Temporal disaggregation with heavy tails

ESCoE Conference on Economic Measurement 2019

Paul Labonne [1] Martin Weale [2]

[1]King's College London & ESCoE

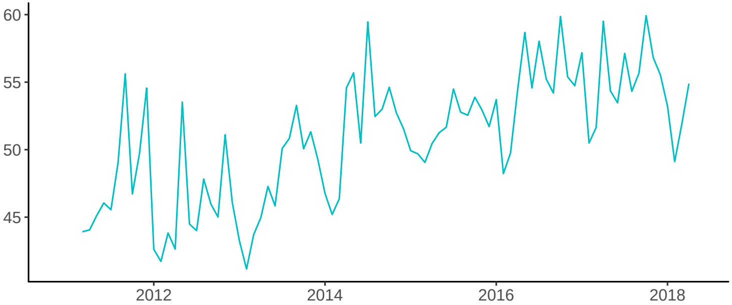
[2]King's College London, ESCoE & Centre for Macroeconomics

Illustration of the VAT-based turnover data

Table: 1: Representation of the quarterly staggers; x = quarterly turnover total

	Month												
	J	F	M	A	M	J	J	A	S	O	N	D	J
Stagger 1			x			x			x			x	
Stagger 2	x			x			x			x			x
Stagger 3		x			x			x			x		

Figure: Rolling quarterly VAT-based turnover figure for industry 493T495



- We want to estimate a monthly seasonally adjusted series from these rolling quarterly figures.

A structural model

- We use a local linear trend model for the monthly seasonally adjusted figures:

$$\begin{aligned} x_t &= \mu_t + e_t, & e_t &\sim \mathrm{N}(0, \sigma_e^2), \\ \mu_{t+1} &= \mu_t + \nu_t + \xi_t, & \xi_t &\sim \mathrm{N}(0, \sigma_\xi^2), \\ \nu_{t+1} &= \nu_t + \zeta_t, & \zeta_t &\sim \mathrm{N}(0, \sigma_\zeta^2), \end{aligned} \tag{1}$$

  where
  - x_t is the seasonally adjusted monthly interpoland;
  - μ_t is a dynamic trend;
  - ν_t is a dynamic slope;
  - and e_t is an irregular component.

- We link this latent model to the rolling quarterly values we observe through an aggregation function.

An approximation of the nonlinear aggregation function

- The exact (nonlinear) aggregation function is

$$y_t = \ln\left[\exp(x_t) + \exp(x_{t-1}) + \exp(x_{t-2})\right] + \gamma_t, \tag{2}$$

where γ_t is a three-month seasonal effect following a dummy seasonal model:

$$\gamma_{t+1} = -\sum_{j=1}^{11} \gamma_{t+1-j} + \omega_t, \qquad \omega_t \sim N(0, \sigma_\omega^2). \tag{3}$$

- Using the mean value theorem, Mitchell et al. (2005) show that we can approximate

$$\sum_{i=0}^{2} \mathrm{h}(x_{t-i}) \approx 3\ \mathrm{h}\left(\frac{\sum_{i=0}^{2} x_{t-i}}{3}\right), \tag{4}$$

where $h(.)$ is a nonlinear transformation.

- Using this approximation we can rewrite (2) as

$$y_t - \ln 3 = \frac{1}{3}x_t + \frac{1}{3}x_{t-1} + \frac{1}{3}x_{t-2} + \delta_t. \tag{5}$$

State space form

- The model can be written in state space form as

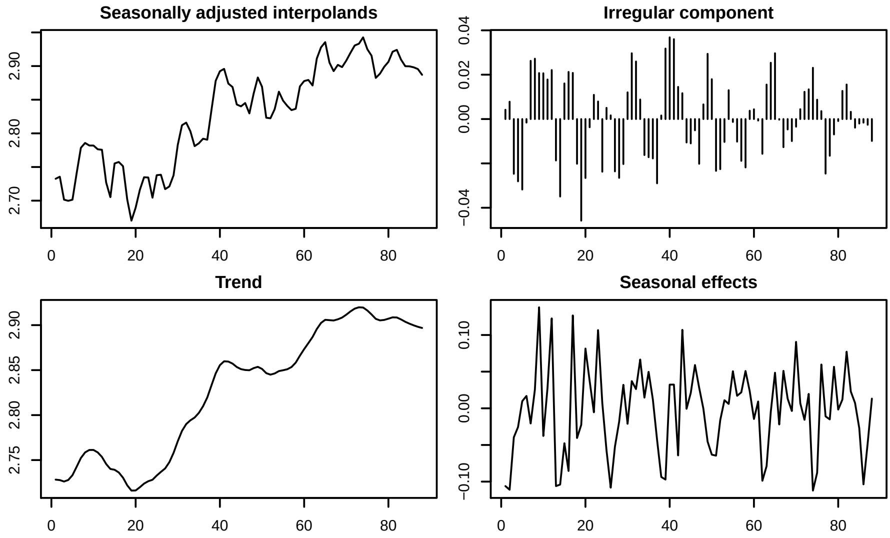
$$\begin{aligned} y_t &= Z\alpha_t, \\ \alpha_{t+1} &= T\alpha_t + R\eta_t, \qquad \eta_t \sim \mathrm{N}(0, Q), \\ \alpha_1 &\sim \mathrm{N}(a_1, P_1). \end{aligned} \tag{6}$$

