12 months. The variable "Fin. L12" refers to opinions on changes in the household financial situation over the past 12 months. The variable "Fin. N12" relates to expectations of changes in the household financial situation over the next 12 months. The variable "Major pur." addresses whether it is a good time for people to make significant purchases.
- STM = structural time series model.

The estimates obtained with the Kalman smoother 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.

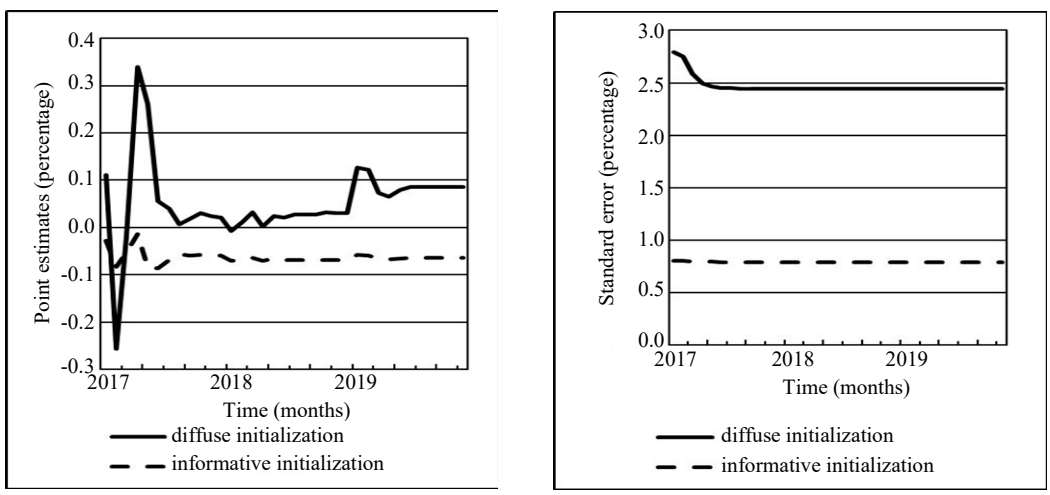
The estimates of the discontinuities obtained with the time series model are instable if only a few observations after the change-over are available. They improve when more data under the new design become available. Figures 4.1 and 4.2 show the real-time estimates for the discontinuities based on the 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, i.e. the extent to which the trend varies over time. As the flexibility of the trend increases, discontinuity estimates become more dependent on local observations close to the period of the change-over to the new survey design. As the trend becomes more time-invariant, the influence of observations further away from the period change-over increases.

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



Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months.

Figure 4.2 Development of the point estimates (left panel) and the standard errors (right panel) of the discontinuities of percentages of the negative answers, Econ. L12



Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months.

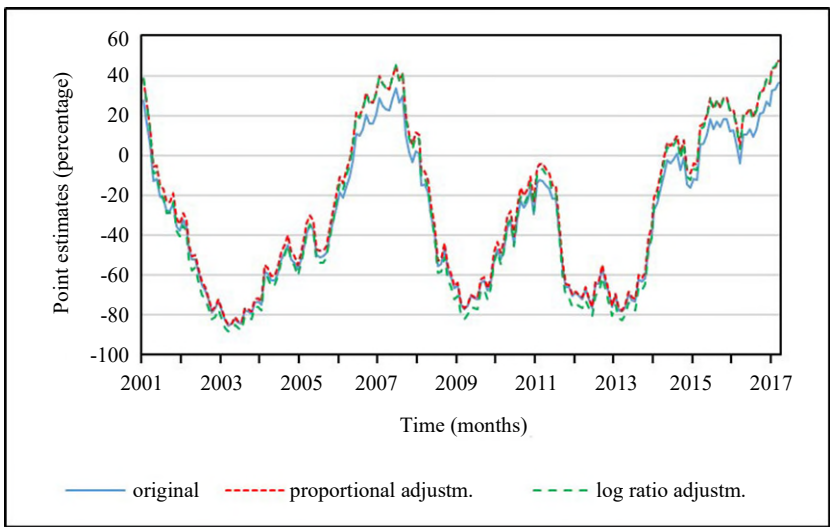
4.4 Results for series adjusted for discontinuities

In this section the effects of the methods to adjust for discontinuities 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)-(4.6) 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 adjustment defined by (4.7) and the additive adjustment in combination with the logratio transformation defined by (4.9) and (4.10). These adjusted

series are used to calculate uninterrupted series for $y_{1,t},\ldots,y_{5,t}$, which are on their turn the input for Model 1 of (3.4)-(3.9).

