Advanced Financial Econometrics

Joachim Grammig, University of Tuebingen

Module IV
Time Series Applications in Finance
Readings:
Brooks (2002), Ch. 6, Hamilton (1994), Ch. 10,11,
Hasbrouck (1991a,b) Hasbrouck (1995), Grammig/Melvin/Schlag (2005)

To analyze the interdependence of three East Asian stock markets, (Tokyo, Singapore and South Korea) we set up a Structural VAR (SVAR)

$$
\begin{aligned}
r_t^T &= k^T && && + \beta_{12}^{(0)} r_t^S + && \beta_{13}^{(0)} r_t^K + && \beta_{11}^{(1)} r_{t-1}^T + \beta_{12}^{(1)} r_{t-1}^S + \beta_{13}^{(1)} r_{t-1}^K + u_t^T \\
r_t^S &= k^S && + \beta_{21}^{(0)} r_t^T && + && \beta_{23}^{(0)} r_t^K + && \beta_{21}^{(1)} r_{t-1}^T + \beta_{22}^{(1)} r_{t-1}^S + \beta_{23}^{(1)} r_{t-1}^K + u_t^S \\
r_t^K &= k^K && + \beta_{31}^{(0)} r_t^T && + \beta_{32}^{(0)} r_t^S + && && \beta_{31}^{(1)} r_{t-1}^T + \beta_{32}^{(1)} r_{t-1}^S + \beta_{33}^{(1)} r_{t-1}^K + u_t^K
\end{aligned}
$$

$$
\underset{(3\times 1)}{\mathbf{y}_t} = \begin{bmatrix} r_t^T \\ r_t^S \\ r_t^K \end{bmatrix} \quad
\underset{(3\times 1)}{\mathbf{k}} = \begin{bmatrix} k^T \\ k^S \\ k^K \end{bmatrix} \quad
\underset{(3\times 3)}{\mathbf{B_0}} = \begin{bmatrix} 1 & -\beta_{12}^{(0)} & -\beta_{13}^{(0)} \\ -\beta_{21}^{(0)} & 1 & -\beta_{23}^{(0)} \\ -\beta_{31}^{(0)} & -\beta_{32}^{(0)} & 1 \end{bmatrix} \quad
\underset{(3\times 1)}{\mathbf{u}_t} = \begin{bmatrix} u_t^T \\ u_t^S \\ u_t^K \end{bmatrix}
$$

$$
\mathbf{B}_0 \mathbf{y}_t = \mathbf{k} + \mathbf{B}_1 \mathbf{y}_{t-1} + \mathbf{u}_t
$$

The innovations of a VAR in primitive form are assumed to be both serially and cross-sectionally uncorrelated (orthogonal/pure/idiosyncratic innovations/shocks)

$$\mathrm{B}_0 \mathbf{y}_t = \mathbf{k} + \mathrm{B}_1 \mathbf{y}_{t-1} + \mathrm{B}_2 \mathbf{y}_{t-2} + \ldots + \mathrm{B}_p \mathbf{y}_{t-p} + \mathbf{u}_t$$

$$\begin{aligned} \mathbb{E}(\mathbf{u}_t) &= \mathbf{0} \\ \mathbb{E}(\mathbf{u}_t \mathbf{u}'_\tau) &= \begin{cases} \mathbf{D} & \text{for} \quad t = \tau \\ \mathbf{0} & \text{otherwise.} \end{cases} \end{aligned}$$

$\mathbf{D}$ diagonal matrix

Writing the VAR in *standard form* „solves“ the system

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \dots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \varepsilon_t$$

$\mathbf{c} = \mathrm{B}_0^{-1}\mathbf{k}$ $(n \times 1)$ vector of constants

$\mathbf{\Phi}_s = \mathrm{B}_0^{-1}\mathrm{B}_s$ $(n \times n)$ matrix of AR coefficients for $s = 1, ..., p$

$\varepsilon_t = \mathrm{B}_0^{-1}\mathbf{u}_t$ $(n \times 1)$ vector generalization of white noise.

The innovations of a VAR in standard form are, by construction, contemporaneusly correlated (composite innovations/shocks)

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \varepsilon_t$$

$$\mathbb{E}(\varepsilon_t) = \mathbb{E}(\mathbf{B}_0^{-1}\mathbf{u}_t) = \mathbf{B}_0^{-1}\mathbb{E}(\mathbf{u}_t) = \mathbf{0}$$

$$\mathbb{E}(\varepsilon_t \varepsilon_t') = \mathbb{E}(\mathbf{B}_0^{-1}\mathbf{u}_t\mathbf{u}_t'[\mathbf{B}_0^{-1}]') \equiv \mathbf{\Omega}$$

$$\mathbb{E}(\varepsilon_t \varepsilon_\tau') = \begin{cases} \mathbf{\Omega} & \text{for} \quad t = \tau \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

The lag operator provides notational convenience

Lag operator:

$L(y_t) = y_{t-1}$, $L^2(y_t) = y_{t-2}, \ldots$

VAR(p) written with lag operator

$$[\mathbf{I}_n - \boldsymbol{\Phi}_1 L - \boldsymbol{\Phi}_2 L^2 - \ldots - \boldsymbol{\Phi}_p L^p]\mathbf{y}_t = \mathbf{c} + \varepsilon_t$$

or

$$\boldsymbol{\Phi}(L)\mathbf{y}_t = \mathbf{c} + \varepsilon_t$$

We take expectations of the endogenous variables

Assuming stationarity: $\mathbb{E}(\mathbf{y}_t) = \boldsymbol{\mu}$

$$\begin{aligned}
\mathbb{E}(\mathbf{y}_t) &= \mathbf{c} + \mathbf{\Phi}_1 \mathbb{E}(\mathbf{y}_{t-1}) + \ldots + \mathbf{\Phi}_p \mathbb{E}(\mathbf{y}_{t-p}) + \mathbb{E}(\boldsymbol{\varepsilon}_t) \\
\boldsymbol{\mu} &= \mathbf{c} + \mathbf{\Phi}_1 \boldsymbol{\mu} + \mathbf{\Phi}_2 \boldsymbol{\mu} + \ldots + \mathbf{\Phi}_p \boldsymbol{\mu} \\
\boldsymbol{\mu} &= \mathbf{c} + [\mathbf{\Phi}_1 + \mathbf{\Phi}_2 + \ldots + \mathbf{\Phi}_p] \boldsymbol{\mu} \\
[\mathbf{I}_n - \mathbf{\Phi}_1 - \mathbf{\Phi}_2 - \ldots - \mathbf{\Phi}_p] \, \boldsymbol{\mu} &= \mathbf{c} \\
[\mathbf{I}_n - \mathbf{\Phi}_1 L - \ldots - \mathbf{\Phi}_p L^p] \, \boldsymbol{\mu} &= \mathbf{c} \\
\mathbf{\Phi}(L) \boldsymbol{\mu} &= \mathbf{c}
\end{aligned}$$

It is convenient to express a VAR in terms of deviations from the means

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t$$

$$\mathbf{\Phi}(L)\boldsymbol{\mu} = \mathbf{c}$$

$$(\mathbf{y}_t - \boldsymbol{\mu}) = \mathbf{\Phi}_1(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \mathbf{\Phi}_2(\mathbf{y}_{t-2} - \boldsymbol{\mu}) + \ldots + \mathbf{\Phi}_p(\mathbf{y}_{t-p} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_t$$

With some additional notation a VAR(*p*) can be rewritten as a VAR(1)

$$(\mathbf{y}_t - \boldsymbol{\mu}) = \boldsymbol{\Phi}_1(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\Phi}_2(\mathbf{y}_{t-2} - \boldsymbol{\mu}) + \ldots + \boldsymbol{\Phi}_p(\mathbf{y}_{t-p} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_t$$

Define:

$$\underset{(np\times 1)}{\boldsymbol{\xi}_t} \equiv \begin{bmatrix} \mathbf{y}_t - \boldsymbol{\mu} \\ \mathbf{y}_{t-1} - \boldsymbol{\mu} \\ \vdots \\ \mathbf{y}_{t-p+1} - \boldsymbol{\mu} \end{bmatrix} \quad \underset{(np\times np)}{\mathbf{F}} \equiv \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 & \boldsymbol{\Phi}_3 & \ldots & \boldsymbol{\Phi}_{p-1} & \boldsymbol{\Phi}_p \\ \mathbf{I}_n & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{I}_n & \mathbf{0} \end{bmatrix} \quad \underset{(np\times 1)}{\mathbf{v}_t} \equiv \begin{bmatrix} \boldsymbol{\varepsilon}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}$$

$$\boldsymbol{\xi}_t = \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t$$

Consider a forward iteration of the VAR(1) system

$$
\begin{aligned}
\boldsymbol{\xi}_t &= \mathbf{F}\boldsymbol{\xi}_{t-1} + \mathbf{v}_t \\
\boldsymbol{\xi}_{t+1} &= \mathbf{F}\boldsymbol{\xi}_t + \mathbf{v}_{t+1} \\
\boldsymbol{\xi}_{t+2} &= \mathbf{F}\boldsymbol{\xi}_{t+1} + \mathbf{v}_{t+2} \\
\boldsymbol{\xi}_{t+3} &= \mathbf{F}\boldsymbol{\xi}_{t+2} + \mathbf{v}_{t+3} = \mathbf{v}_{t+3} + \mathbf{F}(\mathbf{F}\boldsymbol{\xi}_{t+1} + \mathbf{v}_{t+2}) \\
&\vdots \qquad = \mathbf{v}_{t+3} + \mathbf{F}\mathbf{v}_{t+2} + \mathbf{F}^2\boldsymbol{\xi}_{t+1} \\
&\qquad = \mathbf{v}_{t+3} + \mathbf{F}\mathbf{v}_{t+2} + \mathbf{F}^2(\mathbf{F}\boldsymbol{\xi}_t + \mathbf{v}_{t+1}) \\
&\qquad = \mathbf{v}_{t+3} + \mathbf{F}\mathbf{v}_{t+2} + \mathbf{F}^2\mathbf{v}_{t+1} + \mathbf{F}^3\boldsymbol{\xi}_t
\end{aligned}
$$

iterating s times yields:

$$
\boldsymbol{\xi}_{t+s} = \mathbf{v}_{t+s} + \mathbf{F}\mathbf{v}_{t+s-1} + \mathbf{F}^2\mathbf{v}_{t+s-2} + \ldots + \mathbf{F}^{s-1}\mathbf{v}_{t+1} + \mathbf{F}^s\boldsymbol{\xi}_t
$$

To obtain the Vector Moving Average (VMA) representation we focus on the first rows of the system

the first n rows of the system

$$\boldsymbol{\xi}_{t+s} = \mathbf{v}_{t+s} + \mathbf{F}\mathbf{v}_{t+s-1} + \mathbf{F}^2\mathbf{v}_{t+s-2} + \ldots + \mathbf{F}^{s-1}\mathbf{v}_{t+1} + \mathbf{F}^s\boldsymbol{\xi}_t$$

are:

$$\begin{aligned}\mathbf{y}_{t+s} = \boldsymbol{\mu} + \varepsilon_{t+s} + \boldsymbol{\Psi}_1\varepsilon_{t+s-1} + \boldsymbol{\Psi}_2\varepsilon_{t+s-2} + \ldots + \boldsymbol{\Psi}_{s-1}\varepsilon_{t+1} \\ + \mathbf{F}_{11}^{(s)}(\mathbf{y}_t - \boldsymbol{\mu}) + \mathbf{F}_{12}^{(s)}(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \ldots + \mathbf{F}_{1p}^{(s)}(\mathbf{y}_{t-p+1} - \boldsymbol{\mu})\end{aligned}$$

$\mathbf{F}^{(j)}$: $\mathbf{F}$ raised to the j^{th} power

$\mathbf{F}_{11}^{(j)} = \boldsymbol{\Psi}_j$: first n rows and columns 1 through n

$\mathbf{F}_{1p}^{(j)}$: first n rows and columns $(n(p-1)+1)$ through np

Forecast of $\mathbf{y}_{t+s}$ on the basis of $\mathbf{y}_t, \mathbf{y}_{t-1}, \ldots$

$$\hat{\mathbf{y}}_{t+s|t} = \boldsymbol{\mu} + \mathbf{F}_{11}^{(s)}(\mathbf{y}_t - \boldsymbol{\mu}) + \mathbf{F}_{12}^{(s)}(\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \ldots + \mathbf{F}_{1p}^{(s)}(\mathbf{y}_{t-p+1} - \boldsymbol{\mu})$$

Forecast error:

$$\mathbf{y}_{t+s} - \hat{\mathbf{y}}_{t+s|t} = \boldsymbol{\varepsilon}_{t+s} + \boldsymbol{\Psi}_1 \boldsymbol{\varepsilon}_{t+s-1} + \boldsymbol{\Psi}_2 \boldsymbol{\varepsilon}_{t+s-2} + \ldots + \boldsymbol{\Psi}_{s-1} \boldsymbol{\varepsilon}_{t+1}$$

Vector MA(∞) Representation

Eigenvalues of $\mathbf{F}$ inside the unit circle → stationarity of $\{\mathbf{y}_t\}$