  where $\alpha_t = (\mu_t, \mu_{t-1}, \mu_{t-2}, e_t, e_{t-1}, e_{t-2}, \nu_t, \gamma_t, ..., \gamma_{t-10})'$.

- The log-likelihood of model (6) can be evaluated using the Kalman filter.
- Once we have estimated the model we derive the interpolated series through the Kalman smoother.

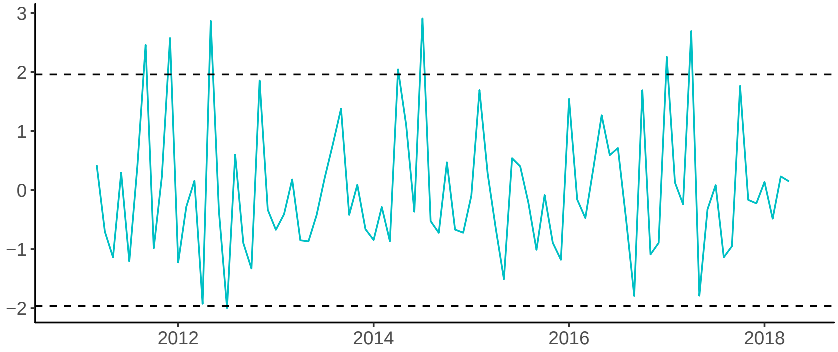
Result from a Gaussian state space model

Unobserved component decomposition



Result from a Gaussian state space model
Analysis of the standardised residuals

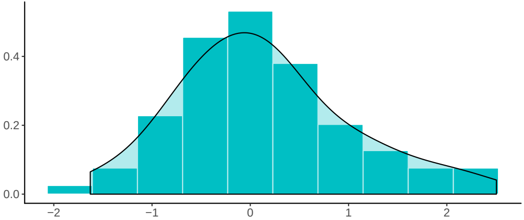
Figure: Standardised observation errors and admissibility region



- We detect eight outlying observations. In a Gaussian state space analysis these observations should be replaced with missing values, and the latter are accommodated easily via the Kalman filter. But we loose all information contained in these months.

Result from a Gaussian state space model
Analysis of the standardised residuals

Figure: Histogram and estimated density of the standardised one-step ahead prediction errors



The normality assumptions can be tested formally using the skewness and kurtosis statistics, respectively S and K, which can also be combined as $\mathrm{N} = n(\frac{S^2}{6} + \frac{(K-3)^2}{24})$.

$$\mathrm{S} \sim N(0, 6/n), \quad \mathrm{K} \sim N(0, 24/n), \quad \text{and} \quad \mathrm{N} \sim \chi^2(2).$$

The Gaussian state space estimation yields

$$S = 0.544, \quad K = 3.135, \quad N = 4.315, \quad z_S = 2.061, \quad z_K = 0.256.$$

A score-driven model
A general location model

- To account for the possible non-Gaussian features of the data and treat outlying observations without discarding them completely we explore a score-driven framework.
- We work from the following dynamic score-driven model:

$$\begin{aligned} y_t &= Z(a_t) + v_t, \\ a_{t+1} &= Ta_t + TAs_t, \end{aligned} \tag{7}$$

$$s_t = \frac{\partial \ell_t}{\partial a_t}.k, \tag{8}$$

where $\ell_t = \ln p(y_t|a_t;\Theta)$ is the predictive log-likelihood with Θ a vector of fixed parameters.

