Multivariate Time Series Analysis

Junye Li

Finance Department
ESSEC Business School

Vector Autoregression VAR

- So far, we have focused mostly on models where y depends on past y's.
- More generally, we might be interested in considering models for more than one variable.
- If we only care about forecasting one series but want to use information from another series, we can estimate an ARMA model and include additional explanatory variables.

- For example, if y_t is the series of interest, but we think x_t might be useful when we estimate the model

$$y_t = \beta_0 + \beta_1 y_{t-1} + \gamma x_{t-1} + \varepsilon_t$$

- This model can be estimated by least squares. Our dependent variable is y_t and the independent variables are y_{t-1} and x_{t-1} ;
- Once the model is fitted, the forecasting can be implemented.

- The 1-step-ahead forecast:

$$E(y_{t+1} \mid F_t) = \beta_0 + \beta_1 E(y_t \mid F_t) + \gamma E(x_t \mid F_t) = \beta_0 + \beta_1 y_t + \gamma x_t$$

- A joint model for y_t and x_t is required if we are interested in multi-step-ahead forecasts, or if we are interested in feedback effects from one process to the other

$$E(y_{t+2} \mid F_t) = \beta_0 + \beta_1 E(y_{t+1} \mid F_t) + \gamma E(x_{t+1} \mid F_t)$$

What do we use here?

- We need a model for x_t as well: Multivariate!!!

Weak Stationarity and Cross-Correlation

- Consider we have k-dimensional multivariate time series $\mathbf{y}_t = (y_{1t}, y_{2t}, \ldots, y_{kt})$.
- The series is weakly stationary if its first and second moments are time invariant.
- Define its mean vector and covariance matrix:

$$\boldsymbol{\mu} = E[\mathbf{y}_t]; \quad \boldsymbol{\Gamma}_0 = E[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})']$$

where the expectation is taken element by element over the joint distribution.

- Let ***D*** be a *k* x *k* diagonal matrix consisting of the standard deviations of y_{it} for i = 1, 2, …, k. The concurrent, or lag-zero, cross-correlation matrix is

$$\boldsymbol{\rho}_0 = [\rho_{ij}(0)] = \mathbf{D}^{-1}\,\boldsymbol{\Gamma}_0\,\mathbf{D}^{-1}$$

- More specifically, the *(i, j)th* element of $\boldsymbol{\rho}_0$ is

$$\rho_{ij}(0) = \frac{\Gamma_{ij}(0)}{\sqrt{\Gamma_{ii}(0)\Gamma_{jj}(0)}} = \frac{\operatorname{cov}(y_{it}, y_{jt})}{std(y_{it})\,std(y_{jt})}$$

- An important topic in multivariate time series analysis is the lead-lag relationships between component series.
- We define the lag-l cross-covariance matrix is

$$\boldsymbol{\Gamma}_l = E[(\mathbf{y}_t - \boldsymbol{\mu}) (\mathbf{y}_{t-l} - \boldsymbol{\mu})']$$

- Therefore, the (i, j)th element of $\boldsymbol{\Gamma}_l$ is the covariance between y_{it} and y_{jt-l}.
- For a weakly stationary series, the cross-covariance matrix $\boldsymbol{\Gamma}_l$ is a function of l, not time index t.

- The lag-l cross-correlation matrix is

$$\boldsymbol{\rho}_l = [\rho_{ij}(l)] = \mathbf{D}^{-1}\, \boldsymbol{\Gamma}_l\, \mathbf{D}^{-1}$$

- The cross-correlation matrix $\boldsymbol{\rho}_l$ of a weakly stationary multivariate time series contain the following information:
  - The diagonal element $\rho_{ii}(l)$ are the autocorrelation function of y_{it}
  - The off-diagonal element $\rho_{ij}(0)$ measures the concurrent linear relationship between y_{it} and y_{jt}.
  - For $l > 0$, the off-diagonal $\rho_{ij}(l)$ measures the linear dependence of y_{it} on the past value y_{jt-l}

Multivariate Portmanteau Test

- The univariate Ljung-Box statistic has been generalized to the multivariate case.
- The test is used to test that there are no auto- and cross-correlations in the vector series $\mathbf{y}_t$.
- This statistic assumes the form

$$Q_k(m) = T^2 \sum_{l=1}^{m} \frac{1}{T-l} tr\left(\hat{\Gamma}_l' \hat{\Gamma}_0^{-1} \hat{\Gamma}_l \hat{\Gamma}_0^{-1}\right)$$

which follows a chi-squared distribution with the degree of freedom of $k^2 m$.

The VAR(1) Model

- If $Q_k(m)$ rejects the null hypothesis, we then build a multivariate model to study the lead-lag relationship between components.
- Suppose we have two variables and we consider the joint model as follows:

$$x_t = \beta_0^x + \beta_1^x x_{t-1} + \beta_2^x y_{t-1} + \varepsilon_t^x$$
$$y_t = \beta_0^y + \beta_1^y x_{t-1} + \beta_2^y y_{t-1} + \varepsilon_t^y$$

- This is the simplest bivariate VAR(1) model.

- Each equation is like an AR(1) model with one other explanatory variable.
- Each equation depends on its own lag and the lag of the other variable.
- We also have now two error terms, one for each equation: ε^y_t and ε^x_t .
- In matrix form:

