Modelling the UK Labour Force Survey using a Structural Time Series Model

Julian Chow[1], *Paul Smith, Gary Brown, Ping Zong*[2]
UK Office for National Statistics

Denise Britz do N. Silva[3]
Instituto Brasileiro de Geografia e Estatística – IBGE - Brazil

Abstract

The UK Labour Force Survey (LFS) is a quarterly survey based on a rotating panel design, which generates about 80% sample overlap between two successive quarters and can induce correlated sampling errors and rotation group bias (RGB). The standard time series analysis packages fail to account for the effect of the sampling error autocorrelation (SEA) and the use of these procedures may produce spurious trends. This paper represents the first attempt to set up a structural time series model for the UK Labour Force Survey series that accounts for the correlation between panel estimates. The proposed model is a first step towards the development of a state-space model to estimate the rotation group effects and discontinuities that may arise due to the transferring of the LFS to the Integrated Household Survey.

Key Words: Labour Force Series, Panel Survey, Rotating Group Bias, Univariate Structural Time Series Model.

1 Introduction

The UK Labour Force Survey (LFS) is a quarterly survey based on a rotating panel design. Each quarter the LFS surveys five panels and each panel is composed of approximately 12000 households. Individuals stay in the sample for five consecutive waves with 20% rotation so that four panels are resurveyed each quarter and one panel is new. The implication for this rotation pattern is that it induces 80% sample overlap between two successive quarters and overlaps over more than two quarters. As a result, the survey sampling errors are correlated over time.

The standard time series analysis packages fail to account for the effect of this sampling error autocorrelation (SEA). Hence, the use of these procedures may produce spurious trends that reflect an underlying trend-cycle induced by the survey rotation pattern. This issue was observed and studied by Hausman and Watson (1985); illustrated empirically by Pfeffermann, Feder and Signorelli (1991, 1998) and Tiller (1992); as well as highlighted by Smith (1999):

> *"When a time series of population values is estimated from a survey the sampling errors complicate the analysis. Complex rotation patterns give rise to complex covariance structures which are superimposed on the covariance structure of the time series and should be taken account of in any analysis."*

In order to decouple the SEA effect from the time series evolution of the true population quantity of interest, the time series model for the survey estimate is defined as the combination of two distinct models. One to describe the evolution of the unobservable population quantities over time and the other that represents the time series relationship between the sampling errors of the sample estimates.

---

[1] Julian Chow (chowj@un.org) works in Department of Economic and Social Affairs, United Nations.
[2] Paul Smith (Paul.Smith@ons.gsi.gov.uk), Gary Brown (Gary.brown@ons.gsi.gov.uk) and Ping Zong (ping.zong@ons.gsi.gov.uk) work in Office for National Statistics UK.
[3] Denise Britz do N. Silva (denise.silva@ibge.gov.br) works in Instituto Brasileiro de Geografia e Estatistica – IBGE- Brazil. .

Scott, Smith and Jones (1977) were the first to propose time series models for the sampling error process taking into account the survey rotation pattern. They examined single-stage and two-stage overlapping surveys and used different ARMA models for representing the sampling error process to take into account the dependence of the autovariance structure of the sampling errors on the pattern of the overlap. In addition, they provided an interesting discussion about which ARMA models would be appropriate under different survey designs. Regarding the sampling error process, they suggested an AR(1) would be reasonable for completely overlapping surveys. For partially overlapping surveys, in which the units are rotated out of the sample after q occasions, a moving-average process of order q (*MA(q)*) was recommended because the autocorrelation function for such models is zero for lags greater than q.

Another problem with the rotating panel design of the LFS is that the number of times respondents have been exposed to the survey may affect the data reported, resulting in systematic differences between panels in one time period. This phenomenon is known as rotation group bias (RGB) and is well documented in the literature (Bailar (1975), Kumar and Lee (1983), Binder and Hidiroglou (1988) and Pfeffermann (1991)).

This paper chronicles progress in a project to set up a Multivariate State-Space Model (MSSM) to account for SEA and RGB in quarterly LFS unemployment estimates. Attention is focused on the development of a Univariate State-Space Model (USSM) as a first step in separating any spurious trend or other effect that may arise from SEA due to sample overlap. The proposal of a MSSM to correct for RGB will be presented in a later paper.