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

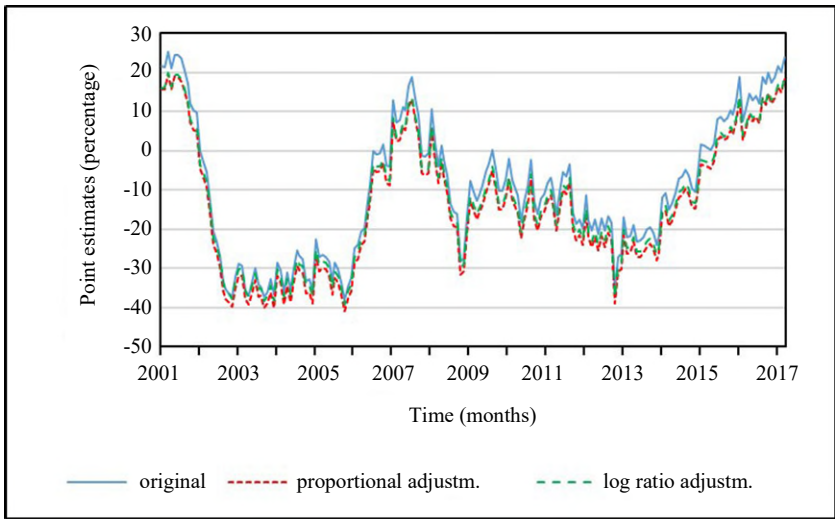
Figure 4.3 Comparison of the adjusting methods for the indicator Econ. L12. Estimates of discontinuities are based on STM with an informative initialization



Notes: - The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months.
- STM = structural time series model.

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

Figure 4.4 Comparison of the adjusting methods for the indicator of Major pur. Estimates of the discontinuities are based on STM with informative initialization

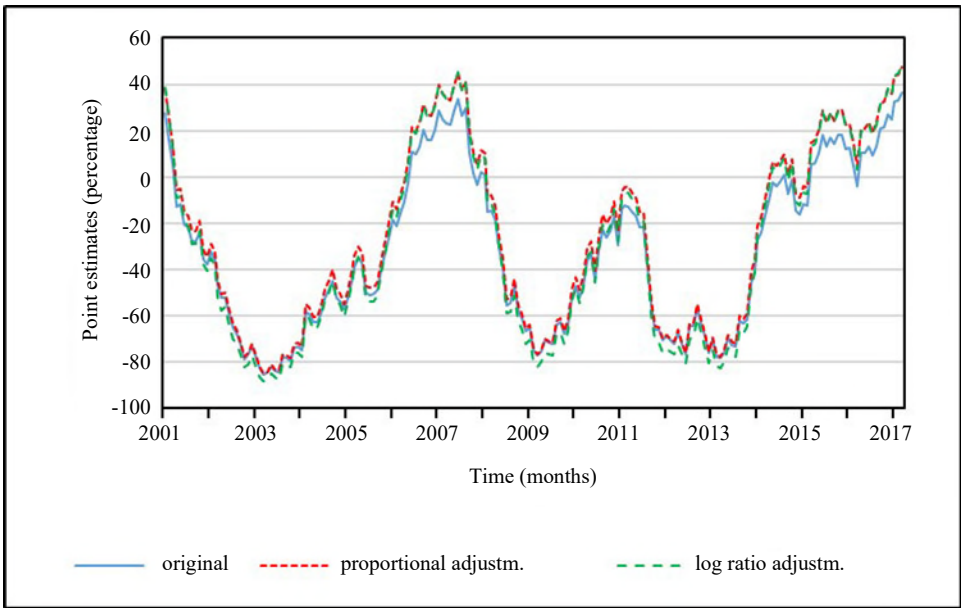


Notes: - The variable "Major pur." relates to the question of whether it is a good time for people to make significant purchases.
- STM = structural time series model.

From these results there is a slight preference for the proportional adjustment, since the logratio transformation sometimes leads to larger adjustments in a different direction as observed during the parallel run. A disadvantage of the proportional method, however, is that the adjusted 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 $\hat{\Delta}(p_1^+) =$ -10 and $\hat{\Delta}(p_1^-) =$ 20, that the proportional adjustments leads to outcomes smaller than -100 in the most negative periods. In this case the logratio transformation to adjust series would be preferred.

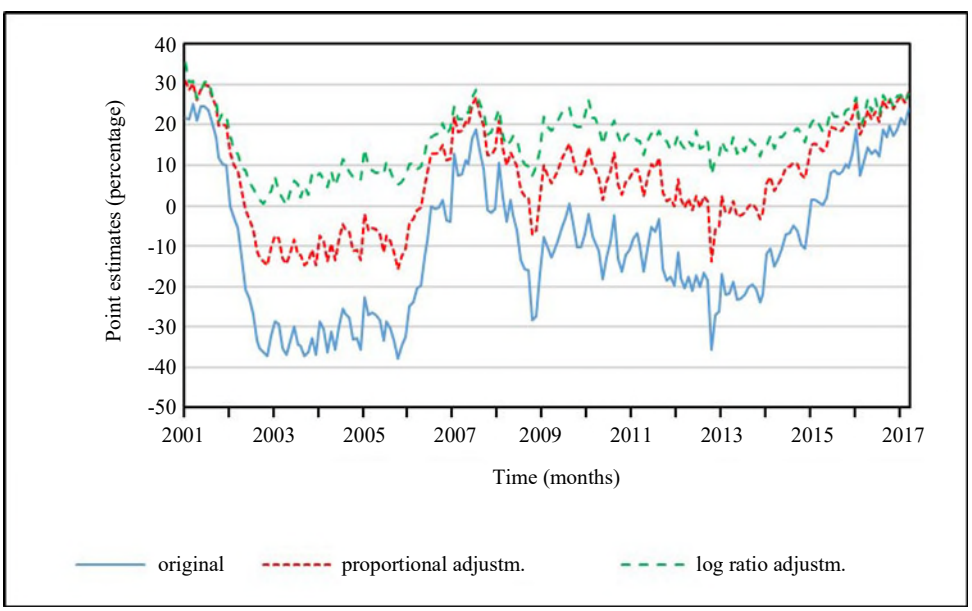
Figure 4.6 shows that also with realistic values of the discontinuities the logratio transformation could result in large adjustments. Here the two methods are applied to the indicator series of Major pur. with arbitrarily, but realistic, chosen values for discontinuities of $\hat{\Delta}(p_5^+) =$ -5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

