Labour Force Survey Data on Unemployment: Identifying Outliers

Discussion paper 05009

Reinier Bikker, Remko Hijman, Peter Kee
and George van Leeuwen

The views expressed in this paper are those of the authors
and do not necessarily reflect the policies of Statistics Netherlands

Statistics Netherlands

Voorburg/Heerlen, November 2005

Explanation of symbols

.	= data not available
*	= provisional figure
x	= publication prohibited (confidential figure)
–	= nil or less than half of unit concerned
–	= (between two figures) inclusive
0 (0,0)	= less than half of unit concerned
blank	= not applicable
2003–2004	= 2003 to 2004 inclusive
2003/2004	= average of 2003 up to and including 2004
2003/'04	= crop year, financial year, school year etc. beginning in 2003 and ending in 2004

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

Publisher
Statistics Netherlands
Prinses Beatrixlaan 428
2273 XZ Voorburg
The Netherlands

Printed by
Statistics Netherlands - Facility Services

Cover design
WAT ontwerpers, Utrecht

Information
E-mail: infoservice@cbs.nl

Where to order
E-mail: verkoop@cbs.nl

Internet
http://www.cbs.nl

Key figure: X-10
ISSN: 1572-0314
Production code: 6008305009

Statistics Netherlands

LABOUR FORCE SURVEY DATA ON UNEMPLOYMENT: IDENTIFYING OUTLIERS

Summary: This paper relates unemployment outcomes from the Dutch labour force survey to the number of individuals registered at Dutch employment offices. Both series evolve rather similar, despite the substantial differences in underlying populations. The empirical relationship is quantified with a state space model. This model is subsequently used to generate forecasts for the unemployed labour force. By comparing the forecasted values with survey outcomes, outliers can be identified.

JEL classification: C32, C53.

Keywords: time-series models, forecasting.

Acknowledgements: We are grateful to Paul Knottnerus and Kees Zeelenberg for providing very valuable comments on earlier drafts.

1. Introduction

Each month Statistics Netherlands reports the average number of persons who were part of the unemployed labour force (*ULF*) during the previous three months. The overall unemployment figures are decomposed according to sex and three age categories. Statistics Netherlands defines *ULF* as the number of persons between 15 and 65 years of age who are not employed or who are employed for less than 12 hours a week, who want employment for at least 12 hours a week, who can start at short notice, and who display active search effort. Especially this last criterion implies that unemployment can be measured only by a survey among individuals. Statistics Netherlands uses its labour force survey (LFS) for this purpose.

The 95% confidence intervals for the LFS-results can be assessed quite accurately provided that these outcomes are not distorted by the presence of extreme values. As is well known, these outliers may lead to erroneous conclusions about the development of unemployment. It is therefore important to identify outliers in the measurement of ULF and to accommodate them when appropriate. Outlier detection may be based on the comparison of unemployment changes as measured by the LFS with changes in related indicators, like the number of persons without work who are registered at Dutch employment offices (CWI; Centres for Work and Income) or the number of unemployment benefits (*WW*).

Since August 2004, Statistics Netherlands employs a state space model to assess the plausibility of the data on *ULF* before publication. The model specifies the empirical relationship between the LFS-data and the data from the CWI, and generates forecasts for the most recent LFS-figures on unemployment. Outliers are then identified by comparing the model predictions with actual observations. We define outliers as LFS-outcomes which fall outside the 95% confidence intervals of the forecasted values.

This paper describes the methodology of the outlier detection applied by Statistics Netherlands. Its outline is as follows. Section 2 describes the CWI-data on registered individuals without work. The relationship between these data and the outcomes from the LFS is discussed in section 3. Section 4 presents the methodology of identifying outliers in the survey data. Finally, section 5 deals with the way in which detected outliers are handled and section 6 concludes.

2. The CWI registration data

Each month Statistics Netherlands receives a file from the CWI with data on registered individuals without work (*RWW*) during the last month. Those classified as *RWW* are registered at CWI, engaged in job search, and do not work or work for less than twelve hours a week. Availability for work of more than twelve hours a week does not play a role in defining *RWW*.

Due to administrative delays it is unlikely that the registration data for *RWW* are always up-to-date. Especially finding work by registered individuals seems to be reported not always in time. If the delay in deregistration on average is higher than the delay in registration, *RWW* lags behind reality more during favourable market conditions (more deregistration than registration).

The data on *ULF* and *RWW* only partly relate to the same individuals. On average nearly 50% of the persons included in the *ULF*-population in 2003 were also included in the *RWW*-population. As a percentage of the *RWW*-population, this overlap was only 32% on average. Among the individuals who belong to *ULF* but not to *RWW* are many persons looking for temporary work, like school children and students searching for a holiday job. Also many individuals who are not eligible for a social allowance, for example school-leavers and women who want to return to work, belong to this category. Therefore, the yearly pattern of both series differs most clearly in summer, when young people usually enter the labour market. Among the individuals who belong to *RWW* but not to *ULF* are those who are not available for employment of twelve hours a week or more. Also included in this non-overlapping category are individuals who do not search actively, for example because they are discouraged.

