PhD Masterclass Time Series Econometrics

Marc K. Francke
m.k.francke@uva.nl

Finance Group
University of Amsterdam Business School
Faculty of Economics and Business

Ortec Finance, Amsterdam

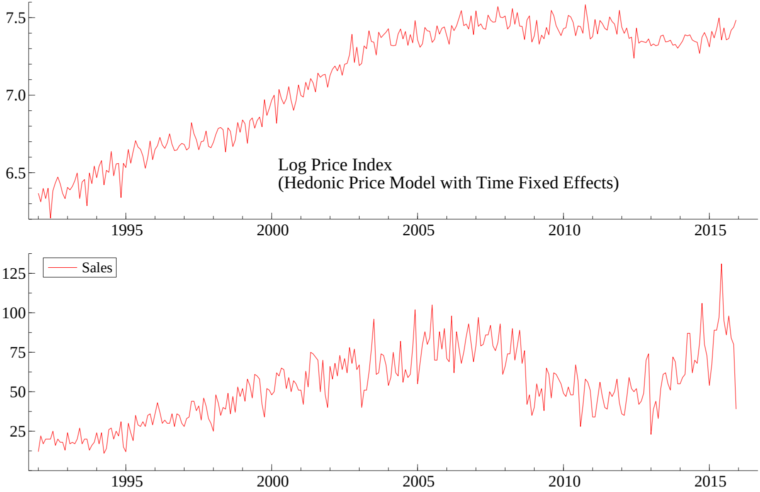
June 2016

ERES 23th Annual Conference Regensburg 2016

Time Series

Thanks Kai Ming Lee for many of the slides

Examples


7.5
7.0
6.5
Log Price Index
(Hedonic Price Model with Time Fixed Effects)
1995
2000
2005
2010
2015
Sales
125
100
75
50
25
1995
2000
2005
2010
2015

Classical Decomposition

A basic model for representing a time series is the additive model

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \qquad t = 1, \ldots, n,$$

also known as the Classical Decomposition.

$$\begin{aligned} y_t &= \text{observation}, \\ \mu_t &= \text{slowly changing component (trend)}, \\ \gamma_t &= \text{periodic component (seasonal)}, \\ \varepsilon_t &= \text{irregular component (disturbance)}. \end{aligned}$$

Local Level Model

- Components can be
  - deterministic functions of time (e.g. polynomials), or
  - stochastic processes;
- Examples
  - Deterministic: linear trend

$$y_t = \delta_0 + \delta_1 t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{NID}(0, \sigma^2_\varepsilon)$$

  - Stochastic: Random Walk plus Noise, or *Local Level* model:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \mathcal{NID}(0, \sigma^2_\varepsilon) \\ \mu_{t+1} &= \mu_t + \eta_t, & \eta_t &\sim \mathcal{NID}(0, \sigma^2_\eta), \end{aligned}$$

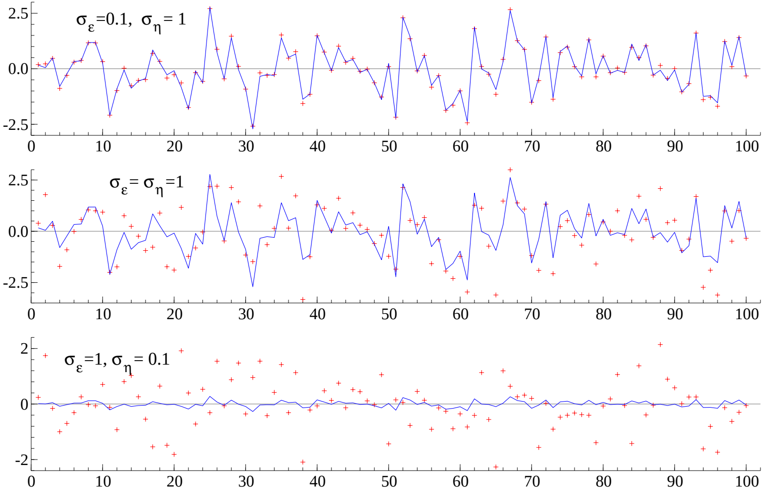
- Initial condition: $\mu_1 \sim \mathcal{N}(a, P)$;
- The disturbances ε_t, η_s are independent for all s, t;
- LL is a simple instance of a *Structural Time Series Model (STSM)* or *Unobserved Components Model (UCM)*.

Local Level Model

$$
\begin{aligned}
y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \mathcal{NID}(0, \sigma_\varepsilon^2) \\
\mu_{t+1} &= \mu_t + \eta_t, & \eta_t &\sim \mathcal{NID}(0, \sigma_\eta^2), \\
\mu_1 &\sim \mathcal{N}(a, P)
\end{aligned}
$$

- The level μ_t and the error term ε_t are unobserved;
- Parameters: $a, P, \sigma_\varepsilon^2, \sigma_\eta^2$;
- Trivial special cases:
  - $\sigma_\eta^2 = 0 \implies y_t \sim \mathcal{NID}(\mu_1, \sigma_\varepsilon^2)$ (White Noise with constant level);
  - $\sigma_\varepsilon^2 = 0 \implies y_{t+1} = y_t + \eta_t$ (pure Random Walk);

Simulated LL Data



Properties of the LL model

$$
\begin{aligned}
y_t &= \mu_t + \varepsilon_t, \qquad && \varepsilon_t \sim \mathcal{NID}(0, \sigma_\varepsilon^2), \\
\mu_{t+1} &= \mu_t + \eta_t, \qquad && \eta_t \sim \mathcal{NID}(0, \sigma_\eta^2),
\end{aligned}
$$

- First difference is stationary:

$$
\Delta y_t = \Delta\mu_t + \Delta\varepsilon_t = \eta_{t-1} + \varepsilon_t - \varepsilon_{t-1}.
$$

- Dynamic properties of Δy_t:

$$
\begin{aligned}
\mathrm{E}(\Delta y_t) &= 0, \\
\gamma_0 &= \mathrm{E}(\Delta y_t \Delta y_t) = \sigma_\eta^2 + 2\sigma_\varepsilon^2, \\
\gamma_1 &= \mathrm{E}(\Delta y_t \Delta y_{t-1}) = -\sigma_\varepsilon^2, \\
\gamma_\tau &= \mathrm{E}(\Delta y_t \Delta y_{t-\tau}) = 0 \quad \text{for } \tau \geq 2.
\end{aligned}
$$