Figure 4.5 Comparison of the adjusting methods for the indicator of Econ. L12 and arbitrarily chosen values for the discontinuities, namely: $\hat{\Delta}(p_1^+) = -10$ and $\hat{\Delta}(p_1^-) = 20$



Note: The variable "Econ. L12" relates to opinions on changes in the general economic situation of the country over the past 12 months.

Figure 4.6 Comparison of the adjusting methods for the indicator of Major pur. and arbitrarily chosen values for the discontinuities, namely: $\hat{\Delta}(p_5^+) = -5$ and $\hat{\Delta}(p_5^-) = -10$



Note: The variable "Major pur." relates to the question of whether it is a good time for people to make significant purchases.

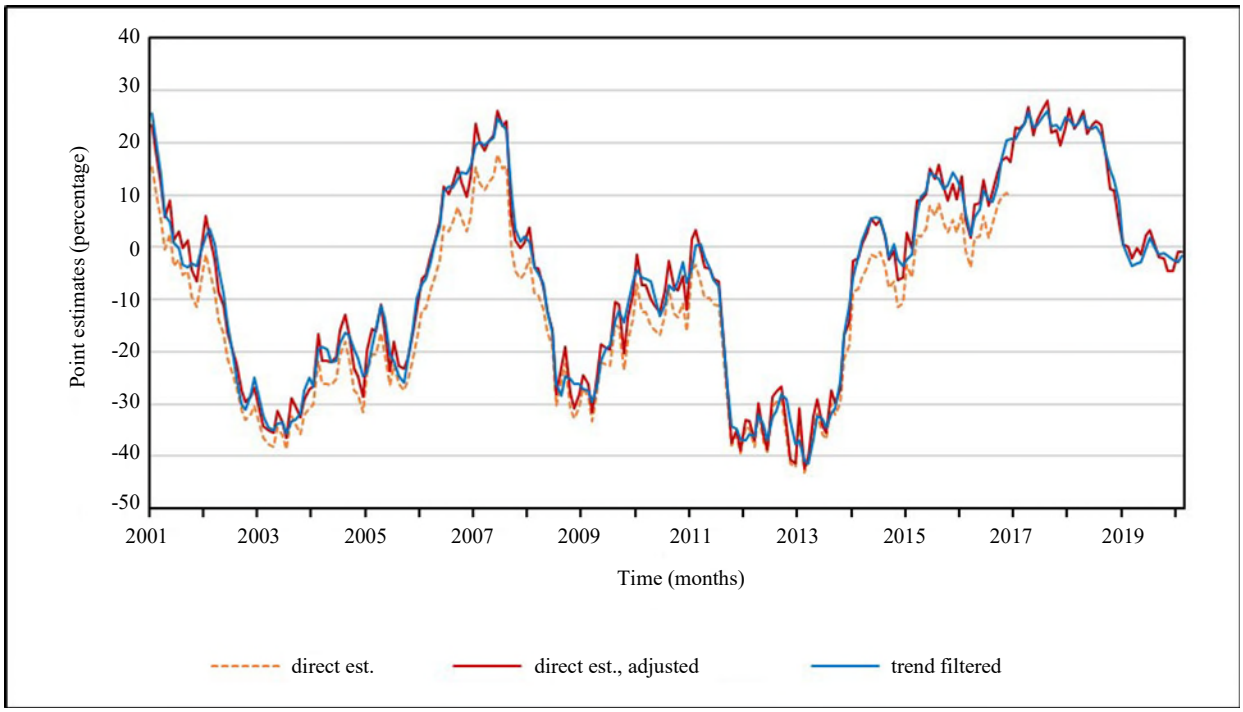
5. Official publications based on structural time series modelling

