Discussion Paper

Bootstrapping standard errors of estimates based on structural time series models

The views expressed in this paper are those of the author(s) and do not necessarily reflect the policies of Statistics Netherlands

2017 | 24

Sabine Krieg,
Jan van den Brakel

Content

1. Introduction 4

2. Background LFS 5

3. Method 7

4. Results 11

5. Estimation of up-to-date standard errors 21

6. Conclusion 22

References 23

Appendix A: details structural time series model 24

Appendix B: More results 26

Summary
Statistics Netherlands applies a structural time series model to compute monthly figures about the labour force. Standard errors of these estimates can be computed by well-known analytic formulas. The monthly figures are also used to derive quarterly and yearly figures and differences between two time periods. For most of these derived figures no analytic formula for standard errors is available. Therefore, a bootstrap algorithm proposed by Pfeffermann and Tiller (2005) is tested to derive standard errors for these figures.

Keywords
Structural time series modelling, bootstrap, mean squared error, labour force

Acknowledgement
The views expressed in this paper are those of the authors and do not necessarily reflect the views of Statistics Netherlands. The authors are thankful to the constructive input and comments received from Marc Smeets and Martijn Souren.

1. Introduction

The purpose of national statistical institutes is to produce reliable macro-economic indicators. Some of the most important indicators are monthly, quarterly and annual figures about the labour force, which are generally obtained with the Labour Force Survey (LFS). National statistical institutes often apply rotating panel designs for their LFS. Since 2010, Statistics Netherlands applies a structural time series model to produce the monthly figures about the labour force (van den Brakel and Krieg, 2009, 2010, 2012, 2014). These model-based figures have two important advantages compared with the traditional design-based or model-assisted estimation techniques, like the general regression (GREG) estimator (Särndal et al., 1992). First, the standard errors are substantially smaller, and second, the time series model can easily deal with the systematic differences between population parameter estimates that are based on the observations obtained in the subsequent waves of a panel. This phenomenon is known in the literature as rotation group bias (RGB), Bailar (1975).

The theory of structural time series models also provides analytic formulas for the standard errors of the model estimates (Harvey, 1989, Durbin and Koopman, 2012). Statistics Netherlands uses the model estimates also to derive other figures like quarterly and yearly figures and filtered estimates for period-to-period change. The standard approach to obtain estimates and standard errors of linear combinations of state variables over different time periods is to keep the required states from preceding periods in the state vector and compute this linear combination including its standard error with the standard Kalman filter recursion (de Jong and MacKinnon, 1988). A consequence of this approach is that the states from preceding periods are updated with the information that becomes available afterwards. Statistics Netherlands, however, does not have a revision policy for the monthly labour force figures. As a consequence this standard approach underestimates the standard error of quarterly and annual means and period-to-period differences. Consequently, until now, Statistics Netherlands publishes the point estimates for these figures without information about the standard error. By the best of our knowledge, there are no analytic expressions for the standard error of linear combinations of filtered states observed at different time periods that correctly account for the correlation between them. Therefore bootstrapping is investigated in this paper as an alternative.

Bootstrap algorithms for structural time series models are proposed in the literature (for example Pfeffermann and Tiller, 2005, or Rodrígez and Ruiz, 2012). There, the problem to be solved is slightly different: the analytic formulas for the standard errors assume that the variance parameters of the structural time series model are known, which is generally not true. By replacing them by maximum likelihood estimates, the standard error is underestimated. A correction for this underestimation is possible by the bootstrap procedures proposed by Pfeffermann and Tiller (2005) and Rodrígez and Ruiz (2012). This problem effects the Dutch LFS, but the underestimation is quite small (Bollineni-Balabay e.a., 2017). In the present paper, bootstrapping is applied since no analytic formula for the standard error is available for the aforementioned model estimates of the Dutch LFS. One version of the approach of Pfeffermann and Tiller (2005) can be adapted for this situation.

The paper is organised as follows. First, some background information about the Dutch LFS is given in Section 2. Then the method of bootstrapping structural time series models is discussed in Section 3, both for the general case and for the specific situation of the Dutch LFS. Results are described in Section 4. The computation time for bootstrapping is large, which means that it is not possible to apply this method in the production of the official publications for all monthly target variables, at least not with the standard computers of Statistics Netherlands. A less computational intensive approximation is shortly described in Section 5. This is not worked out yet, but this should be possible in the near future. In Section 6, the conclusions can be found.

2. Background LFS

The objective of the Dutch LFS is to provide reliable information about the Dutch labour force. The target population of the LFS consists of the non-institutionalised population aged 15 years and over residing in the Netherlands. The sampling frame is a list of all known occupied addresses in the Netherlands, which is derived from the municipal basic registration. Each month a stratified two-stage cluster design is used to select a sample of addresses. Strata are formed by geographical regions. Municipalities are considered as primary sampling units and addresses as secondary sampling units. All households residing at an address, up to a maximum of three, are included in the sample and can be regarded as the ultimate sampling units. Different subpopulations are oversampled to improve the accuracy of the official releases, for example addresses with persons registered at the employment office and subpopulations with low response rates.

Since October 1999, the LFS is conducted as a rotating panel design. Until June 2010, data in the first wave were collected by means of computer assisted personal interviewing (CAPI). The respondents are re-interviewed four times at quarterly intervals by means of computer assisted telephone interviewing (CATI). During these re-interviews a condensed questionnaire is applied to establish changes in the labour market position of the respondents. Participation of households with the Dutch LFS is on a voluntary basis.

In January 2010 the data collection in the first wave changed from CAPI to a mix of CAPI and CATI. This survey redesign resulted in discontinuities that required an extension of the time series model, see Van den Brakel and Krieg (2012). In April 2012, the data collection in the first wave changed again, now to a mix of internet, CATI and CAPI, which resulted again in discontinuities and required another extension of the model. The mode of the re-interviews is not changed, i.e. it is still CATI.

Discontinuities induced by these redesigns are quantified by conducting the first wave under the old and new design in parallel for a period of six months, both with the regular monthly sample size. This enables a direct estimate for the discontinuities for the main parameters in the first wave. In 2012, the field capacity was sufficient to also conduct the second wave in parallel. It was, however, not possible to conduct the other subsequent waves in parallel, too. Possible discontinuities are quantified using an intervention approach.

Key parameters of the LFS are the employed, unemployed and total labour force, which are defined as population totals. Another important parameter is the unemployment rate, which is defined as the ratio of the unemployed labour force over the total labour force. Monthly estimates for these parameters are produced at the national level as well as a breakdown in six domains that is based on the cross classification of gender and three age classes. Monthly estimates are obtained with the following estimation procedure. Each month data are collected in five independent waves. The GREG estimator is applied to produce five independent estimates for a target parameter. Inclusion probabilities reflect the sampling design described above as well as the different response rates between geographic regions. The weighting scheme is based on a combination of different socio-demographic categorical variables. This results in five series of monthly GREG estimates for each target parameter, which are the input for the multivariate structural time series model described in Section 3. With this model reliable estimates for the population parameters are obtained by taking advantage of the sample information observed in preceding periods. The model also accounts for the discontinuities caused by the redesigns and for the RGB and autocorrelation induced by the rotating panel design. Since 2010 this approach is applied to produce official monthly figures about the labour force, Van den Brakel and Krieg (2009, 2012).

The monthly model estimates are also used for the computation of quarterly and yearly figures. Statistics Netherlands aims to publish consistent estimates, i.e. for example the quarterly estimate for the unemployed labour force is the mean of the three monthly figures. The aim of this paper is to estimate the standard error of these estimates. Furthermore, by publishing monthly, quarterly and yearly figures, automatically differences between two time periods can be computed. The standard error of these differences is also estimated in this paper. The following estimates are considered:

- Quarterly figures
- Yearly figures
- The difference of month *t* and month *t-1*
- The difference of month *t* and month *t-3*
- The difference of month *t* and month *t-12*
- The difference of quarter *q* and quarter *q-1*
- The difference of year *a* and year *a-1*

In the paper, monthly figures are also considered in order to compare the estimates of the standard error by bootstrapping and by the analytic formula.

These figures are estimated both for the trend and for the signal (sum of trend and seasonal effect), for the employed, unemployed and the total labour force. This is published on national level and for the breakdown in six subpopulations (based on age class and gender).

The quarterly and yearly figures are about the according calendar periods. In this paper, all rolling quarterly figures and all rolling annually figures are considered.

When figures about the current month *t* are computed based on a structural time series, the information up to and including month *t* is used. When new information (about month *t+1* etc.) becomes available, the estimates for month *t* and the months before can be updated using the new information. These updated estimates are more accurate. Nevertheless, Statistics Netherlands does not wish to revise the figures about the labour force, as this will delay the production of quarterly figures too much. Therefore interest is focused on the standard errors of differences between filtered estimates. No exact analytic formulas are available for these figures. This fact is an important aspect in the computation of the standard errors.

3. Method

3.1 Bootstrapping of structural time series models

A structural time series model consists of two equations, the measurement equation and the state equation. The measurement equation describes how the time series $\boldsymbol{y}_t$ depends on the unknown state variables $\boldsymbol{u}_t$:

$$\boldsymbol{y}_t = \boldsymbol{Z}_t\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_t \sim N(0, \boldsymbol{\Sigma}_t), E(\boldsymbol{\varepsilon}_t, \varepsilon_t') = 0, \ t = 1, \dots, T \qquad (3.1)$$

The state equation describes how the state variable vector $\boldsymbol{u}_t$ evolve over time:

$$\boldsymbol{u}_t = \boldsymbol{G}_t\boldsymbol{u}_{t-1} + \boldsymbol{\eta}_t, \boldsymbol{\eta}_t \sim N(0, \mathbf{Q}_t), E(\boldsymbol{\eta}_t, \boldsymbol{\eta}_t') = 0, \ t = 1, \dots, T \qquad (3.2)$$

with $E(\boldsymbol{\varepsilon}_t, \boldsymbol{\eta}_s') = 0$.

