A bivariate model of electricity loads and prices

Katarzyna Maciejowska

The Energy Finance Christmas Workshop

Wroclaw, 19-20.12.2011

- When should we use multivariate models?.
- A bivariate $VAR(p)$ model.
- Empirical example
  - A $VAR(p)$ model of energy prices and loads.
  - Identification of the structure of the VAR model.
  - Should we use bivariate model (formal testing)?
  - Comparison of the predicting performance of competing specifications.
- Summary

- When should we use multivariate models?.
- A bivariate $VAR(p)$ model.
- Empirical example
  - A $VAR(p)$ model of energy prices and loads.
  - Identification of the structure of the VAR model.
  - Should we use bivariate model (formal testing)?
  - Comparison of the predicting performance of competing specifications.
- Summary

- When should we use multivariate models?.
- A bivariate $VAR(p)$ model.
- Empirical example
  - A $VAR(p)$ model of energy prices and loads.
  - Identification of the structure of the VAR model.
  - Should we use bivariate model (formal testing)?
  - Comparison of the predicting performance of competing specifications.
- Summary

- When should we use multivariate models?.
- A bivariate $VAR(p)$ model.
- Empirical example
  - A $VAR(p)$ model of energy prices and loads.
  - Identification of the structure of the VAR model.
  - Should we use bivariate model (formal testing)?
  - Comparison of the predicting performance of competing specifications.
- Summary

- When should we use multivariate models?.
- A bivariate $VAR(p)$ model.
- Empirical example
  - A $VAR(p)$ model of energy prices and loads.
  - Identification of the structure of the VAR model.
  - Should we use bivariate model (formal testing)?
  - Comparison of the predicting performance of competing specifications.
- Summary

- When should we use multivariate models?.
- A bivariate $VAR(p)$ model.
- Empirical example
  - A $VAR(p)$ model of energy prices and loads.
  - Identification of the structure of the VAR model.
  - Should we use bivariate model (formal testing)?
  - Comparison of the predicting performance of competing specifications.
- Summary

- When should we use multivariate models?.
- A bivariate $VAR(p)$ model.
- Empirical example
    - A $VAR(p)$ model of energy prices and loads.
    - Identification of the structure of the VAR model.
    - Should we use bivariate model (formal testing)?
    - Comparison of the predicting performance of competing specifications.
- Summary

- When should we use multivariate models?.
- A bivariate $VAR(p)$ model.
- Empirical example
  - A $VAR(p)$ model of energy prices and loads.
  - Identification of the structure of the VAR model.
  - Should we use bivariate model (formal testing)?
  - Comparison of the predicting performance of competing specifications.
- Summary

- When should we use multivariate models?.
- A bivariate $VAR(p)$ model.
- Empirical example
  - A $VAR(p)$ model of energy prices and loads.
  - Identification of the structure of the VAR model.
  - Should we use bivariate model (formal testing)?
  - Comparison of the predicting performance of competing specifications.
- Summary

When should we use multivariate models

When should we use multivariate models?

- When we are interested in joined modeling of a behavior of a set of variables (especially, if we want to take into account both short and long term relations).
- When we expect that residuals are correlated (for example: shocks that influence electricity loads and prices are not independent).
- When we use models that depend on a state variable (nonlinear MN, MS models). Multivariate models may improve the estimation efficiency of state parameters and help to give economical interpretation to states.

When should we use multivariate models

When should we use multivariate models?

- When we are interested in joined modeling of a behavior of a set of variables (especially, if we want to take into account both short and long term relations).
- When we expect that residuals are correlated (for example: shocks that influence electricity loads and prices are not independent).
- When we use models that depend on a state variable (nonlinear MN, MS models). Multivariate models may improve the estimation efficiency of state parameters and help to give economical interpretation to states.

When should we use multivariate models

When should we use multivariate models?

- When we are interested in joined modeling of a behavior of a set of variables (especially, if we want to take into account both short and long term relations).
- When we expect that residuals are correlated (for example: shocks that influence electricity loads and prices are not independent).
- When we use models that depend on a state variable (nonlinear MN, MS models). Multivariate models may improve the estimation efficiency of state parameters and help to give economical interpretation to states.

When should we use multivariate models

When should we use multivariate models?

- When we are interested in joined modeling of a behavior of a set of variables (especially, if we want to take into account both short and long term relations).
- When we expect that residuals are correlated (for example: shocks that influence electricity loads and prices are not independent).
- When we use models that depend on a state variable (nonlinear MN, MS models). Multivariate models may improve the estimation efficiency of state parameters and help to give economical interpretation to states.