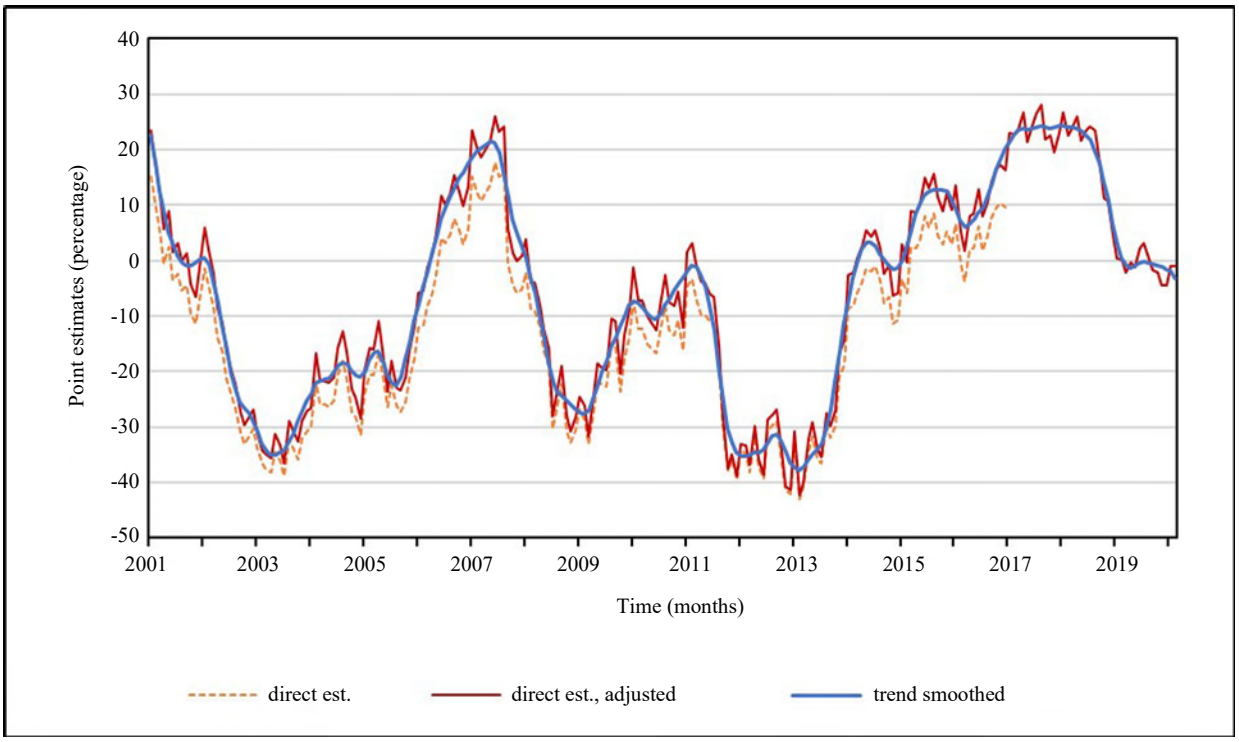
Since 2017 the time series model (3.4)-(3.9) is implemented in the production process of the Dutch CS. The baseline series $\hat{y}_{1,t},\ldots,\hat{y}_{5,t}$ are the input series for the model. The model used for production equals Model 1 assuming zero correlations for sampling errors. (Accounting for correlation between sampling errors is based on comments of the referees, after the methodology was implemented in production.) These series observed for the years until December 2016 are adjusted to the level of the new design following the approach described in Section 4.2. The filtered trend estimates are published and replace the former seasonally adjusted figures. These series of filtered estimates, adjusted for the discontinuity, start in 1986.

In Figure 5.1 the original direct estimates and the direct estimates adjusted for the discontinuity for consumer confidence are compared with the filtered and smoothed trend estimates obtained with time series model (3.4)-(3.9). 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 adjusted 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 adjusted to the level observed under the new design, since time series model (3.4)-(3.9) uses the five adjusted baseline indices as input series. As a result uninterrupted trend series are obtained for the period starting in 1986. 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, Souren and Krieg (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, is described in van den Brakel, Krieg and Smeets (2023).

Figure 5.1 Comparison of the direct estimates (original and adjusted) with filtered trend estimates as published for 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 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 adjust for discontinuities induced by a change-over to a new survey design in January 2017.

The time series in the selected SUTSE model are contemporaneously correlated through the disturbance terms of the slope of the trend, the population irregular terms and the sampling errors. 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. However, the latter are not published by Statistics Netherlands, as this requires a continuous revision of recently 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 in which the discontinuities in the percentages lead to discontinuities in the five basic indices depends on the percentage distribution across these three categories. Since these distributions vary considerably over time, a time varying method to adjust 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 adjusted 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.

Adjusted series of the five baseline indices are derived from the adjusted 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, starting in 1986. With the introduction of the new design, series are revised only once, providing data users with a long, uninterrupted time series. 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.

Acknowledgements

The authors would like to thank two anonymous referees and an Associate Editor for carefully reading and providing useful comments on a former draft of this manuscript.