→ Vector MA(∞) Representation

$$\boldsymbol{\xi}_t = \sum_{i=0}^{\infty} \mathbf{F}^i \mathbf{v}_{t-i}$$

First n rows:

$$\begin{aligned}
\mathbf{y}_t &= \boldsymbol{\mu} + \boldsymbol{\varepsilon}_t + \boldsymbol{\Psi}_1 \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\Psi}_2 \boldsymbol{\varepsilon}_{t-2} + \boldsymbol{\Psi}_3 \boldsymbol{\varepsilon}_{t-3} + \ldots \\
\mathbf{y}_t &= \boldsymbol{\mu} + \left[\mathbf{I}_n + \boldsymbol{\Psi}_1 L + \boldsymbol{\Psi}_2 L^2 \ldots\right] \boldsymbol{\varepsilon}_t \\
\mathbf{y}_t &= \boldsymbol{\mu} + \boldsymbol{\Psi}(L) \boldsymbol{\varepsilon}_t
\end{aligned}$$

Combining results shows how VAR and MA coefficients are related

$$\mathbf{\Phi}(L)\mathbf{y}_t = \mathbf{c} + \boldsymbol{\varepsilon}_t \qquad \mathbf{\Phi}(L)\boldsymbol{\mu} = \mathbf{c} \qquad \mathbf{y}_t = \boldsymbol{\mu} + \mathbf{\Psi}(L)\boldsymbol{\varepsilon}_t$$

$$\begin{aligned}
\mathbf{\Phi}(L)[\boldsymbol{\mu} + \mathbf{\Psi}(L)\boldsymbol{\varepsilon}_t] &= \mathbf{c} + \boldsymbol{\varepsilon}_t \\
\mathbf{\Phi}(L)\boldsymbol{\mu} + \mathbf{\Phi}(L)\mathbf{\Psi}(L)\boldsymbol{\varepsilon}_t &= \mathbf{c} + \boldsymbol{\varepsilon}_t \\
\mathbf{c} + \mathbf{\Phi}(L)\mathbf{\Psi}(L)\boldsymbol{\varepsilon}_t &= \mathbf{c} + \boldsymbol{\varepsilon}_t \\
\underbrace{[\mathbf{\Phi}(L)\mathbf{\Psi}(L)]}_{\mathbf{I}_n}\boldsymbol{\varepsilon}_t &= \boldsymbol{\varepsilon}_t
\end{aligned}$$

The VMA coefficients can be recursively computed from the VAR coefficients

$$\mathbf{I}_n = \Psi(L)\Phi(L)$$

$$\mathbf{I}_n = (\mathbf{I}_n + \Psi_1 L + \Psi_2 L^2 + \ldots)(\mathbf{I}_n - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p)$$

$$\mathbf{I}_n = \mathbf{I}_n + (\Psi_1 - \Phi_1)L + (\Psi_2 - \Phi_1\Psi_1 - \Phi_2)L^2 + \ldots$$

$$\Rightarrow \quad \mathbf{\Psi_1} = \mathbf{\Phi_1}$$

$$\mathbf{\Psi_2} = \mathbf{\Phi_1\Psi_1} + \mathbf{\Phi_2}$$

general for L^s $s = 1, 2, \ldots$:

$$\mathbf{\Psi_s} = \mathbf{\Phi_1\Psi_{s-1}} + \mathbf{\Phi_2\Psi_{s-2}} + \ldots + \mathbf{\Phi_p\Psi_{s-p}}$$

The Impulse-Response Function gives the response of the system to one unit shocks in the ε

$$\mathbf{y}_{t+s} = \boldsymbol{\mu} + \varepsilon_{t+s} + \boldsymbol{\Psi}_1 \varepsilon_{t+s-1} + \boldsymbol{\Psi}_2 \varepsilon_{t+s-2} + \ldots + \boldsymbol{\Psi}_s \varepsilon_t + \ldots + \ldots$$

$$\frac{\partial \mathbf{y}_{t+s}}{\partial \varepsilon_t'} = \boldsymbol{\Psi}_{\mathbf{s}}$$

Sequence of $\boldsymbol{\Psi}_1$, $\boldsymbol{\Psi}_2$,...: Impulse-Response Function

e.g. response of $y_{i,t+s}$ to a one-time impulse in $\varepsilon_{j,t}$ with all other variables dated t or earlier held constant: $\frac{\partial y_{i,t+s}}{\partial \varepsilon_{jt}} = \psi_s[i,j]$

This numerical example shows how to obtain the VMA coefficients from VAR(2) parameters

s	Φ_s			Ψ_s		
1	-0.029	0.034	0.035	-0.029	0.034	0.035
	0.007	0.195	0.044	0.007	0.195	0.044
	0.027	0.090	0.060	0.027	0.090	0.060
2	-0.071	-0.024	0.020	-0.069	-0.015	0.022
	-0.050	-0.062	0.016	-0.047	-0.020	0.028
	0.005	-0.016	0.004	0.006	0.008	0.013
3				0.003	-0.005	-0.002
				-0.008	-0.016	0.003
				-0.006	-0.004	0.004
⋮		⋮			⋮	
10				0.000	0.000	0.000
				0.000	0.000	0.000
				0.000	0.000	0.000

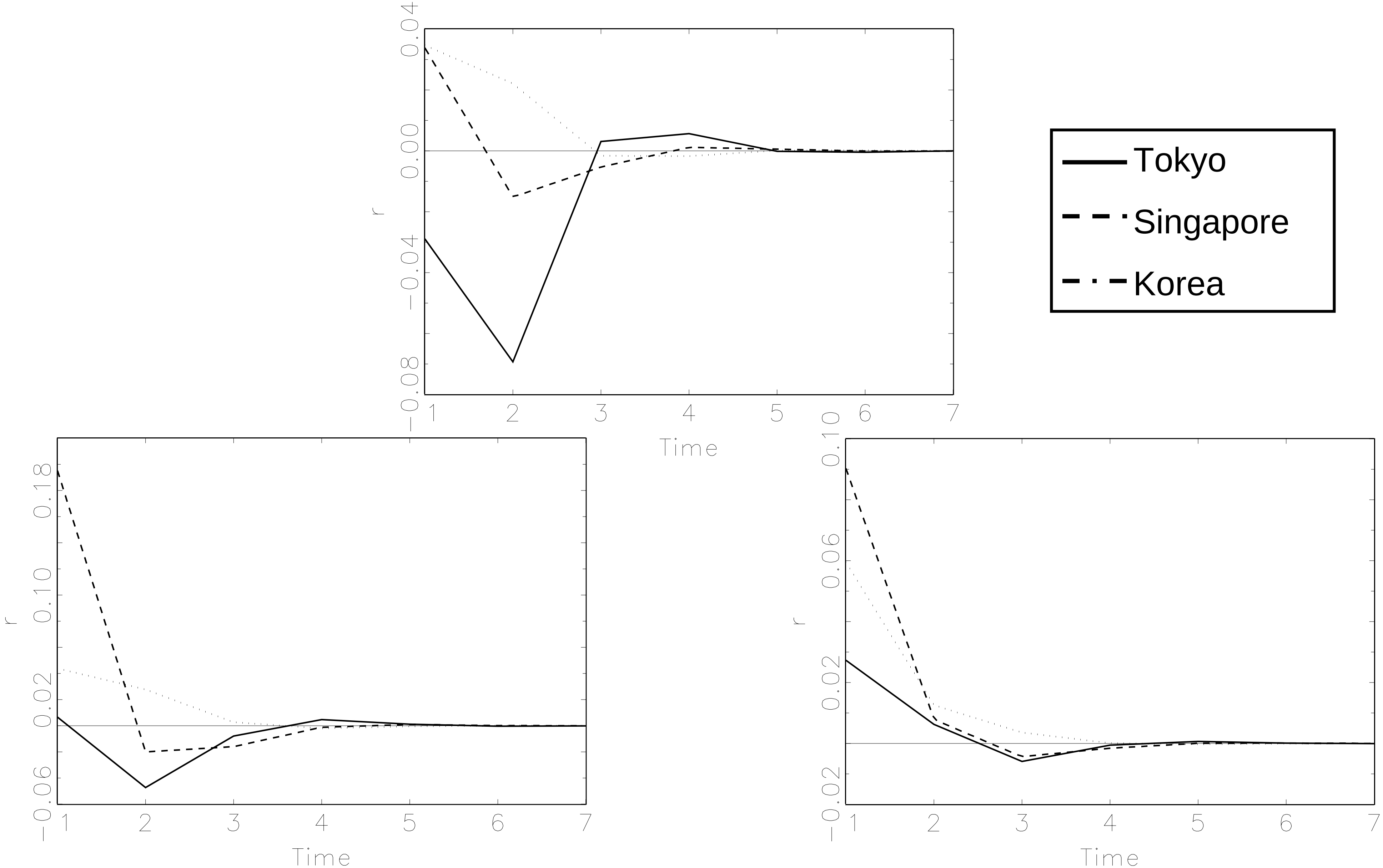
This numerical example shows how to obtain the VMA coefficients from the VAR(2) parameters

$$
\begin{aligned}
\Psi_1 &= \Phi_1 \\
&= \begin{bmatrix} -0.029 & 0.034 & 0.035 \\ 0.007 & 0.195 & 0.044 \\ 0.027 & 0.090 & 0.060 \end{bmatrix} \\
\Psi_2 &= \Phi_1 \Psi_1 + \Phi_2 \\
&= \begin{bmatrix} -0.029 & 0.034 & 0.035 \\ 0.007 & 0.195 & 0.044 \\ 0.027 & 0.090 & 0.060 \end{bmatrix} \cdot \begin{bmatrix} -0.029 & 0.034 & 0.035 \\ 0.007 & 0.195 & 0.044 \\ 0.027 & 0.090 & 0.060 \end{bmatrix} \\
&+ \begin{bmatrix} -0.071 & -0.024 & 0.020 \\ -0.050 & -0.062 & 0.016 \\ 0.005 & -0.016 & 0.004 \end{bmatrix} \\
&= \begin{bmatrix} -0.069 & -0.015 & 0.022 \\ -0.047 & -0.020 & 0.028 \\ 0.006 & 0.008 & 0.013 \end{bmatrix}
\end{aligned}
$$

This numerical example shows how to obtain the VMA coefficients from the VAR(2) parameters

$$
\begin{aligned}
\Psi_3 &= \mathbf{\Phi_1 \Psi_2 + \Phi_2 \Psi_1} \\
&= \begin{bmatrix} -0.029 & 0.034 & 0.035 \\ 0.007 & 0.195 & 0.044 \\ 0.027 & 0.090 & 0.060 \end{bmatrix} \cdot \begin{bmatrix} -0.069 & -0.015 & 0.022 \\ -0.047 & -0.020 & 0.028 \\ 0.006 & 0.008 & 0.013 \end{bmatrix} \\
&+ \begin{bmatrix} -0.071 & -0.024 & 0.020 \\ -0.050 & -0.062 & 0.016 \\ 0.005 & -0.016 & 0.004 \end{bmatrix} \cdot \begin{bmatrix} -0.029 & 0.034 & 0.035 \\ 0.007 & 0.195 & 0.044 \\ 0.027 & 0.090 & 0.060 \end{bmatrix} \\
&= \begin{bmatrix} 0.003 & -0.005 & -0.002 \\ -0.008 & -0.016 & 0.003 \\ -0.006 & -0.004 & 0.004 \end{bmatrix}
\end{aligned}
$$

The plots show a graphical representation of the VMA coefficients



To obtain the idiosyncratic shocks from the composite shocks we need the structural parameters, the matrix B_0

covariance matrix of ε_t:

$$\mathbb{E}(\varepsilon_t \varepsilon_t') = \mathbf{\Omega}$$

relation between shocks in VAR and SVAR: $\varepsilon_t = \mathbf{B}_0^{-1}\mathbf{u}_t$

$$\begin{aligned}\mathbb{E}(\varepsilon_t \varepsilon_t') &= \mathbf{B}_0^{-1}\mathbb{E}(\mathbf{u}_t \mathbf{u}_t')\left[\mathbf{B}_0^{-1}\right]' \\ &= \mathbf{B}_0^{-1}\mathbf{D}\left[\mathbf{B}_0^{-1}\right]'\end{aligned}$$

To identify the structural parameters B_0, we decompose the variance covariance matrix of composite innovations (Choleski-Dekomposition)

$$\mathbf{\Omega} = \mathbf{ADA}'$$

$$= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{21} & 1 & 0 & \dots & 0 \\ a_{31} & a_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 1 \end{bmatrix} \cdot \begin{bmatrix} d_1 & 0 & 0 & \dots & 0 \\ 0 & d_2 & 0 & \dots & 0 \\ 0 & 0 & d_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & d_n \end{bmatrix} \cdot \begin{bmatrix} 1 & a_{21} & a_{31} & \dots & a_{n1} \\ 0 & 1 & a_{32} & \dots & a_{n2} \\ 0 & 0 & 1 & \dots & a_{n3} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$\mathbf{\Omega}$: real symmetric positive definite matrix

$\mathbf{A}$: lower triangular matrix with ones along the principal diagonal

$\mathbf{D}$: diagonal matrix with positive elements

The idiosyncratic innovations can then be backed out from the composite innovations

Define $\mathbf{A} = \mathbf{B}_0^{-1}$

$$\mathbb{E}(\varepsilon_t \varepsilon_t') = \mathbf{\Omega} = \mathbf{ADA}'$$

Construct from $\mathbf{Au}_t = \varepsilon_t$: $\mathbf{u}_t \equiv \mathbf{A}^{-1}\varepsilon_t$ with variance

$$\begin{aligned} \mathbb{E}(\mathbf{u}_t \mathbf{u}_t') &= [\mathbf{A}^{-1}]\mathbb{E}(\varepsilon_t \varepsilon_t')[\mathbf{A}^{-1}]' \\ &= [\mathbf{A}^{-1}]\mathbf{\Omega}[\mathbf{A}']^{-1} \\ &= [\mathbf{A}^{-1}]\mathbf{ADA}'[\mathbf{A}']^{-1} \\ &= \mathbf{D} \end{aligned}$$

This implies: $\mathbb{E}(\mathbf{u}_{it}\mathbf{u}_{jt}') = 0 \; i \neq j$

The numerical example shows the decomposition of the variance covariance matrix in the present application

Example:

$$\mathbf{\Omega} = \mathbf{ADA}'$$

$$\begin{bmatrix} 1.79 & 0.62 & 0.16 \\ 0.62 & 1.99 & 0.28 \\ 0.16 & 0.28 & 2.67 \end{bmatrix} = \begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.34 & 1.00 & 0.00 \\ 0.09 & 0.13 & 1.00 \end{bmatrix} \begin{bmatrix} 1.79 & 0.00 & 0.00 \\ 0.00 & 1.78 & 0.00 \\ 0.00 & 0.00 & 2.63 \end{bmatrix} \begin{bmatrix} 1.00 & 0.34 & 0.09 \\ 0.00 & 1.00 & 0.13 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$$

$\mathbf{\Omega}$ and $\mathbf{D}$ multiplied by 10000.