- We use the generalised asymmetric Student-t distribution of Zhu and Galbraith (2010) which can be skewed and have distinct tail parameters. The predictive log-likelihood for model (7) is

$$\begin{aligned} \ell_t = &-\ln\sigma - \frac{\nu_1+1}{2}\ln\left[1+\frac{1}{\nu_1}\left(\frac{y_t - Z(a_t)}{2\alpha\sigma K(\nu_1)}\right)^2\right] 1(y_t \leq Z(a_t)) \\ &- \frac{\nu_2+1}{2}\ln\left[1+\frac{1}{\nu_2}\left(\frac{y_t - Z(a_t)}{2(1-\alpha)\sigma K(\nu_2)}\right)^2\right] 1(y_t > Z(a_t)) \end{aligned} \tag{9}$$

- We work from the following dynamic score-driven model:

$$\begin{aligned} y_t &= Z(a_t) + v_t, \\ a_{t+1} &= Ta_t + TAs_t, \end{aligned}$$

$$s_t = \frac{\partial \ell_t}{\partial a_t}.k,$$

- The score vector is

$$\frac{\partial \ell_t}{\partial a_t} = \frac{\partial Z(a_t)}{\partial a_t} z_t, \tag{10}$$

where z_t is

$$\begin{aligned} z_t = &\frac{\nu_1 + 1}{1 + \frac{1}{\nu_1}\left(\frac{y_t - Z(a_t)}{2\alpha\sigma K(\nu_1)}\right)^2} \cdot \frac{y_t - Z(a_t)}{\nu_1(2\alpha\sigma K(\nu_1))^2} \; 1(y_t \leq Z(a_t)) \\ &+ \frac{\nu_2 + 1}{1 + \frac{1}{\nu_2}\left(\frac{y_t - Z(a_t)}{2(1-\alpha)\sigma K(\nu_2)}\right)^2} \cdot \frac{y_t - Z(a_t)}{\nu_2(2(1-\alpha)\sigma K(\nu_2))^2} \; 1(y_t > Z(a_t)). \end{aligned} \tag{11}$$

A score-driven model
A general location model

- We work from the following dynamic score-driven model:

$$y_t = Z(a_t) + v_t,$$
$$a_{t+1} = Ta_t + TAs_t,$$

- We re-write the scaled score vector as

$$s_t = \frac{\partial \ell_t}{\partial a_t} z_t.k = \frac{\partial \ell_t}{\partial a_t} u_t,$$

where u_t can be seen as the scaled prediction error. When the distribution is Gaussian we have $u_t = v_t$. Hence the distance between v_t and u_t indicates the extend of the nonlinear weighting. u_t is also important in the iterative smoothing procedure.

- We estimate the vector of unknown fixed parameter $\Theta = (\nu_1, \nu_2, \sigma, \alpha, \text{diag}(A), a_1')'$ by maximising the log-likelihood such that

$$\hat{\Theta} = \underset{\Theta}{\text{argmax}} \sum_{t=3}^{N} \ell_t.$$

A score-driven model
Estimation results

- Gaussian model

$$\alpha = 0.50, \quad \hat{\sigma} = 0.128, \quad \nu_1 = 250.0,$$
$$\nu_2 = 250.0, \quad A_{11} = 0.596, \quad \hat{L} = 133.237.$$

- Symmetric tails

$$\alpha = 0.50, \quad \hat{\sigma} = 0.072, \quad \hat{\nu}_1 = 2.001,$$
$$\hat{\nu}_2 = 2.001, \quad A_{11} = 1.614, \quad \hat{L} = 142.301.$$

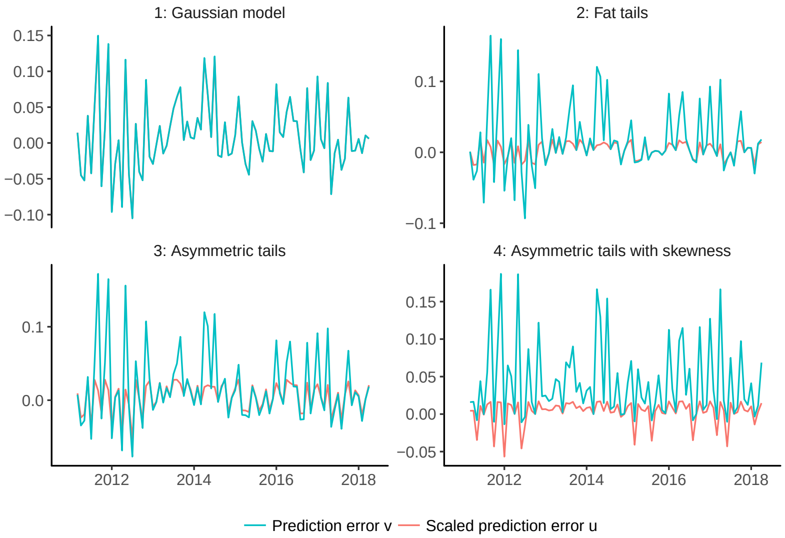
- Asymmetric tails

$$\alpha = 0.50, \quad \hat{\sigma} = 0.0861, \quad \hat{\nu}_1 = 341.887,$$
$$\hat{\nu}_2 = 2.001, \quad A_{11} = 1.356, \quad \hat{L} = 143.484.$$