3. The relationship between the series

December 2004, Statistics Netherlands reported an average of 470,200 individuals for *ULF* during September – November 2004. *RWW* consisted of 696,300 individuals in the month of November 2004. *RWW* lies at a structurally higher level than *ULF*, as can be seen from the graphical presentation in Figure 1. The evolution of the two time series is closely related. We observe in the graph that the difference is rather constant over time. The increase in *ULF* since the spring of 2001 and its decrease mid 2004 are reflected in *RWW* almost at the same time.

The graph also shows the monthly counts of the number of individuals claiming unemployment benefit. This *WW*-series evolves at a structurally lower level (310,400 at the end of September 2004) than the two other series. It provides practically no information in addition to that already provided by the *RWW*-series (the correlation coefficient between *WW* and *RWW* equals 0.98). The reason is the overlap between both populations: benefit receivers are obliged to register at the CWI. Because of this overlap and because *WW*-data become available several months after the *RWW*-data, they are not included in our analysis.

The relationship between *ULF* and *RWW* is plotted in Figure 2. The graph indicates that there are in fact two linear relationships: one for the months January 1998 – December 1999, and another with a steeper slope for the months January 2000 – October 2004. What lies behind this picture is not exactly clear. External factors, like the transformation of the old-style employment office into the CWI in 2001 or the slowing of the economy since 2000, could play a role. A different observation

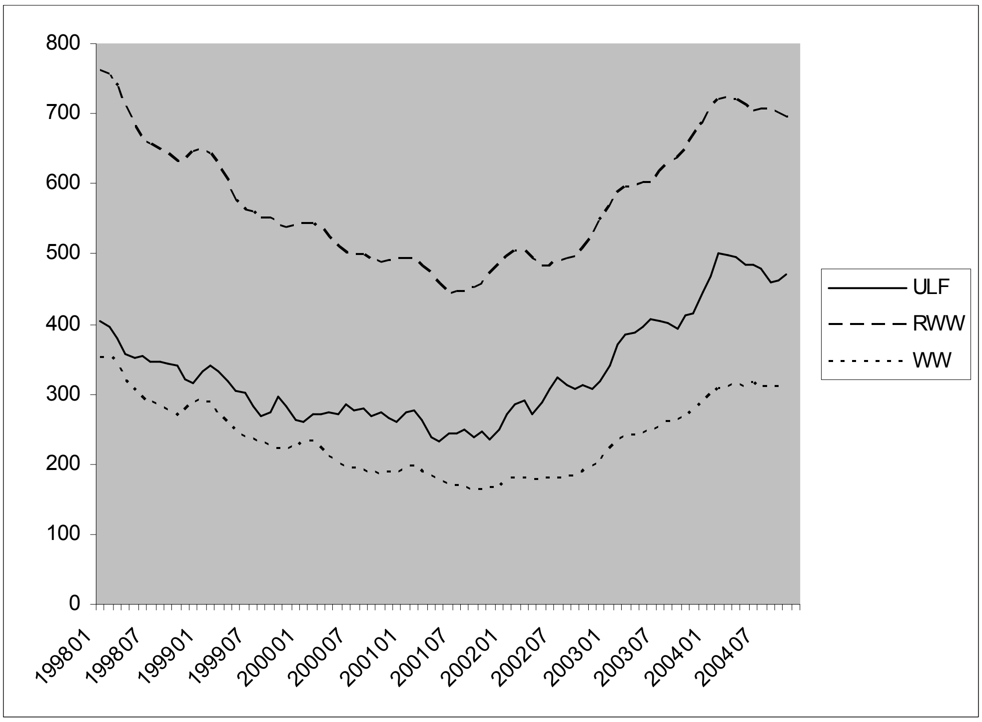

Figure 1. Time series ULF, RWW and WW.

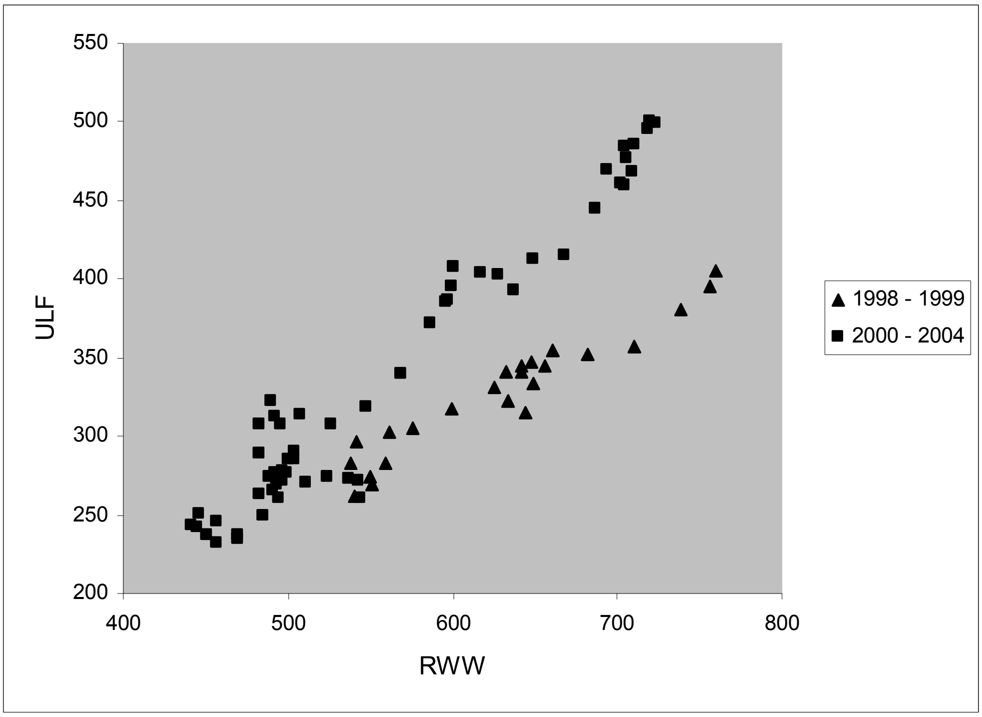

Figure 2. Scatter plot ULF- and RWW-levels.

methodology applied by Statistics Netherlands since 2000 might be responsible as well. Additional research on this issue is necessary to identify the underlying causes.