The composite shocks are generated as linear combinations of the pure innovations

$$\mathbf{A} \cdot \mathbf{u}_t = \varepsilon_t$$

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{21} & 1 & 0 & \dots & 0 \\ a_{31} & a_{32} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \\ \vdots \\ u_{nt} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}$$

Thus, $u_{1t} = \varepsilon_{1t}$ and $u_{jt} = \varepsilon_{jt} - a_{j1}u_{1t} - a_{j2}u_{2t} - \ldots - a_{j,j-1}u_{j-1,t}$

$\Rightarrow$ variable ORDERING matters!

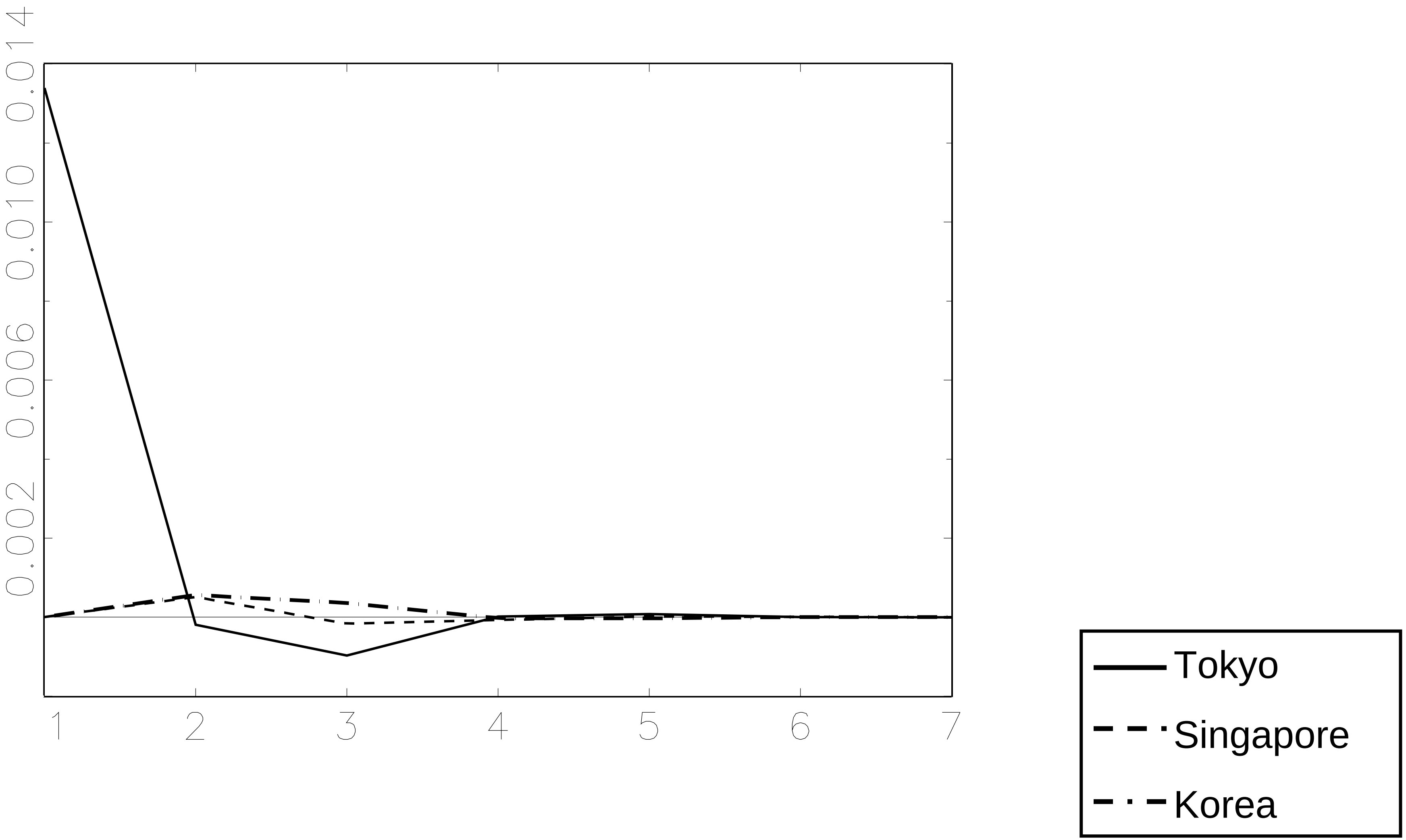
In most applications in economics and finance you want to trace a shock in the pure innovation

$$\frac{\partial \hat{\mathbb{E}}(\mathbf{y}_{t+s} | y_{jt}, y_{j-1,t}, \ldots, y_{1t}, \mathbf{x}_{t-1})}{\partial u_{jt}} = \mathbf{\Psi}_s \mathbf{a}_j$$

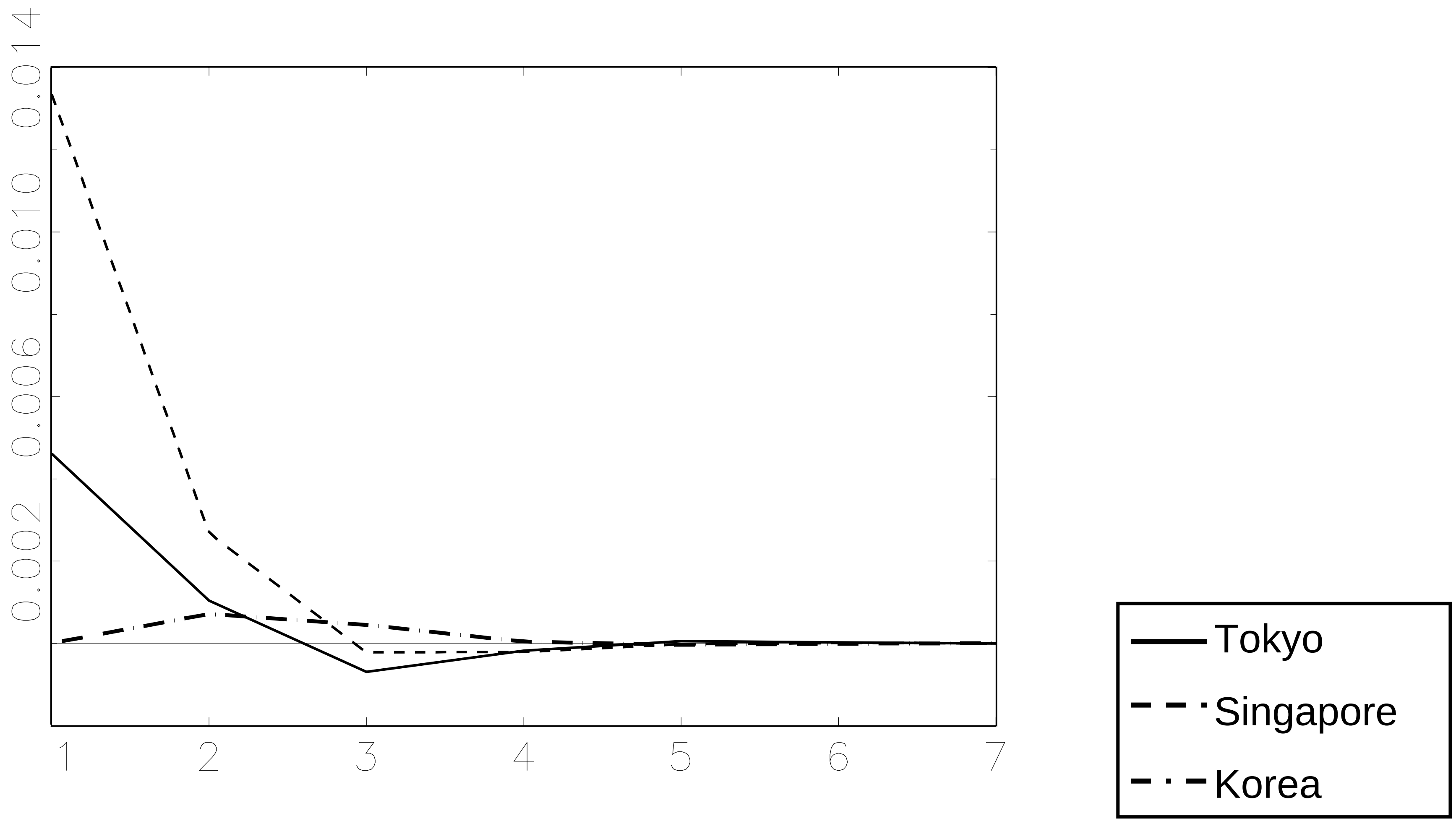
with $\mathbf{a}_j$ as the jth column of $\mathbf{A}$

$\Rightarrow$ orthogonalized impulse response function

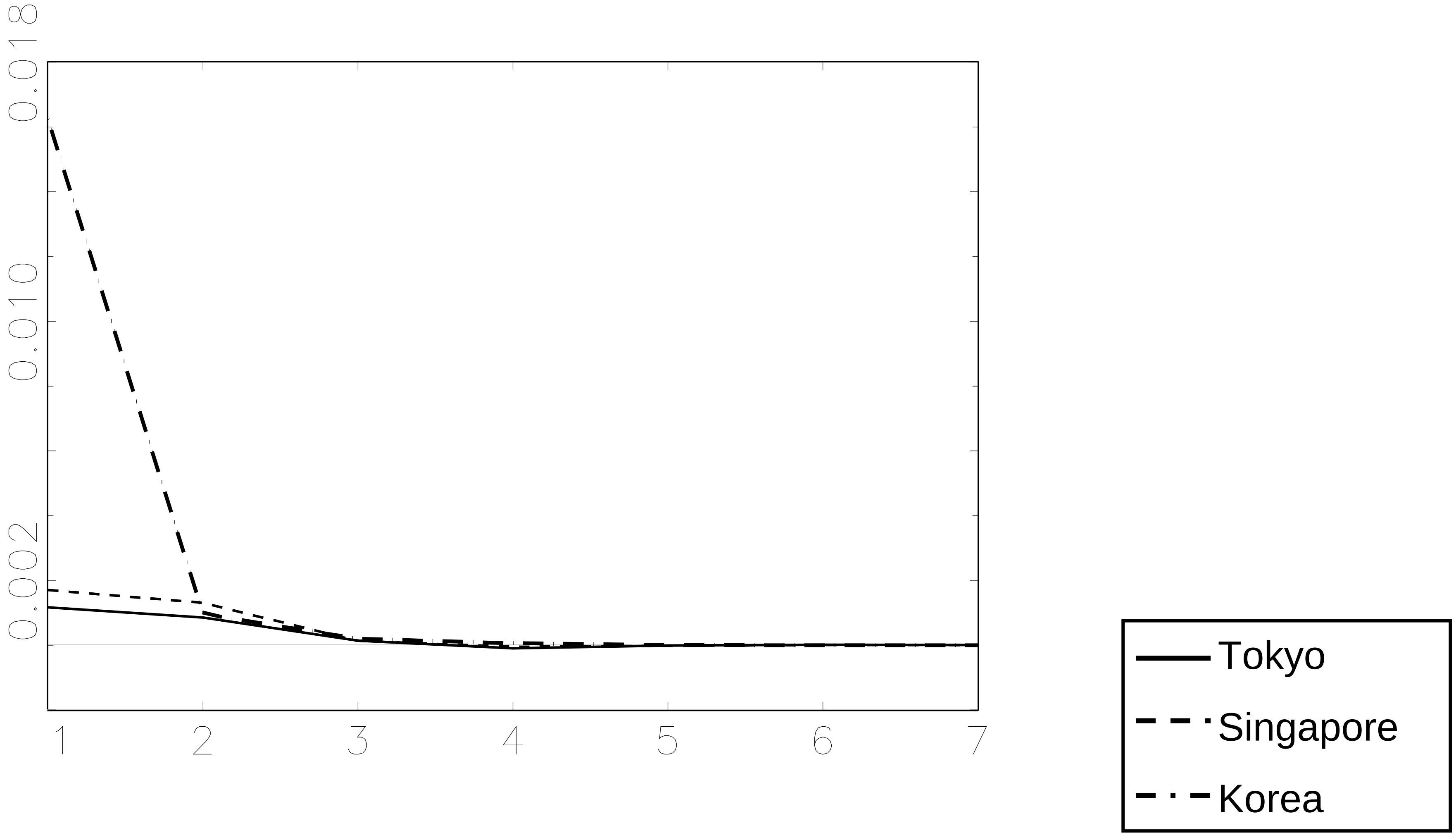
Orthogonalized impulse response function of Tokyo to one standard deviation shock in the SVAR(2) with Cholesky Ordering: Tokyo Singapore Korea