The general way to estimate a structural time series model is to express it in the so-called state-space representation and apply the Kalman filter to obtain optimal estimates for the state variables, see e.g. Durbin and Koopman (2012). Estimates for state variables for period *t* based on the information available up to and including period *t* are referred to as the filtered estimates. The filtered estimates of past state vectors can be updated, if new data become available. This procedure is referred to as smoothing and results in smoothed estimates that are based on the completely observed time series. The variances $\boldsymbol{\Sigma}_t$ and $\mathbf{Q}_t$ depend on a parameter vector $\boldsymbol{\Lambda}$, which is generally unknown, and is therefore estimated using a maximum likelihood procedure before the Kalman filter can be applied. The parameters in $\boldsymbol{\Lambda}$ are generally called hyperparameters.

In the present paper the analysis is conducted with software developed in OxMetrics in combination with the subroutines of SsfPack 3.0, see Doornik (2009) and Koopman e.a. (1999, 2008).

In the literature of bootstrapping structural time series models (for example Pfeffermann and Tiller, 2005), the focus is on the computation of the standard error of linear combinations of state variables, i.e. $\hat{\alpha}_t = \boldsymbol{l}_t'\hat{\boldsymbol{u}}_t$, where $\boldsymbol{l}_t$ is a known vector that defines the linear combination of the estimated state variables at period t. In the analytic formula for the standard errors of $\hat{\alpha}_t$ it is assumed that $\boldsymbol{\Lambda}$ is known, which results in an underestimation of the standard errors when the estimates $\widehat{\boldsymbol{\Lambda}}$ are used. The aim of Pfeffermann and Tiller (2005) is to correct for this underestimation. In the present paper, another problem is considered: For the computation of quarterly and yearly figures and for the computation of differences between time periods more than one month t is involved. The standard approach is to include the required state variables from preceding periods $t' < t$ in the vector $\boldsymbol{u}_t$ and use the standard Kalman filter recursions to obtain mean square error estimates for $\hat{\alpha}_t$. This is the approach followed by de Jong and MacKinnon (1988). In this case the filtered estimates for period t' are updated with the information that became available after period t'. However, Statistics Netherlands combines the filtered estimates of state variables for different time periods. This means that interest is focussed on linear combinations between filtered estimates of state variables at different time periods, for example $\boldsymbol{l}_t'\hat{\boldsymbol{u}}_{t|t} - \boldsymbol{l}_t'\hat{\boldsymbol{u}}_{t'|t'}$ instead of $\boldsymbol{l}_t'\hat{\boldsymbol{u}}_{t|t} - \boldsymbol{l}_t'\hat{\boldsymbol{u}}_{t'|t}$ in the case of the difference between two time periods, where $\hat{\boldsymbol{u}}_{t|\tau}$ denote the estimate for the state vector $\boldsymbol{u}_t$ at period t based on the observation obtained until period τ. If $\tau = t$, then $\hat{\boldsymbol{u}}_{t|\tau}$ is the filtered estimate for $\boldsymbol{u}_t$. If τ equals the last period observed in the series, then $\hat{\boldsymbol{u}}_{t|\tau}$ is the smoothed estimate for $\boldsymbol{u}_t$.
At the best of our knowledge, no analytic MSE formula is applicable for most of the estimates which are published by Statistics Netherlands. Only for the filtered monthly figures that contain state variables for the same period and for all smoothed estimates an analytic MSE formula is applicable. Results based on this formula are considered in this paper to compare the analytic MSE estimates and MSE estimates based on the bootstrap as far as possible.

Let $\hat{\alpha}_{t_1,\dots,t_k} = \boldsymbol{l}'(\hat{\boldsymbol{u}}_{t_1|t_1}', \dots, \hat{\boldsymbol{u}}_{t_k|t_k}')'$ denote the parameter of our interest. The parametric bootstrap procedure for α_t is explained first, as it is worked out by Pfeffermann and Tiller (2005).

1. A large number B of bootstrap state vector series $\boldsymbol{u}_t^b$ and observations $\boldsymbol{y}_t^b, b = 1, \dots, B$ is generated using the model. For the hyperparameters $\boldsymbol{\Lambda}$ the estimates $\widehat{\boldsymbol{\Lambda}}$ based on the original series $\boldsymbol{y}_t$ are used. Using the bootstrap state vectors, $\alpha_t^b = \boldsymbol{l}_t'\boldsymbol{u}_t^b$ can be computed.
2. Then the hyperparameters $\boldsymbol{\Lambda}$ are estimated again, using the observations $\boldsymbol{y}_t^b$. Based on this estimate $\widehat{\boldsymbol{\Lambda}}^b$ and the bootstrap observations $\boldsymbol{y}_t^b$, the Kalman filter is applied to compute estimates $\hat{\mathbf{u}}_t^b$ for the state vector and estimates $\hat{\alpha}_{t,1}^b = \boldsymbol{l}_t'\hat{\mathbf{u}}_t^b$ for α_t. Furthermore, bootstrap estimates for α_t can be computed by applying the Kalman filter using the observations $\boldsymbol{y}_t^b$ and the original hyperparameter estimates $\widehat{\boldsymbol{\Lambda}}$. They are called $\hat{\alpha}_{t,2}^b$.
3. The MSE of $\hat{\alpha}_t$ can be estimated by one of the following formulas:

$$MSE(\hat{\alpha}_t) = MSE^B_{t,1} + 2P_t(\widehat{\mathbf{\Lambda}}) - \overline{P^B_t} \quad (3.3)$$
$$MSE(\hat{\alpha}_t) = MSE^B_{t,2} + P_t(\widehat{\mathbf{\Lambda}}) - \overline{P^B_t} \quad (3.4)$$

with $P_t(\widehat{\mathbf{\Lambda}})$ the analytic MSE of the original estimates $\hat{\alpha}_t$ and $\overline{P^B_t} = \frac{1}{B}\sum_{b=1}^{B} P_t(\widehat{\mathbf{\Lambda}}^b)$ the mean over all analytic MSEs of the bootstrap estimates $\hat{\alpha}^b_{t,1}$. Furthermore

$$MSE^B_{t,1} = \frac{1}{B}\sum_{b=1}^{B}[\hat{\alpha}^b_{t,1} - \hat{\alpha}^b_{t,2}\,]^2$$

and

$$MSE^B_{t,2} = \frac{1}{B}\sum_{b=1}^{B}[\hat{\alpha}^b_{t,1} - \alpha^b_t\,]^2.$$

Formula (3.3) is largely based on the analytic MSE $P_t(\widehat{\mathbf{\Lambda}})$ with a bootstrap bias correction $P_t(\widehat{\mathbf{\Lambda}}) - \overline{P^B_t}$. Furthermore, the term $MSE^B_{t,1}$ corrects for the uncertainty of the hyperparameters in $\widehat{\mathbf{\Lambda}}$. Formula (3.4) is largely based on the bootstrap MSE ($MSE^B_{t,2}$). Pfeffermann and Tiller (2005) prove that the correction $P_t(\widehat{\mathbf{\Lambda}}) - \overline{P^B_t}$ in (3.4) is necessary for a bias of order $O(\frac{1}{T^2})$. For a detailed discussion of the formulas (3.3) and (3.4) see Pfeffermann and Tiller (2005).

3.2 The situation of the Dutch LFS

The multivariate structural time series model is described in detail in Van den Brakel and Krieg (2009, 2012, 2014). Here a short summary of the model is given. More details can also be found in Appendix A.

Let $y_{j,t}$ denote the GREG estimate for the unknown population parameter, say θ_t, based on the *j*-th wave observed at time *t*, *j*= 1,…,5. Due to the applied rotation pattern, each month a vector $\mathbf{Y}_t = (y_{1,t}, y_{2,t}, y_{3,t}, y_{4,t}, y_{5,t})'$ is observed. As a result, a five dimensional time series with GREG estimates for both the monthly employed and unemployed labour force is obtained. This vector can be modelled as

$$\mathbf{Y}_t = \mathbf{1}_5\theta_t + \boldsymbol{\lambda}_t + \mathbf{\Delta}^{(1)}_t\mathbf{B}^{(1)} + \mathbf{\Delta}^{(2)}_t\mathbf{B}^{(2)} + \mathbf{e}_t. \quad (3.5)$$

with $\mathbf{1}_5$ a five dimensional vector with each element equal to one, $\boldsymbol{\lambda}_t = (\lambda_{1,t}, \lambda_{2,t}, \lambda_{3,t}, \lambda_{4,t}, \lambda_{5,t})'$ a vector with time dependent components that account for the RGB, $\mathbf{\Delta}^{(b)}_t = diag(\delta^{(b)}_{t,1}, \delta^{(b)}_{t,2}, \delta^{(b)}_{t,3}, \delta^{(b)}_{t,4}, \delta^{(b)}_{t,5}), b = 1{,}2$ diagonal matrices with dummy variables that change from zero to one at the moment that the survey changes from the first to the second and from the second to the third design, $\mathbf{B}^{(b)} = (B^{(b)}_1, B^{(b)}_2, B^{(b)}_3, B^{(b)}_4, B^{(b)}_5)'$, *b*=1,2 five dimensional vectors with time independent regression coefficients, and $\mathbf{e}_t = (e^1_t, e^2_t, e^3_t, e^4_t, e^5_t)'$ the corresponding survey errors for each wave estimate.

The population parameter θ_t in (1) can be decomposed in a smooth trend component L_t, a trigonometric seasonal component S_t, and an irregular component ε_t, i.e.

$$\theta_t = L_t + S_t + \varepsilon_t. \quad (3.6)$$

To identify the model component for the RGB, it is assumed that the first wave is unbiased, i.e.
$\boldsymbol{\lambda}_t = (0, \lambda_{2,t}, \lambda_{3,t}, \lambda_{4,t}, \lambda_{5,t})'$. For the other four components a random walk is assumed. Autocorrelation between survey errors due to panel overlap is modelled

with an AR(1) model. The model for the survey errors also account for heteroscedastic sampling variance by using the standard errors of $\mathbf{Y}_t$ as a priori information in the measurement equation. For more details of the models for L_t, S_t, ε_t, $\mathbf{e}_t$ and $\boldsymbol{\lambda}_t$ see Appendix A.