Properties of the LL model

- The ACF of Δy_t is

$$\rho_1 = \frac{-\sigma_\varepsilon^2}{\sigma_\eta^2 + 2\sigma_\varepsilon^2} = -\frac{1}{q+2}, \qquad q = \sigma_\eta^2/\sigma_\varepsilon^2,$$
$$\rho_\tau = 0, \qquad \tau \geq 2.$$

- q is called the *signal-noise ratio*;
- The model for Δy_t is MA(1) with restricted parameters such that

$$-1/2 \leq \rho_1 \leq 0$$

  i.e., y_t is ARIMA(0,1,1);

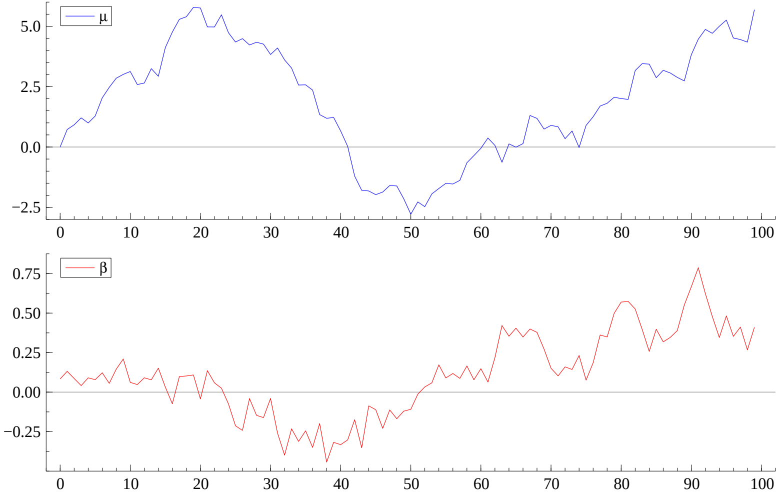
Local Linear Trend Model

The LLT model extends the LL model with a slope:

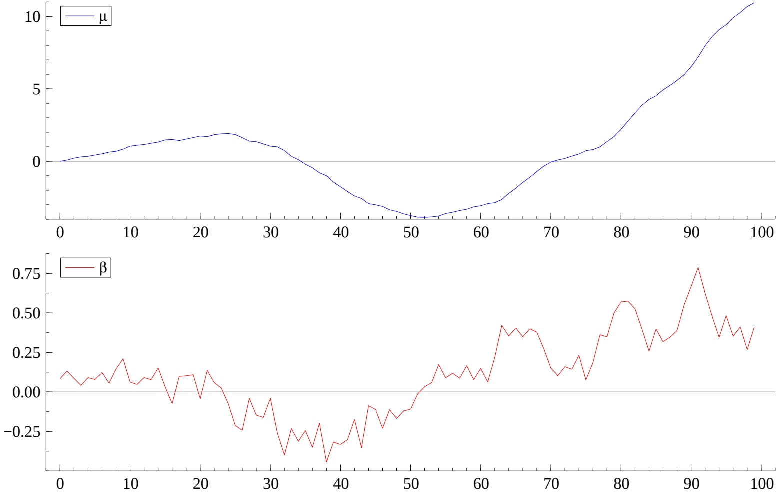
$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \mathcal{NID}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \beta_t + \mu_t + \eta_t, & \eta_t &\sim \mathcal{NID}(0, \sigma_\eta^2), \\ \beta_{t+1} &= \beta_t + \xi_t, & \xi_t &\sim \mathcal{NID}(0, \sigma_\xi^2). \end{aligned}$$

- All disturbances are independent at all lags and leads;
- Initial distributions β_1, μ_1 need to specified;
- Special cases
  - If $\sigma_\xi^2 = 0$ the trend is a random walk with constant drift β_1; (For $\beta_1 = 0$ the model reduces to a Local Level model.)
  - If additionally $\sigma_\eta^2 = 0$ the trend is a straight line with slope β_1 and intercept μ_1;
  - If $\sigma_\xi^2 > 0$ but $\sigma_\eta^2 = 0$, the trend is a smooth curve, or an Integrated Random Walk;

Trend and Slope in LLT Model



Trend and Slope in Integrated Random Walk Model



Seasonal Effects

We have seen specifications for μ_t in the basic model

$$y_t = \mu_t + \gamma_t + \varepsilon_t.$$

Now we will consider the seasonal term γ_t. Let s denote the number of ‘seasons’ in the data:

- $s = 12$ for monthly data,
- $s = 4$ for quarterly data,
- $s = 7$ for daily data when modelling a weekly pattern.

Dummy Seasonal

The simplest way to model seasonal effects is by using dummy variables. The effect summed over the seasons should equal zero:

$$\gamma_{t+1} = -\sum_{j=1}^{s-1} \gamma_{t+1-j}.$$

To allow the pattern to change over time, we introduce a new disturbance term:

$$\gamma_{t+1} = -\sum_{j=1}^{s-1} \gamma_{t+1-j} + \omega_t, \qquad \omega_t \sim \mathcal{NID}(0, \sigma_\omega^2).$$

The expectation of the sum of the seasonal effects is zero.

Trigonometric Seasonal

Defining γ_{jt} as the effect of season j at time t, an alternative specification for the seasonal pattern is

$$\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{jt},$$

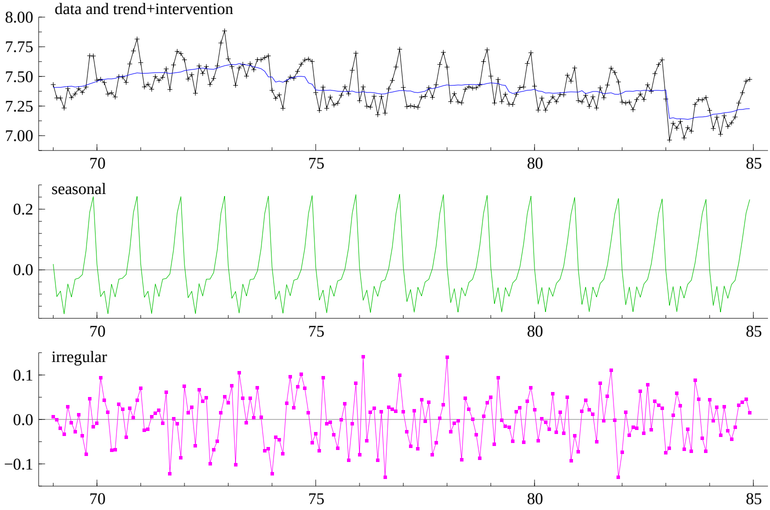
$$\gamma_{j,t+1} = \gamma_{jt} \cos \lambda_j + \gamma^*_{jt} \sin \lambda_j + \omega_{jt},$$

$$\gamma^*_{j,t+1} = -\gamma_{jt} \sin \lambda_j + \gamma^*_{jt} \cos \lambda_j + \omega^*_{jt},$$

$$\omega_{jt}, \omega^*_{jt} \sim \mathcal{NID}(0, \sigma^2_\omega), \qquad \lambda_j = 2\pi j/s.$$

