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 23\textsuperscript{th} Annual Conference Regensburg 2016
Time Series

Unobserved Component Models

Linear Gaussian State Space Models

Examples Programs

Thanks Kai Ming Lee for many of the slides
Examples

Log Price Index
(Hedonic Price Model with Time Fixed Effects)

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

\[ y_t = \mu_t + \gamma_t + \varepsilon_t, \quad t = 1, \ldots, n, \]

also known as the Classical Decomposition.

- \( y_t \) = observation,
- \( \mu_t \) = slowly changing component (trend),
- \( \gamma_t \) = periodic component (seasonal),
- \( \varepsilon_t \) = irregular component (disturbance).
Local Level Model

- Components can be
  - deterministic functions of time (e.g. polynomials), or
  - stochastic processes;
- Examples
  - Deterministic: linear trend
    \[ y_t = \delta_0 + \delta_1 t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2_\varepsilon) \]
  - Stochastic: Random Walk plus Noise, or Local Level model:
    \[ y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2_\varepsilon) \]
    \[ \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma^2_\eta), \]
- Initial condition: \[ \mu_1 \sim \mathcal{N}(a, P); \]
- The disturbances \( \varepsilon_t, \eta_s \) 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

\[ y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2_{\varepsilon}) \]
\[ \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma^2_{\eta}) \]
\[ \mu_1 \sim \mathcal{N}(a, P) \]

- The level \( \mu_t \) and the error term \( \varepsilon_t \) are unobserved;
- Parameters: \( a, P, \sigma^2_{\varepsilon}, \sigma^2_{\eta} \);
- Trivial special cases:
  - \( \sigma^2_{\eta} = 0 \implies y_t \sim \mathcal{N}(\mu_1, \sigma^2_{\varepsilon}) \) (White Noise with constant level);
  - \( \sigma^2_{\varepsilon} = 0 \implies y_{t+1} = y_t + \eta_t \) (pure Random Walk);
Simulated LL Data

\[ \sigma_\varepsilon = 0.1, \quad \sigma_\eta = 1 \]

\[ \sigma_\varepsilon = \sigma_\eta = 1 \]

\[ \sigma_\varepsilon = 1, \quad \sigma_\eta = 0.1 \]
Properties of the LL model

\[ y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \]
\[ \mu_{t+1} = \mu_t + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma^2), \]

- First difference is stationary:
\[ \Delta y_t = \Delta \mu_t + \Delta \varepsilon_t = \eta_{t-1} + \varepsilon_t - \varepsilon_{t-1}. \]

- Dynamic properties of \( \Delta y_t \):
\[ E(\Delta y_t) = 0, \]
\[ \gamma_0 = E(\Delta y_t \Delta y_t) = \sigma^2 + 2\sigma^2, \]
\[ \gamma_1 = E(\Delta y_t \Delta y_{t-1}) = -\sigma^2, \]
\[ \gamma_\tau = E(\Delta y_t \Delta y_{t-\tau}) = 0 \quad \text{for} \quad \tau \geq 2. \]
Properties of the LL model

- The ACF of $\Delta y_t$ is
  \[
  \rho_1 = \frac{-\sigma_\varepsilon^2}{\sigma_\eta^2 + 2\sigma_\varepsilon^2} = -\frac{1}{q + 2},
  \]
  \[q = \frac{\sigma_\eta^2}{\sigma_\varepsilon^2},\]
  \[
  \rho_\tau = 0, \quad \tau \geq 2.
  \]

- $q$ is called the \textit{signal-noise ratio};

- The model for $\Delta y_t$ 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:

\[
\begin{align*}
    y_t &= \mu_t + \varepsilon_t, \\
    \mu_{t+1} &= \beta_t + \mu_t + \eta_t, \\
    \beta_{t+1} &= \beta_t + \xi_t,
\end{align*}
\]

\[
\begin{align*}
    \varepsilon_t &\sim \mathcal{NID}(0, \sigma_\varepsilon^2), \\
    \eta_t &\sim \mathcal{NID}(0, \sigma_\eta^2), \\
    \xi_t &\sim \mathcal{NID}(0, \sigma_\xi^2).
\end{align*}
\]

- All disturbances are independent at all lags and leads;
- Initial distributions \( \beta_1, \mu_1 \) need to specified;
- Special cases
  - If \( \sigma_\xi^2 = 0 \) the trend is a random walk with constant drift \( \beta_1 \);
    (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 \( \beta_1 \) and intercept \( \mu_1 \);
  - 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

\[ \mu \] (solid blue line)

\[ \beta \] (dashed red line)
Seasonal Effects

We have seen specifications for $\mu_t$ in the basic model

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

Now we will consider the seasonal term $\gamma_t$. 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:

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

The expectation of the sum of the seasonal effects is zero.
**Trigonometric Seasonal**

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

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

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

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

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

- 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

- **Data and trend+intervention**
- **Seasonal**
- **Irregular**
Cycles

We can extend the basic model with cycle $\psi_t$

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

where $\psi_t$ can be deterministic

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

or stochastic

$$\begin{align*}
\psi_{t+1} &= \rho [\psi_t \cos \lambda + \psi^*_t \sin \lambda] + \kappa_t, \\
\psi^*_{t+1} &= \rho [-\psi_t \sin \lambda + \psi^*_t \cos \lambda] + \kappa^*_t,
\end{align*}$$

$$\kappa_t, \kappa^*_t \sim \mathcal{NID}(0, \sigma^2).$$
State Space Model: a more general class of models

Linear Gaussian state space model is defined in three parts:

→ State equation:

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

→ Observation equation:

\[ y_t = Z_t \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{NID}(0, H_t), \]

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

Notice that

- \( \zeta_t \) and \( \varepsilon_s \) independent for all \( t, s \), and independent from \( \alpha_1 \);
- observation \( y_t \) can be multivariate;
- state vector \( \alpha_t \) 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 $\alpha_t$ 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:

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

Put regressors in \( Z_t \),

\[
T_t = I, \quad R_t = I,
\]

Result is regression model with coefficient \( \alpha_t \) 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:

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

with 2 \times 1 state vector \( \alpha_t \) and system matrices:

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

- \( Z_t \) 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:

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

with $2 \times 1$ state vector $\alpha_t$ and system matrices:

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

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

\[ \alpha_t = \begin{bmatrix} y_t \\ \phi_2 y_{t-1} + \theta \zeta_t \end{bmatrix} \]
\[ Z_t = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad H_t = 0, \]
\[ T_t = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix}, \quad R_t = \begin{bmatrix} 1 \\ \theta \end{bmatrix}, \quad Q_t = \sigma^2 \]

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 \( y_t = \mu_t + \varepsilon_t: \)

\[
\begin{align*}
\alpha_t &= \mu_t, \\
T_t &= 1, \\
R_t &= 1, \\
Q_t &= \sigma_\eta^2, \\
Z_t &= 1, \\
H_t &= \sigma_\varepsilon^2.
\end{align*}
\]

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

\[
\begin{align*}
\alpha_t &= \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix}, \\
T_t &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \\
R_t &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\
Q_t &= \begin{bmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\xi^2 \end{bmatrix}, \\
Z_t &= \begin{bmatrix} 1 & 0 \end{bmatrix}, \\
H_t &= \sigma_\varepsilon^2.
\end{align*}
\]
UC models in State Space Form

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

LLT model with season: \( \Delta \mu_{t+1} = \beta_t + \eta_t \), \( \Delta \beta_{t+1} = \xi_t \), \( S(L) \gamma_{t+1} = \omega_t \) 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}, \\
Q_t = \begin{bmatrix} \sigma^2_\eta & 0 & 0 \\
0 & \sigma^2_\xi & 0 \\
0 & 0 & \sigma^2_\omega \end{bmatrix}, \\
R_t = \begin{bmatrix} 1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \end{bmatrix}, \\
Z_t = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \end{bmatrix}, \\
H_t = \sigma^2_\varepsilon.
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
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.
The unobserved state $\alpha_t$ can be estimated from the observations with the *Kalman filter*. Define $Y_t = \{y_1, \ldots, y_t\}$, $a_{t+1} = \mathbb{E}(\alpha_{t+1} | Y_t)$, $P_{t+1} = \text{Var}(\alpha_{t+1} | Y_t)$.

$$v_t = y_t - Z_t a_t,$$

$$F_t = Z_t P_t Z_t' + H_t,$$

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

$$a_{t+1} = T_t a_t + K_t v_t,$$

$$P_{t+1} = T_t P_t T_t' + R_t Q_t R_t' - K_t F_t K_t',$$

for $t = 1, \ldots, n$ 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

\[
a_{t+1} = \mathbb{E}(\alpha_{t+1} \mid Y_t), \quad P_{t+1} = \text{Var}(\alpha_{t+1} \mid Y_t);
\]

- The prediction error is

\[
v_t = y_t - \mathbb{E}(y_t \mid Y_{t-1})
\]

\[
= y_t - \mathbb{E}(Z_t \alpha_t + \varepsilon_t \mid Y_{t-1})
\]

\[
= y_t - Z_t \mathbb{E}(\alpha_t \mid Y_{t-1})
\]

\[
= y_t - Z_t a_t;
\]