The discontinuities are modelled with intervention variables $\delta_{t,j}^{(b)}, b = 1,2, j = 1, \dots, 5$ with

$$\delta_{t,j}^{(b)} = \begin{cases} 0 & \text{if} \quad t < \tau_{b,j} \\ 1 & \text{if} \quad t \geq \tau_{b,j} \end{cases}. \tag{3.7}$$

where $\tau_{1,j}$ ($\tau_{2,j}$) denotes the moment that wave j changes from the first to the second (from the second to the third) survey design. Under the assumption that (3.6) correctly models the evolution of the population variable, the regression coefficients in $\mathbf{B}^{(b)}, b = 1,2$ can be interpreted as the systematic effect of the redesign on the level of the series observed in the five waves. Based on the parallel period, direct estimates for $B_1^{(1)}$, $B_1^{(2)}$, and $B_2^{(2)}$ are available. These estimates are used as input of the model without taking the sampling errors of these values into account. For the other regression coefficients in $\mathbf{B}^{(b)}, b = 1,2$, no information is available in advance. These coefficients are estimated as state variables of the time series model. More details about modelling the discontinuities can be found in Van den Brakel and Krieg (2012).

Bootstrap series using the model can be generated using a special function of Ssfpack3.0 (ssfRecursion). Here, parametric bootstrapping is applied, which means that a normal distribution for the disturbances is assumed. The function ssfRecursion cannot generate the data which is collected in parallel for the direct estimates of $B_1^{(1)}$, $B_1^{(2)}$, and $B_2^{(2)}$. Instead, these estimates are part of the matrix $\boldsymbol{Z}_t$. For the other discontinuities, only $\delta_{t,j}^{(b)}$ is part of the matrix $\boldsymbol{Z}_t$. The smoothed estimates for the state vector $\boldsymbol{u}_t$ for the first time period $t = 1$ are used as state variables for $\boldsymbol{u}_1^b$ as well. The further development of the bootstrap state variables $\boldsymbol{u}_t^b, t = 2, \dots, T$ depends on the generated random numbers according to $\widehat{\boldsymbol{\Lambda}}$. Note that in these generated bootstrap series, the values for $y_{1,t}, y_{2,t}, y_{3,t}, y_{4,t}, y_{5,t}$ are quite often smaller than zero. This is not logical as the series are about the number of employed or unemployed persons in the Netherlands. It is assumed that negative numbers do not affect the variability of the differences between the bootstrapped $\alpha_{t_1,\dots,t_k}^b$ and its estimates. Note furthermore that the model estimates for the discontinuities $B_2^{(1)}, \dots, B_5^{(1)}, B_3^{(2)}, B_4^{(2)}$ and $B_5^{(2)}$ are part of $\boldsymbol{u}_1$ and therefore also of $\boldsymbol{u}_1^b$. This way, all discontinuities are the same in all bootstrap series.

The bootstrap procedure proposed by Pfeffermann and Tiller (2005), see step 1 – 3 in Section 3.1, has to be adapted to estimate the MSE of $\hat{\alpha}_{t_1,\dots,t_k}$ instead of $\hat{\alpha}_t$. As the entire series are bootstrapped, the changes in step 1 and 2 are straightforward, i.e. $\alpha_{t_1,\dots,t_k}^b$ and $\hat{\alpha}_{t_1,\dots,t_k,1}^b$ have to be computed. The value $\hat{\alpha}_{t_1,\dots,t_k,2}^b$ is not needed in step 3.

Both formulas of step 3, (3.3) and (3.4), are not applicable, as they involve the analytic MSE which is not available. Therefore, a less accurate simplification of formula (3.4) is used, which means that the MSE estimate is completely based on the bootstrap:

$$MSE\left(\hat{\alpha}_{t_1,\dots,t_k}\right) = MSE_2^B = \frac{1}{B}\sum_{b=1}^{B}[\hat{\alpha}_{t_1,\dots,t_k,1}^b - \alpha_{t_1,\dots,t_k}^b\,]^2. \quad (3.8)$$

The term $P_t(\widehat{\mathbf{\Lambda}}) - \overline{P_t^B}$ from formula (3.4) is skipped here as this term is based on the analytic MSE.

In this paper, *B*=4000 bootstrap runs are carried out. The computation time for these number of runs is around 4 days on a standard computer of Statistics Netherlands for one target variable. A comparison of bootstrap results based on 2000 and 4000 runs shows that an error of around 5% is possible with only 2000 runs. With 4000 runs, it is likely that there is still some (small) error, but a substantial increase of the number of runs is not realistic. On the other hand, the error is sufficiently small for a good impression of the MSE.

In some of the bootstrap runs, numerical problems in de procedure of maximizing the likelihood can occur. This happened in less than 5% of the cases. These runs are left out in the analysis, which means that slightly less than 4000 bootstrap runs are used.

4. Results

In this section the standard error (SE) estimates, estimated as the root MSE based on bootstrapping, are presented. In Section 4.1, results for parameters where an analytic formula is available are discussed. There, the bootstrap estimates are compared with the analytic formula. Furthermore, the correction term $P_t(\widehat{\mathbf{\Lambda}}) - \overline{P_t^B}$ of formula (3.4) is discussed. As this correction term cannot be computed in most of the cases (since no analytic formula of the MSE is available) it is interesting to see how large this correction could be. In Section 4.2, bootstrap results are discussed for parameters where no correct analytic formula is available, for the unemployed labour force on national level. Some of the results are shown in the Appendix B. The results for other variables (breakdown in subpopulations, employed and total labour force) are shown in less detail in Section 4.3.

The results in the figures in Section 4.1 are presented as follows.

- The bootstrap estimate for the standard error based on formula (3.8). This is called bootstrap SE in the figures and shown in green.
- The term $\sqrt{P_t(\widehat{\mathbf{\Lambda}})}$, which is the standard error based on the analytic formula from the standard Kalman filter recursion, abbreviated as St. analytic SE in the figures and shown in black.
- The term $\sqrt{\overline{P_t^B}}$, which is the square root of the mean over the variances based on the analytic formula, shortly called bootstrap mean of the standard analytic standard error and abbreviated as BM St. analytic SE in the figures and shown in red.

In Section 4.2, only the first two results are shown in the figures. In this section standard error of linear combinations of states from different time periods are calculated and in the analytic formula, the approach proposed by de Jong and MacKinnon (1988) is applied.

4.1 Comparison bootstrap and analytic formula

Figure 4.1 shows the results for the standard errors of the smoothed monthly estimates of the trend of the unemployed labour force (national level).

The first conclusion is that the bootstrap estimate based on approximation (3.8) is quite close to the analytic result. So the bootstrap result is plausible. Second, the difference between $P_t(\widehat{\boldsymbol{\Lambda}})$ and $\overline{P_t^B}$ is small, which means that omitting the correction term $P_t(\widehat{\boldsymbol{\Lambda}}) - \overline{P_t^B}$ effects the results only slightly. Third, there are some small differences between the bootstrap-SE and $\sqrt{P_t(\widehat{\boldsymbol{\Lambda}})}$. In some periods, the bootstrap-SE is slightly larger than $\sqrt{P_t(\widehat{\boldsymbol{\Lambda}})}$, in other periods, both values are almost exactly equal. Theoretically, the bootstrap-SE should be slightly larger, as $\sqrt{P_t(\widehat{\boldsymbol{\Lambda}})}$ does not take the estimation of the hyperparameters into account. However, the accuracy of the bootstrap-SE-estimates is limited by the number of bootstrap-runs. It is likely that the fluctuations of these estimates over time would decrease with a substantial larger number of bootstrap-runs, but this is not tested due to the large computation time. It is tested that the fluctuations are larger with less bootstrap runs.

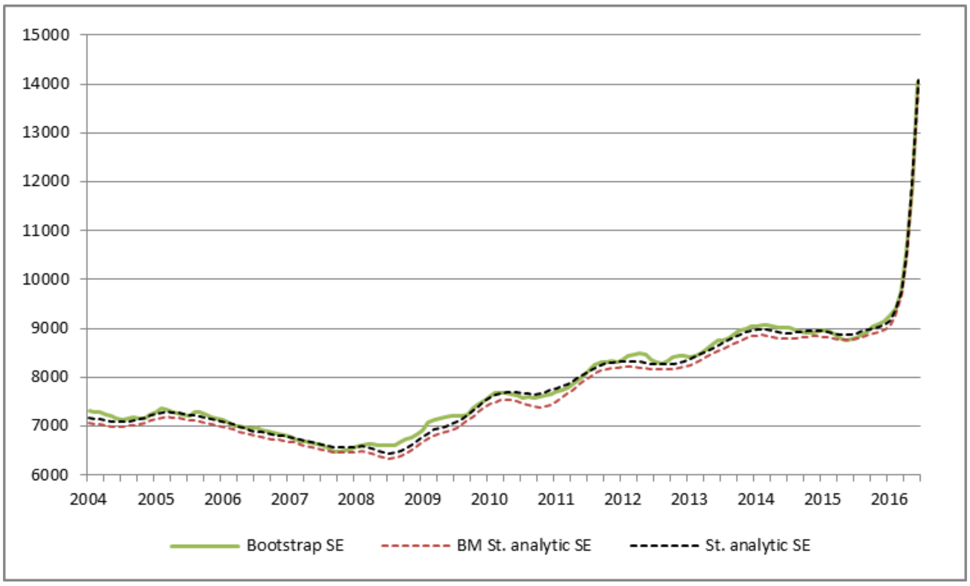

Figure 4.1: Standard error estimates for monthly trend, smoothed, unemployed labour force, national level

Figure 4.2 shows similar results for the filtered monthly estimates of the trend of the unemployed labour force (national level). As mentioned before, for the filtered monthly estimates the analytic MSE formula is a correct estimate. The results are similar as for the smoothed estimates, except that the estimates are more volatile. This makes the comparison of the three lines less easy. In some months in 2010/2011, the results based on the analytic formula are larger than the results based on the bootstrap. This is the period of the first discontinuity. A possible explanation of this difference is the way the discontinuities are processed in the bootstrap.