Orthogonalized impulse response function of Singapore to one standard deviation shock in the SVAR(2) with Cholesky Ordering: Tokyo Singapore Korea



Orthogonalized impulse response function of Korea to one standard deviation shock in the SVAR(2) with Cholesky Ordering: Tokyo Singapore Korea



To attribute information shares to the markets we consider a decomposition of the Mean Squared Forecast Error

$$\mathbf{y}_{t+s} - \hat{\mathbf{y}}_{t+s|t} = \varepsilon_{t+s} + \mathbf{\Psi}_1 \varepsilon_{t+s-1} + \mathbf{\Psi}_2 \varepsilon_{t+s-2} + \ldots + \mathbf{\Psi}_{s-1} \varepsilon_{t+1}$$

$$\begin{aligned} MSE(\hat{\mathbf{y}}_{t+s|t}) &= \mathbb{E}[(\mathbf{y}_{t+s} - \hat{\mathbf{y}}_{t+s|t})(\mathbf{y}_{t+s} - \hat{\mathbf{y}}_{t+s|t})'] \\ &= \mathbf{\Omega} + \mathbf{\Psi}_1 \mathbf{\Omega} \mathbf{\Psi}_1' + \mathbf{\Psi}_2 \mathbf{\Omega} \mathbf{\Psi}_2' + \ldots + \mathbf{\Psi}_{s-1} \mathbf{\Omega} \mathbf{\Psi}_{s-1}' \end{aligned}$$

The Choleski ordering allows a decomposition of the variance of the composite innovations into the contributions of the pure innovations

$$\varepsilon_t = \mathbf{A}\mathbf{u}_t = \mathbf{a}_1 u_{1t} + \mathbf{a}_2 u_{2t} + \ldots + \mathbf{a}_n u_{nt}$$

$$\boldsymbol{\Omega} = \mathbb{E}(\varepsilon_t \varepsilon_t') = \mathbf{A} \cdot \mathbb{E}(\mathbf{u}_t \mathbf{u}_t') \cdot \mathbf{A}' = \mathbf{A}\mathbf{D}\mathbf{A}'$$

$$= \mathbf{a}_1 \mathbf{a}_1' \cdot \mathsf{Var}(u_{1t}) + \mathbf{a}_2 \mathbf{a}_2' \cdot \mathsf{Var}(u_{2t}) + \ldots + \mathbf{a}_n \mathbf{a}_n' \cdot \mathsf{Var}(u_{nt})$$

We can also decompose the MSE of the s-step ahead forecast

$$
\begin{aligned}
MSE(\hat{\mathbf{y}}_{t+s|t}) &= \mathbb{E}[(\mathbf{y}_{t+s} - \hat{\mathbf{y}}_{t+s|t})(\mathbf{y}_{t+s} - \hat{\mathbf{y}}_{t+s|t})'] \\
&= \mathbf{\Omega} + \mathbf{\Psi}_1 \mathbf{\Omega} \mathbf{\Psi}_1' + \mathbf{\Psi}_2 \mathbf{\Omega} \mathbf{\Psi}_2' + \ldots + \mathbf{\Psi}_{s-1} \mathbf{\Omega} \mathbf{\Psi}_{s-1}'
\end{aligned}
$$

$$
\underset{(n\times n)}{MSE(\hat{\mathbf{y}}_{t+s|t})} = \sum_{j=1}^{n} \{\text{Var}(u_{jt}) \cdot [\mathbf{a}_j \mathbf{a}_j' + \mathbf{\Psi}_1 \mathbf{a}_j \mathbf{a}_j' \mathbf{\Psi}_1' + \mathbf{\Psi}_2 \mathbf{a}_j \mathbf{a}_j' \mathbf{\Psi}_2' + \ldots + \mathbf{\Psi}_{s-1} \mathbf{a}_j \mathbf{a}_j' \mathbf{\Psi}_{s-1}']\}
$$

contribution of the jth orthogonalized innovation to the MSE of the s-period-ahead forecast:

$$
\text{Var}(u_{jt}) \cdot [\mathbf{a}_j \mathbf{a}_j' + \mathbf{\Psi}_1 \mathbf{a}_j \mathbf{a}_j' \mathbf{\Psi}_1' + \mathbf{\Psi}_2 \mathbf{a}_j \mathbf{a}_j' \mathbf{\Psi}_2' + \ldots + \mathbf{\Psi}_{s-1} \mathbf{a}_j \mathbf{a}_j' \mathbf{\Psi}_{s-1}']
$$

The numerical example illustrates the decomposition of the variance covariance matrix of the composite shocks (MSE 1 step forecast)

$$\underset{(n\times n)}{MSE(\hat{\mathbf{y}}_{t+1|t})} = \text{Var}(u_t^T)\cdot[\mathbf{a}_1\mathbf{a}_1'] + \text{Var}(u_t^S)\cdot[\mathbf{a}_2\mathbf{a}_2'] + \text{Var}(u_t^K)\cdot[\mathbf{a}_3\mathbf{a}_3']$$

$$\underset{(n\times n)}{MSE(\hat{\mathbf{y}}_{t+1|t})} = 1.79\cdot\begin{bmatrix}1.000 & 0.344 & 0.087\\ 0.344 & 0.119 & 0.030\\ 0.087 & 0.030 & 0.008\end{bmatrix} + 1.78\cdot\begin{bmatrix}0.000 & 0.000 & 0.000\\ 0.000 & 1.000 & 0.127\\ 0.000 & 0.127 & 0.016\end{bmatrix}$$

$$+\ 2.63\cdot\begin{bmatrix}0.000 & 0.000 & 0.000\\ 0.000 & 0.000 & 0.000\\ 0.000 & 0.000 & 1.000\end{bmatrix}$$

$$= \begin{bmatrix}1.793 & 0.618 & 0.157\\ 0.618 & 1.994 & 0.281\\ 0.157 & 0.281 & 2.674\end{bmatrix}$$

$\text{Var}(u_t^T)$, $\text{Var}(u_t^S)$, $\text{Var}(u_t^K)$ and $MSE(\hat{\mathbf{y}}_{t+1|t})$ taken times 10000

The numerical example illustrates the decomposition of the MSE of the two step forecast

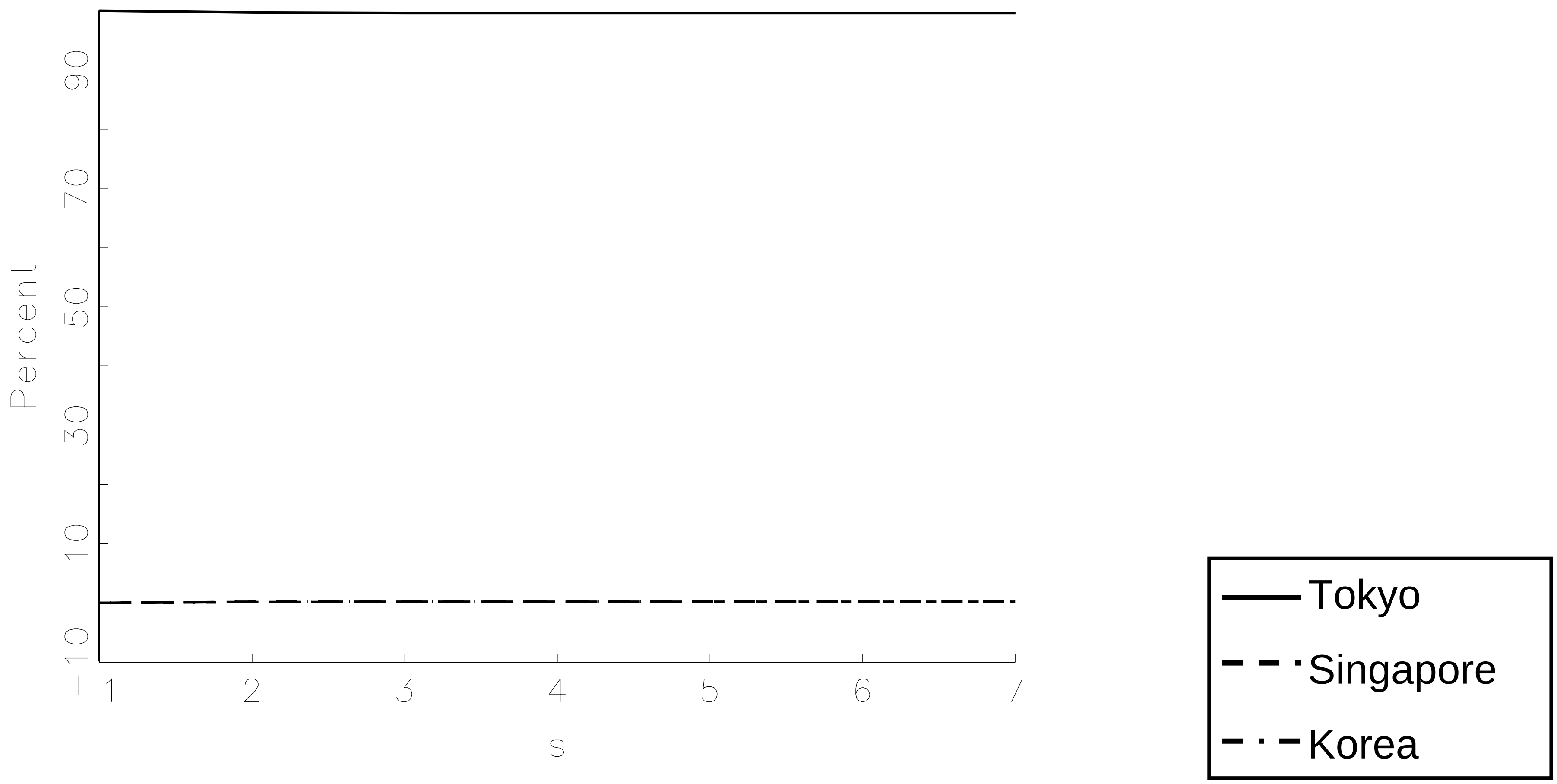
$$\underset{(n\times n)}{MSE(\hat{\mathbf{y}}_{t+2|t})} = \mathsf{Var}(u_t^T)[\mathbf{a}_1\mathbf{a}_1' + \mathbf{\Psi}_1\mathbf{a}_1\mathbf{a}_1'\mathbf{\Psi}_1'] + \mathsf{Var}(u_t^S)[\mathbf{a}_2\mathbf{a}_2' + \mathbf{\Psi}_1\mathbf{a}_2\mathbf{a}_2'\mathbf{\Psi}_1'] + \mathsf{Var}(u_t^K)[\mathbf{a}_3\mathbf{a}_3' + \mathbf{\Psi}_1\mathbf{a}_3\mathbf{a}_3'\mathbf{\Psi}_1']$$

$$\underset{(n\times n)}{MSE(\hat{\mathbf{y}}_{t+2|t})} = 1.79 \cdot \begin{bmatrix} 1.000 & 0.343 & 0.086 \\ 0.343 & 0.125 & 0.035 \\ 0.086 & 0.035 & 0.012 \end{bmatrix} + 1.78 \cdot \begin{bmatrix} 0.001 & 0.008 & 0.004 \\ 0.008 & 1.040 & 0.147 \\ 0.004 & 0.147 & 0.026 \end{bmatrix} + 2.63 \cdot \begin{bmatrix} 0.001 & 0.002 & 0.002 \\ 0.002 & 0.002 & 0.003 \\ 0.002 & 0.003 & 1.004 \end{bmatrix}$$