4. Outlier detection

Outliers are identified by comparing the outcomes for *ULF* from the LFS with forecasts generated by a time-series model. This model describes the empirical relationship between *ULF* and *RWW*. For the one-step-ahead prediction of the three-monthly average ULF_T (the average for the months $T-1$, T and $T+1$) we use observations of *ULF* and *RWW* available from 1998, up to and inclusive ULF_{T-1} and RWW_{T+1}. Forecasts for the total unemployed labour force are obtained by aggregating the predictions for the six subgroups about which Statistics Netherlands reports: men and women in the age categories 15–24, 25–44 and 45–64. Finally, the mutually consistent out-of-sample forecasts for the aggregated number and the disaggregated numbers are compared with realised values.

4.1 Forecasting with an error correction model

We first use a cointegration model because standard tests show that the time series are non-stationary and cointegrated. The employed error correction model (see e.g. Greene 2000, pp. 733-735) consists of two successive and independent OLS-regressions: a long term equilibrium relationship in levels between *ULF* and *RWW* (the cointegration equation) and a short term relationship in first differences (describing fluctuations around the steady state).

The regression results (not presented) for the long term relationship confirm a significantly different slope since 2000 (see Figure 2). Furthermore, the constant term in the equation is significantly higher since the change in labour market conditions in April 2001. This is probably caused by the asymmetric delays in deregistration and registration at the CWI (see section 2).

Many outliers for the various subgroups are identified when confronting the one-month-ahead forecasts generated by the error correction model for the months January 2002 – October 2004 with the estimates from the LFS. In almost all these cases, however, the designation as an outlier is implausible in view of the historical development of the relevant series. Further investigation reveals that forecasted values produced by this model are very sensitive to the length of the estimation period. With the available time series from 1998, the estimated model coefficients are unstable. We conclude that for the time being, the error correction model cannot be employed for the identification of outliers in the Dutch LFS.

4.2 Forecasting with a state space model

This section presents an alternative approach to deliver the required forecasts for the LFS-figures: the state space methodology (see e.g. Harvey 1989; Durbin and Koopman 2001). In state space time-series models the observations are regarded as made up of distinct components (states) which cannot be observed directly and which are modelled explicitly. Information about the most recent states is generated by a recursive process, which is known as the Kalman filter. The recursions enable us to update our knowledge of the states each time a new observation comes in. The most important advantage of state space models is their flexibility.

4.2.1 Model specification

A state space model is made up of two (systems of) equations: (1) a measurement equation, which relates the states to the observations and (2) a transition equation, which describes the development of the states over time. The general model may be written in the form

$$y_t = Z\alpha_t + G\varepsilon_t, \tag{1}$$

$$\alpha_t = J\alpha_{t-1} + H\eta_t, \tag{2}$$

where y_t is a vector of observations which are available for month *t*, with $t = 1, \ldots, T$, and α_t an unobserved vector called the state vector for the same month. The matrices *Z*, *J*, *G* and *H* are usually referred to as system matrices. *Z* describes the relation between the state variables and the observations and *J* describes the transition of states for month *t - 1* to states for month *t*. *Z* and *J* do not change over time in most applications. The system matrices *G* and *H* usually consist of zeros and ones. Finally, ε_t and η_t are vectors of disturbances, which are serially uncorrelated and uncorrelated with each other in all time periods.

We use the state space methodology for the out-of-sample one-step-ahead prediction of ULF_t on the basis of its past realizations and the available *RWW*-series. When predicting ULF_t figures for *RWW* are available up to $t+1$. Shifting back the *RWW*-figures one month gives the vector of observations

$$y_t \equiv \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} FRWW_t \\ ULF_t \end{bmatrix}, \text{ with } FRWW_t = RWW_{t+1}. \tag{3}$$

The starting point when specifying our model is the idea that the observations on unemployment are made up of three components: trend, seasonal and error term. We extract these components from the monthly *RWW*-figures. This series based on monthly figures is better suited for this purpose than the series of *ULF*, which consists of three-monthly means. Then $FRWW_t$ may be written as

$$FRWW_t = Trend_t + S_{1t} + \varepsilon_{1t}, \tag{4}$$

where $Trend_t$ is a linear trend, S_{1t} is a seasonal and ε_{1t} denotes a normally distributed, serially independent, random disturbance term with mean zero and variance σ_1^2, that is $\varepsilon_{1t} \sim NID(0, \sigma_1^2)$.