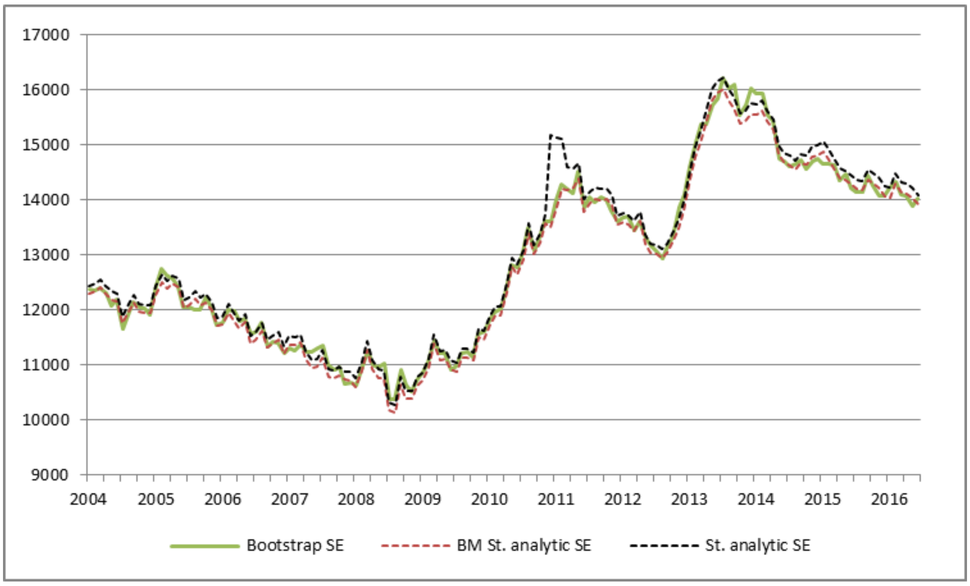

Figure 4.2: Standard error estimates for monthly trend, filtered, unemployed labour force, national level

As a third example, the results for the quarterly smoothed estimates of the trend of the unemployed labour force (national level) are shown in Figure 4.3, obtained as the mean over the three monthly time series model estimates. Again, the results are similar.

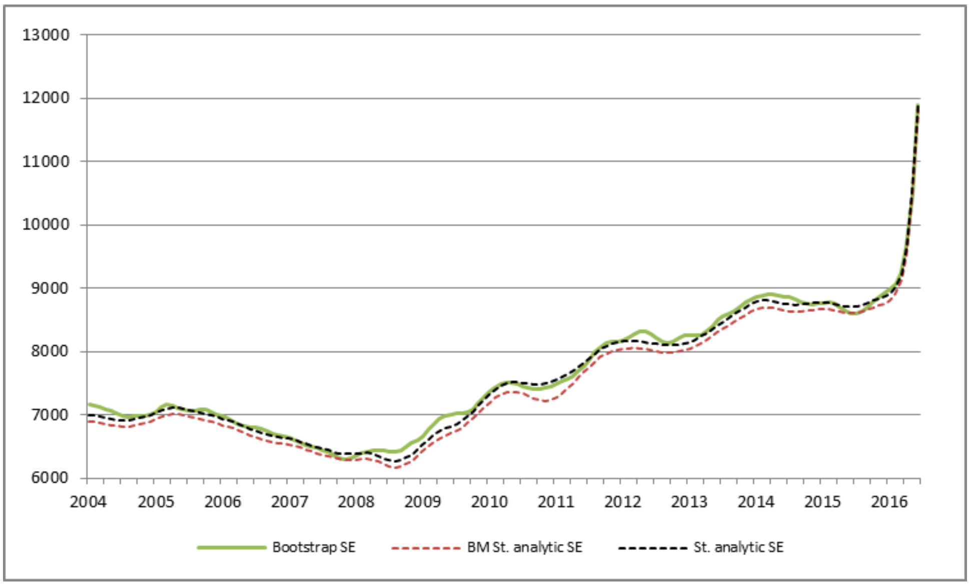

Figure 4.3: Standard error estimates for quarterly trend, smoothed, unemployed labour force, national level

In the examples of this section, formula (3.4) could be applied instead. The figures suggest that the correction term $P_t(\widehat{\boldsymbol{\Lambda}}) - \overline{P_t^B}$ which is omitted in formula (3.8) is small. A short computation shows that the underestimation of the standard error by omitting this correction is around 1% - 1,5% except in the period of 2010/2011 where $P_t(\widehat{\boldsymbol{\Lambda}})$ is larger than the other parameters for some months. In this period, the underestimation is maximal 12% for the filtered trend and between 3% and 4% for the other series.

The results for the other smoothed estimates are not shown, since the focus of this paper is on the filtered estimates where no analytic formula is available. From all cases of the smoothed estimates follows that the bootstrap results and the results from the analytic formula are similar, and the correction term $P_t(\widehat{\boldsymbol{\Lambda}}) - \overline{P_t^B}$ is small.

Results for filtered estimates, unemployed labour force on national level
Figure 4.4 shows the bootstrap-estimates for the standard errors of the quarterly figures for the trend. The estimates based on the analytic formula are also added. This formula computes the standard error of quarterly figures as mean of month *t-2*, *t-1* and *t*, where all estimates are based on the time series observed until month *t*, following the approach of de Jong and MacKinnon (1988). Since in the procedures of Statistics Netherlands the estimates of month *t-2* and *t-1* are not updated, the analytic formula underestimates the true SE. This is clearly shown in Figure 4.4.

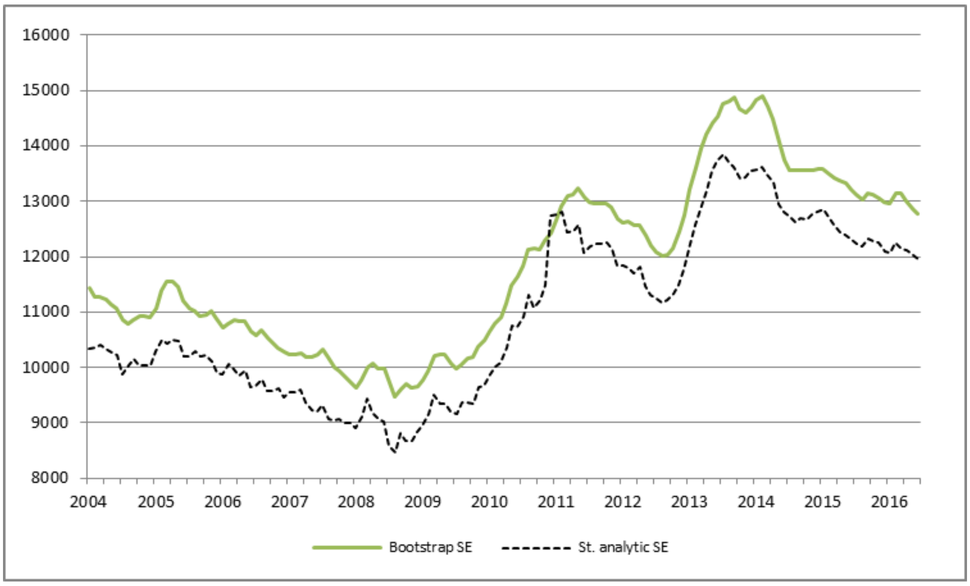

Figure 4.4: Standard error estimates for quarterly trend, filtered, unemployed labour force, national level

Similar results for the other series of the trend are shown in Figure 4.5 – 4.10. In all figures, the estimates based on the analytic formula are smaller than the ones obtained by bootstrapping. So under the assumption that the bootstrap estimates are not seriously biased, the formula underestimates the true standard error, as expected. The amount of this underestimation varies for the different series. The bootstrap estimates and the analytic formula estimates develop similar through time, with some decrease until (around) 2009. Then the estimates increase until (around) 2012. This might be explained partly with the increasing unemployment and partly with to the discontinuity of 2010. After a short period of decrease, the series start increasing again, probably due to the second discontinuity. From around 2014, the series decrease, since the effects of the discontinuities become smaller as more information under the new design becomes available. The maximums due to the discontinuities are not in exactly the same periods for all series. This is caused by the fact that different variables include different sets of months. Also, the maximums are not in exactly the same periods for the results based on the formula and based on bootstrapping. In Figure 4.10 (difference of year *a* and *a-1*), even 3 maximums are found by bootstrapping.