A bivariate $VAR(p)$ model

In bivariate $VAR(p)$ models, it is assumed that the behavior of endogenous variables depends on their past observations and some deterministic components.

$$Y_t = D_t + A_1 Y_{t-1} + A_2 Y_{t-2} + ... + A_p Y_{t-p} + \varepsilon_t \quad (1)$$

where

- p is an order of autoregression.
- Y_t is a 2×1 vector of endogenous variables.
- D_t is a 2×1 vector od deterministic component.
- A_i are 2×2 matrices
- ε_t is a 2×1 vector of residuals. Often ε_t $N(0, \Sigma)$, where Σ is a 2×2 variance-covariance matrix

A bivariate $VAR(p)$ model

When should we use $VAR(p)$ models instead of two $AR(p)$ models (with the deterministic part containing past values of exogenous variables)?

- When residuals are correlated (Σ is not diagonal). Especially, when the model is used for predictions.
- When we are interested in structural analysis (want to give interpretation to residuals and analyze contemporaneous relations of variables of interest).

A bivariate $VAR(p)$ model - structural analysis

The structure of the $VAR(p)$ is defined by matrices A and B:

$$AY_t = D_t + A(L)Y_{t-1} + Bu_t \tag{2}$$

where

- A is a 2×2 matrix that defines the contemporaneous relationship between the endogenous variables
- u_t is a 2×1 vector of structural shocks that are independent (Σ_u is diagonal). For example, u_t could be independent demand and supply shocks that influence electricity loads and prices.
- B is a 2×2 matrix that defines, how structural shocks influence endogenous variables.
- If residuals are normally distributer then $\Sigma = A^{-1}B\Sigma_u(A^{-1}B)'$

Data

The aim of the project is to model the maximum daily electricity prices for Australia (on the example of NSW). Two half-hourly time series (form 01.01.2006-21.09.2010) are used:

- Total demand (loads)
- Price

Data is transformed. For each day (similar Garcia-Ascanio, Mate (2010))

- P_t is a maximum daily price
- L_t is a maximum daily load

Data

The aim of the project is to model the maximum daily electricity prices for Australia (on the example of NSW). Two half-hourly time series (form 01.01.2006-21.09.2010) are used:

- Total demand (loads)
- Price

Data is transformed. For each day (similar Garcia-Ascanio, Mate (2010))

- P_t is a maximum daily price
- L_t is a maximum daily load

Should we use a bivariate model?

Question: Should we use a bivariate or an univariate model?
Solution: A $VAR(p)$ model can be estimated and the diagonality of the variance-covariance matrix can be tested.
If residuals are correlated, it could be investigated, what is the source of the correlation (matrix A or B)?

Should we use a bivariate model?

Question: Should we use a bivariate or an univariate model?
Solution: A $VAR(p)$ model can be estimated and the diagonality of the variance-covariance matrix can be tested.
If residuals are correlated, it could be investigated, what is the source of the correlation (matrix A or B)?

VAR model

A VAR model is fitted

$$Y_t = D_t^s + \sum_{i=1}^{p} A_i^s Y_{t-i} + \varepsilon_t \tag{3}$$

where

- $Y_t = [\ln P_t, \ln L_t]'$
- D_t^s containing a constant and 0/1 variable defining, which day is a working day and which is a weekend.
- $\varepsilon_t = \Sigma_s$
- $p = 16$ (based on sequential LR tests, $p_{max} = 21$
- s defines the quarter of the year (parameter differs between seasons)

VAR model - diagonality test

Two models with unrestricted (1) and diagonal (2) variance-covariance matrix are estimated. The likelihood ratio LR test is used to verify if residuals are uncorrelated:

$$\begin{array}{ll} H_0: & \text{all } \Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4 \text{ are diagonal (2)} \\ H_1: & \text{in at least one quater } \Sigma_s \text{ is not diagonal} \end{array}$$

Results

LR :	df	p-value
$425,50$	4	0

We can reject the null of uncorrelated residuals.

Structural VAR model

In order to estimated the structural parameters it is assumed that:

- There are two structural shocks: supply shock (u_{1t}) and demand shock (u_{2t})
- Loads are contemporarily inelastic (L_t does not depend on P_t and u_{1t}). Hence matrices A and B are lower triangular.
- Diagonals of matrices A and B are ones.

$$A = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$$

- $u_t \sim N(0, \Psi_s)$ with Ψ_s diagonal.
- Loads depends on more lags Y_t then prices (or on a larger set of variables).