2. Univariate State-Space Model without RGB

This section sets up the USSM for the LFS to take into account the SEA. The advantage of the USSM is that it decomposes the unobservable sampling error from the trend and seasonal components from the model. If the variation in the sampling errors is not taken into account, their autocorrelation structure may be absorbed into either the seasonal or the trend components (Pfeffermann, 1991, Silva, 1996, Tiller, 1992).

2.1. Model Setup

Consider the following decomposition for a *quarterly* survey estimate

$$y_t = \theta_t + e_t \qquad (1)$$

where y_t is the design unbiased survey estimate, θ_t is the unknown population quantity and e_t is the survey error. By analogy with a signal extraction approach, θ_t is the signal and e_t is the noise. In this context, one is interested in estimating the unobservable signal θ_t (and the corresponding structural components) based on the past and current observations, $y_1, \ldots, y_t$, in the presence of noise.

The Basic Structural Model (BSM) for the signal θ_t is

$$\theta_t = L_t + S_t + I_t \qquad (1a)$$

$$L_t = L_{t-1} + R_{t-1} + \eta_t^L \qquad (1b)$$

$$R_t = R_{t-1} + \eta_t^R \qquad (1c)$$

$$S_t = -\sum_{j=1}^{3} S_{t-j} + \eta_t^S \qquad (1d)$$

where L_t is the trend component, R_t is the increment for the trend, S_t is the seasonal component, I_t is the irregular component and η_t^L, η_t^R and η_t^S are the error terms associated with the trend, slope and seasonal components respectively.

The time series process for the survey error e_t is defined as

$$e_t = \sum_{j=1}^{p} \phi_j e_{t-j} + \zeta_t \qquad (1e)$$

since the nature of the rotating panel design of the LFS induces autocorrelations of the survey error in the sample estimates. This study assumes the survey errors e_t follow an AR (p) process, where p is the order of autoregressive process.

The complete model for θ_t, e_t, y_t from equations (1a-1e) can then be formulated as a state-space model where the state-vector includes components from both the θ_t and e_t processes.

2.2. State-Space formulation for the signal

The BSM for the signal θ_t can be written in the following state-space form:

$$\theta_t = \mathbf{Z}^\theta \boldsymbol{\alpha}_t^\theta + I_t \qquad (2a)$$

$$\boldsymbol{\alpha}_t^\theta = \boldsymbol{T}^\theta \boldsymbol{\alpha}_{t-1}^\theta + \boldsymbol{G}^\theta \boldsymbol{\eta}_t^\theta \qquad (2b)$$

where

$$\boldsymbol{Z}_t^\theta = \boldsymbol{Z}^\theta = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{\alpha}_t^\theta = \begin{pmatrix} L_t & R_t & S_t & S_{t-1} & S_{t-2} \end{pmatrix}'$$

The transition equation for the signal process is given by

$$\begin{pmatrix} L_t \\ R_t \\ S_t \\ S_{t-1} \\ S_{t-2} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} L_{t-1} \\ R_{t-1} \\ S_{t-1} \\ S_{t-2} \\ S_{t-3} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \eta_t^L \\ \eta_t^R \\ \eta_t^S \end{pmatrix}$$

For each component, the disturbances η_t^L, η_t^R and η_t^S are assumed to be mutually uncorrelated normal deviates with mean zero and variances ${\sigma_L}^2$, σ_R^2, σ_S^2 respectively.

2.3. State-Space Formulation for the Noise (Sampling Errors)

Considering that the LFS has a rotation pattern in which a household is interviewed once every quarter, for five consecutive quarters, the sampling errors contain persistent (although low) autocorrelation which may be absorbed into the seasonal and trend component if not taken into account in the modelling procedure. The following AR (4) order is used for illustrating the state-space formulation of the sampling error process $\{e_t\}$.