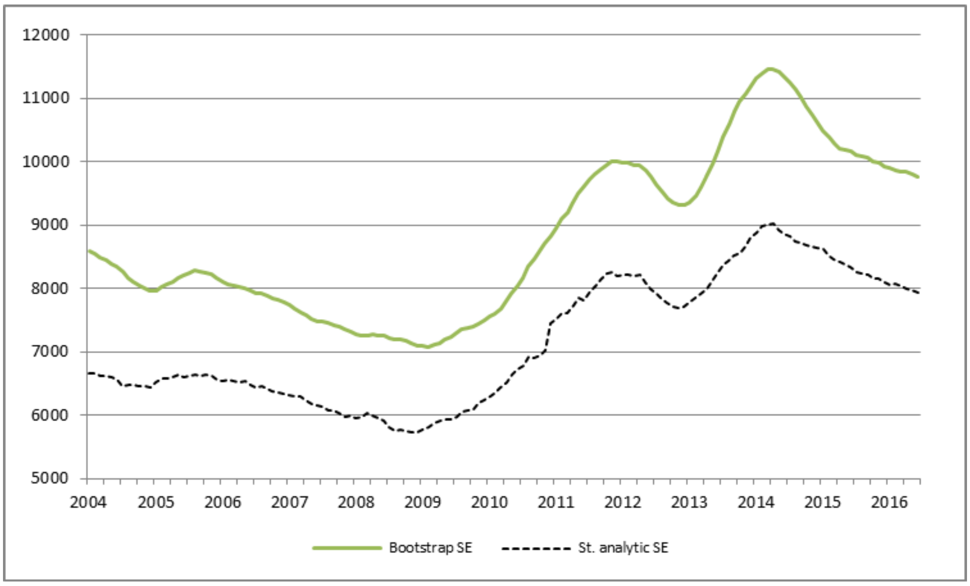

Figure 4.5: Standard error estimates for yearly trend, filtered, unemployed labour force, national level

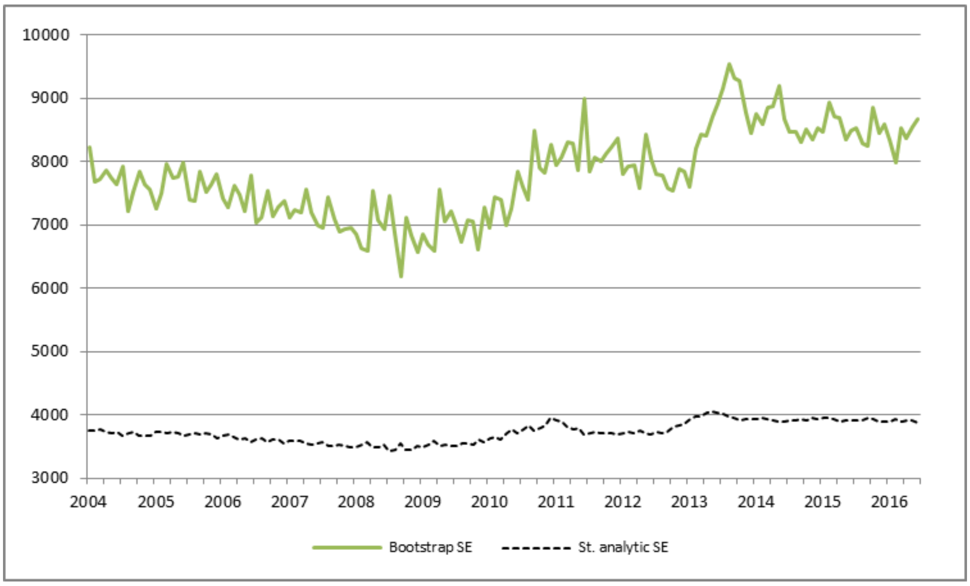

Figure 4.6: Standard error estimates for the difference of the trend of month *t* and *t-1*, filtered, unemployed labour force, national level

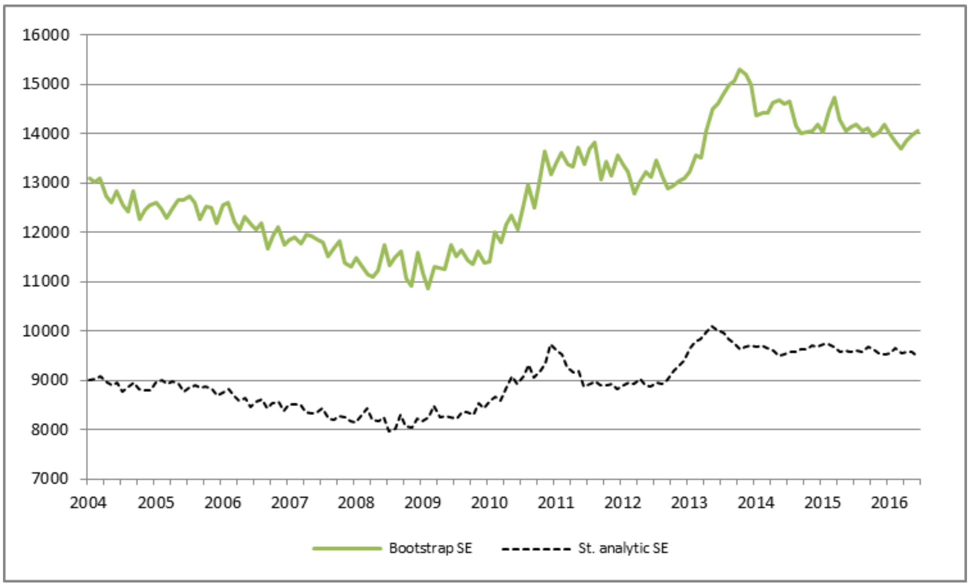

Figure 4.7: Standard error estimates for the difference of the trend of month *t* and *t-3*, filtered, unemployed labour force, national level

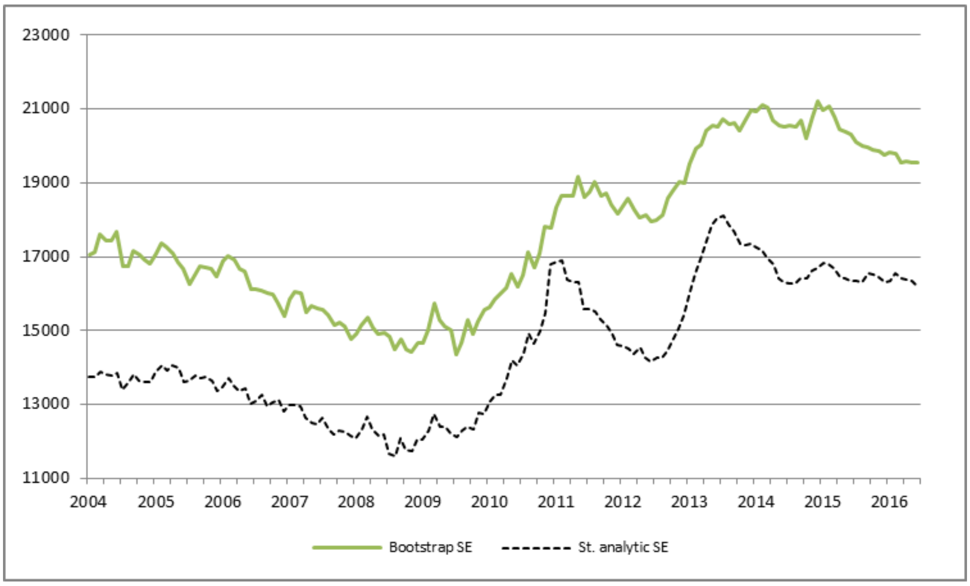

Figure 4.8: Standard error estimates for the difference of the trend of month *t* and *t-12*, filtered, unemployed labour force, national level

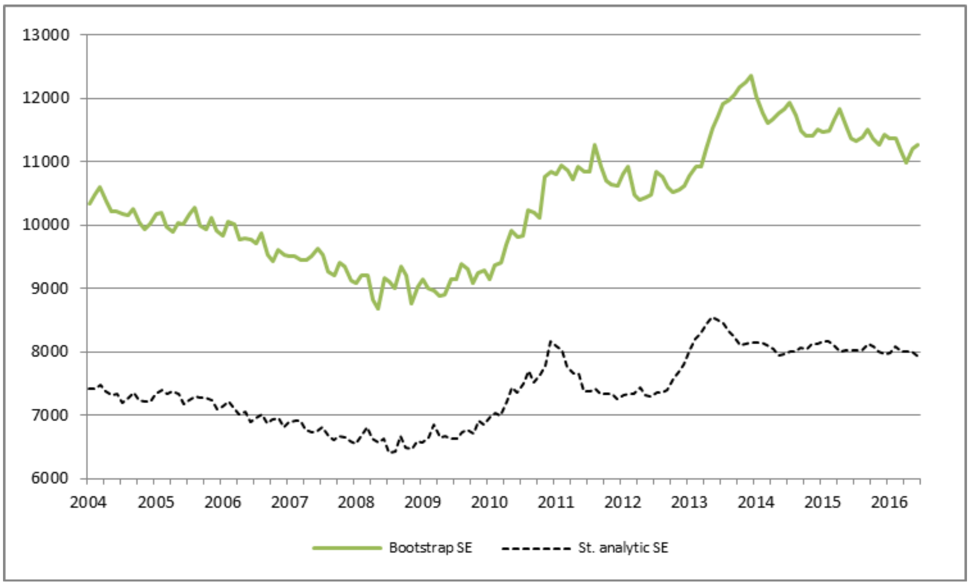

Figure 4.9: Standard error estimates for the difference of the trend of quarter *q* and *q-1*, filtered, unemployed labour force, national level

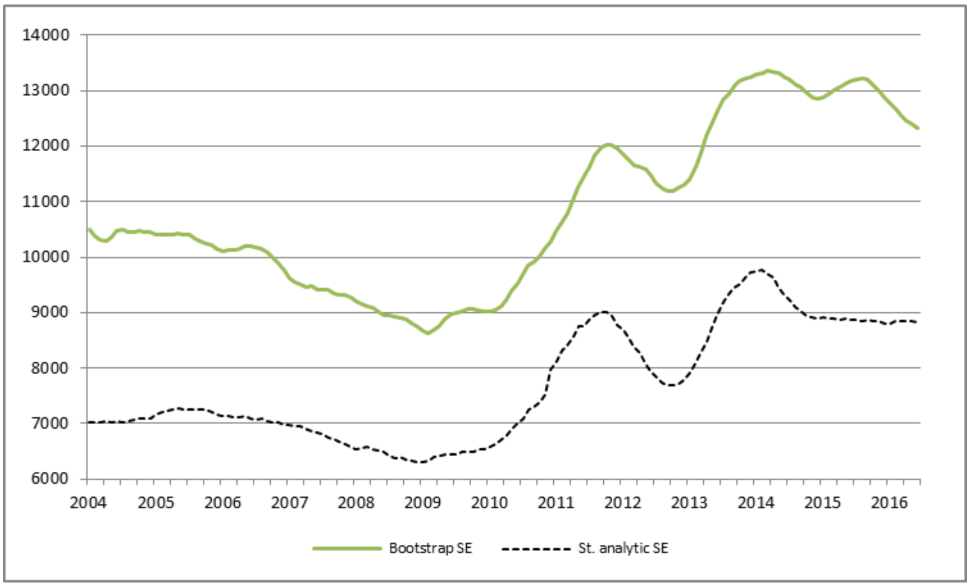

Figure 4.10: Standard error estimates for the difference of the trend of year *a* and *a-1*, filtered, unemployed labour force, national level

Similar results for the signal can be found in the Appendix B. Again, the analytic formula underestimates the standard error (under the abovementioned assumption that the bootstrap results are not seriously biased). For some series, the difference between formula and bootstrap result increase through time for the signal. Note furthermore that for some series, the bootstrap standard error for the trend is slightly smaller than the one for the signal (for example the difference of month t and *t-3*). For other series, it is the other way round (for example the yearly figures). For the difference of month *t* and *t-1*, the difference between the standard error for the trend is much smaller than the one for the signal.