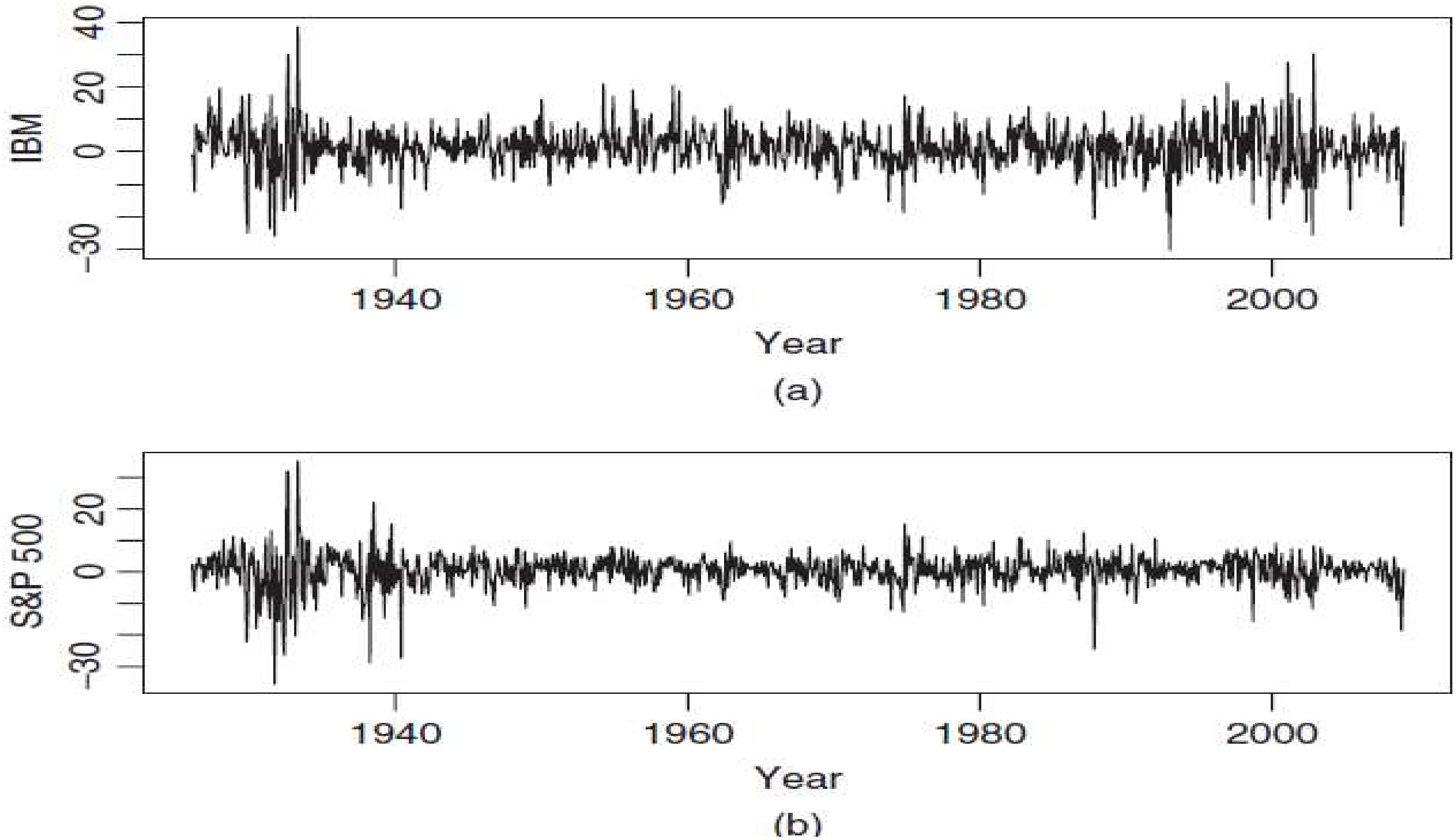
$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \beta_0^x \\ \beta_0^y \end{bmatrix} + \begin{bmatrix} \beta_1^x & \beta_2^x \\ \beta_1^y & \beta_2^y \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^y \end{bmatrix}$$

Assumptions on Error Terms

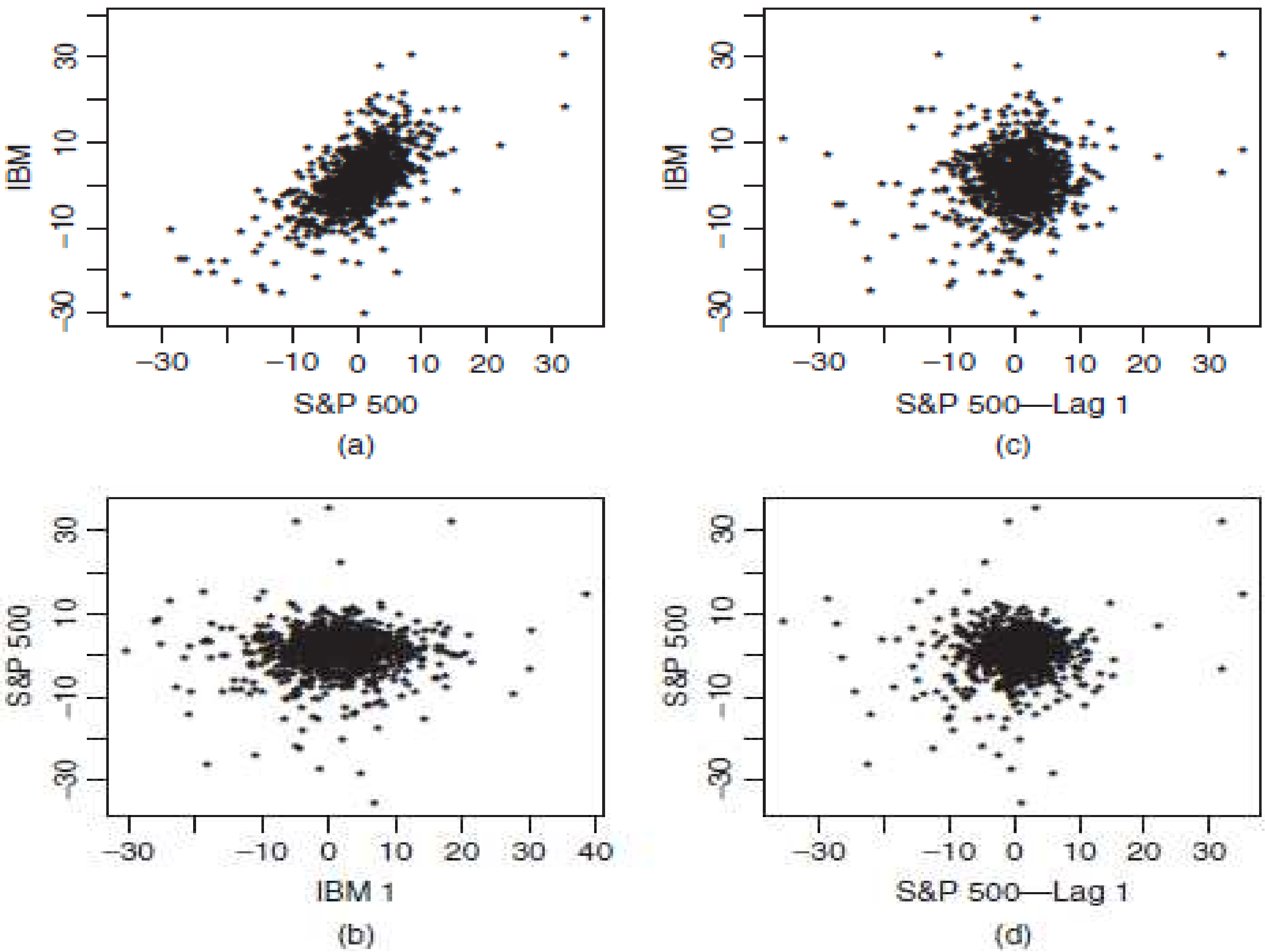
- *Assumption 1*: The errors are uncorrelated over time
  - ε^x_t is uncorrelated with ε^y_{t-j}
  - and ε^y_t is uncorrelated with ε^x_{t-j} for $j > 0$.
- *Assumption 2*: the ε's are iid, but contemporaneously correlated.

$$\Omega = \begin{bmatrix} \underbrace{\sigma^2_{\varepsilon_x}}_{\text{Variance of x}} & \underbrace{\sigma_{\varepsilon_x \varepsilon_y}}_{\text{Contemporaneous covariance}} \\ \sigma_{\varepsilon_x \varepsilon_y} & \underbrace{\sigma^2_{\varepsilon_y}}_{\text{Variance of y}} \end{bmatrix}$$