$$e_t = \boldsymbol{Z}^e \boldsymbol{\alpha}_t^e \qquad (3a)$$

$$\boldsymbol{\alpha}_t^e = \boldsymbol{T}^e \boldsymbol{\alpha}_t^e + \boldsymbol{G}^e \boldsymbol{\eta}_t^e \qquad (3b)$$

where

$$\boldsymbol{Z}_t^\theta = \boldsymbol{Z}^e = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} e_t \\ e_{t-1} \\ e_{t-2} \\ e_{t-3} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 & \phi_3 & \phi_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} e_{t-1} \\ e_{t-2} \\ e_{t-3} \\ e_{t-4} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} (\zeta_t)$$

2.4 Joint State-Space Model for Signal and Noise

Both the signal and noise component can then be incorporated into a USSM, as follows:

$$y_t = \boldsymbol{Z}\boldsymbol{\alpha}_t + I_t \qquad I_t \sim N(0, \sigma_I^2) \qquad (4a)$$

$$\boldsymbol{\alpha}_t = \boldsymbol{T}\boldsymbol{\alpha}_t + \boldsymbol{G}\boldsymbol{\eta}_t \qquad (4b)$$

where (4a) is the measurement equation and (4b) is the transition equation. In the measurement equation (4a), **Z** is a known design matrix, $\boldsymbol{\alpha}_t$ is a vector of unknown state components and I_t is an irregular (disturbance) term.

$$\boldsymbol{Z} = \left[\boldsymbol{Z}^\theta \ , \boldsymbol{Z}^e\right] = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{\alpha}_t^{'} = \left[\boldsymbol{\alpha}_t^\theta, \boldsymbol{\alpha}_t^e\right] = \begin{pmatrix} L_t \\ R_t \\ S_t \\ S_{t-1} \\ S_{t-2} \\ e_t \\ e_{t-1} \\ e_{t-2} \\ e_{t-3} \end{pmatrix}$$

In the transition equation (4b), **T** is the transition matrix and $\boldsymbol{\eta}_t$ is disturbance vector

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{T}^{\theta} & \vdots & \boldsymbol{0}_{5x4} \\ \cdots & \vdots & \cdots \\ \boldsymbol{0}_{4x5} & \vdots & \boldsymbol{T}^{e} \end{bmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \phi_1 & \phi_2 & \phi_3 & \phi_4 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{G}^{\theta} \\ \cdots \\ \boldsymbol{G}^{e} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}^{\theta} & \vdots & \boldsymbol{0}_{5x1} \\ \cdots & \vdots & \cdots \\ \boldsymbol{0}_{4x3} & \vdots & \boldsymbol{G}^{e} \end{bmatrix}$$

$$\boldsymbol{\eta}_t = \begin{bmatrix} \boldsymbol{\eta}_t^{\theta} \\ \cdots \\ \boldsymbol{\eta}_t^{e} \end{bmatrix} = \begin{bmatrix} \eta_t^L \\ \eta_t^R \\ \eta_t^S \\ \zeta_t \end{bmatrix}$$

with $\eta_t^i \sim N(0,\sigma_i^2), i = L,R,S,\zeta$ and $COV(\eta_t^i, \eta_{t-k}^j) = 0 \ \forall \ k > 0 \ ; \ j \neq i$

The variance-covariance matrix for the disturbances is given by:

$$V(\boldsymbol{\eta}_t) = diag\left(\sigma_L^2, \sigma_R^2, \sigma_S^2, \sigma_\zeta^2\right)$$

2.5 Kalman Filter and Maximum Likelihood Estimation

In the model, the unobserved components are estimated by means of Kalman filter recursively. As all components are non-stationary with the exception of the survey errors, the non-stationary components are initialised with a very large variance (ie 100000) relative to the magnitude of the series. The unknown parameters (hyperparameters) such as variances of the disturbances are estimated by using Maximum Likelihood techniques[4].The model was fitted in SAS version 9.1.3 using its matrix language procedure.

3. Sampling Error Calculations

The application of the Kalman filter requires the estimation of the unknown autoregressive (AR) parameters (ϕ_j) of the sampling error process. The AR coefficients were obtained from the Yule-Walker equations based on the estimated SEAs. These, in turn, were estimated using the so-called “Pseudo-Error” approach proposed by Pfeffermann, et al. (1996, 1998).

[4] Readers can refer to Harvey (1989), and Durbin and Koopman (2001) for further technical details about Kalman Filter recursion and Maximum Likelihood estimation.

