

Increasing frequency and improving timeliness of estimates from the UK Labour Force Survey

Duncan Elliott

20 November 2017

Overview

- Current time series outputs
- A time series model for monthly UK unemployment
- Issues in presenting time series information from these models

Presenting time series

- Guidelines on presenting time series
- Typically assume you have seasonally adjusted monthly or quarterly data
- Different types of users will want different levels of detail

Data and metadata reporting and presentation handbook
(OECD, 2007)

ESS guidelines on seasonal adjustment (Eurostat, 2015)

Types of users

- General public (*inquiring citizens*): headlines
- Informed user (*information foragers*): more context and ideas on accuracy
- Analytical user (*expert analysts*): greater metadata, reproducibility

Labour Market Publications

- Regular monthly publications
 - Rolling quarterly estimates (National Statistics)
 - Monthly estimates (Experimental Statistics)
- Time difference between publication date and mid-point of the latest reference period is approximately three months (eg Jul-Sep 2017 estimates published 15 November 2017)

Regular publications

- Headlines for unemployment

 - Seasonally adjusted level

 - Change on the previous quarter

 - Change on the previous year

- Additional information

 - Original time series, seasonally adjusted series

 - Quality reports, methods, metadata, ...

Labour Market November 2017

- Between April to June 2017 and July to September 2017 the number of unemployed people fell
- There were 1.42 million unemployed people
- 59,000 fewer than for April to June 2017
- 182,000 fewer than for a year earlier

Labour Force Survey

- Respondents interviewed five times (waves) at three month intervals
- Approximately 80% overlap between quarters
- Response rates have been declining over time
- Can create monthly wave specific estimates
Simpler calibration groups due to smaller sample size
These are volatile and may have bias eg mode of collection

Issues to address

- Issues to address
 1. Improve the timeliness of the reference period
 2. Improve accuracy of monthly estimates of change
- Develop a time series model to address issues
 - Accounting for survey design
 - Potentially incorporating administrative data

Unemployment input data

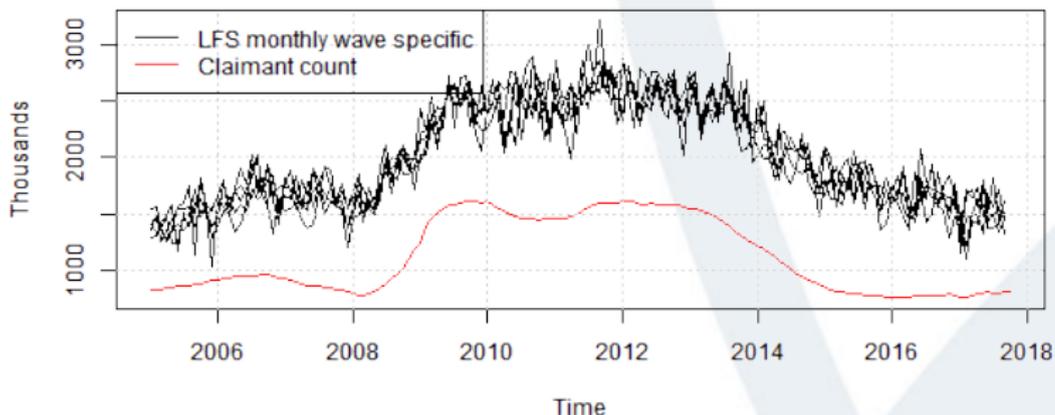


Figure: Monthly wave-specific estimates of UK unemployment for ages 16+ and seasonally adjusted claimant count

Model for unemployment

- Input data: monthly wave-specific time series, sampling errors, claimant count
- Assume population processes for unemployment and claimant count follow basic structural models with correlation in error terms for level and slope
- Bias terms for wave-specific survey errors
- Autoregressive process to model survey error autocorrelation