An Example



40
20
0
-30
IBM
1940
1960
1980
2000
Year
(a)
20
0
-30
S&P 500
1940
1960
1980
2000
Year
(b)


IBM
S&P 500
(a)
IBM
S&P 500—Lag 1
(c)
S&P 500
IBM 1
(b)
S&P 500
S&P 500—Lag 1
(d)

Matrix Notation

- We define

$$\mathbf{y}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix}, \mathbf{v}_t = \begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^y \end{bmatrix}, \ \boldsymbol{\beta}_0 = \begin{bmatrix} \beta_0^x \\ \beta_0^y \end{bmatrix}, \text{ and } \boldsymbol{\beta}_1 = \begin{bmatrix} \beta_1^x & \beta_2^x \\ \beta_1^y & \beta_2^y \end{bmatrix}$$

- Then, our VAR(1) model can be written in

$$\mathbf{y}_t = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \mathbf{y}_{t-1} + \mathbf{v}_t$$

where elements of v_t are iid and the variance-covariance matrix $E(\mathbf{v}_t \mathbf{v}_t') = \Omega$ and $E\left(\mathbf{v}_t \mathbf{v}_{t-j}'\right) = 0$

Moments of VAR(1)

- Taking expectation, we have

$$E[\mathbf{y}_t] = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 E[\mathbf{y}_{t-1}]$$

$$\Rightarrow \boldsymbol{\mu} \equiv E[\mathbf{y}_t] = (\mathbf{I} - \boldsymbol{\beta}_1)^{-1} \boldsymbol{\beta}_0$$

- Using the above result, we can rewrite our model

$$\mathbf{y}_t - \boldsymbol{\mu} = \boldsymbol{\beta}_1 (\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \mathbf{v}_t$$

$$\Rightarrow \tilde{\mathbf{y}}_t = \boldsymbol{\beta}_1 \tilde{\mathbf{y}}_{t-1} + \mathbf{v}_t$$

- Recursive substitution results in

$$\tilde{\mathbf{y}}_t = \mathbf{v}_t + \boldsymbol{\beta}_1 \mathbf{v}_{t-1} + \boldsymbol{\beta}_1^2 \mathbf{v}_{t-2} + \ldots$$

- We then have

$$\operatorname{cov}(\mathbf{y}_t) \equiv \boldsymbol{\Gamma}_0 = \Omega + \boldsymbol{\beta}_1 \Omega \boldsymbol{\beta}'_1 + \boldsymbol{\beta}_1^2 \Omega \boldsymbol{\beta}_1^2{}' + \ldots = \sum_{i=0}^{\infty} \boldsymbol{\beta}_1^i \Omega \boldsymbol{\beta}_1^i{}'$$

- Furthermore,

$$E[\tilde{\mathbf{y}}_t \tilde{\mathbf{y}}'_{t-l}] = \boldsymbol{\beta}_1 E[\tilde{\mathbf{y}}_{t-1} \tilde{\mathbf{y}}'_{t-l}]$$

$$\Rightarrow \boldsymbol{\Gamma}_l = \boldsymbol{\beta}_1 \boldsymbol{\Gamma}_{l-1} \Rightarrow \boldsymbol{\Gamma}_l = \boldsymbol{\beta}^l{}_1 \boldsymbol{\Gamma}_0$$

- and

$$\boldsymbol{\rho}_l = \mathbf{D}^{-1/2} \boldsymbol{\beta} \boldsymbol{\Gamma}_{l-1} \mathbf{D}^{-1/2} = \mathbf{D}^{-1/2} \boldsymbol{\beta} \mathbf{D}^{1/2} \mathbf{D}^{-1/2} \boldsymbol{\Gamma}_{l-1} \mathbf{D}^{-1/2} = \boldsymbol{\gamma} \boldsymbol{\rho}_{l-1}$$

$$\boldsymbol{\gamma} = \mathbf{D}^{-1/2} \boldsymbol{\beta} \mathbf{D}^{1/2}$$

Questions

- How do we interpret the dynamics?
  - There are feedback effects in that each series affects the other series dynamics.

- How do we estimate the model?

- How do we forecast the model?

Interpreting VAR(1) Dynamics

- For a VAR(1),

$$\mathbf{y}_t = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \mathbf{y}_{t-1} + \mathbf{v}_t$$

- If we substitute for $\mathbf{y}_{t-1}$, we get

$$\begin{aligned}\mathbf{y}_t &= \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \left(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \mathbf{y}_{t-2} + \mathbf{v}_{t-1} \right) + \mathbf{v}_t \\ &= \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1^2 \mathbf{y}_{t-2} + \boldsymbol{\beta}_1 \mathbf{v}_{t-1} + \mathbf{v}_t\end{aligned}$$

- Following the same fashion, k-1 times, we get

$$\mathbf{y}_t = \boldsymbol{\beta}_0^* + \boldsymbol{\beta}_1^k \mathbf{y}_{t-k} + \sum_{j=0}^{k-1} \boldsymbol{\beta}_1^j \mathbf{v}_{t-j} \text{ where } \boldsymbol{\beta}_0^* = \left(\sum_{j=0}^{k-1} \boldsymbol{\beta}_1^j \boldsymbol{\beta}_0 \right)$$

- How much does the future values of **y** change when we increase one element of $\mathbf{y}_{t-k}$ by one unit, and keep all previous $\mathbf{y}_{t-j}$ fixed.

- The answer is obtained by taking the derivative of $\mathbf{y}_t$ with respect to $\mathbf{y}_{t-k}$:

$$\frac{d\mathbf{y}_{t,i}}{d\mathbf{y}_{t-k,j}} = \left[\boldsymbol{\beta}_1^k\right]_{i,j}$$

This means (i,j) element of matrix

Impulse Response Function

- When we plot $\left[\beta_1^k\right]_{i,j}$ as a function of k, we see how future values of variable i are impacted by a one unit change in variable j.
- This is called the impulse response function of variable i to a change in variable j.
- This is the primary method used to understand the implied dynamics of a VAR model.

- If we keep substituting, we will obtain the MA representation:

$$\mathbf{y}_t = \boldsymbol{\beta}_1^t y_0 + \boldsymbol{\beta}_0^* + \sum_{j=0}^{\infty} \boldsymbol{\beta}_1^j \mathbf{v}_{t-j}$$

- The derivative of $\mathbf{y}_t$ with respect to elements of past values of $\mathbf{v}_{t-k}$ is the same as the derivative with respect to $\mathbf{y}_{t-k}$ obtained before

$$\frac{d\mathbf{y}_{t,i}}{d\mathbf{v}_{t-k,j}} = \left[\boldsymbol{\beta}_1^k\right]_{i,j}$$

- So powers of the matrix $\boldsymbol{\beta}_1$ determine how a change in one variable today affects the future values.
- Taking powers accommodates the feedback effects from one equation to the other in the right way.
- But here, we do not take into account correlations between the errors.
  - Changes in one error will be correlated with changes in the other error in the same time period if variance-covariance is not diagonal.