Next, the trend and seasonal components are modelled. For the first component a smoothed trend model seems most appropriate. This means that we shall assume that the trend follows a local level with stochastic drift

$$Trend_t = Trend_{t-1} + \beta_{t-1} \tag{5a}$$

$$\beta_t = \beta_{t-1} + \eta_{\beta t}. \tag{5b}$$

In (5a) the trend level follows an Integrated Random Walk (IRW) in which the level in the previous month is augmented with a stochastic slope component (the change in trend) β. The stochastic character of the trend is represented by $\eta_{\beta t}$, with $\eta_{\beta t} \sim NID(0, \sigma_\beta^2)$. We assume that the seasonal pattern in the *FRWW*-series is stochastically as well. Put differently, we assume that within a year the sum of the seasonal components has expectation zero. Hence we write

$$S_{1t} + S_{1,t-1} + \cdots + S_{1,t-11} = \eta_{S1t}, \tag{6}$$

with $\eta_{S1t} \sim NID(0, \sigma_{S1}^2)$.

Given data availability up to and including month $t-1$, applying Kalman filter recursions to equations (4)–(6) would generate the filtered state $\alpha_{t|t-1}$ for each month t ($t \in \{1, \cdots, T\}$) together with the corresponding filtered signal $\hat{y}_{1t|t-1} = Z_1 \alpha_{t|t-1}$, with Z_1 the first row of matrix Z in (1). Since the main objective of the analysis, however, is forecasting the most recent value for *ULF*, the model has to be extended with an equation linking *FRWW* and *ULF*. Section 3 showed a close relationship between the two series which changes only gradually over time (see Figure 1). As was pointed out before, *ULF* and *FRWW* relate to different populations. The seasonal pattern of both time series varies as well. Then, modelling the population difference between *ULF* and *FRWW* as a random walk, the observation equation for *ULF* can be written in the form

$$ULF_t = Trend_t + \delta_t + S_{2t} + \varepsilon_{2t}, \tag{7a}$$

where $Trend_t$ is the trend specified in (5a), ε_{2t} is a disturbance term, with ε_{2t} ~ $NID(0, \sigma_2^2)$, and where the stochastic difference between *ULF* and *FRWW* is

$$\delta_t = \delta_{t-1} + \eta_{\delta t}, \tag{7b}$$

with $\eta_{\delta t} \sim NID(0, \sigma_\delta^2)$, and where the stochastic seasonal pattern in the *ULF*-series is

$$S_{2t} + S_{2,t-1} + \cdots + S_{2,t-11} = \eta_{S2t}, \tag{7c}$$

with $\eta_{S2t} \sim NID(0, \sigma_{S2}^2)$.

The addition of the equations for the state variables δ_t and S_{2t} completes our model. Finally, it should be noted that we took account for a possible correlation of the disturbances of the two measurement equations in the model. This has been achieved by including the covariance between ε_{1t} and ε_{2t} (σ_{12}) as an additional parameter. A full specification in matrix format is provided in the Appendix.

4.2.2 Regression results

The output of the Kalman filter consists of estimates for the state variables $Trend_t$, β_t, $S_{1t}, \ldots, S_{1,t-10}$, δ_t and $S_{2t}, \ldots, S_{2,t-10}$. Table 1 presents the estimation results for the final states, that is for the most recent state variables. The negative slope component for the *FRWW*-trend indicates that the number of persons registered at the CWI is decreasing, though β_T is not statistically significant at the 5% level. Also notice, that the seasonal pattern in the two series differs substantially.

The fact that the filtering process will produce valuable insights into the development of the various components of the observations is an important advantage. This knowledge can be used in assessing the plausibility of new observations, but also to provide an analytical picture when presenting new unemployment numbers. This is illustrated by the figures below. Figure 3 shows the

Table 1. Estimation results for final states, $\alpha_{T|T-1}$, for T = October 2004.

State variable	Estimate	*t*-value[a]
$Trend_T$	700,552	176.27
Slope component trend, β_T	–5,663	–1.66
Difference *ULF* and trend, δ_T	–240,382	–28.54
S_{1t}:		
Seasonal *FRWW* month T	–1,958	–1.20
Seasonal *FRWW* month T-1	–11,015	–6.76
Seasonal *FRWW* month T-2	–11,189	–7.32
Seasonal *FRWW* month T-3	–11,970	–7.86
Seasonal *FRWW* month T-4	–12,251	–8.03
Seasonal *FRWW* month T-5	–15,869	–10.38
Seasonal *FRWW* month T-6	–7,149	–4.68
Seasonal *FRWW* month T-7	5,582	3.66
Seasonal *FRWW* month T-8	17,462	11.34
Seasonal *FRWW* month T-9	21,731	13.92
Seasonal *FRWW* month T-10	18,288	11.75
S_{2t}:		
Seasonal *ULF* month T	–2,935	–0.65
Seasonal *ULF* month T-1	–8,595	–2.07
Seasonal *ULF* month T-2	–5,258	–1.35
Seasonal *ULF* month T-3	4,299	1.11
Seasonal *ULF* month T-4	5,540	1.43
Seasonal *ULF* month T-5	–1,273	–0.33
Seasonal *ULF* month T-6	1,467	0.38
Seasonal *ULF* month T-7	12,080	3.11
Seasonal *ULF* month T-8	16,999	4.35
Seasonal *ULF* month T-9	1,326	0.34
Seasonal *ULF* month T-10	–8,496	–2.16
Variance ε_{1t}, σ_1^2	2,231	2.17
Variance ε_{2t}, σ_2^2	9,407	7.89
Covariance ε_{1t} and ε_{2t}, σ_{12}	–0.0	–0.50
Variance $\eta_{\beta t}$, σ_β^2	4,441	4.03
Variance $\eta_{\delta t}$, σ_δ^2	49,371	42.02
Variance η_{S1t}, σ_{S1}^2	126	0.12
Variance η_{S2t}, σ_{S2}^2	2,500	2.34
Log Likelihood	–725.67	

[a] Estimate divided by standard deviation.