4.2 Results for other target variables

The bootstrap computation is carried out also for the other 20 target variables, i.e. the breakdown in six domains (age class x gender) and for the employed labour force and for the total labour force (both on national level and for the breakdown in six domains). The results for the other variables are quite similar as for the unemployed labour force on national level, with mostly a similar development through time and a similar relative underestimation by the formula. Therefore, only the most important results are summarized in Table 4.1.

	quarter	year	month to month (diff1)	month to month (diff3)	month to month (diff12)	quarter to quarter	year to year
WL, NL							
T, boot	13000	9900	8500	14000	19800	11300	12800
T, form	12200	8100	3900	9600	16400	8000	8800
T+S, boot	12800	9700	12300	15500	19400	11500	12400
T+S, form	12000	8100	9900	12100	16400	8700	8800
WL, m1524							
T, boot	4100	3400	2000	3400	5600	2800	4000
T, form	3900	2800	800	2000	4500	1800	2600
T+S, boot	4100	3300	4200	4800	5500	3200	3800
T+S, form	3900	2800	3800	4000	4600	2600	2600
WL, m2544							
T, boot	5000	4000	3000	5100	7700	4200	5200
T, form	4700	3200	1300	3300	6200	2800	3400
T+S, boot	4900	3900	5600	6300	7500	4400	5000
T+S, form	4600	3200	5000	5200	6300	3500	3400
WL, m4574							
T, boot	4800	4100	2400	4000	6900	3300	5000
T, form	4600	3300	900	2300	5400	2100	3300
T+S, boot	4800	4000	3400	4500	6800	3400	4900
T+S, form	4600	3300	2300	2900	5400	2300	3300
WL, w1524							

T, boot	4200	3500	1900	3200	5700	2700	4200
T, form	4000	2900	700	1800	4400	1700	2700
T+S, boot	4200	3500	3100	3900	5600	2900	4100
T+S, form	4000	2900	2300	2700	4400	2000	2700
WL, w2544							
T, boot	4800	4000	2400	4100	7000	3400	5000
T, form	4600	3200	900	2400	5600	2200	3300
T+S, boot	4800	3900	3700	4700	6900	3600	4900
T+S, form	4600	3200	2700	3300	5600	2500	3300
WL, w4574							
T, boot	4000	3500	1700	3000	5300	2500	3900
T, form	3900	2900	600	1500	4100	1400	2600
T+S, boot	4100	3500	3700	4000	5300	2800	3800
T+S, form	3900	2900	3200	3200	4100	2100	2600
WZ, NL							
T, boot	21900	17900	12400	20700	31700	16900	21900
T, form	20700	14600	5500	13600	25600	11600	15000
T+S, boot	21700	17700	17300	21900	31300	16900	21600
T+S, form	20600	14600	13000	15500	25600	12000	15000
WZ, m1524							
T, boot	7400	6500	3300	5600	9900	4600	7400
T, form	7200	5300	1200	3100	7700	2800	4900
T+S, boot	7400	6400	5800	6500	9800	4900	7200
T+S, form	7200	5300	4800	4600	7700	3400	4900
WZ, m2544							
T, boot	7800	6600	4000	6700	11000	5500	7800
T, form	7500	5500	1600	4000	8700	3600	5200
T+S, boot	7800	6600	6000	7400	10900	5600	7700
T+S, form	7500	5500	4700	5100	8700	3900	5200
WZ, m4574							
T, boot	10900	9800	4500	7700	14200	6400	10900

T, form	10500	7900	1500	4100	10500	3700	7000
T+S, boot	10900	9700	8800	9100	14100	6800	10800
T+S, form	10400	7900	7700	6600	10600	4600	7000
WZ, w1524							
T, boot	7300	6200	3500	5900	9900	4900	7300
T, form	7000	5100	1300	3400	8000	3100	4800
T+S, boot	7200	6000	7400	7800	9800	5400	7000
T+S, form	7000	5100	6700	6400	8100	4200	4800
WZ, w2544							
T, boot	7800	6600	3700	6300	10800	5200	7900
T, form	7400	5300	1300	3600	8500	3200	5200
T+S, boot	7700	6500	7800	8100	10700	5700	7600
T+S, form	7400	5300	6900	6500	8600	4300	5200
WZ, w4574							
T, boot	10600	9700	3800	6500	12700	5500	9800
T, form	10300	8200	1000	2900	8600	2700	6400
T+S, boot	10500	9600	7400	7300	12700	5700	9700
T+S, form	10300	8200	6300	4300	8600	3200	6400
Total labour force, NL							
T, boot	21800	17900	11600	19500	31500	16000	22300
T, form	20600	14300	4800	12300	25400	10700	14800
T+S, boot	21500	17500	20100	23000	31000	16800	21700
T+S, form	20500	14300	17000	17800	25600	12500	14800
Total labour force, m1524							
T, boot	7000	6100	2900	4900	9000	4100	6800
T, form	6700	5100	1000	2600	6900	2400	4400
T+S, boot	6900	6000	5200	5900	8900	4400	6700
T+S, form	6700	5100	4100	4000	6900	3000	4400
Total labour force, m2544							
T, boot	6200	5300	2900	4900	8300	4000	6100
T, form	5900	4400	1100	2800	6600	2500	4100

T+S, boot	6200	5200	4900	5700	8300	4300	5900
T+S, form	5900	4400	4100	4200	6700	3000	4100
Total labour force, m4574							
T, boot	10100	9400	3500	6000	12000	5100	9500
T, form	9800	7800	1000	2700	8100	2600	6000
T+S, boot	10100	9300	7800	7600	12000	5600	9400
T+S, form	9800	7800	7000	5400	8200	3600	6000
Total labour force, w1524							
T, boot	6700	5800	2900	4900	8800	4100	6600
T, form	6400	4700	1000	2600	6800	2400	4300
T+S, boot	6700	5600	6900	7300	8800	4900	6400
T+S, form	6500	4700	6200	6000	7000	3800	4400
Total labour force, w2544							
T, boot	7000	6000	3200	5400	9500	4500	7000
T, form	6700	4900	1100	2900	7400	2700	4600
T+S, boot	7000	5800	7500	7600	9500	5200	6700
T+S, form	6800	4900	6900	6300	7600	4000	4600
Total labour force, w4574							
T, boot	9200	8500	2800	4900	9800	4100	7700
T, form	9100	7500	700	1900	6200	1800	5000
T+S, boot	9300	8400	8200	7300	9900	5000	7600
T+S, form	9100	7500	7600	5700	6400	3500	5000

Table 4.1: mean standard error over July 2015 – June 2016, for trend (T) and signal (T+S), computed by bootstrapping (boot) and by analytic formula (form)

5. Estimation of up-to-date standard errors

In this paper, standard errors are computed by bootstrapping for the series until June 2016. The computation time of the bootstrap procedure is around 4 days for one of the target variables. This is too long to implement the bootstrap procedure in the production process of official monthly figures for 21 series. Fortunately, the standard errors of different time periods are correlated. Therefore, it is possible to predict the standard errors for the near future. A simple and naïve approach is to assume that the standard errors of future months are equal to the available estimates of the past, for example the mean estimate of the last 12 months.

The figures in Section 4.2 show that the bootstrap estimates and the estimates based on the analytic formula develop more or less into the same direction. This means that the analytic formula can be used to predict the bootstrap estimates. A simple but synthetic approach is to correct the estimate based on the analytic formula with the difference or the ratio between bootstrap and analytic formula, assuming that this correction is constant over time.

A more sophisticated way is to combine the standard errors obtained with the bootstrap and the analytic formula in a time series model and predict the bootstrap standard errors for new observations using the standard errors estimated with the analytic formula. Two approaches can be distinguished. A univariate structural time series model for the series of the bootstrap standard errors can be developed, using the series of the standard errors observed with the analytic formula as an auxiliary series. The regression coefficient can be made time dependent. It is also possible to combine the series of standard errors obtained with the bootstrap and the analytic formula in a bivariate structural time series model and model the correlation between the disturbance terms of the underlying components, e.g. the trend. With both models the bootstrap standard error can be predicted with the analytic approximations.

A motivated choice between these options requires further research.

Nevertheless it is advisable to repeat the bootstrap procedure every once in a while. How often this is necessary, depends on the desired accuracy of the standard error estimates and the quality of the predictions described above. Note that the predictions are probably less accurate in case another redesign causes discontinuities again.

6. Conclusion

Statistics Netherlands publishes quarterly and yearly figures about the labour force which are derived from the monthly figures. These monthly estimates are based on a structural time series model. For the quarterly and yearly figures, no correct analytic formula for standard errors is available. Also for differences between time periods, no correct standard error formula is known. Therefore, parametric bootstrapping is tested in this paper for the estimation of standard errors for these figures. This bootstrap is a simplification of the parametric bootstrap proposed by Pfeffermann and Tiller (2005). The simplification implies that a bias correction on a standard bootstrap is left out, since this bias correction requires a correct analytic approximation of the standard error for each bootstrap replica. In further research a double bootstrap might be investigated to approximate this term. This option, however, will not be feasible in practice since it heavily increases the required computation time.