3.1 Pseudo-Error Approach

In the case of a rotating panel survey, the model for the sampling errors can be identified by analysing the pseudo-errors given by: $\tilde{e}_t^{(k)} = y_t^{(k)} - y_t$, where $y_t^{(k)}$ is the observed survey estimate of panel *k* at time *t*. That is, $y_t^{(k)}$ is a panel estimate based on data from a single time period *t* and only includes values from a set of households that join and leave the survey at the same time (a panel *k*)[5]. If there is no rotation bias, it follows that:

$$
\begin{aligned}
\tilde{e}_t^{(k)} &= y_t^{(k)} - y_t = y_t^{(k)} - \frac{1}{K}\sum_{k=1}^{K} y_t^{(k)} \\
&= (y_t^{(k)} - \theta_t) - \frac{1}{K}\sum_{k=1}^{K} (y_t^{(k)} - \theta_t) \\
&= e_t^{(k)} - \frac{1}{K}\sum_{k=1}^{K} e_t^{(k)} = e_t^{(k)} - e_t \quad ,
\end{aligned}
$$

where $e_t^{(k)}$ is the unobserved sampling error of panel *k* at time *t* . Thus, contrasts in $\tilde{e}_t^{(k)}$ are in fact functions of the panel sampling errors only. Assuming that the sampling errors are uncorrelated if the panels do not overlap, and also that the autocorrelation structure of $\{ e_t^{(k)} \}$ depends on the lag but not on the panel, it can be shown that $CORR(\tilde{e}_t^{(k)}, \tilde{e}_{t-h}^{(k)}) = CORR(e_t^{(k)}, e_{t-h}^{(k)})$.

Models for the panel sampling errors can be specified by applying simple time series model identification procedures to the various pseudo-error series, $\{ \tilde{e}_t^{(k)} \},\ k = 1,\dots,K$. Hence, after generating a pseudo-error series, its autocorrelation function can be estimated using time series procedures from any standard statistical software. Note, however, that this course of action depends on the restrictive assumption that the autocorrelation structure of the sampling errors does not vary between panels.

To overcome this problem, Pfeffermann, Bell and Signorelli (1996) proposed a method that allows for different panel autocorrelation structures and showed that the autocorrelation function ρ_h of the sampling errors can be obtained as:

$$
\rho_h = \frac{\sum_{k=1}^{K} C_h^{(k)}}{\sum_{k=1}^{K} C_0^{(k)}} = \frac{(K^2 - K)\ COV(e_{t-h}, e_t)}{\sqrt{(K^2 - K)\ COV(e_t, e_t)\ (K^2 - K)\ COV(e_{t-h}, e_{t-h})}}
$$

where $COV(\tilde{e}_{t-h}^{(k)}, \tilde{e}_t^{(k)}) = C_h^{(k)}$.

Therefore, according to Pfeffermann et al. (1996, 1998), the SEA can be estimated based on the autocovariance function of pseudo-errors, which can be observed and computed for each

---

[5] For the UK LFS, *k=5* since in each quarter the sample is composed of five rotating panels.

panel. Silva (1996) details the proof and provides the pseudo-error approach for multivariate time series.

In a practical situation, an estimate for the autocorrelation function ρ_h is computed using the sample autocorrelation function of the pseudo-error series. In addition, using the Yule-Walker equations (see Wei,1993, p.135) and the estimated autocorrelation function, the analyst can obtain estimates for the partial autocorrelation function of the sampling error process as well as preliminary estimates for the parameters of AR(p) time series model.

4. Empirical Results

4.1. Sampling Error Autocorrelation Function

The quarterly unemployment rate series from the UK LFS in the period from 1992 Q2 to 2008 Q3 was used in this study. As there are five waves in the sampling design of the LFS and the respondent in a particular wave is completely rotated out after five periods (t, t+1, t+2, t+3, t+4), AR models up to order 4 were tested for the sampling error. One can note however that a MA(4) could also be used to represent the sampling error underlying process as suggested by Scott, Smith and Jones (1977). The reasons for testing AR(p) models first is the correspondence between low order AR models with higher order MA models and also that other authors have already showed empirically that AR models with low orders are adequate to model sampling errors in the case partial overlapping surveys (see Brakel and Krieg, 2009 and Silva, 1996 and 2001)

Table 1 shows the results of autocovariance, autocorrelation and autoregressive parameter estimates under separate panels. The standard error of the estimated autocorrelations is equal to $1/\sqrt{T}$ where T is the length of the series. Therefore, only correlations with an absolute value larger that 0.246 are significantly different from zero at a 5% significance level. In this case, only a few estimates are significant (when looking at the panel/wave estimates separately). The estimated autocorrelation for wave/panel 1 may be more reliable assuming that no rotation group bias impacts the responses of the households in the first interview.