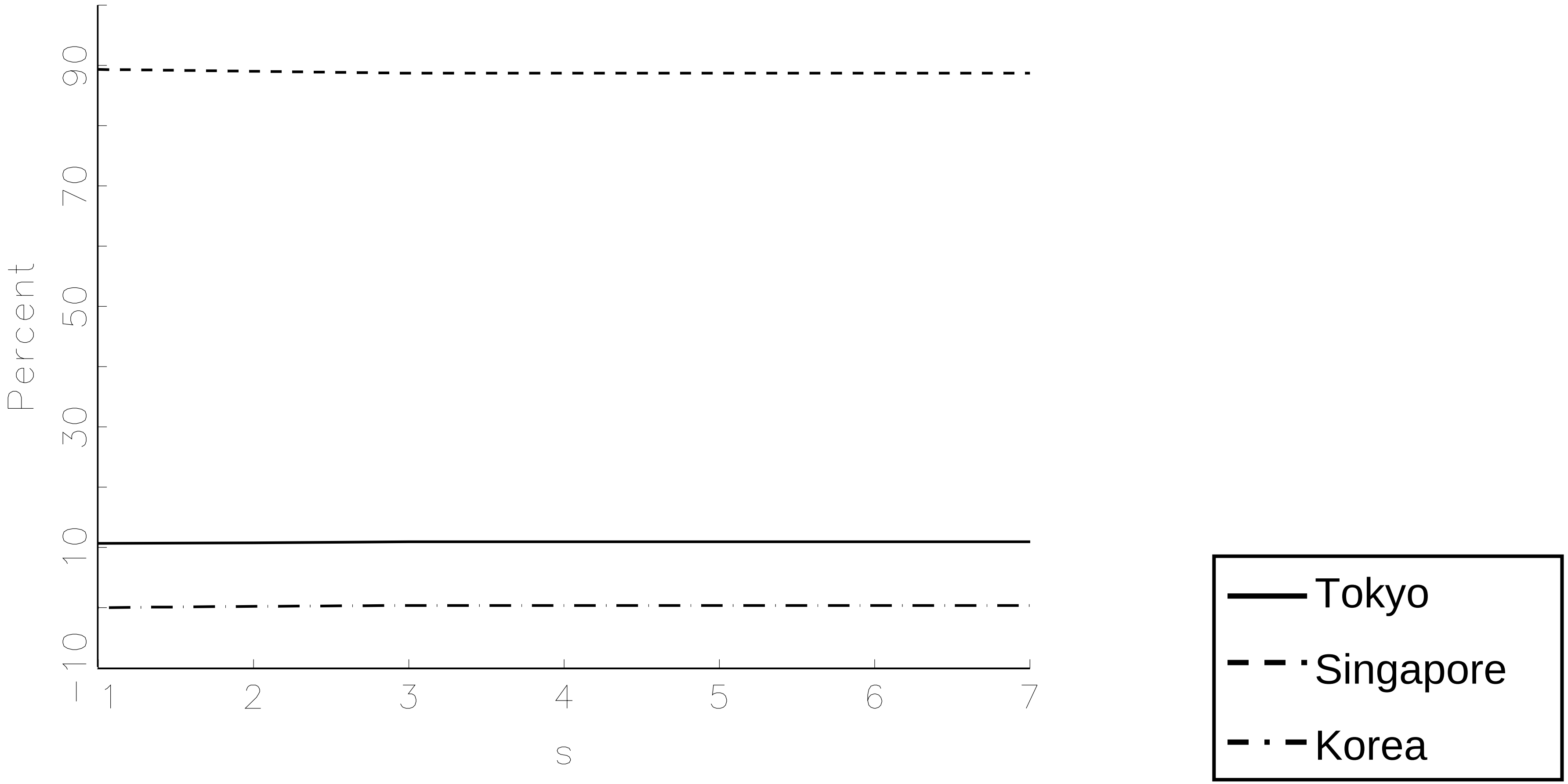
$$= \begin{bmatrix} 1.799 & 0.633 & 0.167 \\ 0.633 & 2.082 & 0.332 \\ 0.167 & 0.332 & 2.707 \end{bmatrix}$$

$\mathsf{Var}(u_t^T), \mathsf{Var}(u_t^S), \mathsf{Var}(u_t^K)$ and $MSE(\hat{\mathbf{y}}_{t+2|t})$ taken times 10000

Variance Decomposition of Tokyo
Cholesky Ordering: Tokyo Singapore Korea

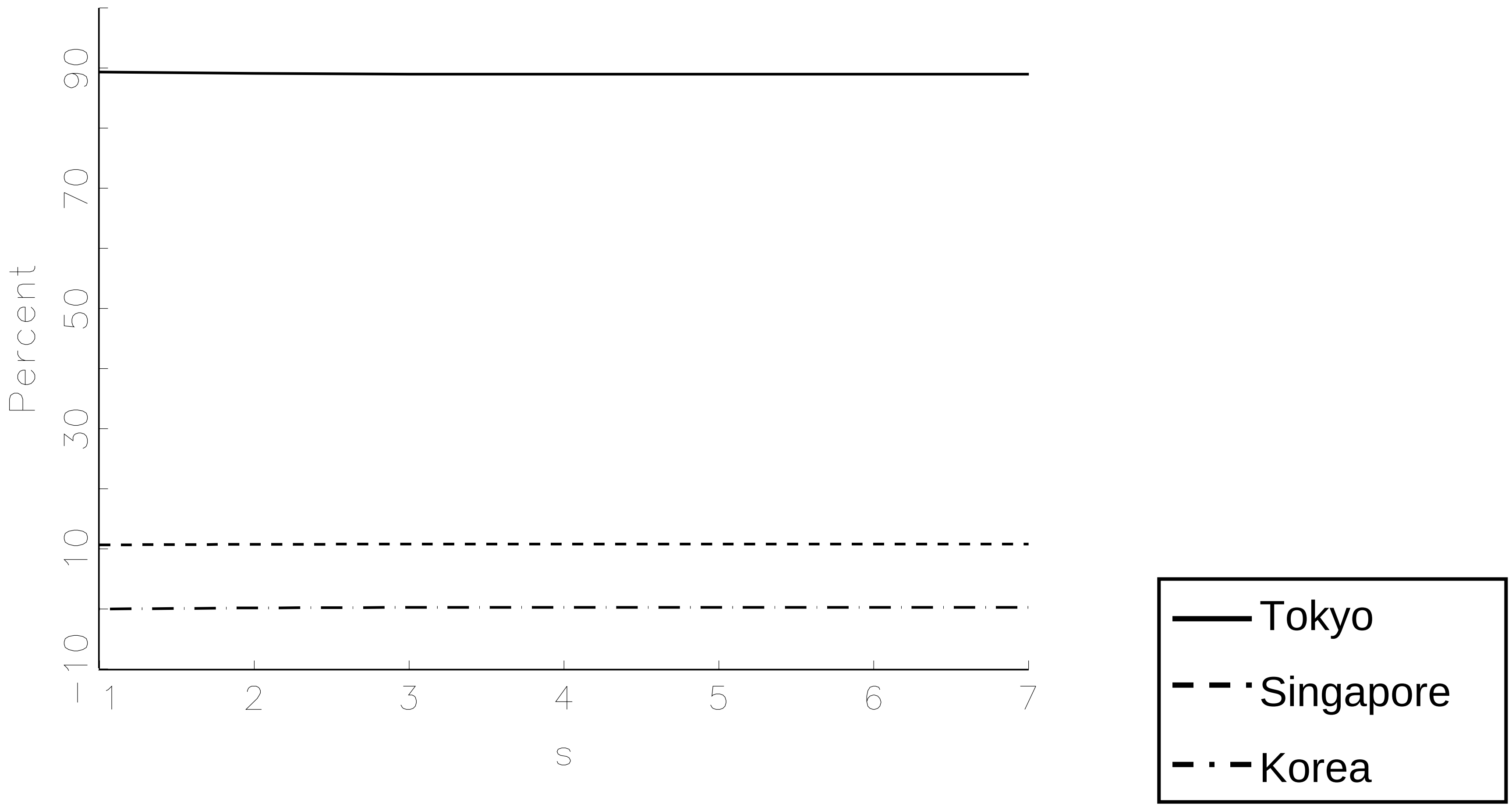


Variance Decomposition of Singapore
Cholesky Ordering: Tokyo Singapore Korea

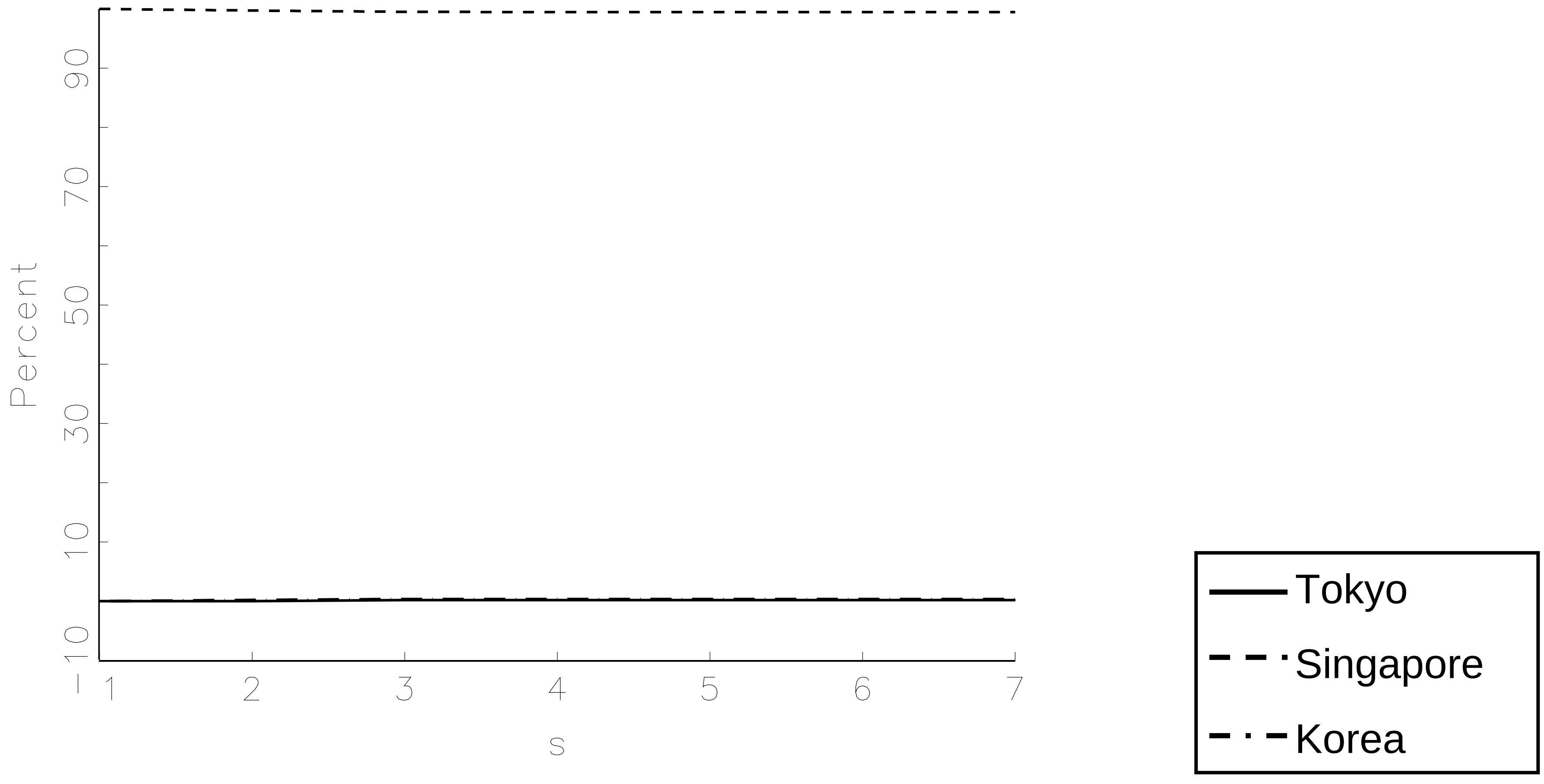


Variance Decomposition of Korea
Cholesky Ordering: Tokyo Singapore Korea

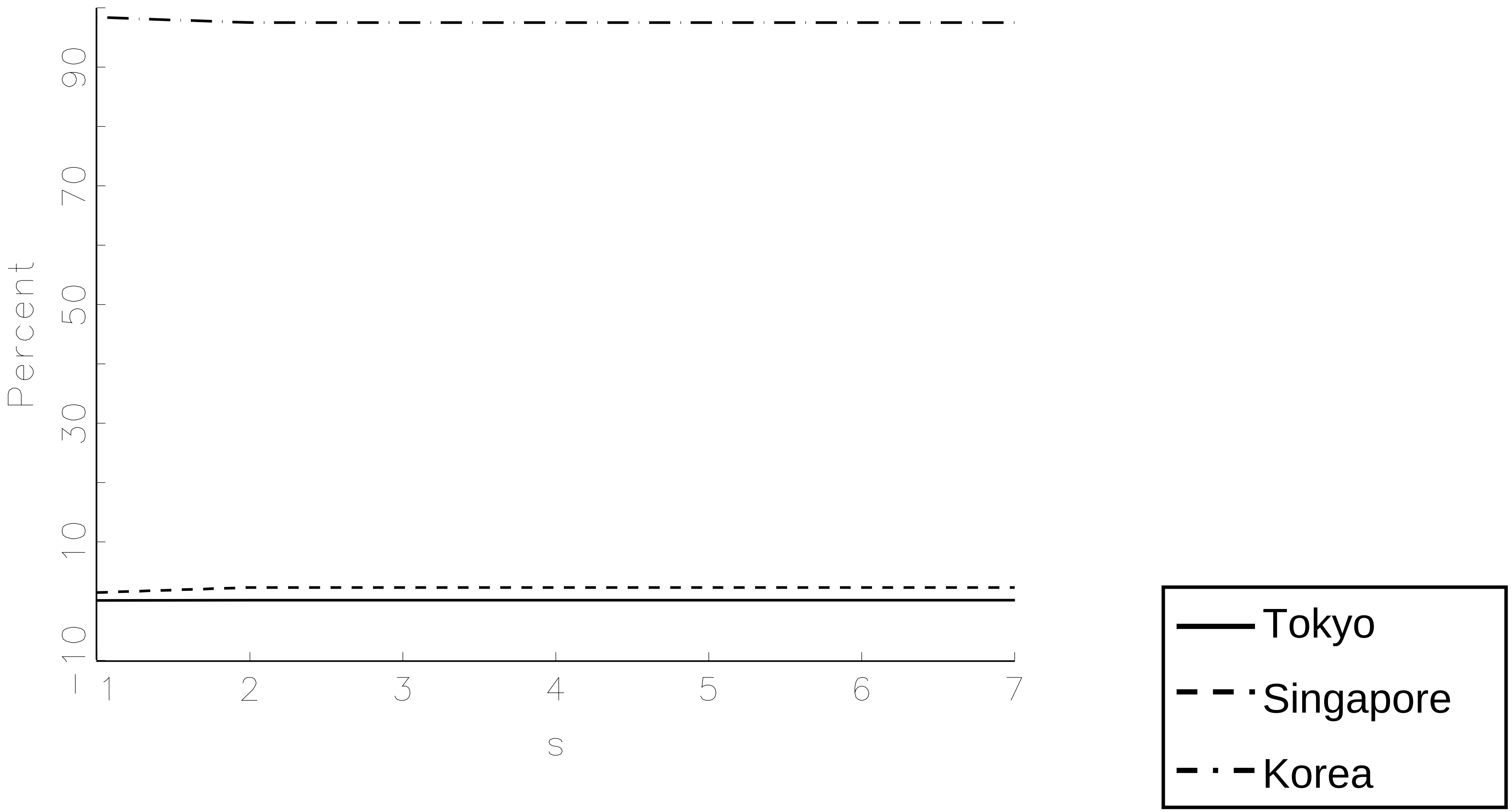
Variance Decomposition of Tokyo
Cholesky Ordering: Singapore Tokyo Korea