Appendix

Generalized regression estimator for the Consumer Survey

Parameters for the CS are defined as means. Let y denote the unknown population mean of a parameter of interest. Let y_k denote the observation obtained from respondent k for parameter y, $\mathbf{x}_k = (x_{k,1},\ldots,x_{k,l})'$ an l-dimensional vector with auxiliary variables for respondent k and n the number of respondents observed in the sample. Furthermore, $\mathbf{x}$ denote an l-dimensional vector with the true population means of the auxiliary variables $\mathbf{x}_k$. The GREG estimator is based on the following linear regression model: $y_k = \mathbf{b}'\mathbf{x}_k + \varepsilon_k$, with $\mathbf{b}$ an l-dimensional vector with regression coefficients and ε_k independently distributed residuals with $\mathrm{E}(\varepsilon_k) = 0$ and $\mathrm{Var}(\varepsilon_k) = \sigma^2$. Then the generalized regression (GREG) estimator for y is defined as

$$\hat{y}^r = \hat{y}^d + \hat{\mathbf{b}}'(\mathbf{x} - \hat{\mathbf{x}}^d),$$

with $\hat{y}^d$ and $\hat{\mathbf{x}}^d$ the unweighted sample estimates for the population means of y and the auxiliary variables $\mathbf{x}$, i.e. $\hat{y}^d = \frac{1}{n}\sum_{k=1}^{n} y_k$, and $\hat{\mathbf{x}}^d = \frac{1}{n}\sum_{k=1}^{n} \mathbf{x}_k$. Furthermore, an estimator for $\mathbf{b}$ is given by $\hat{\mathbf{b}} = \left(\sum_{k=1}^{n} \mathbf{x}_k \mathbf{x}_k'\right)^{-1} \sum_{k=1}^{n} \mathbf{x}_k y_k$. The auxiliary variable used in the GREG estimator for the CS is defined by the following weighting scheme:

Income(5) * Householdcomp.(2) + Urbanisation(5) * Householdcomp.(2) + Housholdsize(5) + HouseOwner(2) * Householdcomp.(2) + Age(7) + EducationLevel(4) + HouseholdType(4) + Gender(2) * Householdcomp.(2)

with:

- *Income(5)*: disposable household income of the household to which the person belongs classified into five quantiles, i.e. lowest 20%, 20-40%, 40-60%, 60-80%, highest 20%
- *Householdcomp.(2)*: head of household is single versus head of household is living with a partner
- *Urbanisation(5)*: urbanization level municipality in 5 categories, i.e. very urban, urban, moderately urban, little urban or rural
- *Housholdsize(5)*: household with 1, 2, 3, 4 or 5 or more persons
- *HouseOwner(2)*: person living in rented or owner-occupied accommodation
- *Age(7)*: age category in seven categories, i.e. 15-24 years, 25-34 years, 35-44 years, 45-54 years, 55-64 years, 65-74 years or 75 years and over
- *EducationLevel(4)*: education level in four categories, i.e. low, medium, high or unknown
- *HouseholdType(4)*: household type to which a person belongs in four categories, i.e. couple without children, couple with children, single-parent household, other household including singles
- *Gender(2)*: gender in 2 categories, i.e. men or women

Derivation of (3.10)

The derivation is provided for a fraction. It is understood that this applies also for parameters defined as percentages. For notational convenience, the subscript t is dropped in the following derivations. Since $\hat{y}_i = \hat{p}_i^+ - \hat{p}_i^-$, it follows that

$$\mathrm{Var}\left(\hat{y}_i\right) = \mathrm{Var}\left(\hat{p}_i^+\right) + \mathrm{Var}\left(\hat{p}_i^-\right) - 2\mathrm{Cov}\left(\hat{p}_i^+, \hat{p}_i^-\right). \tag{A.1}$$

An estimator for the variance under simple random sampling for a sample mean is given by

$$\mathrm{Var}\left(\hat{p}_i^+\right) = \frac{1}{n}\frac{1}{n-1}\sum_{k=1}^{n}\left(y_{i,k}^+ - \overline{y}_i^+\right)^2.$$

Since $y_{i,k}^+ = 1$ if respondent k answers question i positive and $y_{i,k}^+ = 0$ otherwise, it follows that ${y_{i,k}^+}^2 = y_{i,k}^+$, $\overline{y}_i^+ = \hat{p}_i^+$ and

$$\mathrm{Var}\left(\hat{p}_i^+\right) = \frac{1}{n}\frac{1}{n-1}\sum_{k=1}^{n}\left(y_{i,k}^+ - 2y_{i,k}^+\hat{p}_i^+ + \hat{p}_i^{+2}\right) = \frac{1}{n}\frac{1}{n-1}n\hat{p}_i^+\left(1-\hat{p}_i^+\right) \approx \frac{1}{n}\hat{p}_i^+\left(1-\hat{p}_i^+\right).$$

In a similar it follows that $\mathrm{Var}\left(\hat{p}_i^-\right) \approx \frac{1}{n}\hat{p}_i^-\left(1-\hat{p}_i^-\right)$. An estimator for the covariance under simple random sampling for two sample means is given by

$$\mathrm{Cov}\left(\hat{p}_i^+, \hat{p}_i^-\right) = \frac{1}{n}\frac{1}{n-1}\sum_{k=1}^{n}\left(y_{i,k}^+ - \overline{y}_i^+\right)\left(y_{i,k}^- - \overline{y}_i^-\right).$$