Table 1: Autocovariance, autocorrelation and autoregressive parameter estimates under separate panels .

Wave	Parameter Estimates	Lag			
		1	2	3	4
1	Autocovariance	0.033	0.032	0.006	0.013
	Autocorrelations	0.467	0.442	0.077	0.174
	Partial autocorrelations	0.467	0.286	-0.284	0.162
	Autoregressive parameters	0.461	0.319	-0.351	0.162
2	Autocovariance	-0.015	0.009	-0.014	0.012
	Autocorrelations	-0.412	0.234	-0.366	0.310
	Partial autocorrelations	-0.412	0.077	-0.298	0.086
	Autoregressive parameters	-0.331	-0.033	-0.267	0.086
3	Autocovariance	-0.006	0.000	-0.003	-0.007
	Autocorrelations	-0.109	-0.004	-0.054	-0.125
	Partial autocorrelations	-0.109	-0.016	-0.057	-0.140
	Autoregressive parameters	-0.119	-0.025	-0.073	-0.140
4	Autocovariance	-0.011	-0.002	-0.001	-0.003
	Autocorrelations	-0.202	-0.036	-0.016	-0.049
	Partial autocorrelations	-0.202	-0.080	-0.042	-0.068

	Autoregressive parameters	-0.224	-0.096	-0.057	-0.068
5	Autocovariance	0.008	0.001	0.011	-0.002
	Autocorrelations	0.113	0.008	0.164	-0.024
	Partial autocorrelations	0.113	-0.004	0.165	-0.063
	Autoregressive parameters	0.125	-0.025	0.173	-0.063
Aggregate	Autocovariance	0.000	0.002	0.000	0.000
	Autocorrelations	0.028	0.132	-0.003	0.042
	Partial autocorrelations	0.028	0.131	-0.010	0.025
	Autoregressive parameters	0.026	0.128	-0.011	0.025

Moreover, when analysing the aggregate value that corresponds to the published unemployment series there is no evidence of SEA despite the sample overlap. This may be due to the time interval between interviews (three months so lag 2 refers to interviews that take place six months apart).

In addition, as the UK LFS sample design is unclustered there is no restriction to replace one household by another in the same area (sort of similar household) when the original one is rotated out of the sample.

Although the autocorrelations of the aggregate series are not significantly different from zero, a decision to fit an AR (1) model for the sampling errors was taken in order to test the modelling procedure and to prepare the work for the next stage of modelling the panel series concurrently (via a multivariate model) to account for RGB and discontinuities.

Table 2 presents the estimates of the AR coefficients for models of orders 1 to 4. Instead of using the pseudo-error approach to estimate the AR coefficients, they could be treated as hyperparameters and be estimated using maximum likelihood in the state-space model. However, recent experiences from Statistics Netherlands (Brakel and Krieg, 2009) suggested that the model does not work very well if the AR parameters are estimated within the state-space model. Therefore, for this first experiment of fitting a state-space model to the UK series, the pseudo-error approach was chosen: the coefficients and disturbance variance of the sampling error process were hence calculated outside the Kalman Filter procedure.

Table 2: Parameter Estimates of AR (1 – 4)

AR Orders	AR(1) Coefficient	AR(2) Coefficient	AR(3) Coefficient	AR(4) Coefficient
AR(1)	0.0282	n.a	n.a	n.a
AR(2)	0.0245	0.1314	n.a	n.a
AR(3)	0.0258	0.1316	-0.0099	n.a
AR(4)	0.0260	0.1282	-0.0106	0.0254

Table 3 summarises several diagnostics statistics for models with different orders. The results show that AR (1) performs best because it has the lowest forecast error according to all prediction measures. Therefore, an AR (1) model was chosen under the pseudo-error approach for the USSM.