- Without the disturbance, the trigonometric specification is identical to the deterministic dummy specification.
- The autocorrelation in the trigonometric specification lasts through more lags: changes occur in a smoother way;

Seatbelt Law



Cycles

We can extend the basic model with cycle ψ_t

$$y_t = \mu_t + \gamma_t + \psi_t + \varepsilon_t,$$

where ψ_t can be deterministic

$$\psi_t = A\cos(\lambda t + B)$$

or stochastic

$$\begin{aligned}
\psi_{t+1} &= \rho\left[\psi_t \cos\lambda + \psi t^* \sin\lambda\right] + \kappa_t, \\
\psi_{t+1}^* &= \rho\left[-\psi_t \sin\lambda + \psi_t^* \cos\lambda\right] + \kappa_t^*, \\
\kappa_t, \kappa_t^* &\sim \mathcal{NID}(0, \sigma_\kappa^2).
\end{aligned}$$

State Space Model: a more general class of models

Linear Gaussian state space model is defined in three parts:

→ State equation:

$$\alpha_{t+1} = T_t \alpha_t + R_t \zeta_t, \qquad \zeta_t \sim \mathcal{NID}(0, Q_t),$$

→ Observation equation:

$$y_t = Z_t \alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, H_t),$$

→ Initial state distribution $\alpha_1 \sim \mathcal{N}(a_1, P_1)$.

Notice that

- ζ_t and ε_s independent for all t, s, and independent from α_1;
- observation y_t can be multivariate;
- state vector α_t is unobserved;
- matrices T_t, Z_t, R_t, Q_t, H_t determine structure of model.

State Space Model

- state space model is linear and Gaussian: therefore properties and results of multivariate normal distribution apply;
- state vector α_t evolves as a VAR(1) process;
- system matrices usually contain unknown parameters;
- estimation has therefore two aspects:
  - measuring the unobservable state (prediction, filtering and smoothing) conditional on unknown parameters;
  - estimation of unknown parameters (maximum likelihood estimation);
- state space methods offer a *unified approach* to a wide range of models and techniques: dynamic regression, ARIMA, UC models, latent variable models, spline-fitting and many ad-hoc filters;
- next, some well-known model specifications in state space form ...

Regression with Time Varying Coefficients

General state space model:

$$\begin{aligned} \alpha_{t+1} &= T_t\alpha_t + R_t\zeta_t, \qquad \zeta_t \sim \mathcal{NID}(0, Q_t), \\ y_t &= Z_t\alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, H_t). \end{aligned}$$

Put regressors in Z_t,

$$T_t = I, \qquad R_t = I,$$

Result is regression model with coefficient α_t following a random walk.

ARMA in State Space Form

Example: AR(2) model $y_{t+1} = \phi_1 y_t + \phi_2 y_{t-1} + \zeta_t$, in state space:

$$\alpha_{t+1} = T_t \alpha_t + R_t \zeta_t, \qquad \zeta_t \sim \mathcal{NID}(0, Q_t),$$
$$y_t = Z_t \alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, H_t).$$

with 2×1 state vector α_t and system matrices:

$$Z_t = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad H_t = 0$$
$$T_t = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix}, \qquad R_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad Q_t = \sigma^2$$

- Z_t and $H_t = 0$ imply that $\alpha_{1t} = y_t$;
- First state equation implies $y_{t+1} = \phi_1 y_t + \alpha_{2t} + \zeta_t$ with $\zeta_t \sim \mathcal{NID}(0, \sigma^2)$;
- Second state equation implies $\alpha_{2,t+1} = \phi_2 y_t$;

ARMA in State Space Form

Example: MA(1) model $y_{t+1} = \zeta_t + \theta\zeta_{t-1}$, in state space:

$$\alpha_{t+1} = T_t\alpha_t + R_t\zeta_t, \qquad \zeta_t \sim \mathcal{NID}(0, Q_t),$$
$$y_t = Z_t\alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{NID}(0, H_t).$$

with 2×1 state vector α_t and system matrices:

$$Z_t = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad H_t = 0$$
$$T_t = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad R_t = \begin{bmatrix} 1 \\ \theta \end{bmatrix}, \qquad Q_t = \sigma^2$$

- Z_t and $H_t = 0$ imply that $\alpha_{1t} = y_t$;
- First state equation implies $y_{t+1} = \alpha_{2t} + \zeta_t$ with $\zeta_t \sim \mathcal{NID}(0, \sigma^2)$;
- Second state equation implies $\alpha_{2,t+1} = \theta\zeta_t$;

ARMA in State Space Form

Example: ARMA(2,1) model

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \zeta_t + \theta\zeta_{t-1}$$

in state space form

$$\alpha_t = \begin{bmatrix} y_t \\ \phi_2 y_{t-1} + \theta\zeta_t \end{bmatrix}$$

$$Z_t = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad H_t = 0,$$

$$T_t = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix}, \qquad R_t = \begin{bmatrix} 1 \\ \theta \end{bmatrix}, \qquad Q_t = \sigma^2$$

All ARIMA(p, d, q) models have a (non-unique) state space representation.

UC models in State Space Form

State space model: $\alpha_{t+1} = T_t\alpha_t + R_t\zeta_t, \quad y_t = Z_t\alpha_t + \varepsilon_t$.