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

\[
\begin{pmatrix}
  x \\
  y \\
  z
\end{pmatrix}
\sim \mathcal{N}
\left( 
\begin{pmatrix}
  \mu_x \\
  \mu_y \\
  0
\end{pmatrix},
\begin{pmatrix}
  \Sigma_{xx} & \Sigma_{xy} & \Sigma_{xz} \\
  \Sigma_{yx} & \Sigma_{yy} & \Sigma_{yz} \\
  \Sigma_{zx} & \Sigma_{zy} & \Sigma_{zz}
\end{pmatrix}
\right)
\]

Then

\[
\begin{align*}
\mathbb{E}(x|y, z) &= \mathbb{E}(x|y) + \Sigma_{xz} \Sigma_{zz}^{-1} z, \\
\text{Var}(x|y, z) &= \text{Var}(x|y) - \Sigma_{xz} \Sigma_{zz}^{-1} \Sigma_{zx},
\end{align*}
\]
Kalman Filter

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

- \( Y_t = \{ Y_{t-1}, y_t \} = \{ Y_{t-1}, v_t \}, \) \( \text{E}(v_t y_{t-j}) = 0, \ j = 1, \ldots, t - 1 \)

- Apply Lemma

\[
\begin{align*}
    x & = \alpha_{t+1}, \quad \text{E}(x|y) = \text{E}(\alpha_{t+1}|Y_{t-1}) = T_t a_t, \\
    y & = Y_{t-1}, \quad \text{Var}(x|y) = \text{Var}(\alpha_{t+1}|Y_{t-1}) = F_t, \\
    z & = v_t, \quad \Sigma_{zz} = \text{Var}(v_t) = F_t, \\
    \Sigma_{xz} &= \text{Cov}(\alpha_{t+1}, v_t) = T_t P_t Z_t', \\
    \text{Cov}(\alpha_{t+1}, v_t) &= \text{Cov}(T_t \alpha_t + R_t \zeta_t, Z_t(\alpha_t - a_t) + \varepsilon_t) = T_t P_t Z_t' \\
    \text{Var}(\alpha_{t+1}|Y_{t-1}) &= T_t P_t T_t' + R_t Q_t R_t'
\end{align*}
\]

- We carry out lemma and obtain the state update

\[
\begin{align*}
    a_{t+1} &= \text{E}(\alpha_{t+1}|Y_{t-1}, y_t) = \text{E}(\alpha_{t+1}|Y_{t-1}, v_t) \\
    &= T_t a_t + T_t P_t Z_t' F_t^{-1} v_t = T_t a_t + K_t v_t; \\
    P_{t+1} &= P_{t+1} = T_t P_t T_t' + R_t Q_t R_t' - K_t F_t K_t'
\end{align*}
\]

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 $\alpha_{t+1}$ 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

- Observation $a_t$
- Filtered level $a_t$
- State variance $P_t$
- Prediction error $v_t$
- Prediction error variance $F_t$
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.

\[
\hat{\alpha}_t = \mathbb{E}(\alpha_t | Y_n), \quad V_t = \text{Var}(\alpha_t | Y_n),
\]

\[
r_t = \text{weighted sum of innovations}, \quad N_t = \text{Var}(r_t),
\]

\[
L_t = T_t - K_t Z_t.
\]

Starting with \( r_n = 0, N_n = 0 \), the smoothing recursions are given by

\[
r_{t-1} = F_t^{-1} v_t + L_t r_t, \quad N_{t-1} = F_t^{-1} + L_t' N_t L_t,
\]

\[
\hat{\alpha}_t = a_t + P_t r_{t-1}, \quad V_t = P_t - P_t N_{t-1} P_t.
\]
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

- Observation: \( a_t \)
- \( P_t \)
Missing Observations, Filter and Smoother
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 $\psi$. The model is completely Gaussian; we estimate by Maximum Likelihood. The loglikelihood of 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 $\ell$. 
Diagnostics

- Null hypothesis: standardised residuals
  \[ v_t / \sqrt{F_t} \sim \mathcal{N}(0, 1) \]
- 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](http://www.doornik.com/download.html)
- **SsfPack** can be freely downloaded from
  [http://www.ssfpack.com/download.html](http://www.ssfpack.com/download.html)
- **Documentation**
- **STAMP**
  *(Structural Time Series Analyser, Modeller and Predictor)*
Hedonic Price Model with Time Fixed Effects (1)

- Model specification
  \[ y_{it} = \mu_t + x_{it}\beta + \varepsilon_{it}, \varepsilon_{it} \sim \mathcal{NID}(0, \sigma^2), \quad t = 1, \ldots, T, \quad i = 1, \ldots, n_t \]

- 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)