- Common solution: take a stand on the way that the shocks propagate.
  - x contemporaneously causes y to change, or the other way around.
- The answer to this question can not be addressed with pure statistics.
  - Economics
- Choosing an order that shocks propagate is equivalent to a choice of orthogonalization
  - Make variance-covariance matrix lower triangular

- There is a natural ordering in the market. For example, the bid and ask prices are posted firstly, and the traders trade at the prevailing bid or ask price:
  - Prices influence trades because prices are set prior to the trade.
  - But trades do not contemporaneously affect prices.
- How can we decompose the trades errors to have a part that is related to the price errors and a part that is completely unrelated?

- Let $\varepsilon_t^x = \gamma \varepsilon_t^y + u_t^x$, and $\varepsilon_t^y = u_t^y$.
- By construction, u^x_t is uncorrelated with u^y_t.
- Changes in u^x_t only affect x, and changes in u^y_t affect both x and y.
- In general, we have

$$\mathbf{v}_t = \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \gamma & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} = \begin{bmatrix} u_{1t} \\ \gamma u_{1t} + u_{2t} \end{bmatrix}$$

- We can choose which variable takes first element, and therefore the direction of causality is imposed.

- Let P denote the lower triangular matrix

$$P = \begin{bmatrix} 1 & 0 \\ \gamma & 1 \end{bmatrix}$$

- So that $\mathbf{v}_t = \begin{bmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{bmatrix} = P \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$. Then we have

$$\mathbf{y}_t = \boldsymbol{\beta}_1^t y_0 + \boldsymbol{\beta}_0^* + \sum_{j=0}^{t-1} \boldsymbol{\beta}_1^j \mathbf{v}_{t-j} \Leftrightarrow \mathbf{y}_t = \boldsymbol{\beta}_1^t y_0 + \boldsymbol{\beta}_0^* + \sum_{j=0}^{t-1} \boldsymbol{\beta}_1^j P \mathbf{u}_{t-j}$$

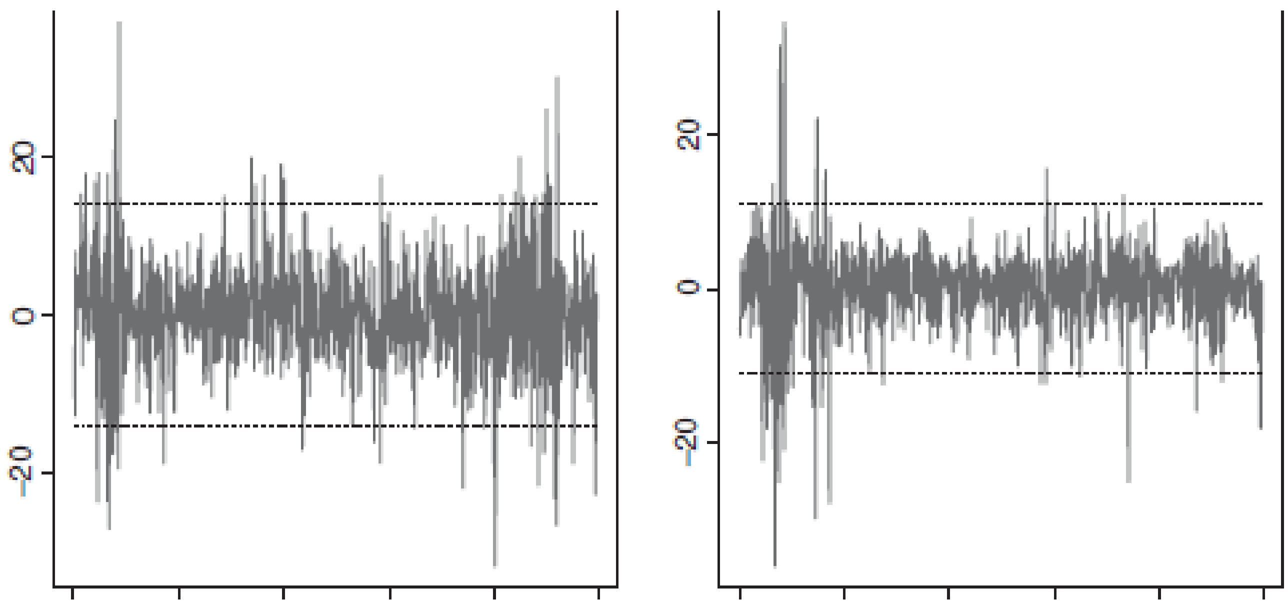
- So P determines how moving one variable in period *t* affects others contemporaneously.
- The powers of $\boldsymbol{\beta}_1$ determine how future values of **y** will change.