LL model $\Delta\mu_{t+1} = \eta_t$ and $y_t = \mu_t + \varepsilon_t$:

$$\alpha_t = \mu_t, \qquad T_t = 1, \qquad R_t = 1, \qquad Q_t = \sigma^2_\eta,$$
$$Z_t = 1, \qquad H_t = \sigma^2_\varepsilon.$$

LLT model $\Delta\mu_{t+1} = \beta_t + \eta_t, \quad \Delta\beta_{t+1} = \xi_t$ and $y_t = \mu_t + \varepsilon_t$:

$$\alpha_t = \begin{bmatrix}\mu_t \\ \beta_t\end{bmatrix}, \qquad T_t = \begin{bmatrix}1 & 1 \\ 0 & 1\end{bmatrix}, \qquad R_t = \begin{bmatrix}1 & 0 \\ 0 & 1\end{bmatrix}, \qquad Q_t = \begin{bmatrix}\sigma^2_\eta & 0 \\ 0 & \sigma^2_\xi\end{bmatrix},$$
$$Z_t = \begin{bmatrix}1 & 0\end{bmatrix}, \qquad H_t = \sigma^2_\varepsilon.$$

UC models in State Space Form

State space model: $\alpha_{t+1} = T_t\alpha_t + R_t\zeta_t, \quad y_t = Z_t\alpha_t + \varepsilon_t$.

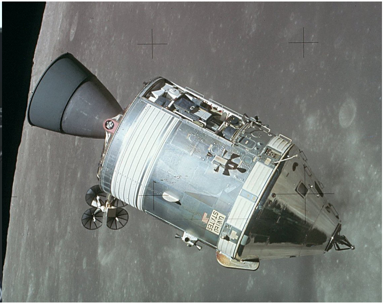
LLT model with season: $\Delta\mu_{t+1} = \beta_t + \eta_t, \quad \Delta\beta_{t+1} = \xi_t$, $S(L)\gamma_{t+1} = \omega_t$ and $y_t = \mu_t + \gamma_t + \varepsilon_t$:

$$\alpha_t = \begin{bmatrix}\mu_t & \beta_t & \gamma_t & \gamma_{t-1} & \gamma_{t-2}\end{bmatrix}',$$

$$T_t = \begin{bmatrix}1 & 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0\\ 0 & 0 & -1 & -1 & -1\\ 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0\end{bmatrix}, \quad Q_t = \begin{bmatrix}\sigma_\eta^2 & 0 & 0\\ 0 & \sigma_\xi^2 & 0\\ 0 & 0 & \sigma_\omega^2\end{bmatrix}, \quad R_t = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix},$$

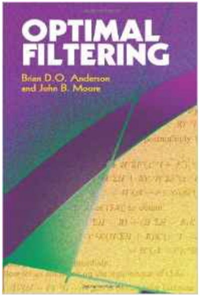
$$Z_t = \begin{bmatrix}1 & 0 & 1 & 0 & 0\end{bmatrix}, \quad H_t = \sigma_\varepsilon^2.$$

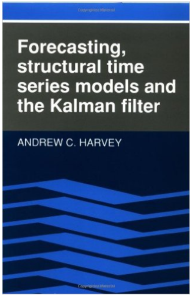
How to estimate state space models?

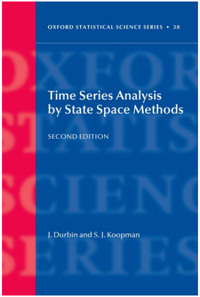


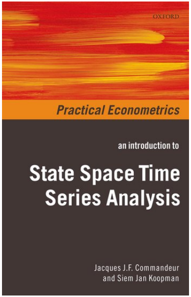
Let us go to rocket science
Use of Kalman filter: Apollo program, NASA Space Shuttle, Navy submarines, unmanned aerospace vehicles

Books on state space models and Kalman filter









Kalman Filter

- The Kalman filter calculates the mean and variance of the unobserved state, given the observations.
- The state is Gaussian: the complete distribution is characterized by the mean and variance.
- The filter is a recursive algorithm; the current best estimate is updated whenever a new observation is obtained.
- To start the recursion, we need a_1 and P_1 ($\alpha_1 \sim \mathcal{N}(a_1, P_1)$), which we assume to be given.
- There are various ways to initialize when a_1 and P_1 are unknown, which we will not discuss here.

Kalman Filter

The unobserved state α_t can be estimated from the observations with the *Kalman filter*. Define $Y_t = \{y_1, \dots, y_t\}, a_{t+1} = \mathrm{E}(\alpha_{t+1} | Y_t), P_{t+1} = \mathrm{Var}(\alpha_{t+1} | Y_t)$.

$$
\begin{aligned}
v_t &= y_t - Z_t a_t, \\
F_t &= Z_t P_t Z_t' + H_t, \\
K_t &= T_t P_t Z_t' F_t^{-1}, \\
a_{t+1} &= T_t a_t + K_t v_t, \\
P_{t+1} &= T_t P_t T_t' + R_t Q_t R_t' - K_t F_t K_t',
\end{aligned}
$$

for $t = 1, \dots, n$ and starting with given values for a_1 and P_1.

Kalman Filter

State space model: $\alpha_{t+1} = T_t\alpha_t + R_t\zeta_t, \quad y_t = Z_t\alpha_t + \varepsilon_t$.

- Writing $Y_t = \{y_1, \ldots, y_t\}$, define

$$a_{t+1} = \mathrm{E}(\alpha_{t+1}|Y_t), \qquad P_{t+1} = \mathrm{Var}(\alpha_{t+1}|Y_t);$$

- The prediction error is

$$\begin{aligned} v_t &= y_t - \mathrm{E}(y_t|Y_{t-1}) \\ &= y_t - \mathrm{E}(Z_t\alpha_t + \varepsilon_t|Y_{t-1}) \\ &= y_t - Z_t\,\mathrm{E}(\alpha_t|Y_{t-1}) \\ &= y_t - Z_t a_t; \end{aligned}$$

- It follows that $v_t = Z_t(\alpha_t - a_t) + \varepsilon_t$ and $\mathrm{E}(v_t) = 0$;
- The prediction error variance is $F_t = \mathrm{Var}(v_t) = Z_t P_t Z_t' + H_t$.

Lemma

The proof of the Kalman filter uses a lemma from multivariate Normal regression theory.