If respondent k answers question i positive, then $y_{i,k}^+ = 1$ and $y_{i,k}^- = 0$. If respondent k answers question i negative, then $y_{i,k}^+ = 0$ and $y_{i,k}^- = 1$. If respondent k answers question i neutral, then $y_{i,k}^+ = 0$ and $y_{i,k}^- = 0$. Therefore $y_{i,k}^+ y_{i,k}^- = 0$ for all k and

$$\begin{aligned}\mathrm{Cov}\left(\hat{p}_i^+, \hat{p}_i^-\right) &= \frac{1}{n}\frac{1}{n-1}\sum_{k=1}^{n}\left(y_{i,k}^+ y_{i,k}^- - y_{i,k}^+\overline{y}_i^- - y_{i,k}^-\overline{y}_i^+ + \overline{y}_i^+\overline{y}_i^-\right)\\ &= \frac{1}{n}\frac{1}{n-1}\left(-2n\overline{y}_i^+\overline{y}_i^- + n\overline{y}_i^+\overline{y}_i^-\right)\\ &= -\frac{1}{n}\frac{1}{n-1}n\overline{y}_i^+\overline{y}_i^- \approx -\frac{1}{n}\hat{p}_i^+\hat{p}_i^-.\end{aligned}$$

Inserting the results for $\mathrm{Var}\left(\hat{p}_i^-\right)$, $\mathrm{Var}\left(\hat{p}_i^-\right)$ and $\mathrm{Cov}\left(\hat{p}_i^+, \hat{p}_i^-\right)$ gives (3.10).

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:

$$\hat{\mathbf{y}}_t = \mathbf{Z}_t\boldsymbol{\alpha}_t, \tag{A.2}$$

with $\mathbf{Z}_t$ a design matrix. There is no measurement error in (A.2) 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:

$$\boldsymbol{\alpha}_t = \boldsymbol{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t, \tag{A.3}$$

$$\begin{aligned}\mathrm{E}(\boldsymbol{\eta}_t) &= 0_p,\\ \mathrm{Cov}(\boldsymbol{\eta}_t, \boldsymbol{\eta}_{t'}) &= \begin{cases}\mathbf{Q}_t & \text{if } t = t'\\ \mathbf{O}_{p\times p} & \text{if } t \neq t',\end{cases}\end{aligned}$$

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. In what follows, it is understood that $\mathrm{diag}(\mathbf{x})$ is an operator that transforms the p-dimensional vector $\mathbf{x}$ into a $p\times p$ diagonal matrix with the elements of $\mathbf{x}$ on the diagonal. In a similar way, $\mathrm{blockdiag}(\mathbf{X},\ldots,\mathbf{Y})$ is an operator that transforms a list of specified matrices $\mathbf{X},\ldots,\mathbf{Y}$, into one large block-diagonal matrix with the specified matrices as block-diagonal elements.

For Model (3.4)-(3.9) it follows that the state space representation is obtained with:

$$\boldsymbol{\alpha}_t = \left(\boldsymbol{\alpha}_t^{L1}, \boldsymbol{\alpha}_t^{L2}, \boldsymbol{\alpha}_t^{L3}, \boldsymbol{\alpha}_t^{L4}, \boldsymbol{\alpha}_t^{L5}, \boldsymbol{\alpha}_t^{S1}, \boldsymbol{\alpha}_t^{S2}, \boldsymbol{\alpha}_t^{S3}, \boldsymbol{\alpha}_t^{S4}, \boldsymbol{\alpha}_t^{S5}, \boldsymbol{\alpha}_t^{I}, \boldsymbol{\alpha}_t^{e}\right)',$$

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

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

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

$$\boldsymbol{\alpha}_t^{e} = \left(e_{1,t}, e_{2,t}, e_{3,t}, e_{4,t}, e_{5,t}\right),$$

$\mathbf{Z}_t = \left(\mathbf{Z}_t^L, \mathbf{Z}_t^S, \mathbf{Z}_t^I, \mathbf{Z}_t^e\right)$ with

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

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

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

$\mathbf{T} = \text{blockdiag}\left(\mathbf{T}^L, \mathbf{T}^S, \mathbf{T}^I, \mathbf{T}^e\right)$ with

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

$$\mathbf{T}^S = \mathbf{I}_5 \otimes \text{blockdiag}\left(\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \mathbf{C}_4, \mathbf{C}_5, 1\right) \text{ with (A.4)}$$

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

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

$\boldsymbol{\eta}_t = \left(\boldsymbol{\eta}_t^{L1}, \boldsymbol{\eta}_t^{L2}, \boldsymbol{\eta}_t^{L3}, \boldsymbol{\eta}_t^{L4}, \boldsymbol{\eta}_t^{L5}, \boldsymbol{\eta}_t^{S1}, \boldsymbol{\eta}_t^{S2}, \boldsymbol{\eta}_t^{S3}, \boldsymbol{\eta}_t^{S4}, \boldsymbol{\eta}_t^{S5}, \boldsymbol{\eta}_t^{I}, \boldsymbol{\eta}_t^{e}\right)$ with