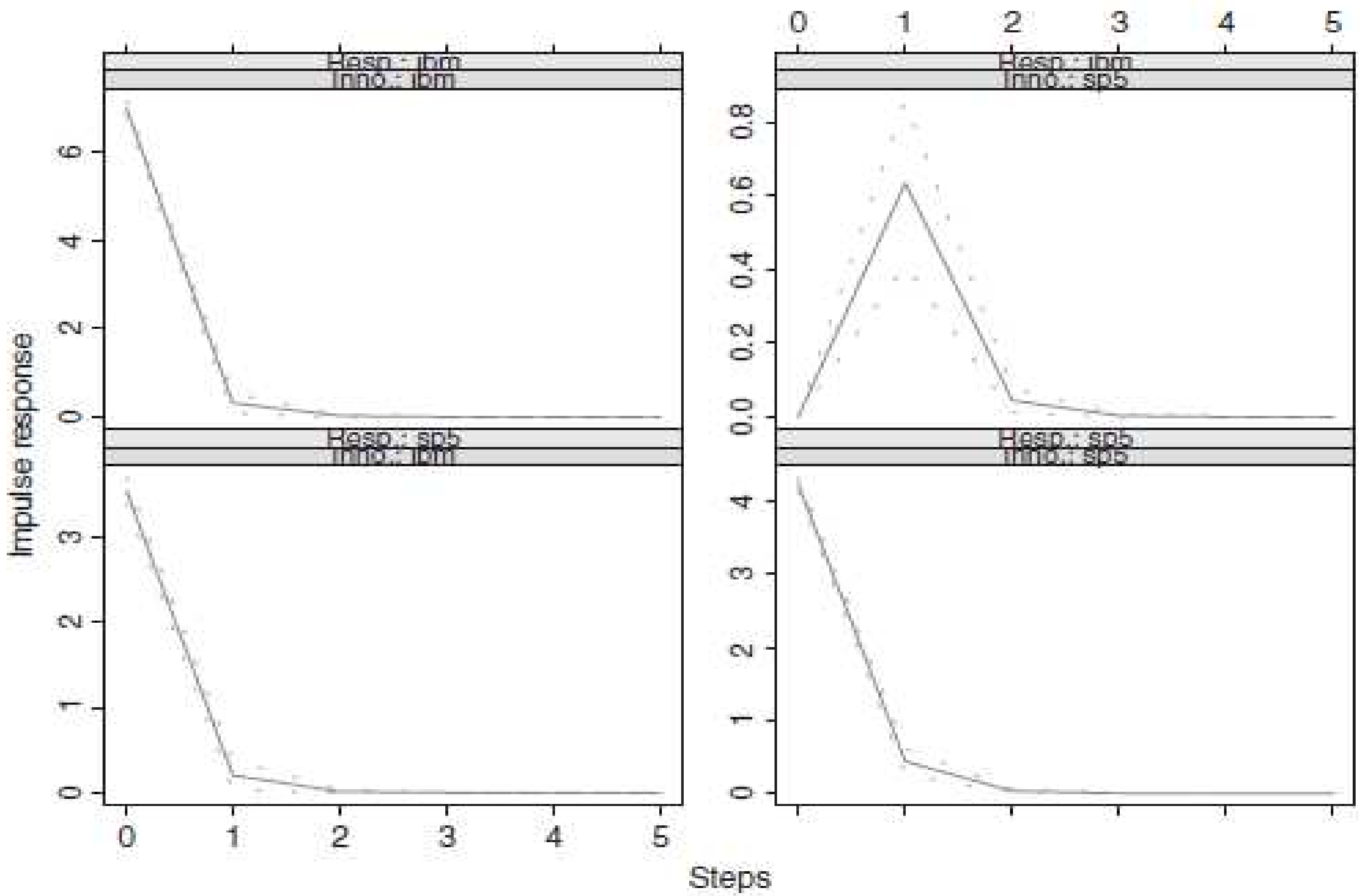
An Example

```
Coefficients:
                    ibm      sp5
 (Intercept)     1.0614   0.4087
   (std.err)     0.2249   0.1773
   (t.stat)      4.7198   2.3053

   ibm.lag1     -0.0320  -0.0223
  (std.err)      0.0413   0.0326
   (t.stat)     -0.7728  -0.6855

   sp5.lag1      0.1503   0.1020
  (std.err)      0.0525   0.0414
   (t.stat)      2.8612   2.4637

Regression Diagnostics:
                    ibm      sp5
     R-squared   0.0101   0.0075
Adj. R-squared   0.0081   0.0055
  Resid. Scale   7.0078   5.5247
```

- The fitted model is

$$\mathrm{IBM}_t = 1.06 - 0.03\mathrm{IBM}_{t-1} + 0.15\mathrm{SP5}_{t-1} + a_{1t},$$
$$\mathrm{SP5}_t = 0.41 - 0.02\mathrm{IBM}_{t-1} + 0.10\mathrm{SP5}_{t-1} + a_{2t}.$$





Resp.: ibm
Inno.: ibm
Resp.: ibm
Inno.: sp5
Resp.: sp5
Inno.: ibm
Resp.: sp5
Inno.: sp5
Impulse response
Steps

VAR(p) Models

- To generalize the idea of VAR(1), we have VAR(p)

$$\mathbf{y}_t = \boldsymbol{\beta}_0 + \sum_{j=1}^{p} \boldsymbol{\beta}_j \mathbf{y}_{t-j} + \mathbf{v}_t$$

- We can still write $\mathbf{y}_t$ as a function of the past values of $\mathbf{v}_t$:

$$\mathbf{y}_t = \boldsymbol{\beta}_1^t y_0 + \boldsymbol{\beta}_0^* + \sum_{j=0}^{t} \psi_j \mathbf{v}_{t-j}$$

- We still need to take a stand on the order the shocks propagate: $\mathbf{y}_t = \boldsymbol{\beta}_1^t y_0 + \boldsymbol{\beta}_0^* + \sum_{j=0}^{t} \psi_j P \mathbf{u}_{t-j}$

How to Estimate VAR(p)

- First, choose p. Can be the same for all variables in **y**.
- Estimate the model equation by equation using OLS just as we did for the univariate models.
- If the model is well specified, the residuals should be uncorrelated:
  - Correlation of residuals of each equation;
  - Cross-correlation of residuals of different equations.

Hypothesis Test

- For a VAR(i), i = 0, 1, 2, …, p, parameters can be estimated by OLS.
- The residual is

$$\hat{\mathbf{v}}_t = \mathbf{y}_t - \hat{\boldsymbol{\beta}}_0 - \hat{\boldsymbol{\beta}}_1 \mathbf{y}_{t-1} - \ldots - \hat{\boldsymbol{\beta}}_i \mathbf{y}_{t-i}$$

- The residual covariance is given by

$$\hat{\boldsymbol{\Omega}} = \frac{1}{T - 2i - 1} \sum_{t=i+1}^{T} \hat{\mathbf{v}}_t \hat{\mathbf{v}}'_t$$

- In general, if you want to compare VAR(i) and VAR(i-1), and test H_0: $\boldsymbol{\beta}_i = 0$, we construct the test statistic

$$M(i) = -(T - k - i - \frac{3}{2}) \ln\left(\frac{\left|\hat{\boldsymbol{\Omega}}_i\right|}{\left|\hat{\boldsymbol{\Omega}}_{i-1}\right|} \right)$$

- |A| denotes the determinants of the Matrix A. This statistic asymptotically follows a chi-square distribution with k^2 degrees of freedom

Forecasting VAR's

- Let $E_t(\mathbf{y}_{t+k})$ denote the k-step ahead forecast of $\mathbf{y}_{t+k}$.
- Then, the 1-step ahead forecast is:

$$E_t\left(\mathbf{y}_{t+1}\right)=\boldsymbol{\beta}_0+\sum_{j=1}^{p}\boldsymbol{\beta}_j E_t\left(\mathbf{y}_{t+1-j}\right)$$

$$E_t\left(\mathbf{y}_{t+1}\right)=\boldsymbol{\beta}_0+\sum_{j=1}^{p}\boldsymbol{\beta}_j \mathbf{y}_{t+1-j}$$

- The 2-step ahead forecast:

$$E_t\left(\mathbf{y}_{t+2}\right)=\boldsymbol{\beta}_0+\sum_{j=1}^{p}\boldsymbol{\beta}_j E_t\left(\mathbf{y}_{t+2-j}\right)$$

$$E_t\left(\mathbf{y}_{t+2}\right)=\boldsymbol{\beta}_0+\boldsymbol{\beta}_1 E_t\left(\mathbf{y}_{t+1}\right)+\sum_{j=2}^{p}\boldsymbol{\beta}_j \mathbf{y}_{t+2-j}$$

- The k-step ahead forecast:

$$E_t\left(\mathbf{y}_{t+k}\right)=\boldsymbol{\beta}_0+\sum_{j=1}^{p}\boldsymbol{\beta}_j E_t\left(\mathbf{y}_{t+k-j}\right)$$

$$E_t\left(\mathbf{y}_{t+k}\right)=\boldsymbol{\beta}_0+\sum_{j=1}^{k-1}\boldsymbol{\beta}_j E_t\left(\mathbf{y}_{t+k-j}\right)+\sum_{j=k}^{p}\boldsymbol{\beta}_j \mathbf{y}_{t+k-j}$$