Lemma Suppose x, y and z are jointly Normally distributed vectors with $\mathrm{E}(z) = 0$ and $\Sigma_{yz} = 0$.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_x \\ \mu_y \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} & \Sigma_{xz} \\ \Sigma'_{xy} & \Sigma_{yy} & 0 \\ \Sigma'_{xz} & 0 & \Sigma_{zz} \end{pmatrix} \right)$$

Then

$$\mathrm{E}(x|y,z) = \mathrm{E}(x|y) + \Sigma_{xz}\Sigma_{zz}^{-1} z,$$
$$\mathrm{Var}(x|y,z) = \mathrm{Var}(x|y) - \Sigma_{xz}\Sigma_{zz}^{-1}\Sigma'_{xz},$$

Kalman Filter

State space model: $\alpha_{t+1} = T_t\alpha_t + R_t\zeta_t, \quad y_t = Z_t\alpha_t + \varepsilon_t.$

- $Y_t = \{Y_{t-1}, y_t\} = \{Y_{t-1}, v_t\}$, $\mathrm{E}(v_t y_{t-j}) = 0$, $j = 1, \ldots, t-1$
- Apply Lemma

$$\begin{array}{lllllll} x & = & \alpha_{t+1}, & \mathrm{E}(x|y) & = & \mathrm{E}(\alpha_{t+1}|Y_{t-1}) & = & T_t a_t, \\ y & = & Y_{t-1}, & \mathrm{Var}(x|y) & = & \mathrm{Var}(\alpha_{t+1}|Y_{t-1}) & = & \\ z & = & v_t, & \Sigma_{zz} & = & \mathrm{Var}(v_t) & = & F_t, \\ & & & \Sigma_{xz} & = & \mathrm{Cov}(\alpha_{t+1}, v_t) & = & T_t P_t Z_t', \end{array}$$

$\mathrm{Cov}(\alpha_{t+1}, v_t) = \mathrm{Cov}(T_t\alpha_t + R_t\zeta_t, Z_t(\alpha_t - a_t) + \varepsilon_t) = T_t P_t Z_t'$

$\mathrm{Var}(\alpha_{t+1}|Y_{t-1}) = T_t P_t T_t' + R_t Q_t R_t'$

- We carry out lemma and obtain the state update

$$\begin{aligned} a_{t+1} &= \mathrm{E}(\alpha_{t+1}|Y_{t-1}, y_t) = \mathrm{E}(\alpha_{t+1}|Y_{t-1}, v_t) \\ &= T_t a_t + T_t P_t Z_t' F_t^{-1} v_t = T_t a_t + K_t v_t; \\ P_{t+1} &= P_{t+1} = T_t P_t T_t' + R_t Q_t R_t' - K_t F_t K_t'. \end{aligned}$$

with $K_t = T_t P_t Z_t' F_t^{-1}$

Kalman Filter

- Conditional on Y_{t-1} the best prediction of y_t is $\mathcal{N}(Z_t a_t, F_t)$
- When the actual observation arrives, the prediction error $(y_t - Z_t\alpha_t)|Y_{t-1}$ is distributed as $\mathcal{N}(v_t, F_t)$
- The best prediction of the new state α_{t+1} is based both on the old estimate a_t and the new information v_t:

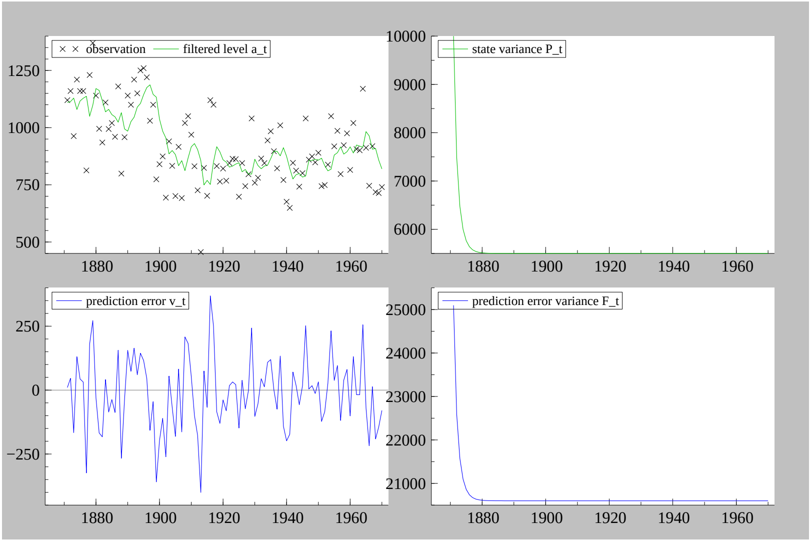
$$\alpha_{t+1}|Y_t \sim \mathcal{N}(a_{t+1} = T_t a_t + K_t v_t, P_{t+1} = T_t P_t T_t' + R_t Q_t R_t' - K_t F_t K_t')$$

- The *Kalman gain*

$$K_t = T_t P_t Z_t' F_t^{-1}$$

is the optimal weighting matrix for the new evidence.

Kalman Filter Illustration



Smoothing

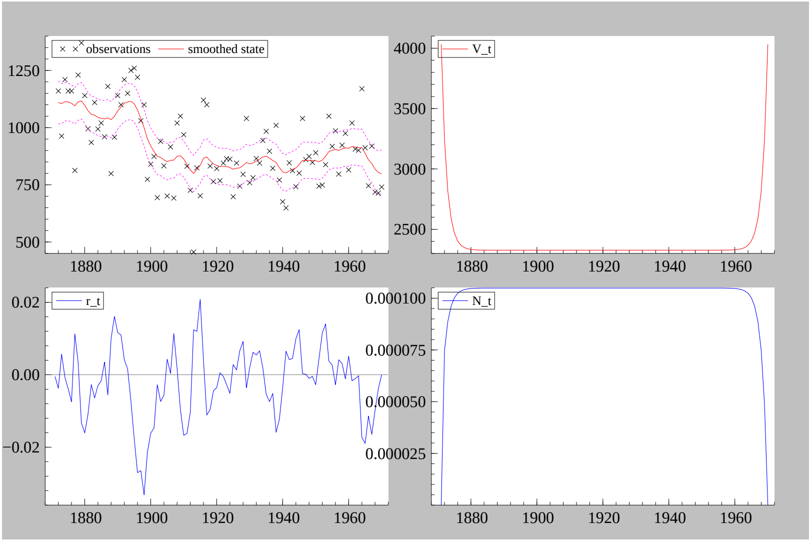
- The filter calculates the mean & variance conditional on Y_t;
- The Kalman smoother calculates the mean and variance conditional on the full set of observations Y_n;
- After the filtered estimates are calculated, the smoothing recursion starts at the last observations and runs until the first.

$$\hat{\alpha}_t = \mathrm{E}(\alpha_t | Y_n), \qquad V_t = \mathrm{Var}(\alpha_t | Y_n),$$
$$r_t = \text{weighted sum of innovations}, \quad N_t = \mathrm{Var}(r_t),$$
$$L_t = T_t - K_t Z_t.$$