Variance Decomposition of Singapore
Cholesky Ordering: Singapore Tokyo Korea



Variance Decomposition of Korea
Cholesky Ordering: Singapore Tokyo Korea



Application of Cointegration Methods in Finance

Internationally cross-listed stock prices during overlapping trading hours: price discovery and exchange rate effects
Journal of Empirical Finance 12 (2005), 139-164

Joachim Grammig , Michael Melvin , Christian Schlag

Overview

- Motivation
- Theoretical background and econometric modeling
- Data and empirical results
- Conclusion

Hypothesis regarding price discovery in international equity trading and empirical tests based on high frequency data

- Simultaneous trading of same asset at different trading venues
- Worldwide competition for liquidity. Viability of securities markets depends on performance of trading mechanisms.
  Efficient capital market: Value-relevant information flows quickly into prices.
- *Q1: Price discovery in home market or at the world's leading trading venue?*
- Bacidore/Sofianos (2000): "**Price discovery takes place at home and NYSE market participants take those prices as given**"
- ⇨"Winner market takes all"-hypothesis (Chowdry and Nanda, RFS 1991): In case of international parallel trading one market will dominate price discovery.
- Kim/Szakmary/Mathur (JBF 2000): Home market dominates price discovery.
  Problem: Aggregation of price dynamics in daily data. Non-simultaneous trading (time zones).
- *Q2: Symmetric reaction of stock prices to exchange rate movements?*

Starting point: 100 % Price discovery in home market

P_t^h : Stock price home market at time t in (log)

P_t^u : Stock price US market in \$ (log)

E_t : \$/ exchange rate (log)

E_t and P_t^h follow random walks

$$E_t = E_{t-1} + \varepsilon_t^e$$

$$P_t^h = P_{t-1}^h + \varepsilon_t^h$$

The US price tracks the home market price:

.

$$P_t^u = P_{t-1}^h + E_{t-1} + \varepsilon_t^u$$

Cointegration between home market price, US price and exchange rate

Arbitrage prevents long run deviations from equilibrium $\Longrightarrow$
log-exchange rate, log-€-Kurs und log-$-Kurs are cointegrated

$$
\begin{aligned}
P_t^h - P_t^u + E_t &= \\
\left[P_{t-1}^h + \varepsilon_t^h - P_{t-1}^h - E_{t-1} - \varepsilon_t^u + E_{t-1} + \varepsilon_t^e\right] &= \\
\varepsilon_t^h - \varepsilon_t^u + \varepsilon_t^e &
\end{aligned}
$$

with cointegrating vector (1 -1 1).

- Only own innovations ε_t^h exert permanent impact on **€** price. (100% information share)
- Only own innovations ε_t^e exert permanent impact on exchange rate. (100% information share)
- $-Preis: Merely transitory influence of own market innovations ε_t^u.
  Only home market and exchange rate innovations permanently impounded in US price.

In a general model the innovations of all three price series contribute to the long run dynamics of the system

One cointegrating relation between €-price, $-price and exchange rate but...

- .. innovations ε_t^h, ε_t^e and ε_t^u may exert permanent effects on all three price series
- .. their importance (the information share) is determined empirically.

Non-stationary VAR using €-price, $-price and exchange rate.

Cointegration between €-price, $-price and exchange rate.

Granger representation theorem ⇨ VECM

Write VECM in VMA representation and simulate VMA parameters

Decompose variance of long run effect of each price series into the effects caused by the innovations of each series.

Variance Share = Information Share

In a general model the innovations of all three price series contribute to the long run dynamics of the system

Assumptions for a general model:

One cointegrating relation between €-price, $-price and exchange rate but...

-.. Innovations ε_t^h, ε_t^e and ε_t^u may exert permanent effects on all three price series.

.. their importance (the information share) is determined empirically.

Estimation of the information shares is based on a VECM

Non-stationary VAR using €-price, $-price and exchange rate.

Cointegration between €-price, $-price and exchange rate.

Granger representation theorem ⇒ VECM

$$\Delta E_t = \beta_1(\alpha_1 P^h_{t-1} - \alpha_2 P^u_{t-1} - \alpha_3 E_t) + \delta_{11}\Delta P^h_{t-1} + \delta_{12}\Delta P^u_{t-1} + \delta_{13}\Delta E_{t-1} + \varepsilon^e_t$$

$$\Delta P^h_t = \beta_2(\alpha_1 P^h_{t-1} - \alpha_2 P^u_{t-1} - \alpha_3 E_t) + \delta_{21}\Delta P^h_{t-1} + \delta_{22}\Delta P^u_{t-1} + \delta_{23}\Delta E_{t-1} + \varepsilon^h_t$$

$$\Delta P^u_t = \beta_3(\alpha_1 P^h_{t-1} - \alpha_2 P^u_{t-1} - \alpha_3 E_t) + \delta_{31}\Delta P^h_{t-1} + \delta_{32}\Delta P^u_{t-1} + \delta_{33}\Delta E_{t-1} + \varepsilon^u_t$$

By simulating the VECM we obtain the weight matrix from which the information shares can be computed

Write VECM in VMA representation: $\begin{bmatrix} \Delta E_t \\ \Delta P_t^h \\ \Delta P_t^u \end{bmatrix} = \begin{bmatrix} \varepsilon_t^e \\ \varepsilon_t^h \\ \varepsilon_t^u \end{bmatrix} + \mathbf{\Psi}_1 \begin{bmatrix} \varepsilon_{t-1}^e \\ \varepsilon_{t-1}^h \\ \varepsilon_{t-1}^u \end{bmatrix} + \mathbf{\Psi}_2 \begin{bmatrix} \varepsilon_{t-2}^e \\ \varepsilon_{t-2}^h \\ \varepsilon_{t-2}^u \end{bmatrix} + \ldots$

$$\mathbf{\Psi} = \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} \\ \psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix} = \mathbf{I} + \mathbf{\Psi}_1 + \mathbf{\Psi}_2 \; + \; \ldots$$

$$\begin{pmatrix} \text{permanent impact on exchange rate} \\ \text{permanent impact on -Price} \\ \text{permanent impact on \$-Price} \end{pmatrix} = \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} \\ \psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix} \times \begin{bmatrix} \varepsilon_t^e \\ \varepsilon_t^h \\ \varepsilon_t^u \end{bmatrix}$$

(follows from Stock/Watson's common trends representation of cointegrated systems)

Economoic common sense: $\psi_{12} = 0$, $\psi_{13} = 0$: Stock prices do not affect exchange rate.

Cointegration implies $\psi_{22} = \psi_{32}$ and $\psi_{23} = \psi_{33}$.

Hasbrouck (1995): Defines the information share of a market as its contribution to the variance of the permanent component of a given price series

Var(perm. impact on exchange rate)

$$= \psi_{11}^2 \mathrm{Var}(\varepsilon_t^e) \;+\; \psi_{12}^2 \mathrm{Var}(\varepsilon_t^h) + \psi_{13}^2 \mathrm{Var}(\varepsilon_t^u)$$

(neglecting contemporaneous correlations)

$$\frac{\psi_{11}^2 \mathrm{Var}(\varepsilon_t^e)}{\psi_{11}^2 \mathrm{Var}(\varepsilon_t^e) + \psi_{12}^2 \mathrm{Var}(\varepsilon_t^h) + \psi_{13}^2 \mathrm{Var}(\varepsilon_t^u)} \equiv \text{Information Share}$$

Hypothesized $\psi_{12} = 0,\ \psi_{13} = 0$

100% of relevant information is generated in exchange rate series itself (Empirically testable)

Information shares for home market and US market?

“Winner market takes all”-hypothesis: One market dominates!

Sofianos’ “home market hypothesis”.

The empirical analysis is based on high frequency data for three NYSE traded German stocks and US/€ exchange rate data

XETRA (electronic trading system of German Stock Exchange) and NYSE (TAQ) bid-ask prices for SAP, Deutsche Telekom (DT) and DaimlerChrysler (DCX).

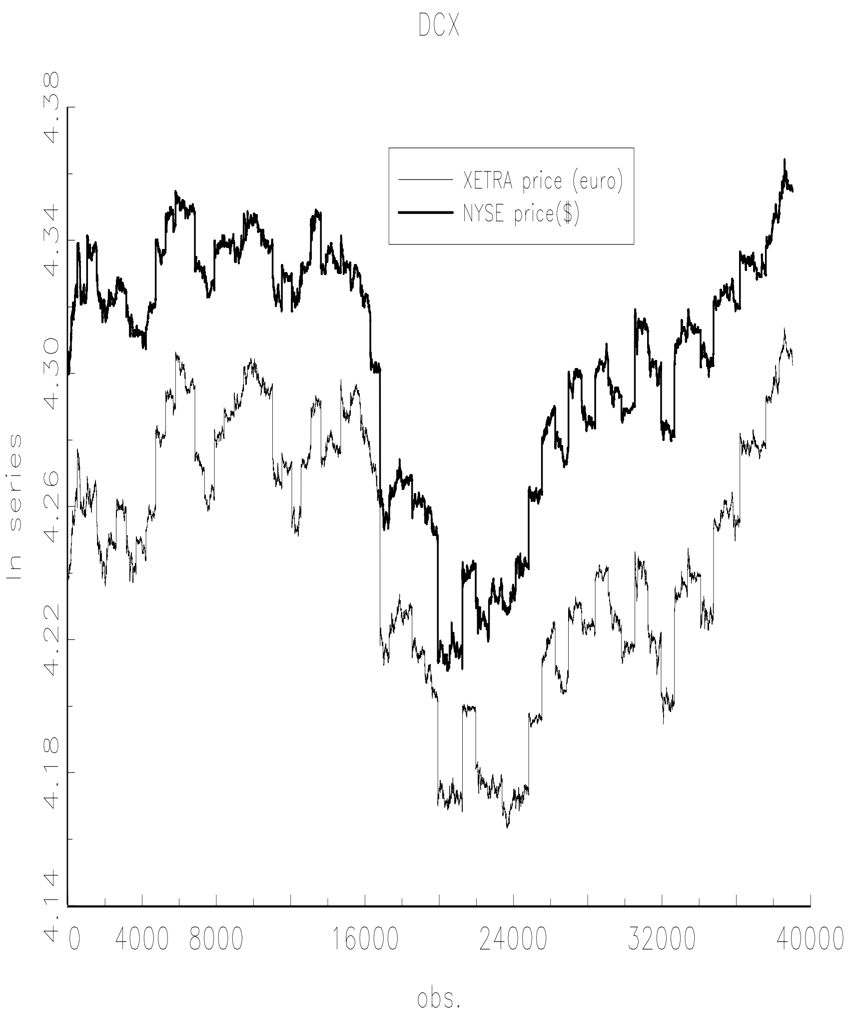
US/€ indicative quotes: Olsen & Associates, Zürich

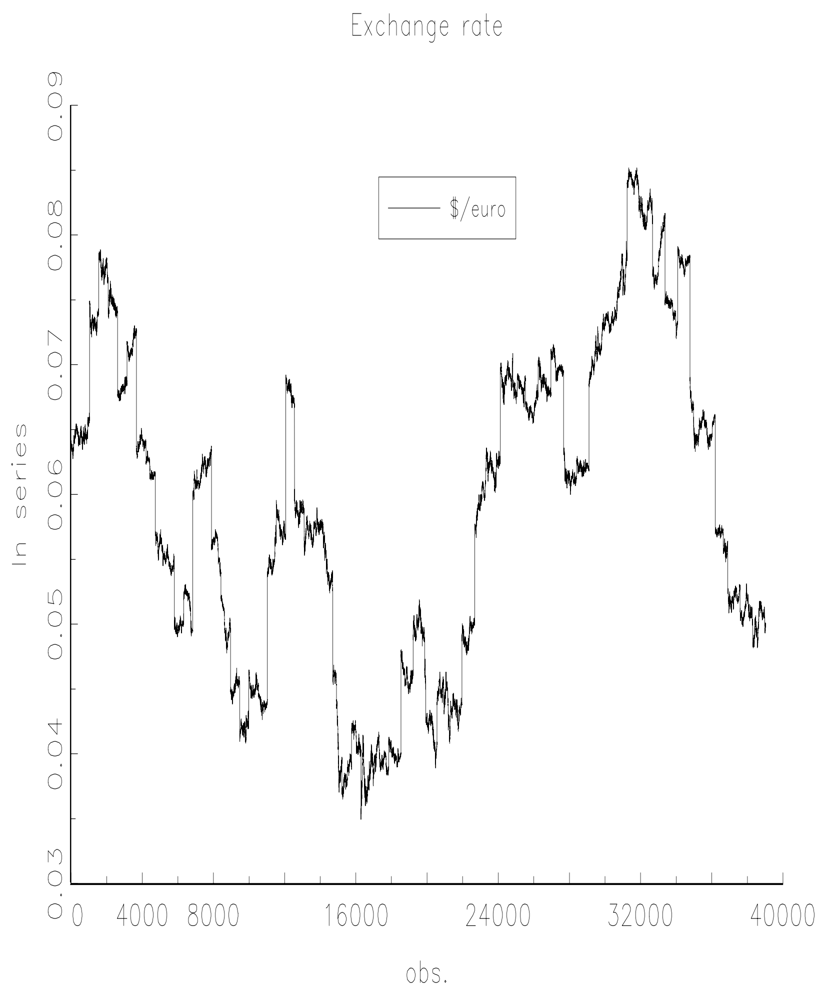
August-Oktober 1999

Mid-quotes from overlapping trading period NYSE-XETRA [GMT 14:30-16(:30)]