The bootstrap results are plausible for two reasons. First, in cases where a correct analytic formula is available, the bootstrap results are very similar to the analytic results. In that case, it is also possible to apply the bootstrap methods applied by Pfeffermann and Tiller (2005). Our results underestimate the standard error estimates based on this paper by around 1 – 1.5%.

The second reason why the results are plausible is because the bootstrap results are larger than an approximation of the standard errors by an analytic formula for which is known that it underestimates the true standard error. The true standard error is not known, therefore, it is not possible to decide whether the bootstrap results are truly accurate. Only a carefully designed simulation study can give the answer, which would, however, computationally almost impossible to carry out.

The computation time of the bootstrap procedure is quite large (around 4 days for one of the 21 target variables). Therefore, it is not possible to repeat the computation every month, at least not for all target variables. It should be possible to predict the standard error with a relatively simple method. The development of this method is left for the near future.

References

Bailar, B.A. (1975). The effects of rotation group bias on estimates from panel surveys. Journal of the American Statistical Association, 70, 23-30.

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

Brakel, J.A. van den and S. Krieg (2009). Estimation of the Monthly Unemployment Rate through Structural Time Series Modelling in Rotating Panel Design. Survey Methodology, 16, 2, 177-190.

Brakel, J.A. van den and S. Krieg (2010). Schatten van maandcijfers over de beroepsbevolking. CBS publicatie, 6 mei 2010, Heerlen. http://www.cbs.nl/NR/rdonlyres/008529E3-597E-497C-BAAC-27574253D803/0/ebbmaandcijfersschatten.pdf

Brakel, J.A. van den and S. Krieg (2012). Dealing with small sample sizes, rotation group bias and discontinuities in a rotating panel design. BPA nr. PPM-2012-11-13-JBRL. Centraal Bureau voor de Statistiek, Heerlen.

Brakel, J.A. van den and S. Krieg (2014). Modelleren van methodebreuken in maandcijfers over de beroepsbevolking. Centraal Bureau voor de Statistiek, Heerlen.

De Jong, P. and M.J. MacKinnon (1988). Covariances for smoothed estimates in State Space models. Biometrika, 75, 601-602.

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

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

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

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

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

Pfeffermann, D. and R. Tiller (2005). Bootsrap Approximation to prediction MSE for State-Space Models with Estimated Parameters. Journal of Time Series Analysis, 26, 893-916.

Rodríguez, A. and E. Ruiz (2012). Bootstrap prediction mean squared errors of unobserved states based on the Kalman filter with estimated parameters. Computational Statistics and Data Analysis, 56, 62-74.

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

Appendix A: details structural time series model

The trend L_t is modelled following the so called smooth trend model:

$$L_t = L_{t-1} + R_{t-1},$$

$$R_t = R_{t-1} + \eta_{R,t}. \qquad \text{(A.1)}$$

with R_t the slope. The disturbances $\eta_{R,t}$ are normally distributed with

$$E(\eta_{R,t}) = 0,$$

$$Cov(\eta_{R,t}, \eta_{R,t'}) = \begin{cases} \sigma_{R,i}^2 & \text{if} \quad t = t' \\ 0 & \text{if} \quad t \neq t' \end{cases}. \qquad \text{(A.2)}$$

The seasonal pattern is modelled using the trigonometric model:

$$S_t = \sum_{l=1}^{6} S_{t,l}, \qquad \text{(A.3)}$$

with

$$S_{t,l} = S_{t-1,l}\cos(h_l) + S_{t-1,l}^{*}\sin(h_l) + \eta_{S,t,l},$$

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

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

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

$$Cov(\eta_{S,t,l}, \eta_{S,,t',l'}) = Cov(\eta_{S,t,l}^{*}, \eta_{S,,t',l'}^{*}) = \begin{cases} \sigma_S^2 & \text{if} \quad t = t' \text{and } l = l' \\ 0 & \text{otherwise} \end{cases}$$

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

The irregular component ε_t contains the unexplained variation and is modelled as a white noise process:

$$E(\varepsilon_t) = 0,$$

$$Cov(\varepsilon_t, \varepsilon_{t'}) = \begin{cases} \sigma_\varepsilon^2 & \text{if} \quad t = t' \\ 0 & \text{if} \quad t \neq t' \end{cases}. \qquad \text{(A.4)}$$

The systematic differences, i.e. the RGB, between the subsequent waves are modelled in (3.5) with $\boldsymbol{\lambda}_t = (\lambda_{1,t}, \lambda_{2,t}, \lambda_{3,t}, \lambda_{4,t}, \lambda_{5,t})'$. The absolute bias in the monthly labour force figures cannot be estimated from the sample data only. Therefore additional restrictions for the elements of $\boldsymbol{\lambda}_t$ are required to identify the model. Here it is assumed that an unbiased estimate for θ_t is obtained with the first wave, i.e. $y_{1,t}$. This implies that the first component of $\boldsymbol{\lambda}_t$ equals zero. The other elements of $\boldsymbol{\lambda}_t$ measure the time dependent differences with respect to the first wave. To this end $\lambda_{j,t}$ are modelled as random walks for *j* = 2,3,4, and 5. As a result it follows that

$$\lambda_{1,t} = 0, \lambda_{j,t} = \lambda_{j,t-1} + \eta_{\lambda,j,t}, j = 2, 3, 4, 5$$

$$E(\eta_{\lambda,j,t}) = 0,$$

$$Cov(\eta_{\lambda,j,t}, \eta_{\lambda,j',t'}) = \begin{cases} \sigma_\lambda^2 & \text{if} \quad t = t' \text{and } j = j' \\ 0 & \text{otherwise} \end{cases}$$

The model for the survey errors takes the estimates for the variances of the GREG estimates $y_{1,t}, y_{2,t}, y_{3,t}, y_{4,t}, y_{5,t}$ into account. It also takes the correlation of the survey errors into account by modelling them as AR(1)-processes. For details see Van den Brakel and Krieg (2009).

Appendix B: More results

Figures for signal, unemployed labour force, national level

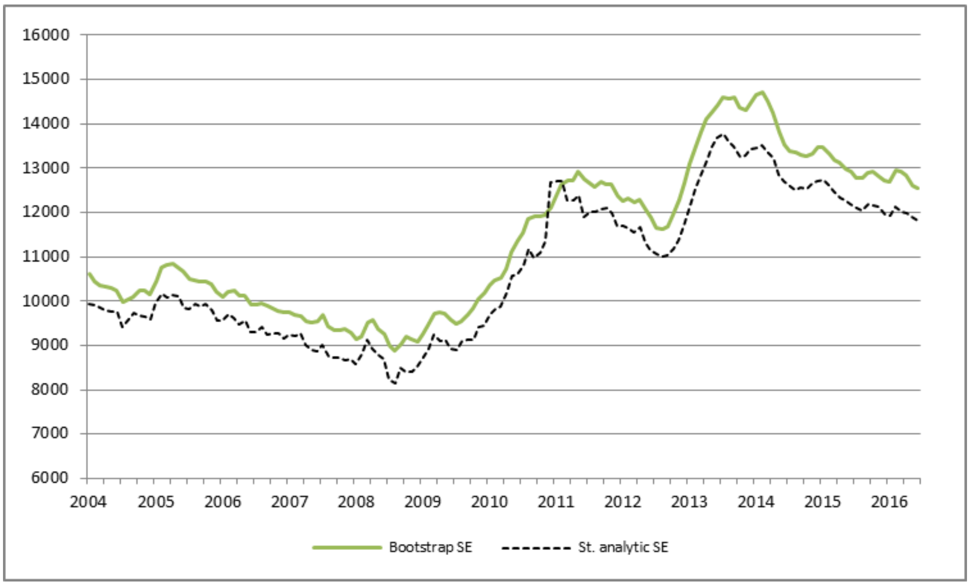

Figure B.1: Standard error estimates for quarterly signal, filtered, unemployed labour force, national level

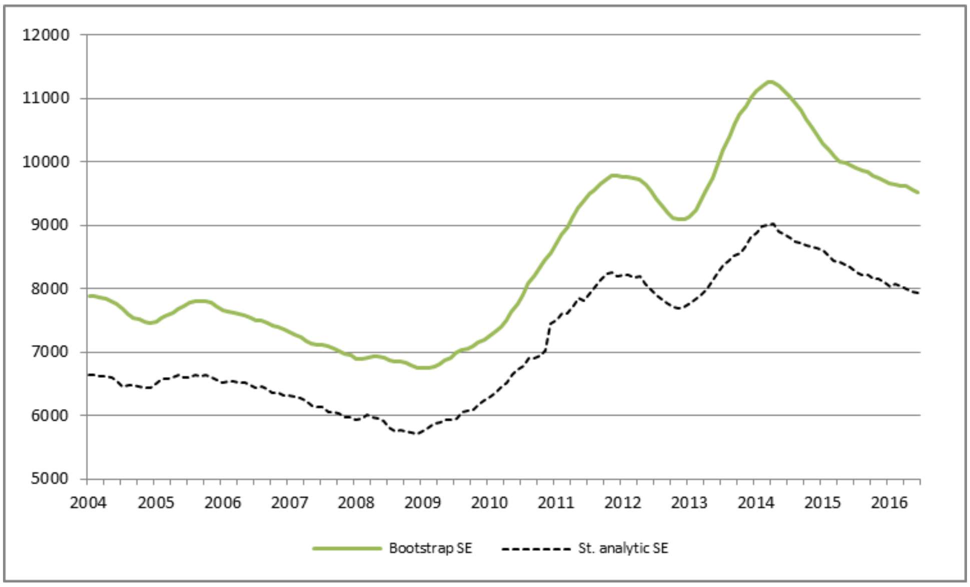

Figure B.2: Standard error estimates for yearly signal, filtered, unemployed labour force, national level