LFS population process

$$y_t^{(i)} = Y_t + \mathbf{e}_t^{(i)}, \quad i = 1, 2, \dots, 5$$

$$Y_t = L_t + S_t,$$

$$L_t = L_{t-1} + R_{t-1} + w_t^L, \quad w_t^L \sim N(0, \sigma_L^2),$$

$$R_t = R_{t-1} + w_t^R, \quad w_t^R \sim N(0, \sigma_R^2),$$

$$S_t = - \sum_{s=1}^{s=11} S_{t-s} + w_t^S, \quad w_t^S \sim N(0, \sigma_S^2)$$

LFS survey error process

$$e_t^{(i)} = b_t^{(i)} + \varepsilon_t^{(i)}$$

$$b_t^{(1)} = - \sum_{i=2}^5 b_t^{(i)}$$

$$b_t^{(i)} = b_{t-1}^{(i)} + w_t^{bi}, \quad w_t^{bi} \sim N(0, \sigma_{bi}^2), \quad \text{for } i = 2, \dots, 5$$

$$\varepsilon_t^{(i)} = k_t^{(i)} \tilde{\varepsilon}_t^{(i)}$$

$$\tilde{\varepsilon}_t^{(i)} = \sum_{j=1}^{i-1} \phi_j^{(i)} \tilde{\varepsilon}_{t-3j}^{(i-j)} + \epsilon_t^{(i)}, \quad \epsilon_t^{(i)} \sim N(0, (1 - \sum_{j=1}^{i-1} (\phi_j^{(i)})^2))$$

Including admin data

Claimant count process

$$\begin{aligned}y_t^{(c)} &= L_t^c + e_t^{(c)}, \\L_t^{(c)} &= L_{t-1}^{(c)} + R_{t-1}^{(c)} + w_t^{L(c)}, \quad w_t^{L(c)} \sim N(0, \sigma_{L^{(c)}}^2), \\R_t^{(c)} &= R_{t-1}^{(c)} + w_t^{R(c)}, \quad w_t^{R(c)} \sim N(0, \sigma_{R^{(c)}}^2),\end{aligned}$$

Where

$$\begin{aligned}\text{COV}(w_t^{L(c)}, w_t^L) &= \rho_L \sqrt{\sigma_{L^{(c)}}^2 \sigma_L^2} \\ \text{COV}(w_t^{R(c)}, w_t^R) &= \rho_R \sqrt{\sigma_{R^{(c)}}^2 \sigma_R^2}\end{aligned}$$

Estimation

- Estimate pseudo survey error autocorrelations ($\phi_j^{(i)}$)
- Put equations into a state space framework
- Hyperparameters estimated using maximum likelihood estimation
- Kalman Filter and Kalman Smoother used for estimation of unobserved components
- Estimation done using `d1m` package in R

Some model outputs

- Estimates of unobserved components from state vector α_t
 - L_t average level,
 - R_t estimate of monthly change in average level,
 - S_t seasonality,
 - $b_t^{(i)}$ wave-specific bias,
 - $\varepsilon_t^{(i)}$ wave-specific error
- Filtered estimates ($\alpha_{t|t}$)
- Smoothed estimates ($\alpha_{t|T}$)
- Standard errors of components

Model based estimates: level

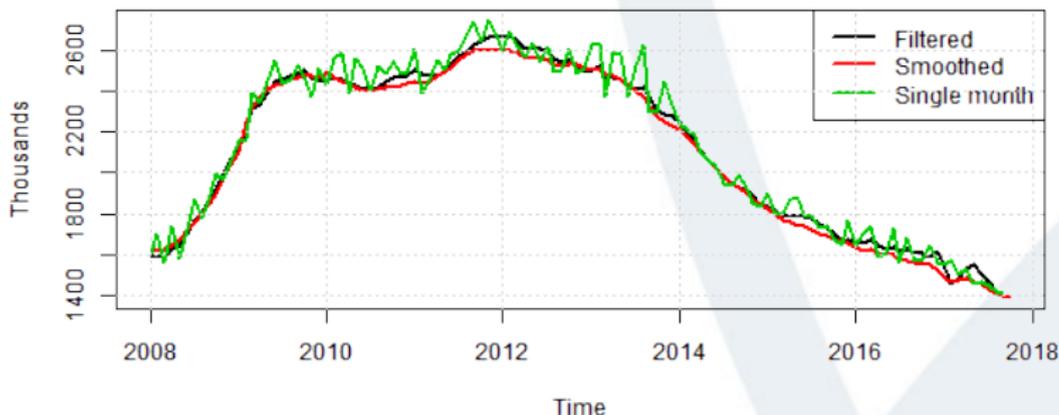


Figure: Filtered, Smoothed and original single month estimates (seasonally adjusted) of UK Unemployment aged 16 plus