<table>
<thead>
<tr>
<th></th>
<th>coef</th>
<th>sd</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HouseSize</td>
<td>0.6609</td>
<td>0.0104</td>
<td>63.29</td>
</tr>
<tr>
<td>LotSize</td>
<td>0.1014</td>
<td>0.0070</td>
<td>14.41</td>
</tr>
<tr>
<td>Construction Period 1931-1944</td>
<td>0.0368</td>
<td>0.0308</td>
<td>1.19</td>
</tr>
<tr>
<td>Construction Period 1945-1959</td>
<td>-0.0975</td>
<td>0.0197</td>
<td>-4.94</td>
</tr>
<tr>
<td>Construction Period 1960-1970</td>
<td>-0.1131</td>
<td>0.0170</td>
<td>-6.65</td>
</tr>
<tr>
<td>Construction Period 1971-1980</td>
<td>-0.0469</td>
<td>0.0199</td>
<td>-2.36</td>
</tr>
<tr>
<td>Construction Period 1981-1990</td>
<td>-0.0349</td>
<td>0.0188</td>
<td>-1.86</td>
</tr>
<tr>
<td>Construction Period 1991-2000</td>
<td>-0.0161</td>
<td>0.0186</td>
<td>-0.87</td>
</tr>
<tr>
<td>Construction Period &gt; 2001</td>
<td>0.0202</td>
<td>0.0209</td>
<td>0.97</td>
</tr>
</tbody>
</table>

$\sigma$  

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of obs.</td>
<td>4100</td>
</tr>
<tr>
<td>Number of regressors</td>
<td>24</td>
</tr>
<tr>
<td>Number of time fixed effects</td>
<td>288</td>
</tr>
</tbody>
</table>
Hedonic Price Model with Time Fixed Effects (3)
Hedonic Price Model with Local Linear Trend (1)

- **Model specification**
  \[ y_{it} = \mu_t + x_{it} \beta + \varepsilon_{it}, \quad \varepsilon_{it} \sim \mathcal{NID}(0, \sigma^2), \]
  \[ \mu_{t+1} = \mu_t + \kappa_t + \eta_t, \quad \eta_t \sim \mathcal{NID}(0, \sigma_{\eta}^2), \]
  \[ \kappa_{t+1} = \kappa_t + \xi_t, \quad \xi_t \sim \mathcal{NID}(0, \sigma_{\xi}^2). \]

- **Model estimation**
  - Split the observations in means and deviations from means
  - Estimate \( \beta \) 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 \( \text{Var}(\hat{\beta}) \) as initial condition for \( \beta \)
    - The total loglikelihood is the sum of the loglikelihood produced by the Kalman filter and the loglikelihood from the OLS part
Hedonic Price Model with Local Linear Trend (2)

<table>
<thead>
<tr>
<th></th>
<th>OLS (Time Fixed Effects)</th>
<th>OLS + KF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coef</td>
<td>sd</td>
</tr>
<tr>
<td>House Size</td>
<td>0.6609</td>
<td>0.0104</td>
</tr>
<tr>
<td>Lot Size</td>
<td>0.1014</td>
<td>0.0070</td>
</tr>
<tr>
<td>Construction Period 1931-1944</td>
<td>0.0368</td>
<td>0.0308</td>
</tr>
<tr>
<td>Construction Period 1945-1959</td>
<td>-0.0975</td>
<td>0.0197</td>
</tr>
<tr>
<td>Construction Period 1960-1970</td>
<td>-0.1131</td>
<td>0.0170</td>
</tr>
<tr>
<td>Construction Period 1971-1980</td>
<td>-0.0469</td>
<td>0.0199</td>
</tr>
<tr>
<td>Construction Period 1981-1990</td>
<td>-0.0349</td>
<td>0.0188</td>
</tr>
<tr>
<td>Construction Period 1991-2000</td>
<td>-0.0161</td>
<td>0.0186</td>
</tr>
<tr>
<td>Construction Period &gt; 2001</td>
<td>0.0202</td>
<td>0.0209</td>
</tr>
<tr>
<td>σ</td>
<td>0.1202</td>
<td></td>
</tr>
<tr>
<td>ση</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>σξ</td>
<td>0.0010</td>
<td></td>
</tr>
<tr>
<td>Number of obs.</td>
<td>4100</td>
<td></td>
</tr>
<tr>
<td>Number of regressors</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Number of time fixed effects</td>
<td>288</td>
<td></td>
</tr>
</tbody>
</table>
Hedonic Price Model with Local Linear Trend (3)
Number of Sales (1)

- Decompose in trend, seasonal, irregular using STAMP
Number of Sales (2)

- **logSales**
- **Level + Intv**
- **logSales-Seasonal**
- **logSales-Irregular**

Graphs showing the trend and components of the number of sales from 1995 to 2015.
Number of Sales (3)

Unobserved Component Models
Linear Gaussian State Space Models
Examples Programs
Questions?