Table 3: Different AR Order Results Comparison

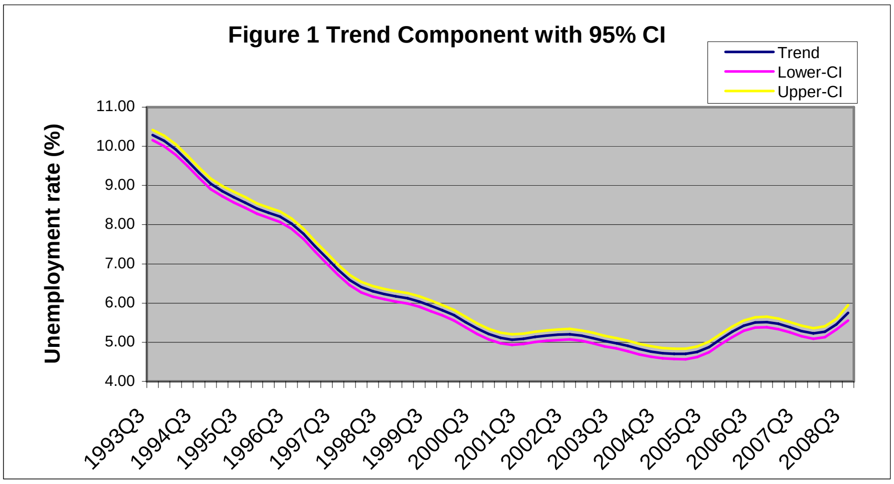
AR Orders	Mean Bias	Mean Abs Bias	Mean Relative Bias	Root-Square relative Error	1-stage ahead Prediction Error variance
AR(1)	0.0019	0.1740	0.1639	0.0011	0.0502
AR(2)	0.0021	0.1750	0.1671	0.0011	0.0512
AR(3)	0.0022	0.1752	0.1667	0.0011	0.0511
AR(4)	0.0023	0.1747	0.1677	0.0011	0.0506

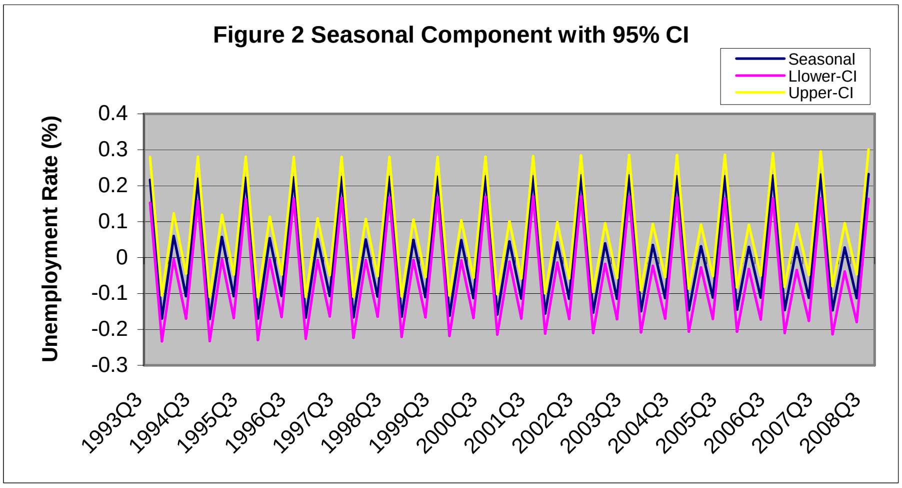
The AR (1) model with the parameter estimated under the pseudo-error approach was then formulated into a state-space form and the variance of the corresponding random disturbance was calculated. The sampling error follows the AR (1) model:

$$e_t = 0.028e_{t-1} + \zeta_t$$

4.3 Results of the USSM for UK Unemployment Rate

Estimates of UK unemployment rate trend and seasonal components as well as their corresponding confidence intervals derived from USSM are presented in the Figures 1 and 2 below.