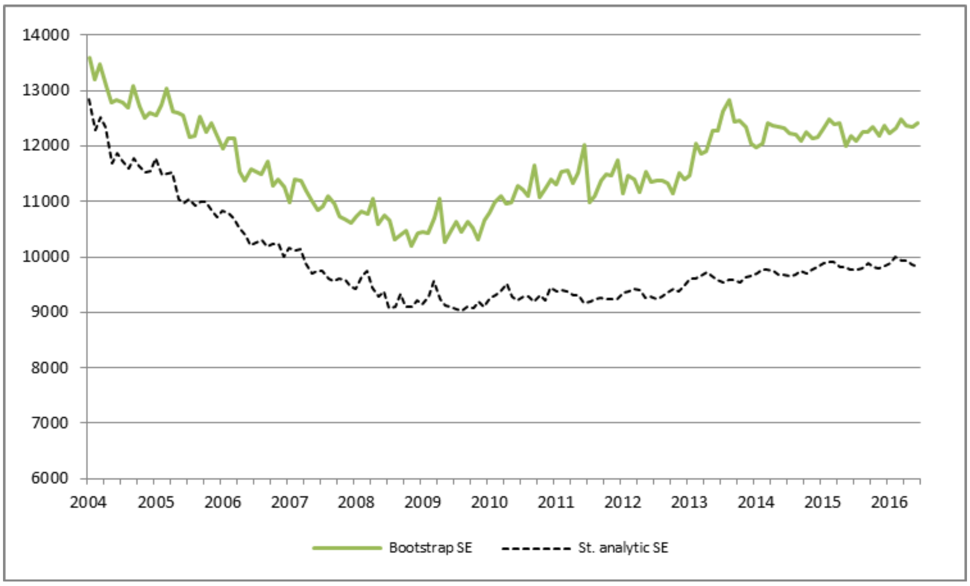

Figure B.3: Standard error estimates for the difference of the signal of month *t* and *t-1*, filtered, unemployed labour force, national level

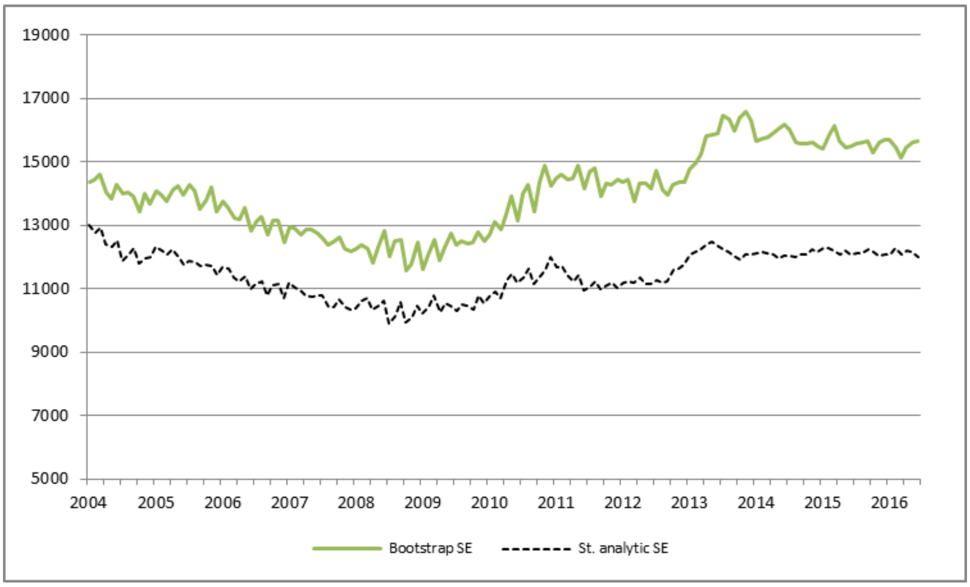

Figure B.4: Standard error estimates for the difference of the signal of month *t* and *t-3*, filtered, unemployed labour force, national level

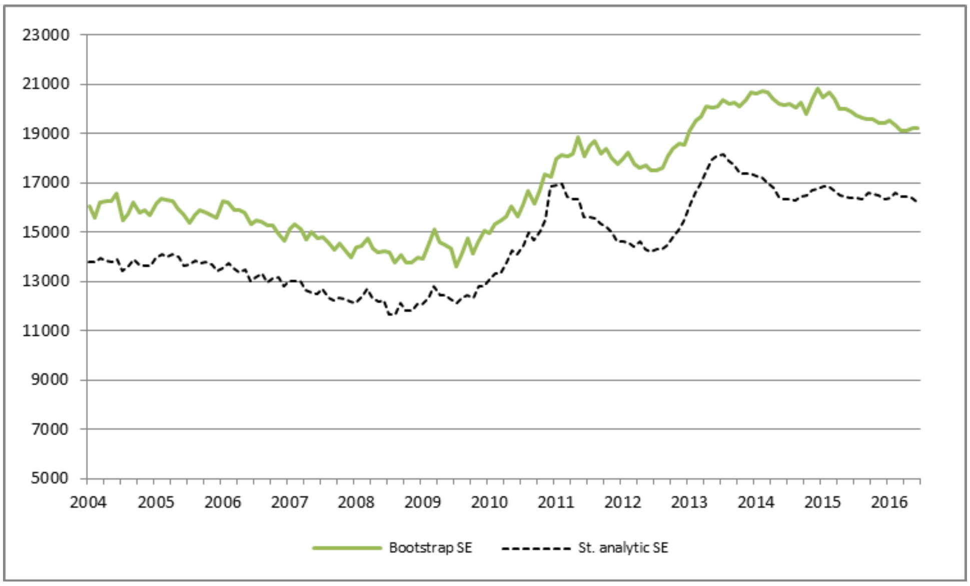

Figure B.5: Standard error estimates for the difference of the signal of month *t* and *t-12*, filtered, unemployed labour force, national level

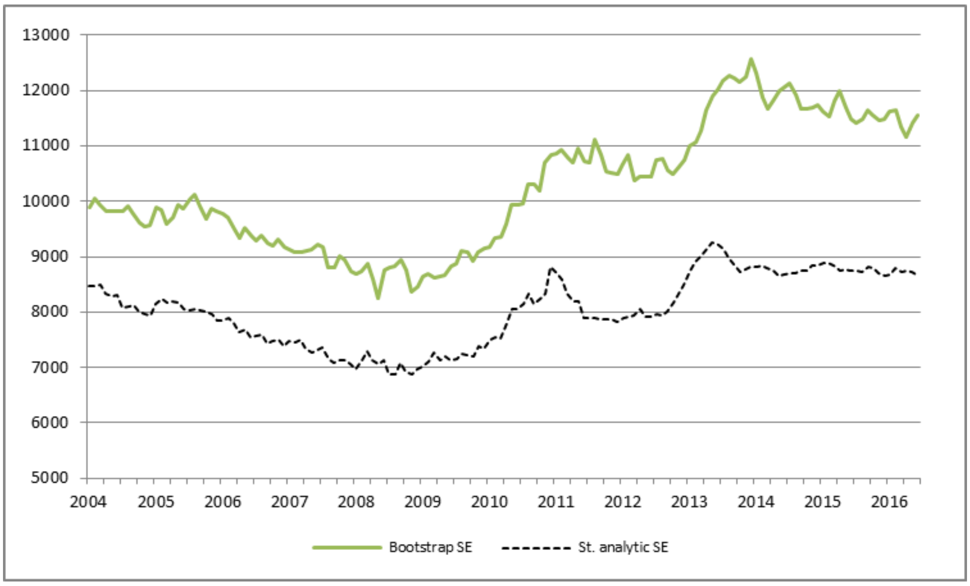

Figure B.6: Standard error estimates for the difference of the signal of quarter *q* and *q-1*, filtered, unemployed labour force, national level

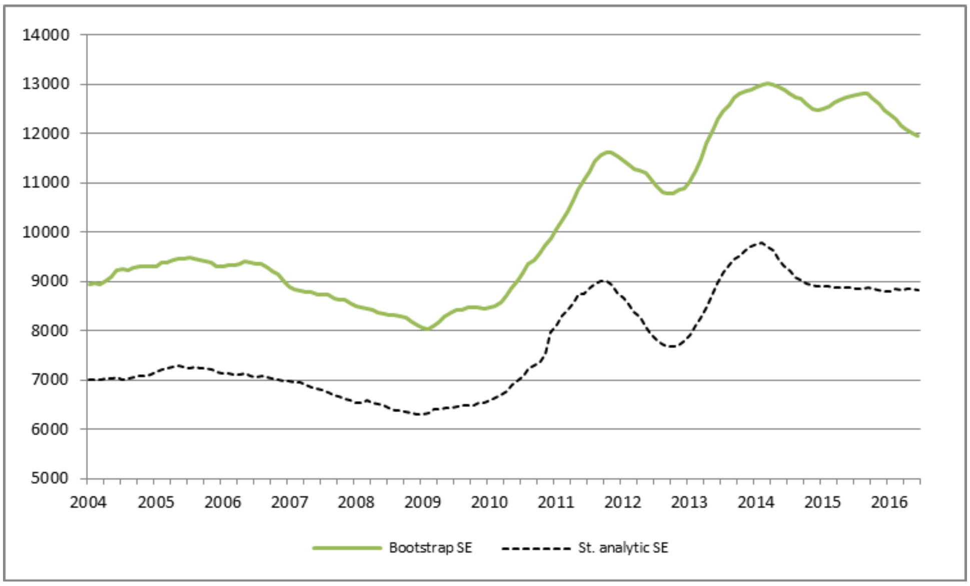

Figure B.7: Standard error estimates for the difference of the signal of year *a* and *a-1*, filtered, unemployed labour force, national level

Explanation of symbols

Empty cell	Figure not applicable
.	Figure is unknown, insufficiently reliable or confidential
*	Provisional figure
**	Revised provisional figure
2015–2016	2015 to 2016 inclusive
2015/2016	Average for 2015 to 2016 inclusive
2015/’16	Crop year, financial year, school year, etc., beginning in 2015 and ending in 2016
2013/’14–2015/’16	Crop year, financial year, etc., 2013/’14 to 2015/’16 inclusive

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

Colofon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
Statistics Netherlands, Studio BCO

Design
Edenspiekermann

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