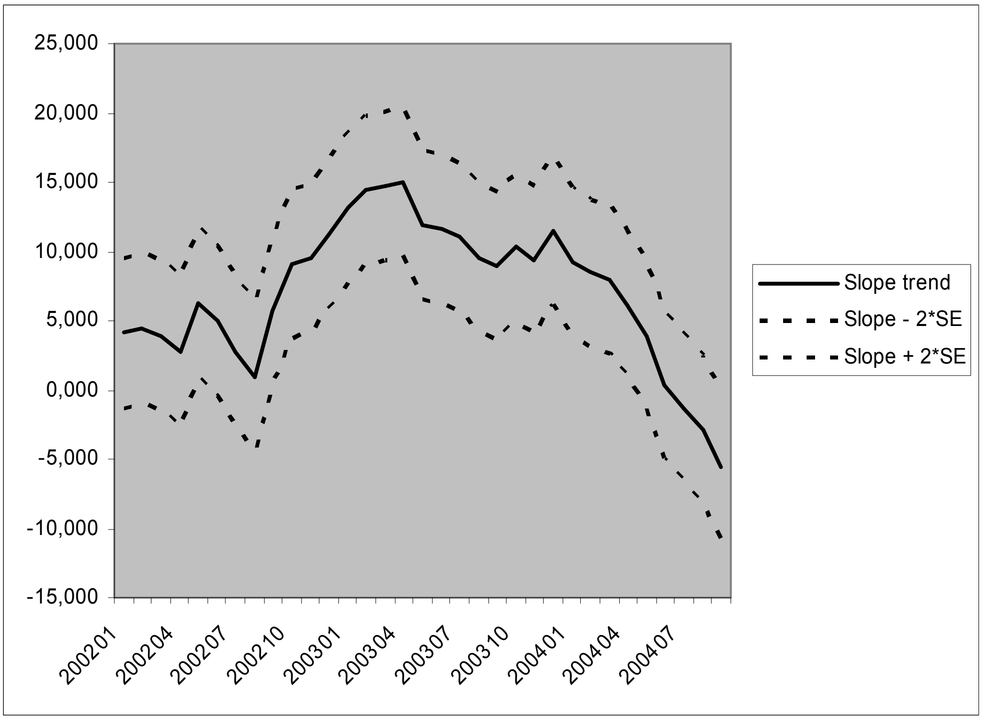

Figure 3. Filtered state of the slope component in FRWW, $\beta_{t|t-1}$.

filtered states of the monthly slope term β_t during January 2002 – October 2004. The dotted lines constitute the 95% confidence interval. We see that the growth rate of the trend after a period of stabilization during the second half of 2003 declines in more recent months. The extent to which these developments are reflected in *ULF* depends on the magnitude of the stochastic differences between *ULF* and *FRWW*. Figure 4 displays the filtered states for the non-seasonal difference δ_t, conditional on the trend level estimates for *FRWW* and the estimated seasonal pattern in *ULF*. The graph shows that this difference followed an inverted U-shaped pattern during 2004. Finally, Figures 5 and 6 show the filtered states of the seasonal component in *FRWW* and *ULF*. There is a clear difference in seasonal patterns. The seasonal component in *FRWW* peaks around the turn of the year. The peak for *ULF* occurs during the months February and March and hence looks like an echo-effect of the peak in the other series. Finally, notice that the seasonal pattern in *FRWW* is much smoother than the one in *ULF*.

4.2.3 Simulations

Our use of the state space model can be illustrated as follows. Before publishing the most recent three-monthly average for *ULF*, we test whether this is an outlier. This is done by first making a forecast for ULF_T (the average for the months $T-1$, T and $T+1$) with the available information. *RWW* is then available for month $T+1$. Shifting back the figures for *RWW* one month gives the series y_{1T}. The most recent definitive figure for *ULF* is ULF_{T-1}. We start with predicting $\alpha_{T|T-1}$: the state-