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

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

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

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

$\mathbf{Q}_t = \text{blockdiag}\left(\mathbf{Q}^L, \mathbf{Q}^S, \mathbf{Q}^I, \mathbf{Q}_t^e\right)$ with

$\mathbf{Q}^L = 10\times 10$ 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 $q_{2i,2i'} = \varsigma_{R,i,i'}^2$ for $i, i' = 1,\ldots,5,$ and all other off-diagonal elements are equal to zero

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

$\mathbf{Q}^I =$ 5×5 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 $i, i' = 1,\ldots,5,$

$\mathbf{Q}_t^e =$ 5×5 matrix with elements defined by (3.9)

The state space representation of Model (4.4)-(4.6) in Section 4 is obtained with measurement equation (A.2) where the observed series $\hat{\mathbf{y}}_t$ is replaced by $\hat{\mathbf{p}}_t = \left(\hat{p}_t^+,\ \hat{p}_t^0,\ \hat{p}_t^-\right)'$ and transition equation (A.3) where:

$$\boldsymbol{\alpha}_t = \left(\boldsymbol{\alpha}_t^{L+}, \boldsymbol{\alpha}_t^{L0}, \boldsymbol{\alpha}_t^{L-}, \boldsymbol{\alpha}_t^{S+}, \boldsymbol{\alpha}_t^{S0}, \boldsymbol{\alpha}_t^{S-}, \boldsymbol{\alpha}_t^{\beta}, \boldsymbol{\alpha}_t^{I}, \boldsymbol{\alpha}_t^{e}\right)',$$

$$\boldsymbol{\alpha}_t^{Li} = \left(L_{i,t}, R_{i,t}\right),\ i = +,0,-,$$

$$\boldsymbol{\alpha}_t^{Si} = \left(S_{i,t,1}, S_{i,t,1}^*, S_{i,t,2}, S_{i,t,2}^*, S_{i,t,3}, S_{i,t,3}^*, S_{i,t,4}, S_{i,t,4}^*, S_{i,t,5}, S_{i,t,5}^*, S_{i,t,6}\right),\ i = +,0,-,$$

$$\boldsymbol{\alpha}_t^{\beta} = \left(\beta_t^+, \beta_t^0, \beta_t^-\right), \quad \boldsymbol{\alpha}_t^{I} = \left(I_t^+, I_t^0, I_t^-\right), \quad \boldsymbol{\alpha}_t^{e} = \left(e_t^+, e, e_t^-\right),$$

$$\mathbf{Z}_t = \left(\mathbf{Z}_t^L\ \mathbf{Z}_t^S\ \mathbf{Z}_t^{\beta}\ \mathbf{Z}_t^I\ \mathbf{Z}_t^e\right) \text{ with}$$

$$\mathbf{Z}_t^L = \mathbf{I}_3 \otimes (1,0),$$

$$\mathbf{Z}_t^S = \mathbf{I}_3 \otimes (1,0,1,0,1,0,1,0,1,0,1),$$

$$\mathbf{Z}_t^{\beta} = x_t \mathbf{I}_3, \quad \mathbf{Z}_t^I = \mathbf{Z}_t^e = \mathbf{I}_3,$$

$$\mathbf{T} = \text{blockdiag}\left(\mathbf{T}^L, \mathbf{T}^S, \mathbf{T}^{\beta}, \mathbf{T}^I, \mathbf{T}^e\right) \text{ with}$$

$$\mathbf{T}^L = \mathbf{I}_3 \otimes \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

$$\mathbf{T}^S = \mathbf{I}_3 \otimes \text{blockdiag}\left(\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \mathbf{C}_4, \mathbf{C}_5, 1\right) \text{ with } \mathbf{C}_i \text{ defined in (A.4)},$$

$$\mathbf{T}^{\beta} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{T}^I = \mathbf{T}^e = \mathbf{O}_{3\times 3},$$

$$\boldsymbol{\eta}_t = \left(\boldsymbol{\eta}_t^{L+}, \boldsymbol{\eta}_t^{L0}, \boldsymbol{\eta}_t^{L-}, \boldsymbol{\eta}_t^{S+}, \boldsymbol{\eta}_t^{S0}, \boldsymbol{\eta}_t^{S-}, \boldsymbol{\eta}_t^{\beta}, \boldsymbol{\eta}_t^{I}, \boldsymbol{\eta}_t^{\varepsilon}\right) \text{ with}$$