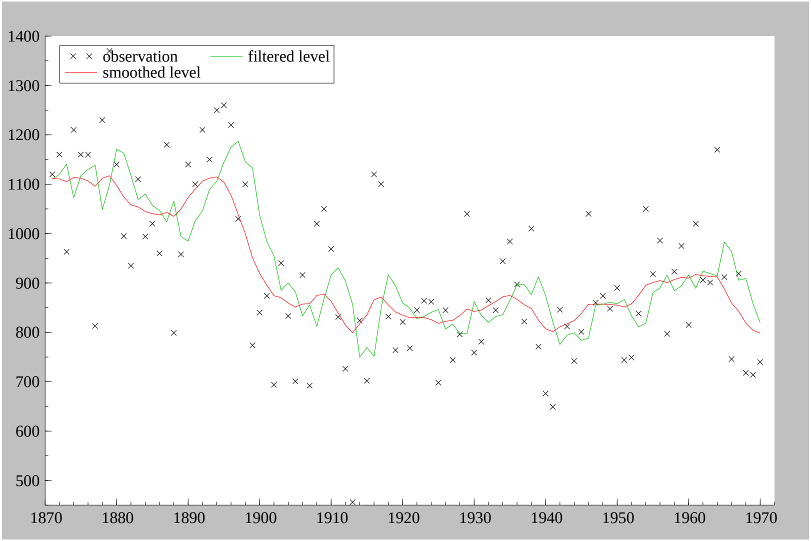
Starting with $r_n = 0$, $N_n = 0$, the smoothing recursions are given by

$$r_{t-1} = F_t^{-1} v_t + L_t r_t, \qquad N_{t-1} = F_t^{-1} + L_t' N_t L_t,$$
$$\hat{\alpha}_t = a_t + P_t r_{t-1}, \qquad V_t = P_t - P_t N_{t-1} P_t.$$

Smoothing Illustration



Filtering and Smoothing



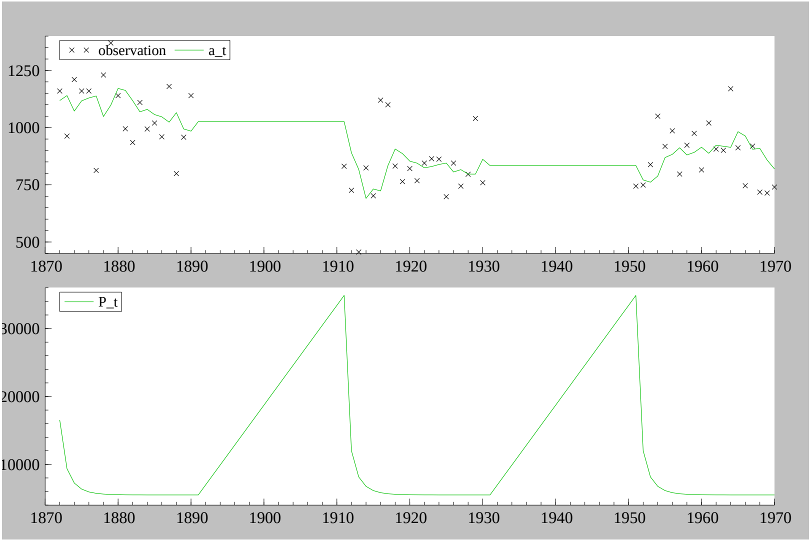
Missing Observations

Missing observations are very easy to handle in Kalman filtering:

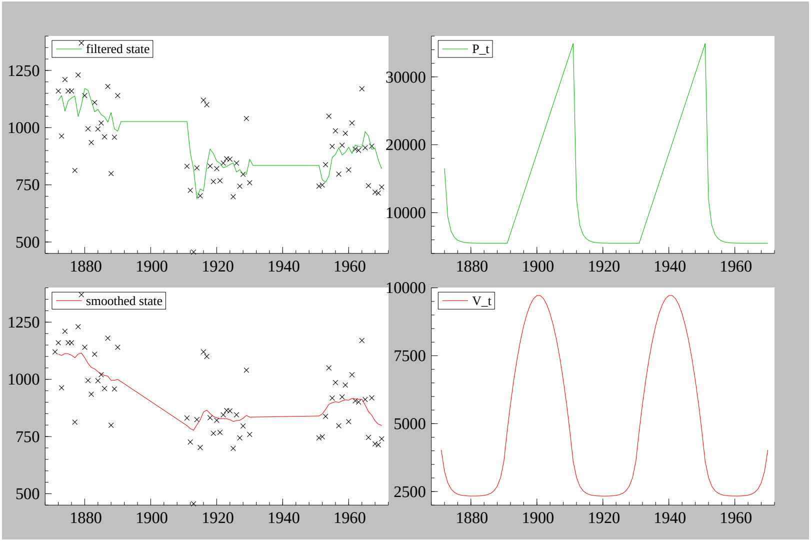
- suppose y_j is missing
- put $v_j = 0, K_j = 0$ and $F_j = \infty$ in the algorithm
- proceed further calculations as normal

The filter algorithm extrapolates according to the state equation until a new observation arrives. The smoother interpolates between observations.

Missing Observations



Missing Observations, Filter and Smoohter

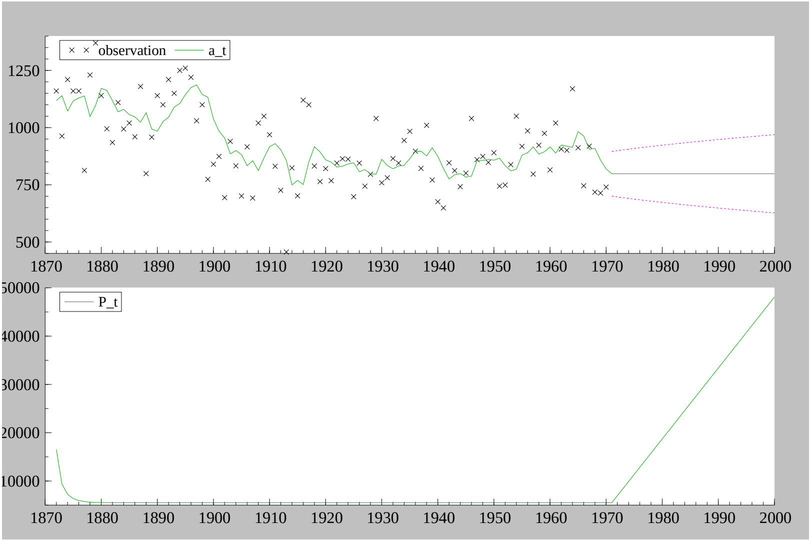


Forecasting

Forecasting requires no extra theory: just treat future observations as missing:

- put $v_j = 0, K_j = 0$ and $F_j = \infty$ for $j = n + 1, \ldots, n + k$
- proceed further calculations as normal
- forecast for y_j is $Z_j a_j$

Forecasting



Parameter Estimation

The system matrices in a state space model typically depends on a parameter vector ψ. The model is completely Gaussian; we estimate by Maximum Likelihood.
The loglikelihood af a time series is

$$\log L = \sum_{t=1}^{n} \log p(y_t | Y_{t-1}).$$

In the state space model, $p(y_t | Y_{t-1})$ is a Gaussian density with mean a_t and variance F_t:

$$\ell = \log L = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{n} \left(\log |F_t| + v_t' F_t^{-1} v_t \right),$$

with v_t and F_t from the Kalman filter. This is called the *prediction error decomposition* of the likelihood. Estimation proceeds by numerically maximising ℓ.

Diagnostics

- Null hypothesis: standardised residuals

$$v_t/\sqrt{F}_t \sim \mathcal{NID}(0, 1)$$

- Apply standard test for Normality, heteroskedasticity, serial correlation;
- A recursive algorithm is available to calculate smoothed disturbances (auxilliary residuals), which can be used to detect breaks and outliers;
- Model comparison and parameter restrictions: use likelihood based procedures (LR test, AIC, BIC).

Ox, SSFPack and STAMP

- Ox can be freely downloaded from
  `http://www.doornik.com/download.html`
- SsfPack can be freely downloaded from
  `http://www.ssfpack.com/download.html`
- Documentation
  `http://www.ssfpack.com/documentation.html`
  and `http://www.ssfpack.com/files/SsfEctJ.pdf`
- STAMP
  (Structural Time Series Analyser, Modeller and Predictor)
  `http://stamp-software.com/`

Hedonic Price Model with Time Fixed Effects (1)

- Model specification

$$y_{it} = \mu_t + x_{it}\beta + \varepsilon_{it}, \varepsilon_{it} \sim \mathcal{NID}(0, \sigma^2), \ t = 1, \dots, T, \ i = 1, \dots, n_t$$

- Model estimation: define $\tilde{y}_{it} = y_{it} - \bar{y}_{.t}$ and $\tilde{x}_{it} = x_{it} - \bar{x}_{.t}$

$$\hat{\beta} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{y}, \quad \text{Var}(\hat{\beta}) = \sigma^2(\tilde{X}'\tilde{X})^{-1},$$

$$\hat{\mu}_t = \bar{y}_{.t} - \bar{x}_{.t}\hat{\beta}, \quad \text{Var}(\hat{\mu}_t) = \sigma^2/n_t + \sigma^2\bar{x}_{.t}(\tilde{X}'\tilde{X})^{-1}\bar{x}'_{.t},$$

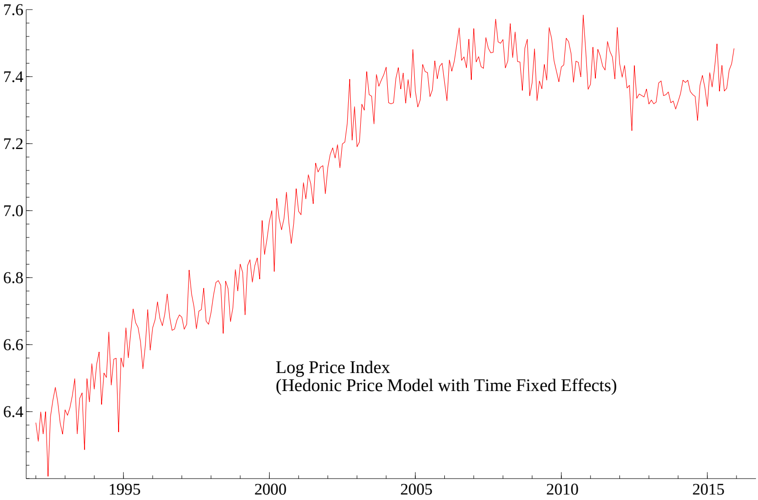
$$-2\ell = \ln(2\pi\sigma^2) + \frac{RSS}{m},$$

$$RSS = (\tilde{y} - \tilde{X}\hat{\beta})'(\tilde{y} - \tilde{X}\hat{\beta}), \quad m = \sum_{t=1}^{T} n_t - T - k,$$

Hedonic Price Model with Time Fixed Effects (2)

	coef	sd	t-value
HouseSize	0.6609	0.0104	63.29
LotSize	0.1014	0.0070	14.41
Construction Period 1931-1944	0.0368	0.0308	1.19
Construction Period 1945-1959	-0.0975	0.0197	-4.94
Construction Period 1960-1970	-0.1131	0.0170	-6.65
Construction Period 1971-1980	-0.0469	0.0199	-2.36
Construction Period 1981-1990	-0.0349	0.0188	-1.86
Construction Period 1991-2000	-0.0161	0.0186	-0.87
Construction Period > 2001	0.0202	0.0209	0.97
σ	0.1202		
Number of obs.	4100		
Number of regressors	24		
Number of time fixed effects	288		

Hedonic Price Model with Time Fixed Effects (3)



Hedonic Price Model with Local Linear Trend (1)

- Model specification

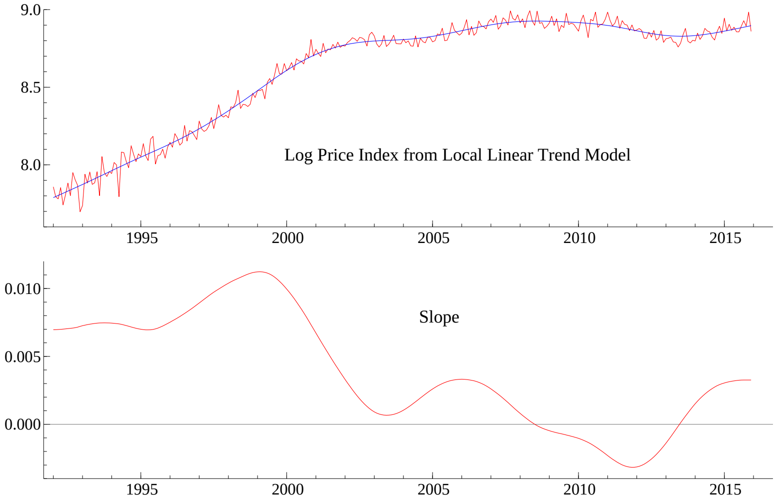
$$\begin{aligned} y_{it} &= \mu_t + x_{it}\beta + \varepsilon_{it}, & \varepsilon_{it} &\sim \mathcal{NID}(0, \sigma^2), \\ \mu_{t+1} &= \mu_t + \kappa_t + \eta_t, & \eta_t &\sim \mathcal{NID}(0, \sigma_\eta^2), \\ \kappa_{t+1} &= \kappa_t + \xi_t, & \xi_t &\sim \mathcal{NID}(0, \sigma_\xi^2). \end{aligned}$$