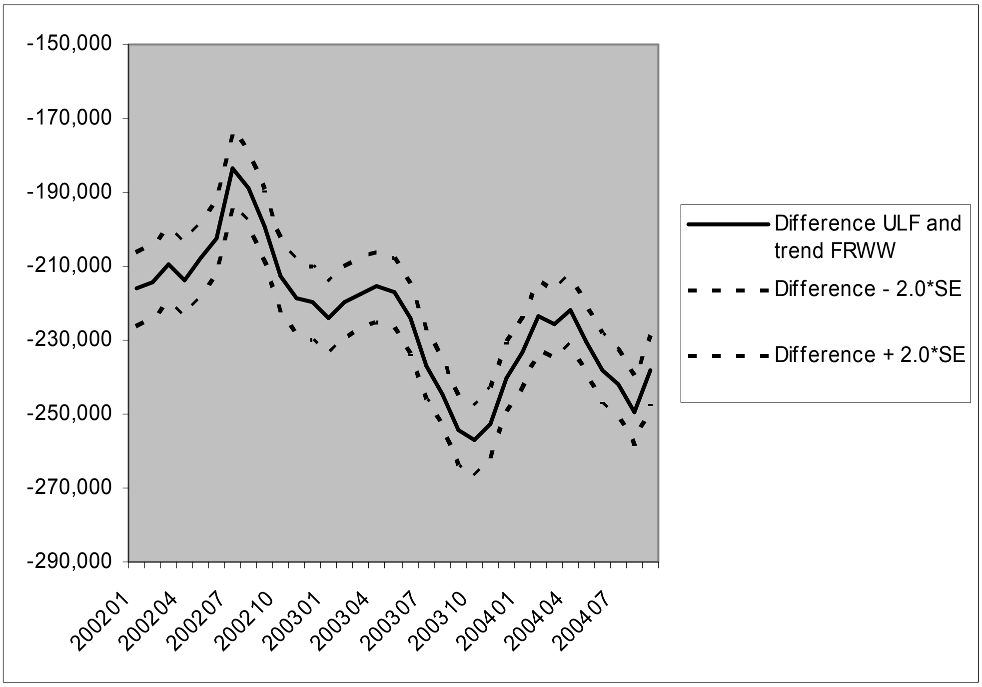

Figure 4. Filtered state of the difference between ULF and trend FRWW, $\delta_{t|t-1}$.

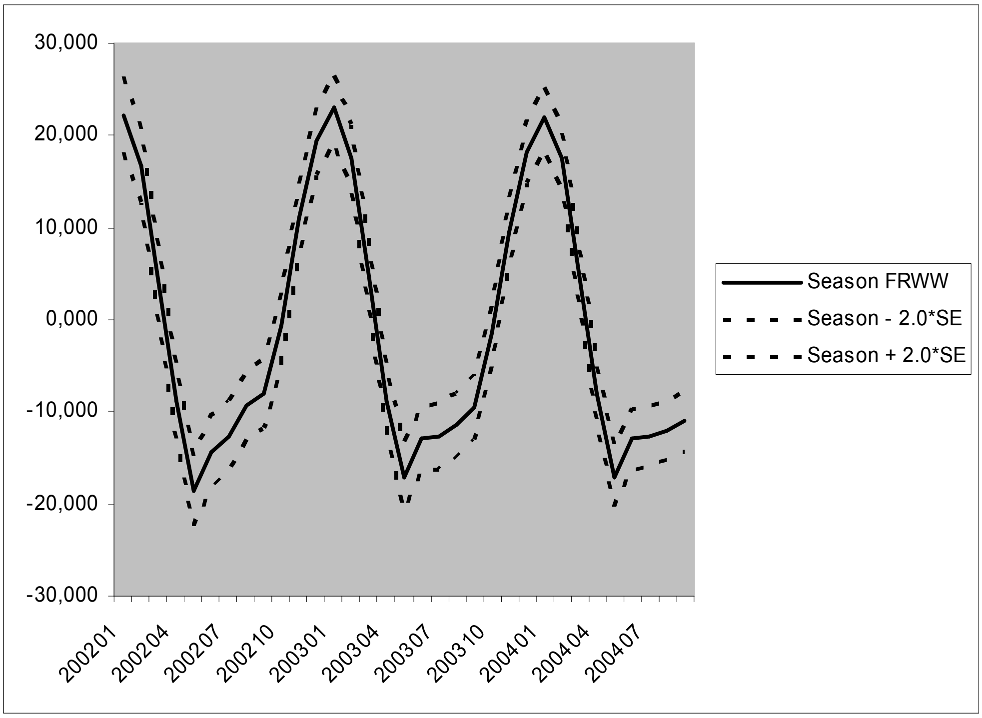

Figure 5. Filtered state of the seasonal component in FRWW, $S_{1t|t-1}$.