Structural VAR model

Under the assumptions:

$$
\begin{array}{lll}
(1) & P_t = & -aL_t + DP_t^s + \sum_{i=1}^{p_1} AP_i^s(L)Y_{t-i} + u_{1t} + bu_{2t} \\
(2) & L_t = & DL_t^s + \sum_{i=1}^{p_2} AL_i^s(L)Y_{t-i} + u_{2t}
\end{array}
$$

The model can be estimated in two steps

- The equation (2) can be estimated and values of the demand shock can be computed ($\hat{u_{2t}}$).
- Estimates $\hat{u_{2t}}$ can be plugged into the equation (1) and the parameters can be estimated.

Specification of the demand equation

Since there are evidence that (2) has a MA(2) component we modify the demand equation

$$L_t = DL_t^s + \sum_{i=1}^{14} AL_i^s(L) Y_{t-i} + u_{2t} + \gamma_1 u_{2t-1} + \gamma_2 u_{2t-2}$$

Then the predicted values are

$$\hat{L}_t = DL_t^s + \sum_{i=1}^{14} AL_i^s(L) Y_{t-i} + \gamma_1 \hat{u}_{2t-1} + \gamma_2 \hat{u}_{2t-2}$$

and $\hat{u}_{2t} = L_t - \hat{L}_t$
The AIC criteria indicate either $p_2 = 14$ or $p_2 = 13$. Hence, $p_2 = 14$ is chosen in further analysis.

Specification of the price equation

We investigate two situations ($p_1 = 12$ based on sequential testing under the assumption $p_1 < 14$):

- L_t is known at the time t (i)

$$P_t = -aL_t + DP_t^s + \sum_{i=1}^{12} AP_i^s(L)Y_{t-i} + u_{1t} + b\hat{u_{2t}}$$

- L_t is not known at the time t (ii)

$$P_t = \tilde{DP}_t^s + \sum_{i=1}^{14} \tilde{AP}_i^s(L)Y_{t-i} + u_{1t} - a(\gamma_1 \hat{u}_{2t-1} + \gamma_2 \hat{u}_{2t-2})$$

There is no need for a bivariate model if:

- (i), $b = 0$
- (ii), $a = 0$

Structural model - results

Results

- L_t is known at the time t (i)

Constrain	LogL	LR	df	p-value
no	$-1229,04$			
$a=0$	$-1240,45$	$22,826$	4	0
$b=0$	$-1232,94$	$7,813$	4	$0,098$

- L_t is not known at the time t (ii)

Constrain	LogL	LR	df	p-value
no	$-1428,33$			
$a=0$	$-1433,68$	$10,696$	4	$0,030$

Results

The results indicates that

- For a model (i), loads are important explanatory variable but demand shocks have only weak influence on prices ($p - value = 0,098$, significant for $\alpha = 0,1$). Hence there are some weak evidence for a bivariate model.
- For a model (ii), lagged demand shocks have a significant impact on the prices and therefore both equations need to be estimated.

Comparison of the prediction performance

Predictions are compared on the basis of the RMSE and MAE

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \hat{u}^2_{2,T_0+i}}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |\hat{u}_{2,T_0+i}|$$

Comparison of the forecast performance

Results for a one day ahead forecast
RMSE

period/model	(i)	(i), $b = 0$	(ii)	(i), $a = 0$
$1Q$	$0,707$	$0,706$	$0,851$	$0,866$
$2Q$	$0,595$	$0,594$	$0,682$	$0,695$
$4Q$	$0,630$	$0,628$	$0,730$	$0,745$

MAE

period/model	(i)	(i), $b = 0$	(ii)	(i), $a = 0$
$1Q$	$0,473$	$0,475$	$0,559$	$0,573$
$2Q$	$0,396$	$0,398$	$0,459$	$0,474$
$4Q$	$0,484$	$0,486$	$0,558$	$0,584$

Summary

Summary

- There are some evidence that demand shocks (u_t) influence contemporaneously Prices.
- The advantages of using a bivariate model are more evident, when Loads are not directly observed at time t.
- Inclusion of estimated demand shocks in the model may improve the forecast accuracy (especially, when Loads are also predicted).

Further research:

- Normality tests reject strongly the null of normality of shocks. Therefore, some modifications of the model should be proposed: shocks could be modeled with MS or a MN.
- The analysis should be extended to other regions/countries.