- Model estimation
  - Split the observations in means and deviations from means
  - Estimate β on deviations data $\tilde{y}$ by OLS
  - Apply Kalman filter on means data $\bar{y}$
    - State vector $\alpha_t = (\beta_t', \mu_t, \kappa_t)'$ (note that $\beta_{t+1} = \beta_t + 0$)
    - Use $\hat{\beta}$ and Var($\hat{\beta}$) as initial condition for β
    - The total loglikelihood is the sum of the lglikelihood produced by the Kalman filter and the loglikelihood from the OLS part

Hedonic Price Model with Local Linear Trend (2)

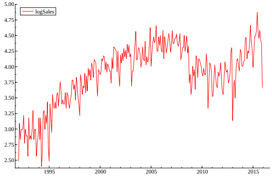
	OLS (Time Fixed Effects)			OLS + KF		
	coef	sd	t-value	coef	sd	t-value
House Size	0.6609	0.0104	63.29	0.6613	0.0105	63.22
Lot Size	0.1014	0.0070	14.41	0.1013	0.0071	14.35
Construction Period 1931-1944	0.0368	0.0308	1.19	0.0374	0.0309	1.21
Construction Period 1945-1959	-0.0975	0.0197	-4.94	-0.0979	0.0198	-4.96
Construction Period 1960-1970	-0.1131	0.0170	-6.65	-0.1127	0.0170	-6.62
Construction Period 1971-1980	-0.0469	0.0199	-2.36	-0.0465	0.0199	-2.34
Construction Period 1981-1990	-0.0349	0.0188	-1.86	-0.0354	0.0188	-1.88
Construction Period 1991-2000	-0.0161	0.0186	-0.87	-0.0166	0.0186	-0.89
Construction Period $>$ 2001	0.0202	0.0209	0.97	0.0202	0.0209	0.97
σ	0.1202					0.1208
σ_η						0.0000
σ_ξ						0.0010
Number of obs.	4100					
Number of regressors	24					
Number of time fixed effects	288					

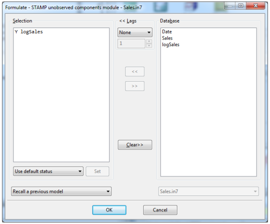
Hedonic Price Model with Local Linear Trend (3)

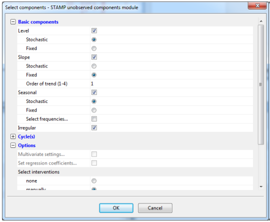


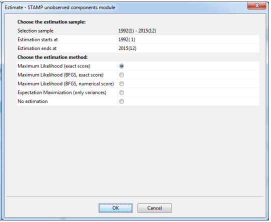
Number of Sales (1)

- Decompose in trend, seasaonal, irregular using STAMP

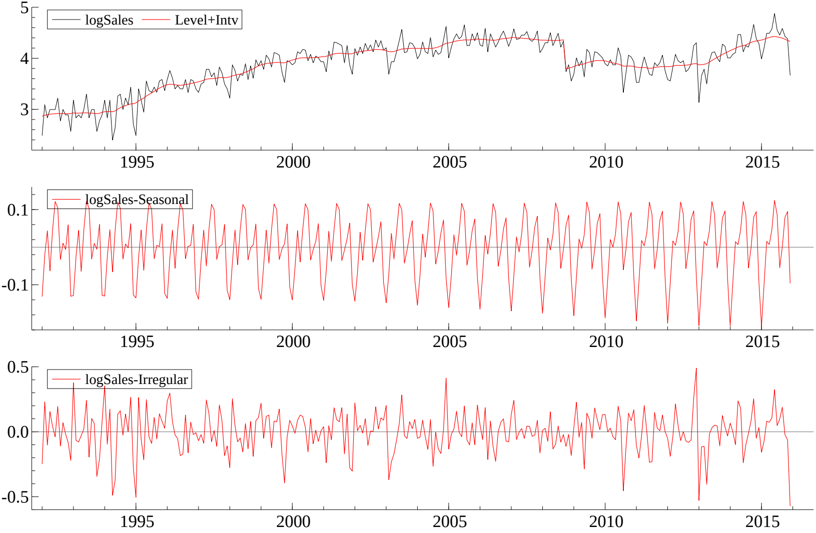




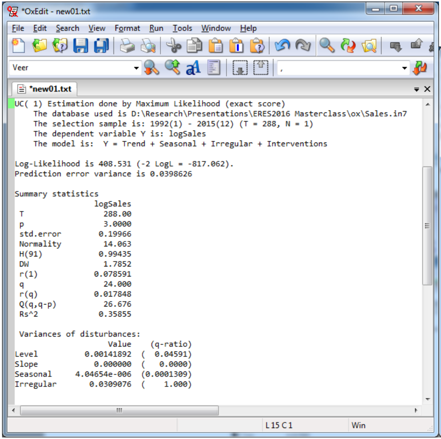


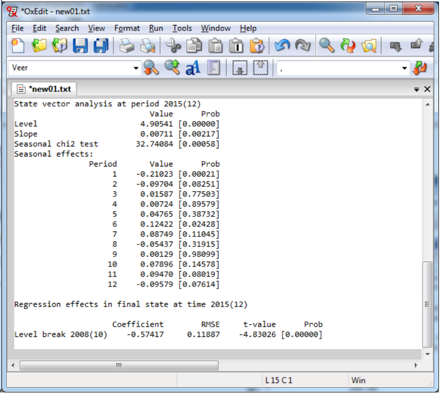


Number of Sales (2)



Number of Sales (3)





Questions?