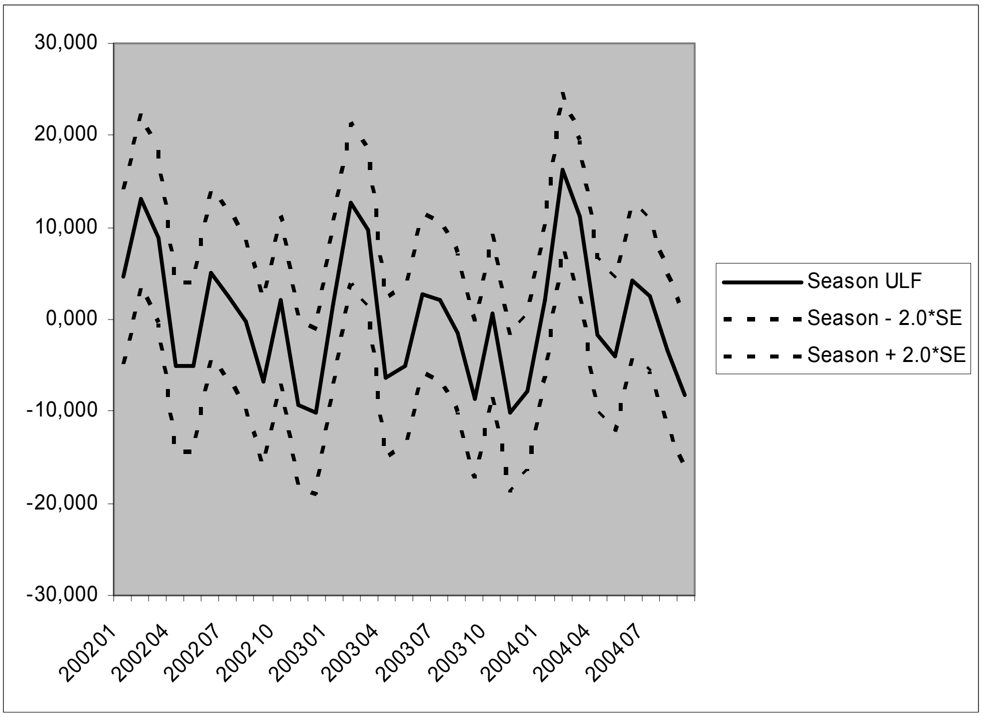

Figure 6. Filtered state of the seasonal component in ULF, $S_{2t|t-1}$.

vector for month T , given the filtered states up to $T-1$. This vector is made up of the forecasts for the components $Trend_{T|T-1}$, $\beta_{T|T-1}$, $S_{1T|T-1}$, $\delta_{T|T-1}$ and $S_{2T|T-1}$. From $\alpha_{T|T-1}$ we obtain $\hat{y}_{2T|T-1}$: the forecast for ULF_T , conditional on the data up to and including $T-1$. Secondly, this one-step-ahead forecast is compared with ULF_T . The survey result for month T is identified as an outlier if it falls outside the 95% confidence interval of the forecasted value.

The state space approach designates only a few outliers in the observations for the subgroups: five for the months January 2002 – October 2004. These observations are located reasonably close to the 95% confidence boundary of the forecasts. Though credible, the outliers therefore cannot be classified as “strong”. Two observations are assigned as outliers in the year 2004, both showing up in the average over the first three months. As can be seen from Table 2, these observations relate to unemployed women in the age categories 15–24 and 45–64. These outliers may be classified as “mild” as they lie just outside the 95% confidence interval of the forecasts.

Figure 7 presents graphically the results for the total unemployed labour force during January 2003 – October 2004. The dotted lines constitute the 95% confidence interval of the forecasts, which are obtained by aggregating the predicted values for the unemployed subpopulations. Only two outliers are detected in the observations on aggregate unemployment: in July 2003 and in December 2003. The realizations which appeared to be turning points, that is *ULF* in 2003:10, 2004:03 and 2004:09,

Table 2. Detecting outliers in the observations for ULF: January-March 2004.

Sex	Age	ULF	Forecast	Lower boundary	Upper boundary	Standardized deviation.[a]
		× 1000				
M+F	15–64	500.2	484.4[b]	465.5[c]	503.4[c]	1.7
M	15–24	64.4	60.6	54.4	66.9	1.2
M	25–44	127.3	125.3	114.0	136.5	0.4
M	45–64	70.1	69.5	64.1	74.9	0.2
F	**15–24**	**57.9**	**48.4**	**40.0**	**56.9**	**2.2**
F	25–44	123.3	128.8	120.5	137.0	-1.3
F	**45–64**	**57.1**	**51.8**	**46.8**	**56.8**	**2.1**

[a]) Difference observation and forecast divided by the standard deviation of the forecast.

[b]) Sum of the forecasts for the strata of ULF.

[c]) Interval equals the square root of the sum of the quadratic intervals for forecasts strata ULF.

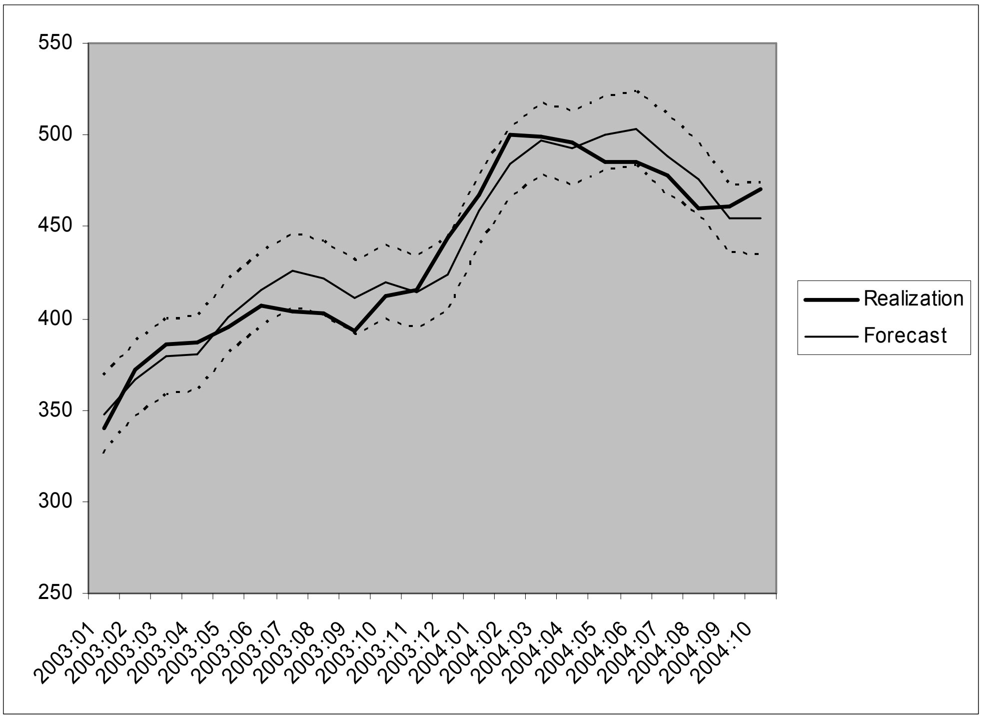

Figure 7. Confronting the forecasts for ULF with observed values.