The results of estimating their trend and seasonal components using the USSM and X-12-ARIMA (US Bureau of Census) are displayed in Figures 3, 4, 5 and 6.[6]

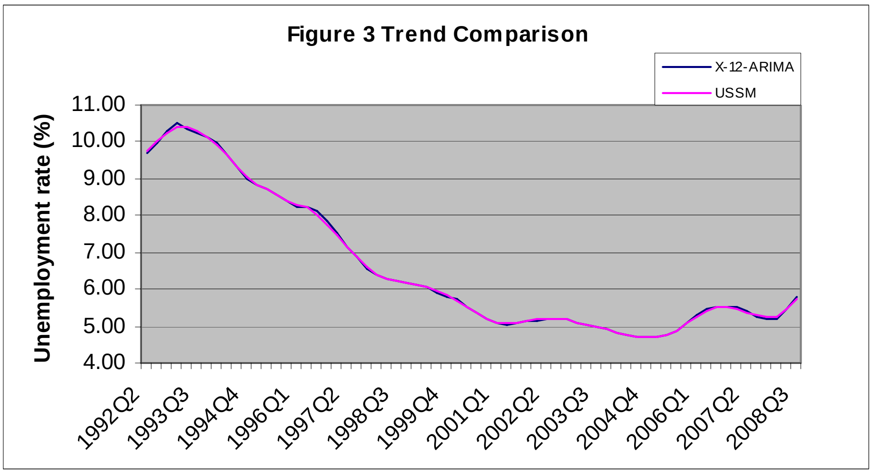


Figure 3 shows that the trend estimates obtained from USSM and from X-12-ARIMA are very close, but there are miner differences at turning points.

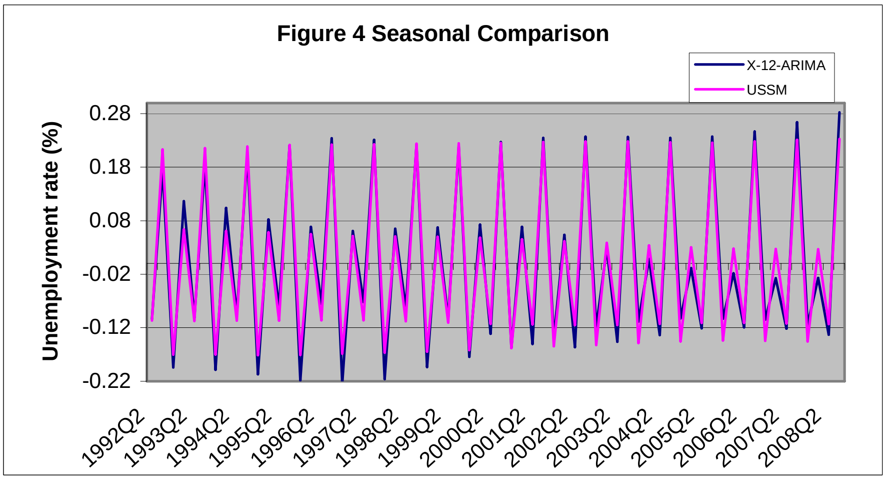


Figure 4 shows the seasonal amplitude estimated from USSM is slightly wider than that estimated from X-12-ARIMA.

[6] X-12-ARIMA is the standard seasonal adjustment package used by ONS.

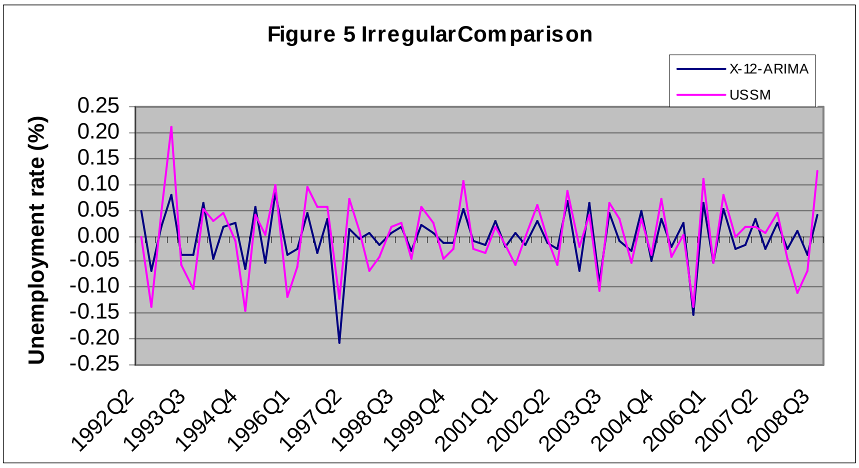


Figure 5 displays the estimated irregular $(e_t + \zeta_t)$ derived from USSM and derived from X-12-ARIMA. The estimated error item $(e_t + \zeta_t)$ derived from USSM accounts for sampling error variation so is in general larger than the estimated irregular derived from X-12-ARIMA.