Spurious Regression

- Generally speaking, it is a bad idea to regress one random walk process on another.
- That is, if y and x both follow random walks, it makes no sense to run the regression

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

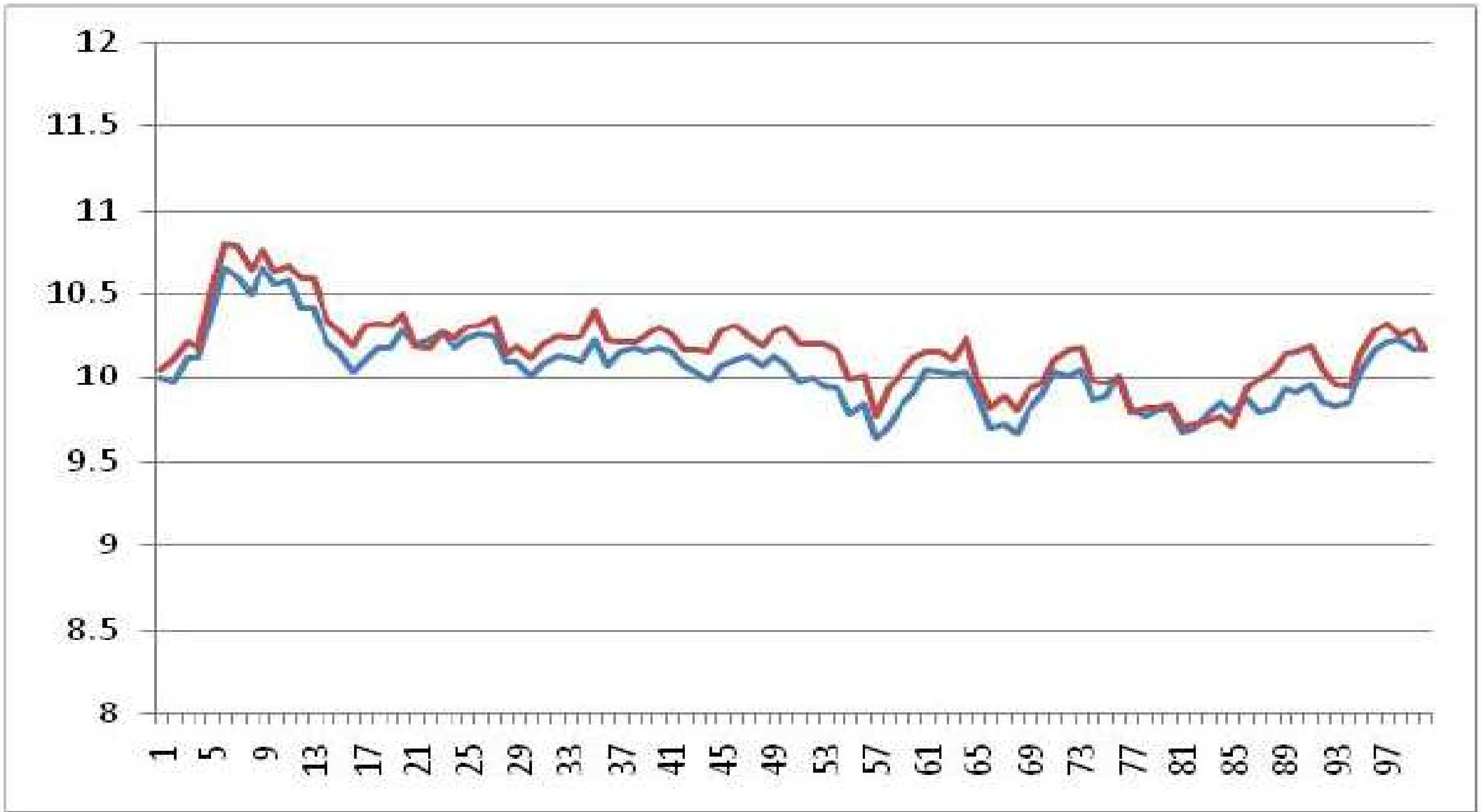
- Take a look at the OLS estimate of β_1:

$$\hat{\beta}_1 = \frac{\operatorname{cov}(x, y)}{\sigma_x^2} = \frac{\sigma_x \sigma_y \rho_{xy}}{\sigma_x^2} = \frac{\sigma_y \rho_{xy}}{\sigma_x} \rightarrow \frac{\infty \rho_{xy}}{\infty}$$

Cointegration

- Based on this observation, Cointegration is introduced.
- Cointegration is a special relationship that two non-mean reverting series can exhibit.
- Sometimes, a pair of series might each follow a random walk, but over the long run their paths are connected:
  - While they wander around over time, the two series can not get far away.
  - For example, Bid/Ask prices

Bid and Ask Prices



Formal Definition

- The series y_t and x_t are said to be cointegrated if both y_t and x_t follow random walks but there exists a linear combination

$$z_t = y_t - \gamma x_t$$

where z is stationary.

- γ describes the cointegrating relationship. In most cases, it is one.

Testing for Cointegration

- Case I: known cointegrating relationship
  - First, test to see that both y_t and x_t have a unit root.
  - Second, create a sequence z_t

$$z_t = y_t - \underbrace{\gamma}_{\substack{known \\ value}} x_t$$

  - Test the series z_t using a standard random walk test (DF).
  - If z_t is stationary, then x and y are cointegrated.

- Case II: unknown cointegrating relationship
  - First, test to see that both y_t and x_t have a unit root.
  - Second, regress y_t on x_t and estimate γ. Create the residual series