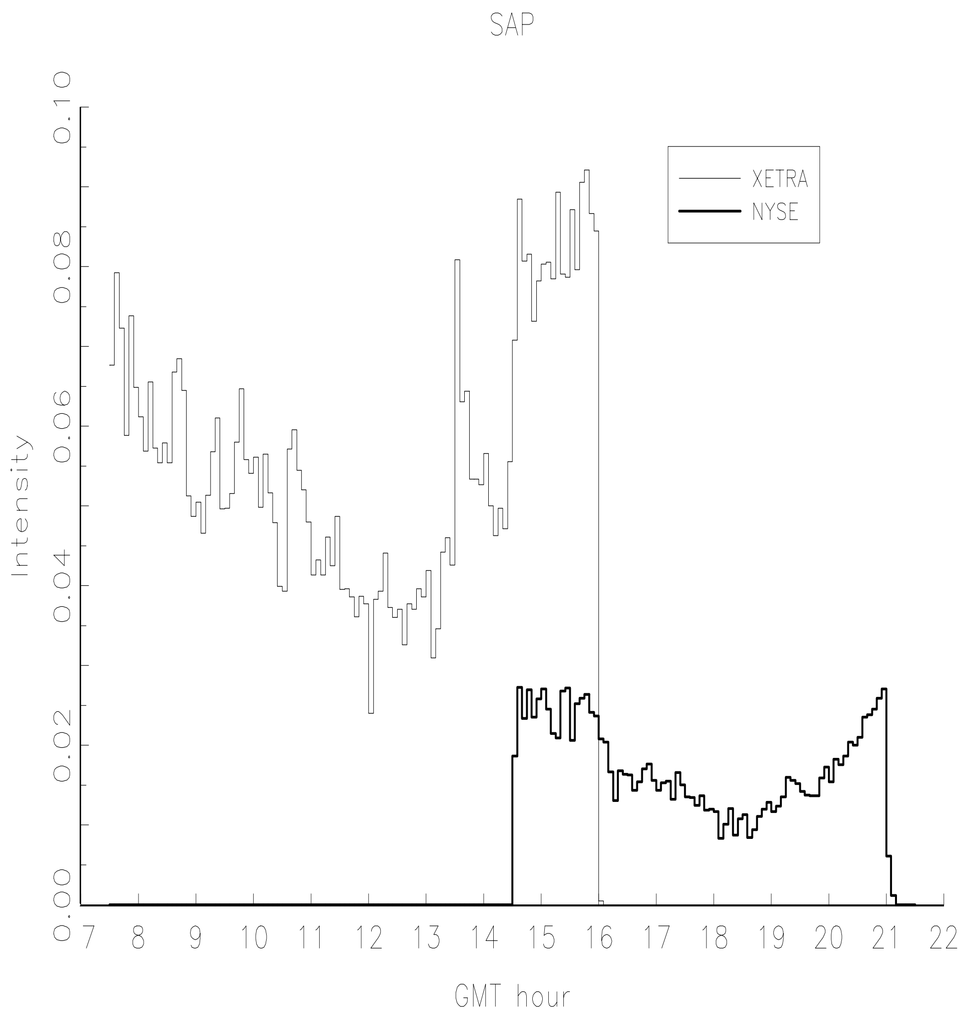
Equally spaced 10 seconds data generated from transactions data.

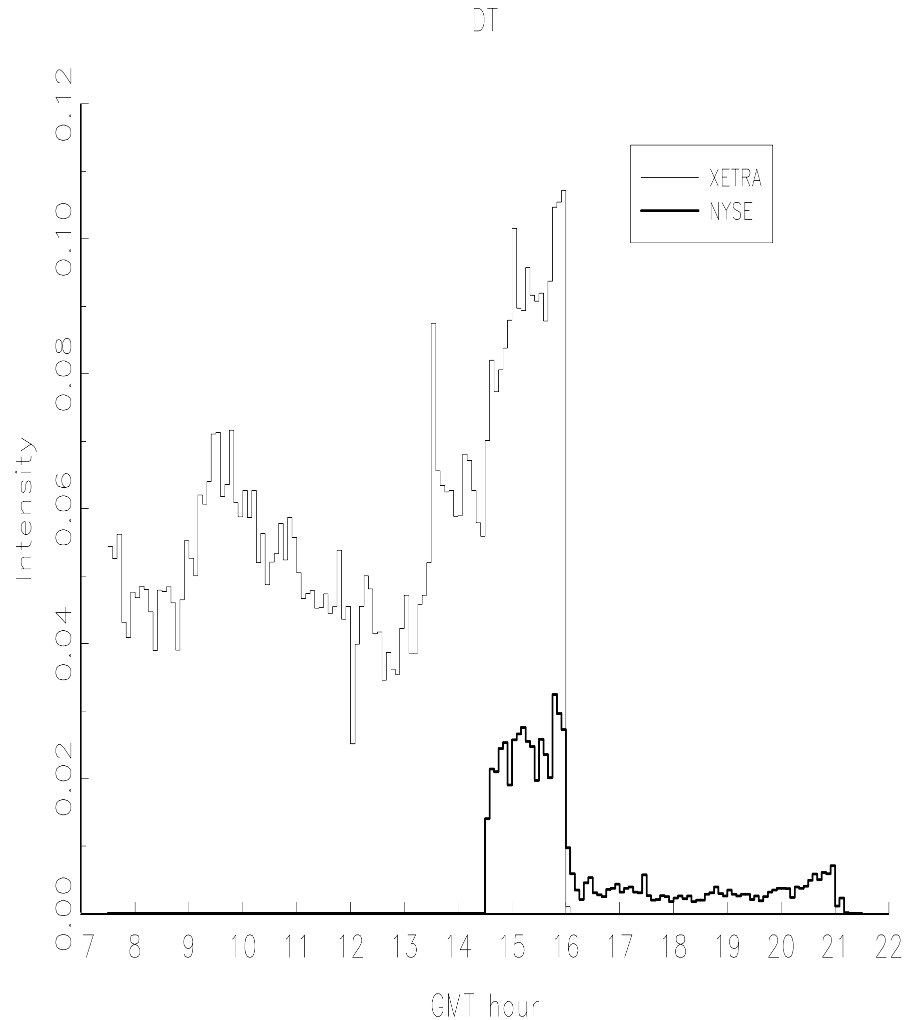
A look at the data





Comparing Deutsche Telekom and SAP one finds significant differences in intra day quoting intensity patterns





The empirical results

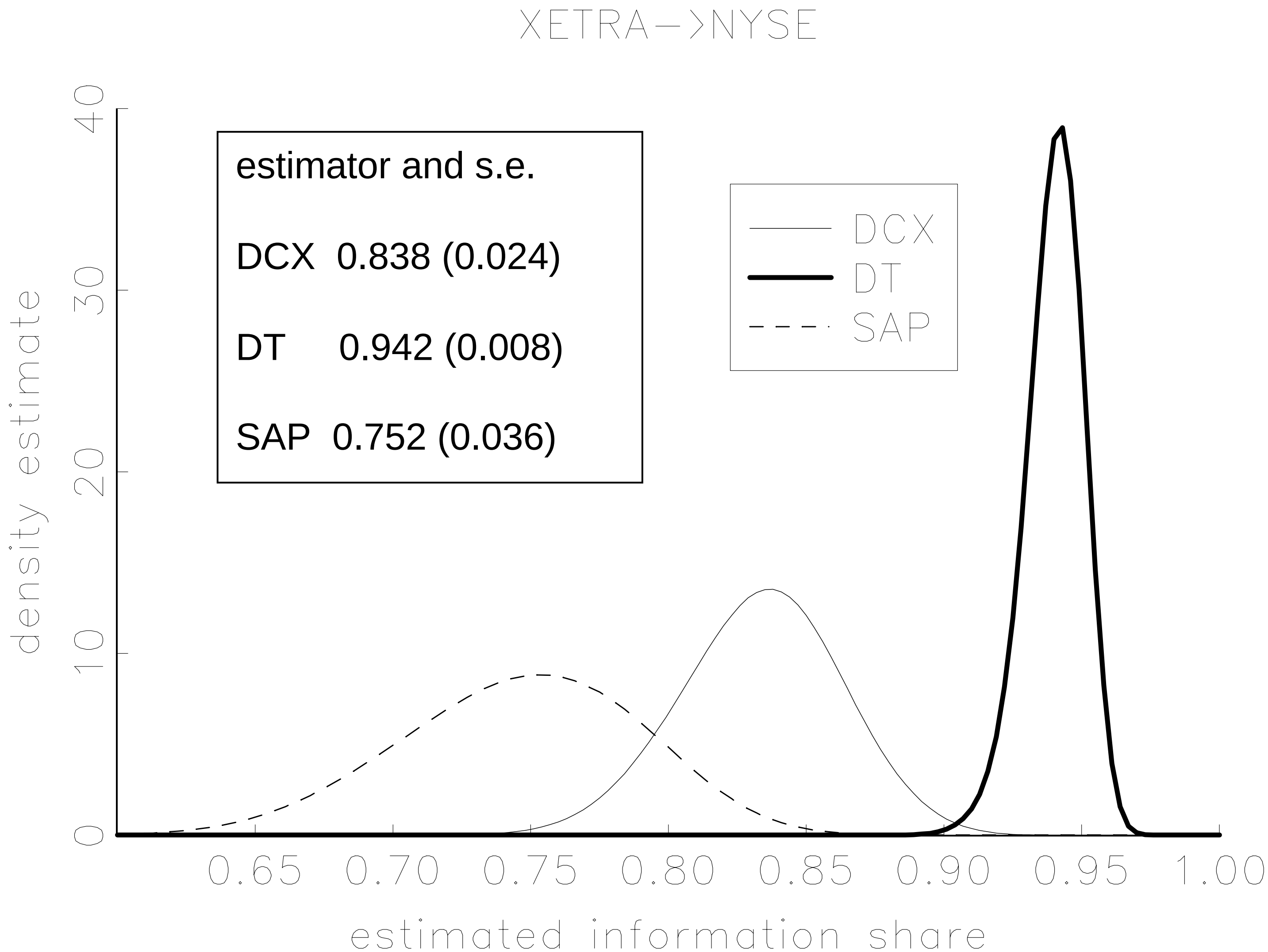
- ⇨ Johansen's method confirms the existence of ONE cointegrating relation between stock prices and exchange rate.
- ⇨ Implies two stochastic trends (efficient stock price and exchange rate).
- ⇨ As expected, no permanent impact of stock prices on exchange rates.
- ⇨ Only the US price incorporates exchange rate shocks. The home market does not react. Unexpected (?) asymmetric effect.
- ⇨ Support for "winner market takes all"-hypothesis.
- ⇨ Support for home market hypothesis, but qualitative differences are obvious:

Deutsche Telekom as "national" stock: Price discovery exclusively in Germany

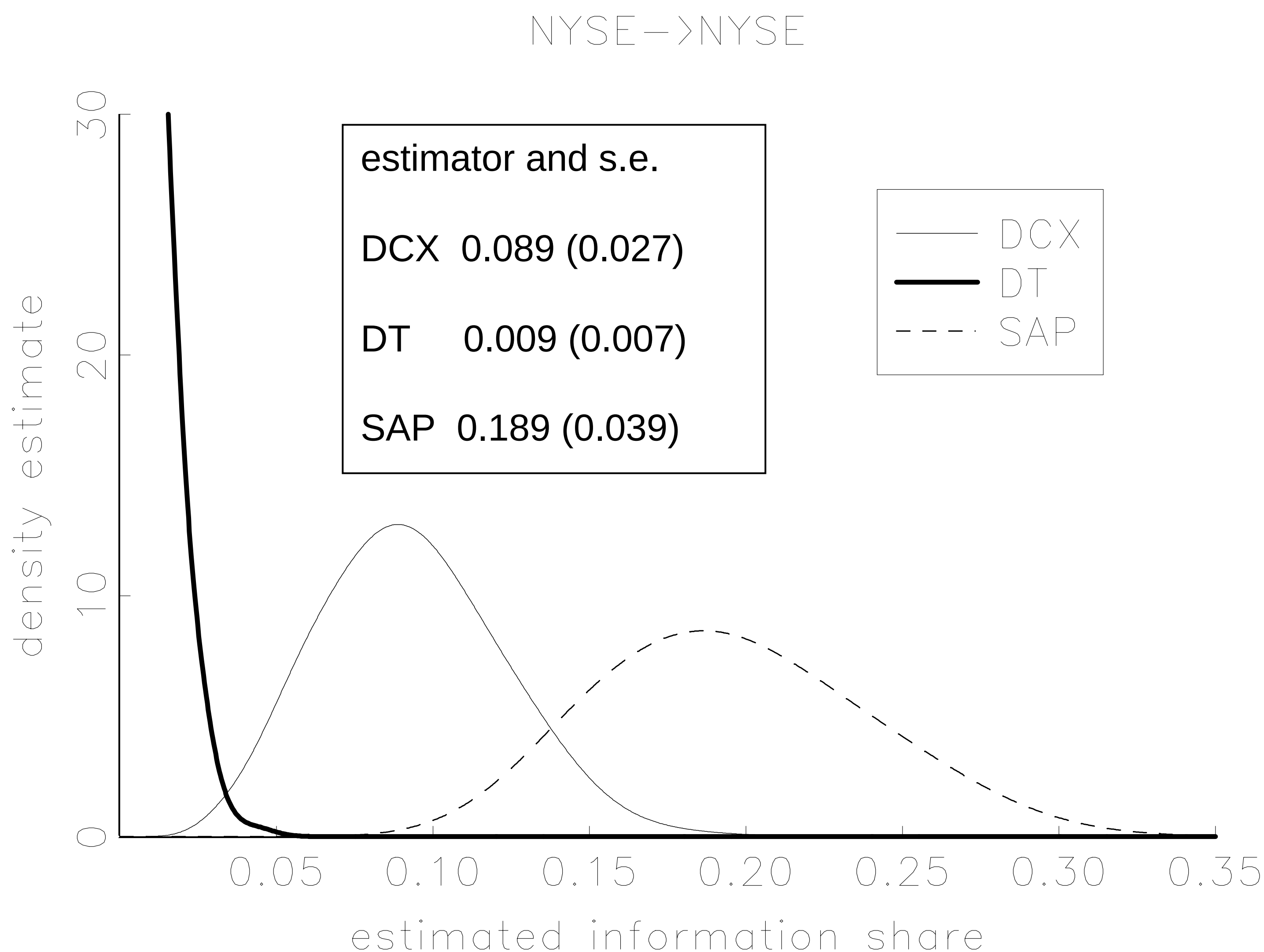
DaimlerChrysler: The larger information share is generated in the German market