Model based estimates: change

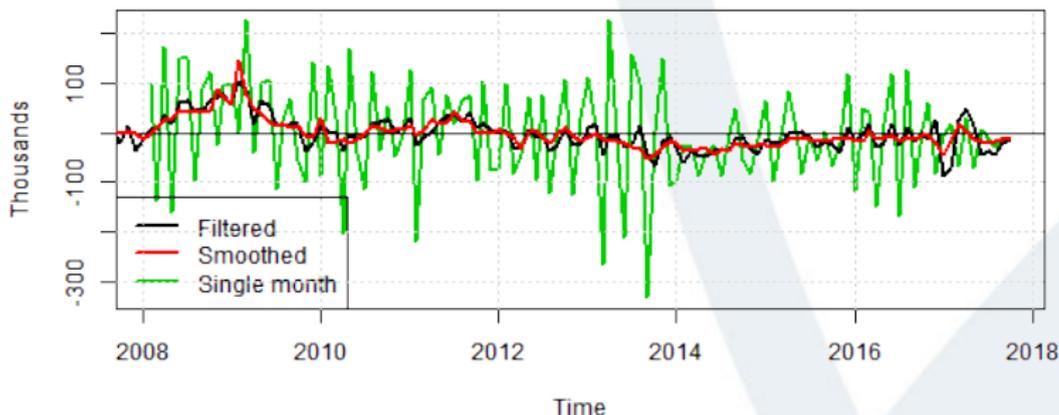


Figure: Month on month change of filtered, smoothed and original single month estimates (seasonally adjusted) of UK Unemployment aged 16 plus

Filtered estimates

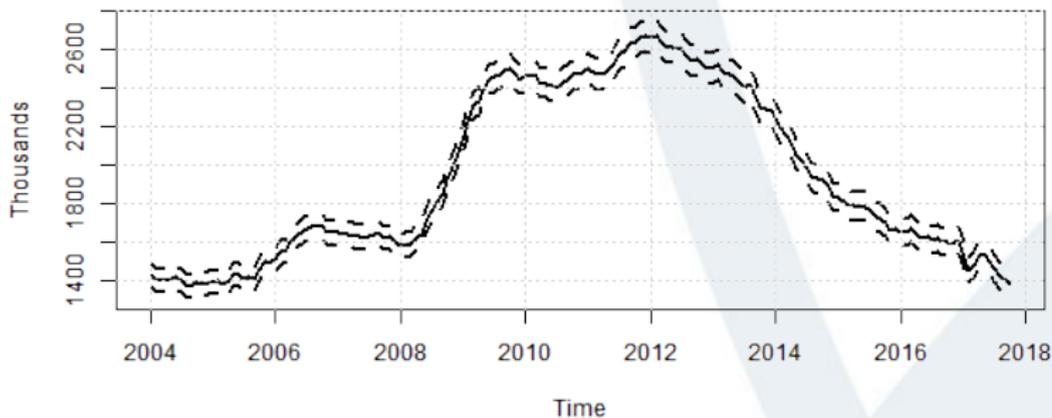


Figure: Filtered trend of UK Unemployment aged 16 plus with 95% confidence interval

Gain in accuracy

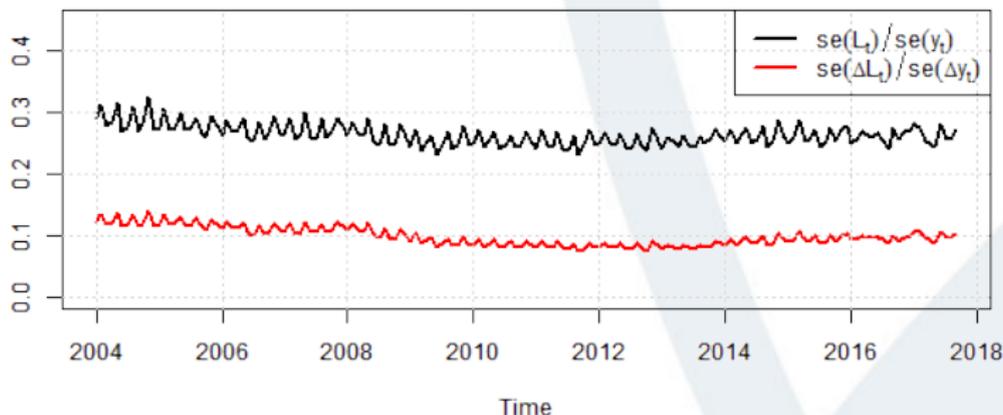


Figure: Standard errors of filtered trend relative design based standard errors for single month estimates of UK Unemployment aged 16 plus

Summary of model

- Greater accuracy, especially of monthly change
- Timeliness improved by two periods
- What should be published?