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

$$\boldsymbol{\eta}_t^{Si} = \left(\eta_{S,i,t,1}, \eta_{S,i,t,1}^*, \ldots, \eta_{S,i,t,5}, \eta_{S,i,t,5}^*, \eta_{S,i,t,6}\right), i = +,0,-,$$

$$\boldsymbol{\eta}_t^{\beta} = 0_3, \quad \boldsymbol{\eta}_t^{I} = \left(I_t^+, I_t^0, I_t^-\right), \quad \boldsymbol{\eta}_t^{e} = \left(e_t^+, e_t^0, e_t^-\right),$$

$$\mathbf{Q}_t = \text{blockdiag}\left(\mathbf{Q}^L, \mathbf{Q}^S, \mathbf{Q}^{\beta}, \mathbf{Q}^I, \mathbf{Q}_t^e\right) \text{ with}$$

$$\mathbf{Q}^L = \text{diag}\left(0, \sigma_{R,+}^2, 0, \sigma_{R,0}^2, 0, \sigma_{R,-}^2\right),$$

$$\mathbf{Q}^S = \text{diag}\left(1_{11} \otimes \sigma_{\text{S},+}^2, 1_{11} \otimes \sigma_{\text{S},0}^2, 1_{11} \otimes \sigma_{\text{S},-}^2\right),$$

$$\mathbf{Q}^{\beta} = \mathbf{O}_{3\times 3},$$

$$\mathbf{Q}^I = \text{diag}\left(\sigma_{I,+}^2, \sigma_{I,0}^2, \sigma_{I,-}^2\right),$$

$$\mathbf{Q}_t^e = 3 \times 3 \text{ matrix with elements defined by } (4.6).$$

References

Bell, W.R. and Hillmer, S.C. (1990). The time series approach to estimation for repeated surveys. *Survey Methodology*, 16(2), 195-215. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/1990002/article/14535-eng.pdf.

Binder, D.A. and Dick, J.P. (1989). Modelling and estimation for repeated surveys. *Survey Methodology*, 15(1), 29-45. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/1989001/article/14579-eng.pdf.

Binder, D.A. and Dick, J.P. (1990). A method for the analysis of seasonal ARIMA models. *Survey Methodology*, 16(2), 239-253. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/1990002/article/14533-eng.pdf.

Blight, B.J.N. and Scott, A.J. (1973). A stochastic model for repeated surveys. *Journal of the Royal Statistical Society, Series B*, 35, 61-66.

Bollineni-Balabay, O., van den Brakel, J.A. and Palm, F. (2017). State space time series modelling of the Dutch Labour Force Survey: Model selection and mean squared errors estimation. *Survey Methodology*, 43(1), 41-67. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2017001/article/14819-eng.pdf.

Boonstra, H.J. and van den Brakel, J.A. (2019). Estimation of level and change for unemployment using structural time series models. *Survey Methodology*, 45(3), 395-425. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2019003/article/00005-eng.pdf.

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

Doran, H.E. (1992). Constraining Kalman filter and smoothing estimates to satisfy time varying restrictions. *Review of Economics and Statistics*, 74 568-572.

Durbin, J. and Koopman, S.J. (2012). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

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

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

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

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

Harvey, A.C. and Durbin, J. (1986). The effects of seat belt legislation on British road casualties: A case study in structural time series modelling. *Journal of the Royal Statistical Society, Series A*, 149, 187-227.

Harvey, A. and Koopman, S.J. (1993). Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association*, 88(424), 1228-1236.

Hungnes, H., Skjerpen, T., Hamre, J.I., Jansen, X.C., Pham, D.Q. and Sandvik, O. (2024). Structural break in the Norwegian Labor Force Survey due to a redesign during a pandemic. *Journal of Official Statistics*, 40, 122-160.

Koopman, S.J., Shephard, N. and Doornik, J.A. (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal, Royal Economic Society*, 2(1), 107-160.

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

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

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

Pfeffermann, D. and Bleuer, S.R. (1993). Robust joint modelling of labour force series of small areas. *Survey Methodology*, 19(2), 149-163. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/1993002/article/14458-eng.pdf.

Pfeffermann, D. and Burck, L. (1990). Robust small area estimation combining time series and cross-sectional data. *Survey Methodology*, 16(2), 217-237. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/1990002/article/14534-eng.pdf.

Pfeffermann, D. and Sikov, A. (2011). Imputation and estimation under nonignorable nonresponse in household surveys with missing covariate information. *Journal of Official Statistics*, 27, 181-209.

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

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

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

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

Sverchkov, M. (2008). A new approach to estimation of response probabilities when missing data are not missing at random. *Proceedings of the Survey Research Methods Section*, American Statistical Assiciation, 867-874.

Tam, S.M. (1987). Analysis of repeated surveys using a dynamic linear model. *International Statistical Review*, 55, 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, 149-166.

van den Brakel, J.A. and Krieg, S. (2015). Dealing with small sample sizes, rotation group bias and discontinuities in a rotating panel design. *Survey Methodology*, 41(2), 267-296. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2015002/article/14231-eng.pdf.

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

van den Brakel, J.A., Krieg, S. and Smeets, M. (2023). *The Methodology of the Dutch Consumer Confidence Survey During the Corona Crisis*. Discussion paper, December 2023. Statistics Netherlands, Heerlen.

van den Brakel, J.A. and Roels, J. (2010). Intervention analysis with state-space models to estimate discontinuities due to a survey redesign. *Annals of Applied Statistics*, 4, 1105-1138.

van den Brakel, J.A., Söhler, E., Daas, P. and Buelens, B. (2017). Social media as a data source for official statistics; the Dutch Consumer Confidence Index. *Survey Methodology*, 43(2), 183-210. Paper available at https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2017002/article/54871-eng.pdf.

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

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