- Asymmetric tails with skewness

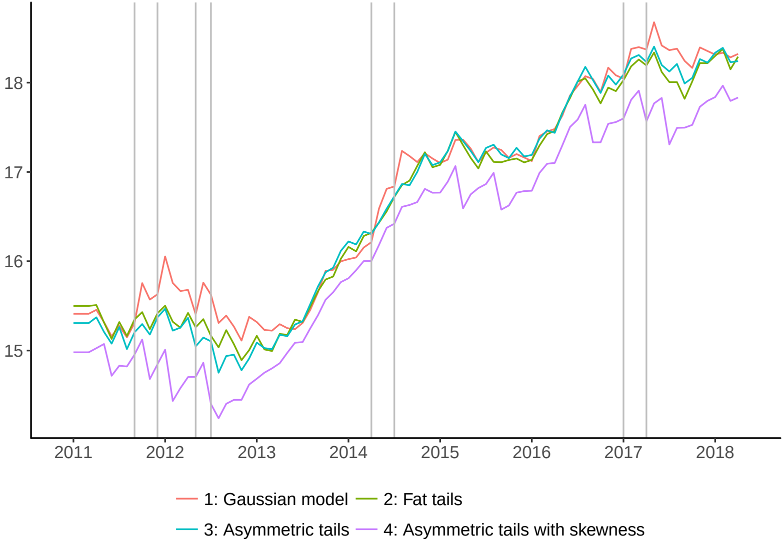
$$\hat{\alpha} = 0.184, \quad \hat{\sigma} = 0.0880, \quad \hat{\nu}_1 = 341.887,$$
$$\hat{\nu}_2 = 4.962, \quad A_{11} = 2.078, \quad \hat{L} = 154.758.$$

A score-driven model
Estimation results



A score-driven model

One-step ahead predictions



- Harvey (2013) proposes a iterative scheme:

  1: Estimate and run the DCS model (7) with the original data y_t;

  2: Generate pseudo observations as $\tilde{y}_t = Z(a_t) + u_t$;

  3: Estimate the Gaussian state space model with the pseudo observations $\tilde{y}_t$ which yield the smoothed estimates $\hat{\alpha}_t$;

  4: Set $a_t = \hat{\alpha}_t$ in (11) which yields new values of u_t;

  5: Generate pseudo-observations as $\tilde{y}_t = Z(\hat{\alpha}_t) + u_t$;

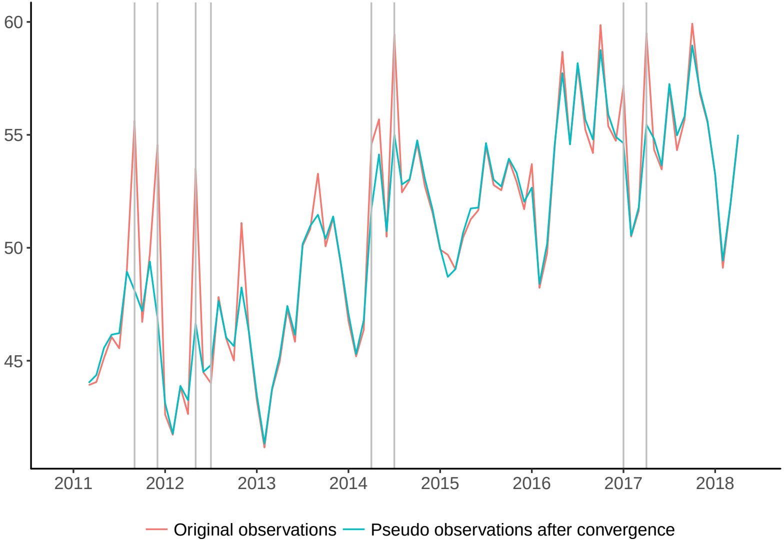
  6: Iterate steps 3 to 5 until convergence.

- Buccheri et al. (2018) propose an approximate smoother derived from the Kalman smoother:

$$\begin{aligned} r_{t-1} &= s_t + (T - TAI_{t|t-1}k)' r_t, \\ \hat{\alpha}_t &= a_t + Ar_{t-1}, \end{aligned} \tag{12}$$

A score-driven model

Pseudo observations from Harvey's smoother



Temporal disaggregation with heavy tails

Figure: Interpolated series in levels. Vertical lines correspond to Gaussian outliers.