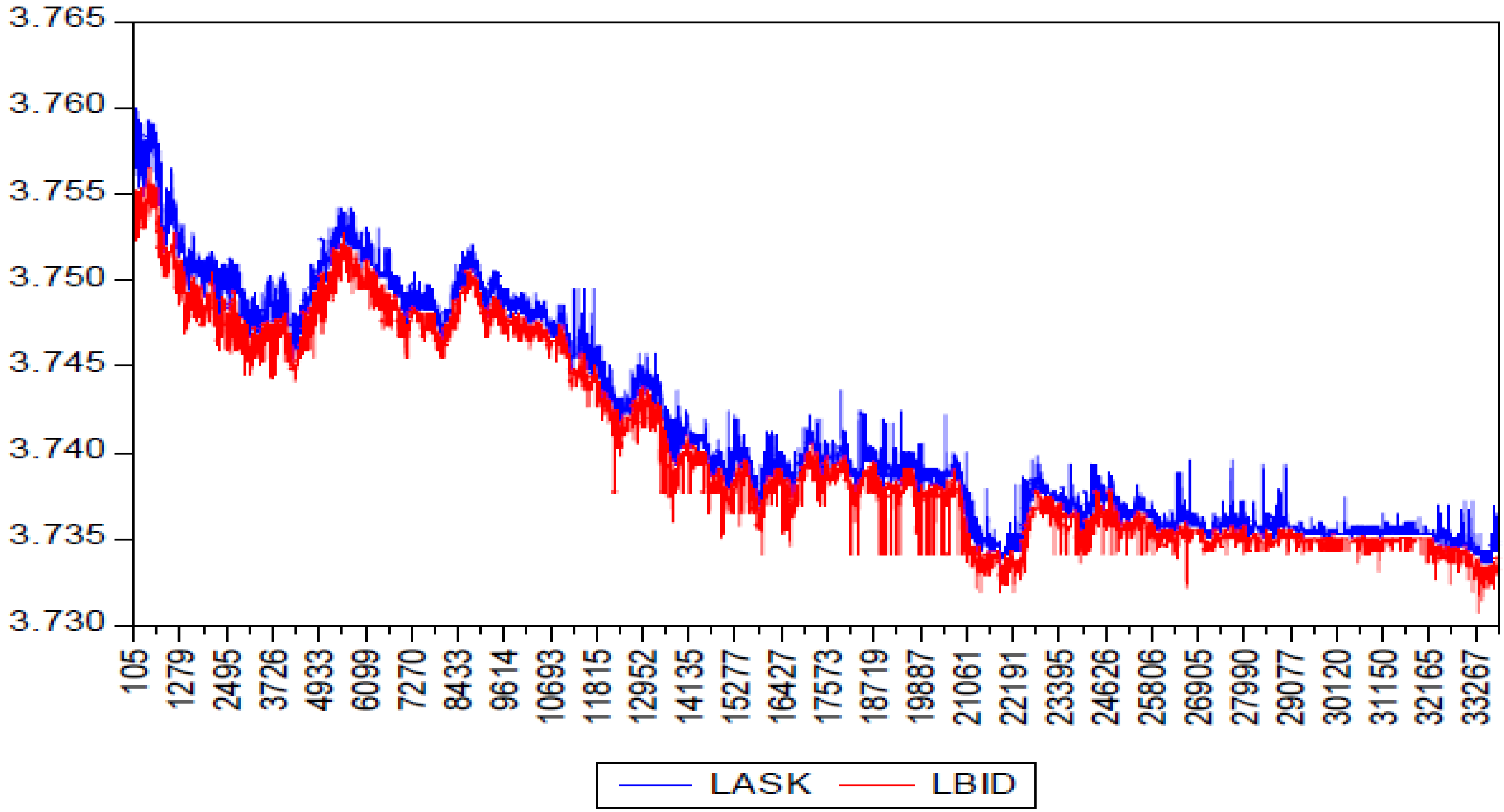
$$y_t = \hat{\gamma}_1 x_t + z_t \Leftrightarrow z_t = y_t - \hat{\gamma}_1 x$$

  - Test to see if z_t has a unit root.
  - We can not use the standard DF.

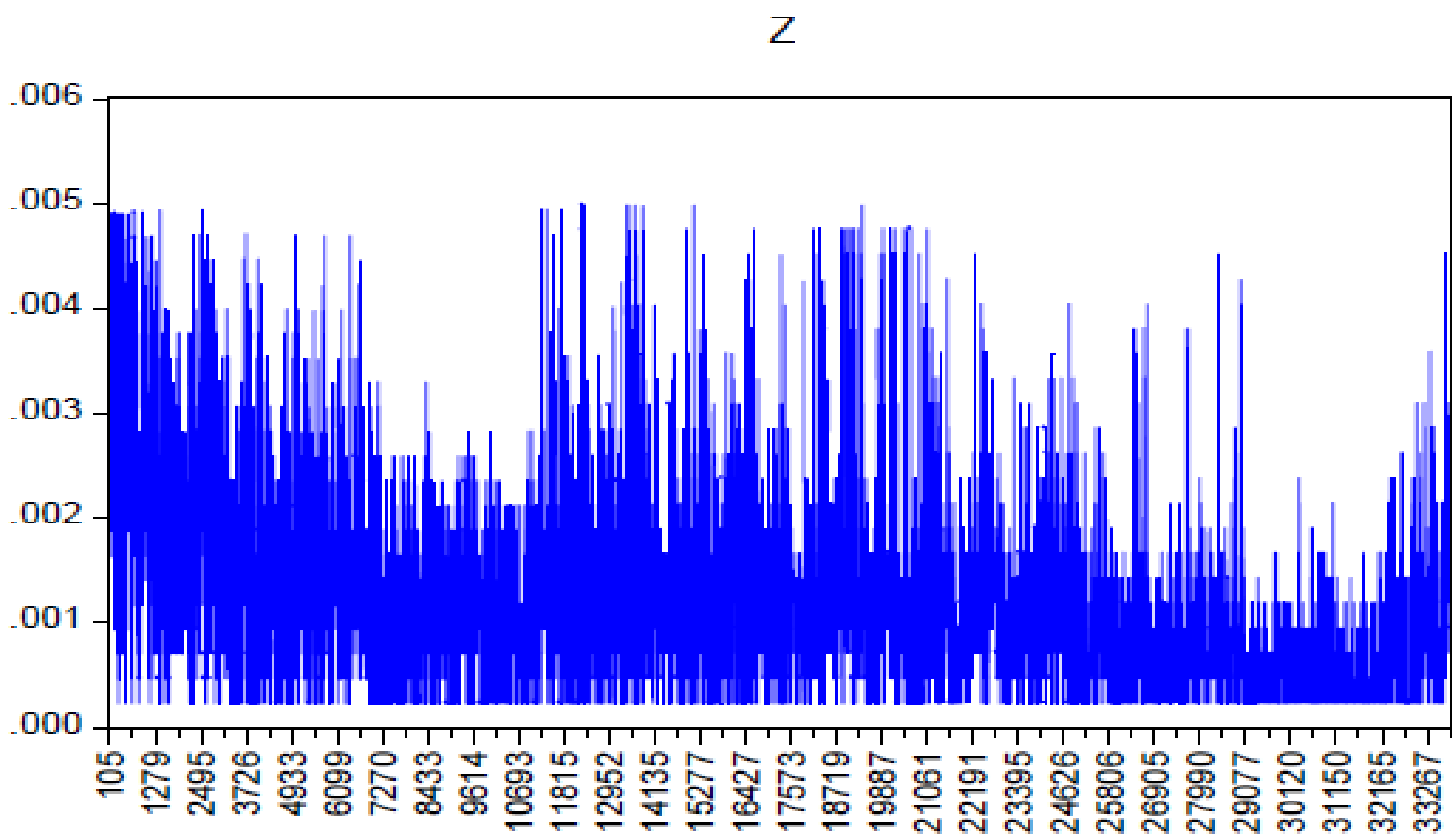
Engle-Granger Test

- The estimated γ used to construct z_t series messes things up.
- The DF-statistic under this construction does not have standard distribution.
- The new distribution and critical values are developed by Engle and Granger and other people using Monte Carlo methods.
- You need to check EG table instead of t table.