What do users want?

From my experience in the worlds of finance sector economics, journalism and politics, **there is immense suspicion of formulae and these convoluted methods**. For all the faults of claimant counts and the LFS, which are widely known, they are relatively easy things to 'sell' to users.

(Simon Briscoe, Financial Times, response to Harvey and Chung (2000) Estimating the Underlying Change in Unemployment in the UK

Guidance on publishing

- Making data meaningful: a guide to presenting statistics (UNECE, 2009)
- Data and metadata reporting and presentation handbook (OECD, 2007)
- Methodology of short-term business statistics: interpretation and guidelines (Eurostat, 2006)
- ESS guidelines on seasonal adjustment (Eurostat, 2015)

Dimensions of quality

1. Relevance
2. Accuracy and reliability
3. Timeliness and punctuality
4. Accessibility and clarity
5. Coherence and comparability

Also consider trade-off between dimensions, user needs and perceptions, cost and respondent burden, confidentiality, transparency and security

Current publications

- Latest level, period on period, and annual changes of SA series
- Time series charts of longer spans
- Standard errors and coefficients of variation
- Revisions
- Quality reports on dimensions of quality
- Datasets of unadjusted and seasonally adjusted time series

Quality: unemployment

- Which components and filtered or smoothed estimates?
relevance, reliability, accessibility and clarity, coherence and comparability
- When and how to revise (additional data, hyperparameters, changes in the model)
reliability, coherence and comparability
- Language to use for outputs: trend, filtered trend, level
accessibility and clarity

Quality: unemployment

- Comparisons with other published data (monthly model-based vs quarterly design-based, developing models for sub-populations and other identity constraints)
coherence and comparability
- Providing additional information from the model including inputs
relevance, confidentiality, transparency

Accessibility v accuracy

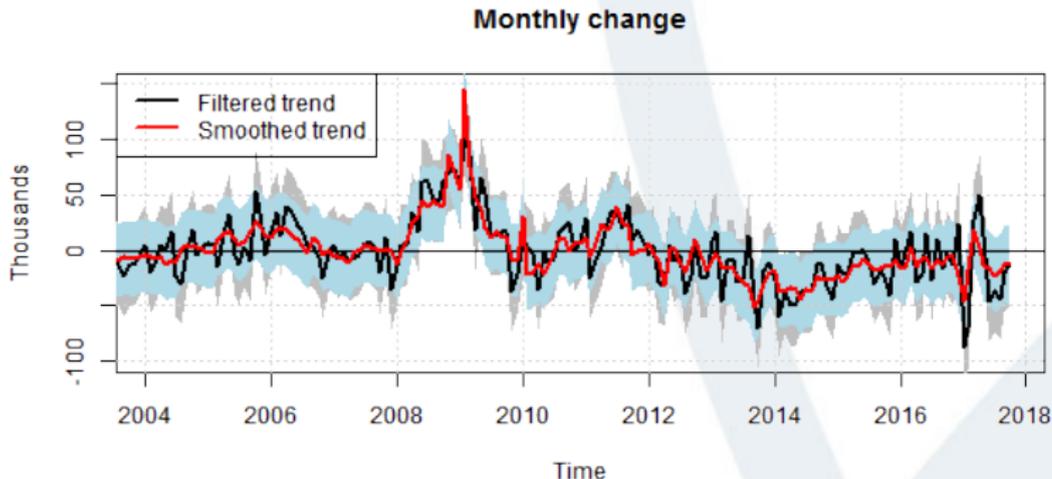


Figure: Monthly changes of filtered and smoothed trend estimates with 95% confidence intervals (grey and light blue respectively)

Proposed outputs

- Methods papers on input data and time series models
- Revisions policy for revising the time series models (model specification, hyperparameters, revisions to input data)
- Filtered trend component with measures of accuracy
- Other outputs of the model and input data made available on request

Summary

- Time series models developed to improve timeliness, periodicity and accuracy of UK unemployment
- Models provides a wealth of information but need careful consideration of what information to publish and how
- Useful case studies for moves towards new data and methods

Thank you!

Acknowledgements:

Alastair Cameron¹, Greg Dixon¹, Matthew Greenaway¹, Bethan Russ¹, Ping Zong¹, Danny Pfeffermann²

¹Office for National Statistics

²University of Southampton