are within the confidence interval. This means that turning points are handled quite well by the model. This is probably due to the fact that the trend in our model accounts for the most recent monthly figure for *RWW*. A trend based on the three-monthly means of *ULF* would attach less weight to this most recent month.

5. Outlier treatment

Outliers may be due to errors in processing the data. The separate search for outliers within the various subgroups facilitates the efficient tracing of this kind of mistakes. If outliers cannot be attributed to processing errors, further investigation has to determine whether they are due to recent developments which had an effect on *ULF* but not (yet) on *RWW*, or vice versa. To avoid explanations which are not convincing, an assessment of likely developments in the *ULF*-series has to be made before the survey results become available. If plausible explanations are absent, outliers are adjusted.

The way of accommodating outliers is as follows. In the case of 95% confidence intervals of realised and forecasted values that overlap, survey outcomes are adjusted up to the confidence boundary of the predictions. In the case of non-overlapping margins, survey outcomes are adjusted up to their own confidence boundaries.

6. Conclusion

The possibility of outliers in the survey results calls for an objective and transparent method to judge the outcomes on unemployment before they are published. We check these figures by confronting them with forecasted values, which are derived from the empirical relationship with register data from the CWI. By employing state space modelling for this purpose, outliers in the survey outcomes are identified in a thorough, advanced, and flexible way. In addition, this methodology provides a quantitative underpinning for the adjustment of outliers.

Appendix

The specification of equations (1) and (2)

I) Measurement equations

$$
\begin{bmatrix} FRWW_t \\ ULF_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & . & . & 0 & 0 & 0 & 0 & . & . & 0 \\ 1 & 0 & 0 & 0 & . & . & 0 & 1 & 1 & 0 & . & . & 0 \end{bmatrix} \begin{bmatrix} Trend_t \\ \beta_t \\ S_{1t} \\ S_{1,t-1} \\ . \\ . \\ S_{1,t-10} \\ \delta_t \\ S_{2t} \\ S_{2,t-1} \\ . \\ . \\ S_{2,t-10} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}
$$

II) Transition equations

$$
\begin{bmatrix} Trend_t \\ \beta_t \\ S_{1t} \\ S_{1,t-1} \\ S_{1,t-2} \\ . \\ S_{1,t-10} \\ \delta_t \\ S_{2t} \\ S_{2,t-1} \\ . \\ . \\ S_{2,t-10} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & . & . & 0 & 0 & 0 & 0 & . & . & 0 \\ 0 & 1 & 0 & 0 & . & . & 0 & 0 & 0 & 0 & . & . & 0 \\ 0 & 0 & -1 & -1 & . & . & -1 & 0 & 0 & 0 & . & . & 0 \\ 0 & 0 & 1 & 0 & . & . & 0 & 0 & 0 & 0 & . & . & 0 \\ 0 & 0 & 0 & 1 & 0 & . & 0 & 0 & 0 & 0 & . & . & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & . & 0 & 0 & 0 & . & . & 0 \\ 0 & 0 & 0 & 0 & . & 1 & 0 & 0 & 0 & 0 & . & . & 0 \\ 0 & 0 & 0 & 0 & . & . & 0 & 1 & 0 & 0 & . & . & 0 \\ 0 & 0 & 0 & 0 & . & . & 0 & 0 & -1 & -1 & . & . & -1 \\ 0 & 0 & 0 & 0 & . & . & 0 & 0 & 1 & 0 & . & . & 0 \\ 0 & 0 & 0 & 0 & . & . & 0 & 0 & 0 & 1 & 0 & . & 0 \\ 0 & 0 & 0 & 0 & . & . & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & . & . & 0 & 0 & 0 & 0 & . & 1 & 0 \end{bmatrix} . \begin{bmatrix} Trend_{t-1} \\ \beta_{t-1} \\ S_{1,t-1} \\ S_{1,t-2} \\ S_{1,t-3} \\ . \\ S_{1,t-11} \\ \delta_{t-1} \\ S_{2,t-1} \\ S_{2,t-2} \\ S_{2,t-3} \\ . \\ S_{2,t-11} \end{bmatrix} + \begin{bmatrix} 0 \\ \eta_{\beta t} \\ \eta_{S1t} \\ 0 \\ 0 \\ . \\ 0 \\ \eta_{\delta t} \\ \eta_{S2t} \\ 0 \\ . \\ . \\ 0 \end{bmatrix}
$$

References

Durbin, J. and Koopman, S. (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

Greene, W. (2000). *Econometric Analysis*, 4th ed. New York: Prentice Hall International.

Harvey, A.C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.