Bid/Ask Prices



Case 1: Create $z_t = y_t - x_t$ series



Augmented DF Test

Augmented Dickey-Fuller Unit Root Test on Z

Null Hypothesis: Z has a unit root
Exogenous: None
Lag Length: 46 (Automatic - based on SIC, maxlag=49)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-496.0697	0.0001
Test critical values:	1% level	-2.565040	
	5% level	-1.940835	
	10% level	-1.616693	

*MacKinnon (1996) one-sided p-values.

P-value

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(Z)
Method: Least Squares
Date: 03/02/10 Time: 23:52
Sample: 105 34000 IF SPD<.005
Included observations: 29532

Unknown Cointegrating Relationship

- Perform Engle-Granger Test

Date: 03/02/10 Time: 23:37
Series: LASK LBID
Sample: 105 34000 IF SPD<.005
Included observations: 29532
Null hypothesis: Series are not cointegrated
Cointegrating equation deterministics: C
Fixed lag specification (lag=14)

Dependent	tau-statistic	Prob.*	z-statistic	Prob.*
LASK	-10.59709	0.0000	-248.0307	0.0000
LBID	-11.31205	0.0000	-287.7100	0.0000

*MacKinnon (1996) p-values.

Intermediate Results:

	LASK	LBID
Rho - 1	-0.129385	-0.145776
Rho S.E.	0.012209	0.012887
Residual variance	3.94E-07	3.49E-07
Long-run residual variance	2.28E-08	2.14E-08
Number of lags	14	14
Number of observations	7973	7973
Number of stochastic trends**	2	2

**Number of stochastic trends in asymptotic distribution

Error Correction Model (ECM)

- If y_t and x_t are cointegrated, then the model for changes in y_t and x_t follows what is called the Error Correction Model (ECM).
- The model is specified as a VAR in changes, but it includes a special term on the right-hand side.
- This special term is given by $z_t = y_t - \gamma x_t$

- Take the first difference of y and x. The ECM is specified as follows:

$$\begin{bmatrix} \Delta y_t \\ \Delta x_t \end{bmatrix} = \Delta \mathbf{y}_t = \underbrace{\boldsymbol{\beta}_0 + \sum_{j=1}^{p} \boldsymbol{\beta}_j \Delta \mathbf{y}_t}_{\text{This is the usual VAR(p) } part.} + \underbrace{\boldsymbol{\alpha} z_{t-1}}_{\text{This is the new part}} + \mathbf{v}_t$$

- So this is just a regular VAR model for the changes in y and x, but it has the error correction term.

An Example

- A simple error correction model for log bid and ask prices is given by

$$\begin{bmatrix} \Delta \ln(ask_t) \\ \Delta \ln(bid_t) \end{bmatrix} = \begin{bmatrix} \beta_0^a \\ \beta_0^b \end{bmatrix} + \begin{bmatrix} \beta_1^a & \beta_1^a \\ \beta_1^b & \beta_1^b \end{bmatrix} \begin{bmatrix} \Delta \ln(ask_{t-1}) \\ \Delta \ln(bid_{t-1}) \end{bmatrix} + \begin{bmatrix} \alpha^a \\ \alpha^b \end{bmatrix} Spd_{t-1} + \begin{bmatrix} \varepsilon_t^a \\ \varepsilon_t^b \end{bmatrix}$$

- Recall $Spd_t = z_t = \ln(ask_t) - \ln(bid_t)$
- So α determine how the bid and ask prices change as a function of the spread

Vector Autoregression Estimates

Vector Autoregression Estimates
Date: 03/03/10 Time: 00:49
Sample: 105 34000 IF SPD<.005
Included observations: 29532
Standard errors in ( ) & t-statistics in [ ]

	DLASK	DLBID
DLASK(-1)	-0.328384	-0.325040
	(0.00254)	(0.00279)
	[-129.208]	[-116.682]
DLBID(-1)	-0.146700	-0.152252
	(0.00146)	(0.00160)
	[-100.306]	[-94.9767]
C	0.000288	-0.000887
	(8.0E-06)	(8.7E-06)
	[36.1691]	[-101.662]
Z(-1)	-0.312163	0.678346
	(0.00136)	(0.00149)
	[-230.091]	[456.172]
R-squared	0.787151	0.927449
Adj. R-squared	0.787129	0.927442
Sum sq. resids	0.037871	0.045497
S.E. equation	0.001132	0.001241
F-statistic	36399.80	125822.7
Log likelihood	158423.4	155714.2
Akaike AIC	-10.72866	-10.54518
Schwarz SC	-10.72754	-10.54406
Mean dependent	-0.000704	0.001281
S.D. dependent	0.002455	0.004608

ECM

Vector Autoregression Estimates

Vector Autoregression Estimates
Date: 03/03/10 Time: 00:52
Sample: 105 34000 IF SPD<.005
Included observations: 29532
Standard errors in ( ) & t-statistics in [ ]

	DLASK	DLBID
DLASK(-1)	-0.466615	-0.024657
	(0.00413)	(0.00768)
	[-113.065]	[-3.21131]
DLBID(-1)	0.005274	-0.482499
	(0.00218)	(0.00406)
	[2.41829]	[-118.923]
C	-0.000738	0.001342
	(1.1E-05)	(2.0E-05)
	[-66.9680]	[65.4865]
R-squared	0.405525	0.416160
Adj. R-squared	0.405485	0.416120
Sum sq. resids	0.105770	0.366131
S.E. equation	0.001893	0.003521
F-statistic	10071.72	10524.09
Log likelihood	143257.3	124922.1
Akaike AIC	-9.701634	-8.459914
Schwarz SC	-9.700791	-8.459072
Mean dependent	-0.000704	0.001281
S.D. dependent	0.002455	0.004608

VAR

- Market forces should force wide spreads to narrow.
- We should expect that a wide spread should lead to an increase in the bid and a decrease in the ask.
- An additional variables can be included in the model

$$\begin{bmatrix} \Delta \ln(ask_t) \\ \Delta \ln(bid_t) \end{bmatrix} = \begin{bmatrix} \beta_0^a \\ \beta_0^b \end{bmatrix} + \begin{bmatrix} \beta_1^a & \beta_1^a \\ \beta_1^b & \beta_1^b \end{bmatrix} \begin{bmatrix} \Delta \ln(ask_{t-1}) \\ \Delta \ln(bid_{t-1}) \end{bmatrix} + \begin{bmatrix} \alpha^a \\ \alpha^b \end{bmatrix} Spd_{t-1} + \begin{bmatrix} \theta^a \\ \theta^b \end{bmatrix} x_{t-1} + \begin{bmatrix} \varepsilon_t^a \\ \varepsilon_t^b \end{bmatrix}$$

- Can be a buy-sell indicator.
- Can be the (signed for buy or sell) size of the previous trade.
- Buys tend to raise both the bid and the ask.
- Sells tend to decrease both the bid and the ask.
- Trade size matters. Large trades have a larger price impact. The effect increases at a decreasing rate.