6. Summary

- This paper describes the use of a state-space model to estimate structural time series components for the UK LFS while taking the SEA into account.
- The advantages of the USSM are as follows.
  1. SEA can be identified and estimated and the sampling error time series process can be modelled using a state-space formulation.
  2. Once the variance-covariance matrix is calculated, the state components, the trend and seasonal effects can be estimated through the Kalman filter. Using the Kalman filter, current estimates can be recursively filtered and past predictors can be updated.

7. Further work

- Take the claimant count into account as an auxiliary variable to improve the USSM.
- Develop the MSSM model to capture RGB effects.
- Further develop the MSSM to model discontinuities that may occur in the future due to survey redesign.

Acknowledgements

The Authors would like to thank Professor Danny Pfeffermann for his constructive comments and suggestions.

References

Binder, D. A. and M. A. Hidiroglou.1988. Sampling in time. In Handbook of Statistics, 6, 187-211. Eds, P.R.Krishnaiah and C.R.Rao. Elsevier Science.

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

Brakel, J. V. D and S. Krieg. 2009. Structural Time Series Modelling of theMonthly Unemployment Rate is a Rotating Panel. Discussion paper (09031). Statistics Netherlands.

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

Harvey, A.C., and G. D. A. Phillips. 1979. The Estimation of Regression Models with Autoregressive-Moving Average Disturbances. Biometrika pp. 49-59. Vol 66.

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

Harvey, A.C., and N. Shephard. 1993. Structual Time Series Models. Handbook of Statistics, Vol, 11.

Pfeffermann, D. 1991. Estimation and Seasonal Adjustment of Population means using Data From Repeated Surveys. Journal of Business and Economic Statistics, April, 1991, Vol.9, No2, pp163-177.

Pfeffermann, D., P. Bell and D. Signorelli. 1996. Labour Force Trend Estimation in Small Areas. Proceedings of the Bureau of Census Annual Research Conference and Technology Interchange. Washington, DC: U.S. Bureau of the Census, 407-431.

Pfeffermann, D., P. Bell, and D. Signorelli 1998. Estimation of Autocorrelations of Survey Error with Application to Trend Estimation in Small Errors. Journal of Business & Economic Statistics. Vol. 16, No.3. 407-443

Pfeffermann, D., and R. Tiller. 2006. Small-Area Estimation With State-Space Models Subject to benchmark Constraints. Journal of the American Statistical Association. vol. 101 no.476.

Pfeffermann, D., 2007,. Time Series 2. Lecture Notes for the MSc Course in Official Statistics, University of Southampton.

Scott, A.J., T. M. F. Smith, and R. G. Jones. 1977. The application of time series methods to the analysis of repeated surveys. *International Statistics Review*. 45, 13-28.

Silva, D.B.N. 1996. Modelling Compositional Time Series from Repeated Surveys. Unpublished Ph.D Thesis. University of Southampton

Silva, D.B.N. and T. M. F. Smith. 2001. Modelling Compositional Time Series from Repeated Surveys. *Survey methodology*, Vol. 27, No. 2, 205-215.

Smith, T. M. F. 1999. Defining parameters of interest in longitudinal studies and some implications for design. *In, Proceedings of the 52nd Session: Invited Papers*. Voorburg, The Netherlands, International Statistical Institute, 4pp. (Bulletin of the International Statistical Institute, VIII).

Tiller, R.B. 1992. Time Series Modeling of Sample Survey Data From the U.S. Current Population Survey. Journal of Official Statistics, 8, 149-166

Wei, W.W.S. 1993. Time Series Analysis: Univariate and Multivariate Methods. Addison-Wesley Publishing Company,Inc.