SAP ("New Economy", significant US-sales): Largest US information share

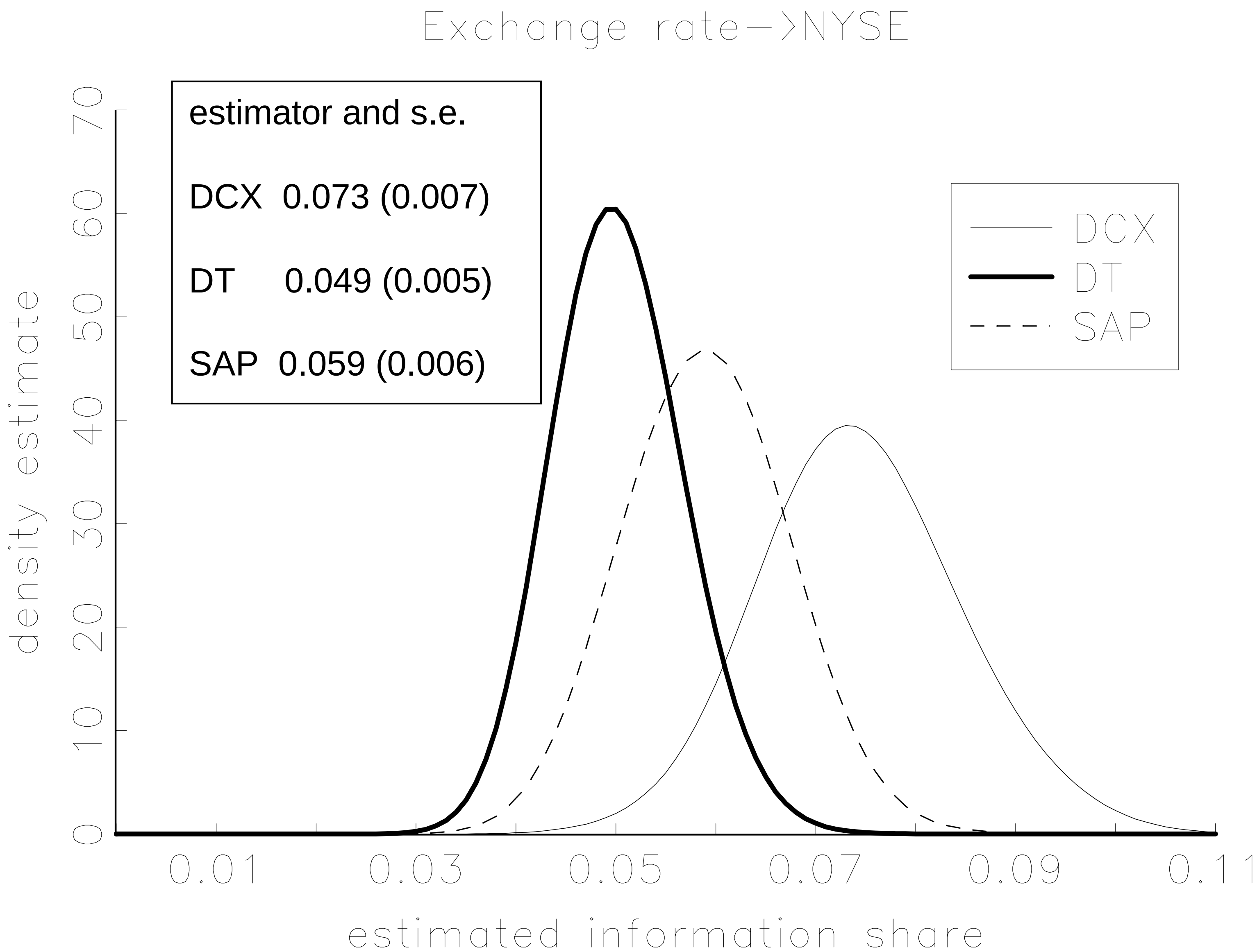
Information of share XETRA innovations w.r.t NYSE price
(Kernel density estimates based on 1000 Bootstrap replications (Li/Maddala, 1997))



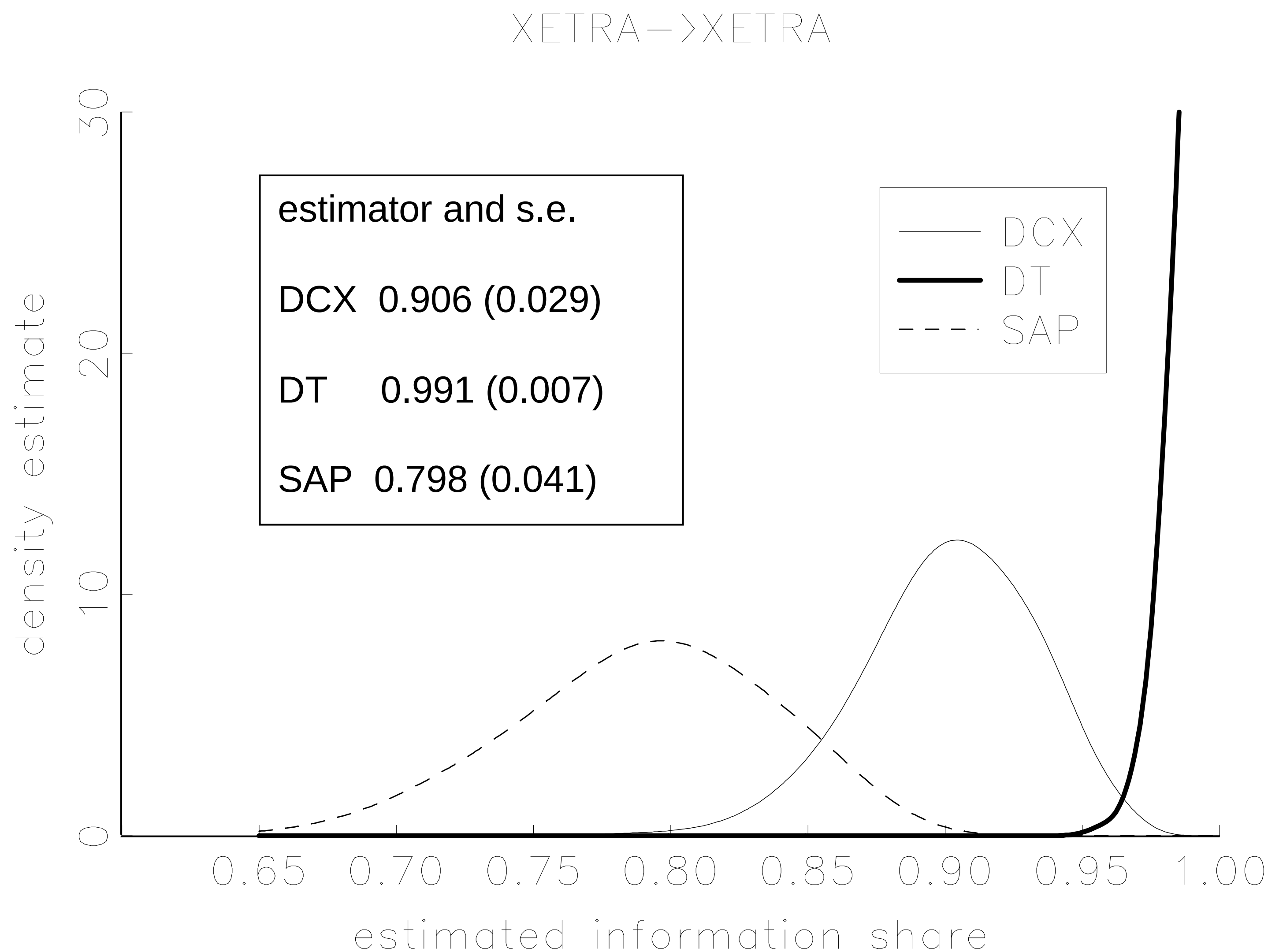
Information share of NYSE innovations w.r.t. NYSE price



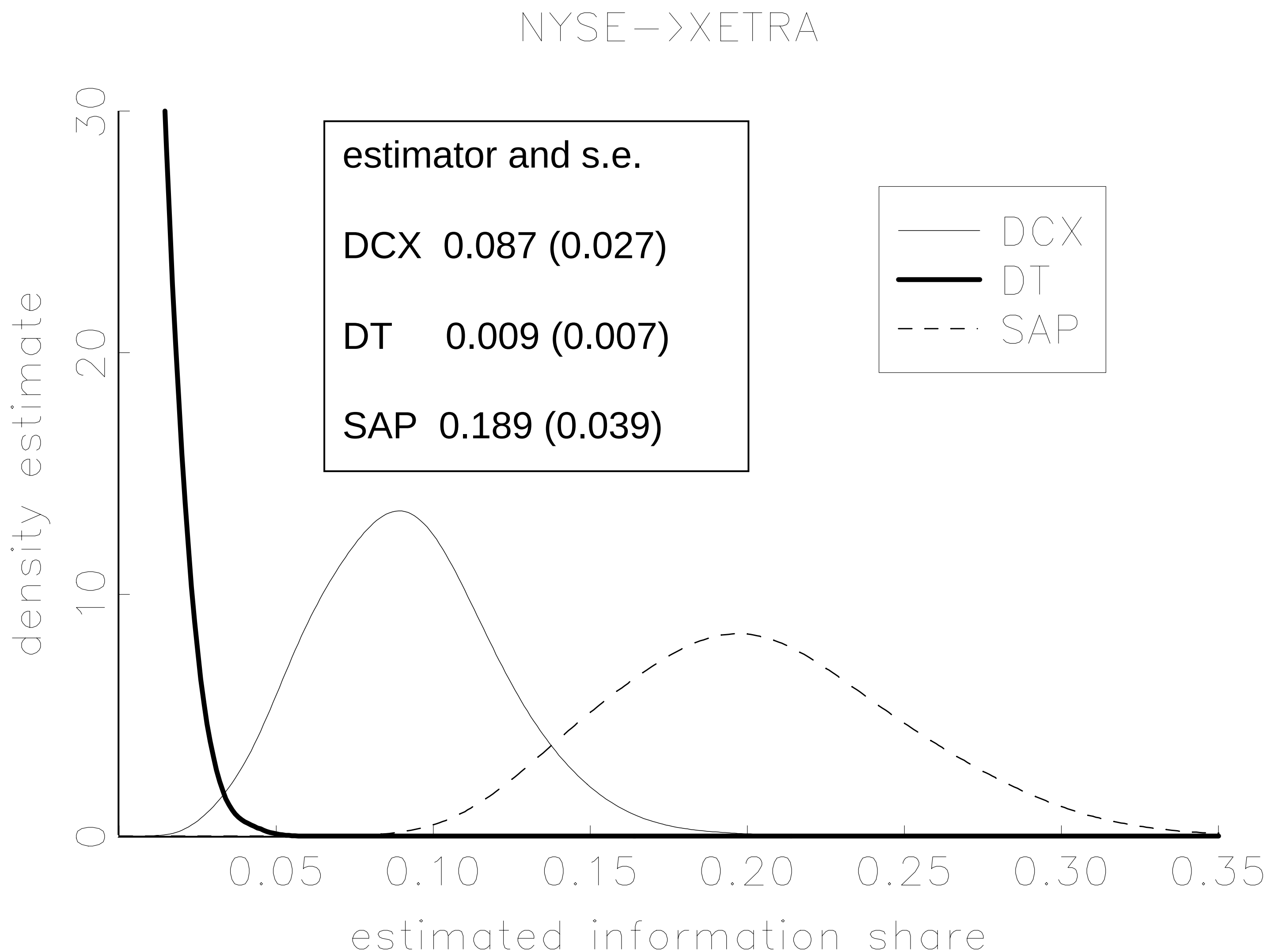
Information share of exchange rate innovations w.r.t NYSE price



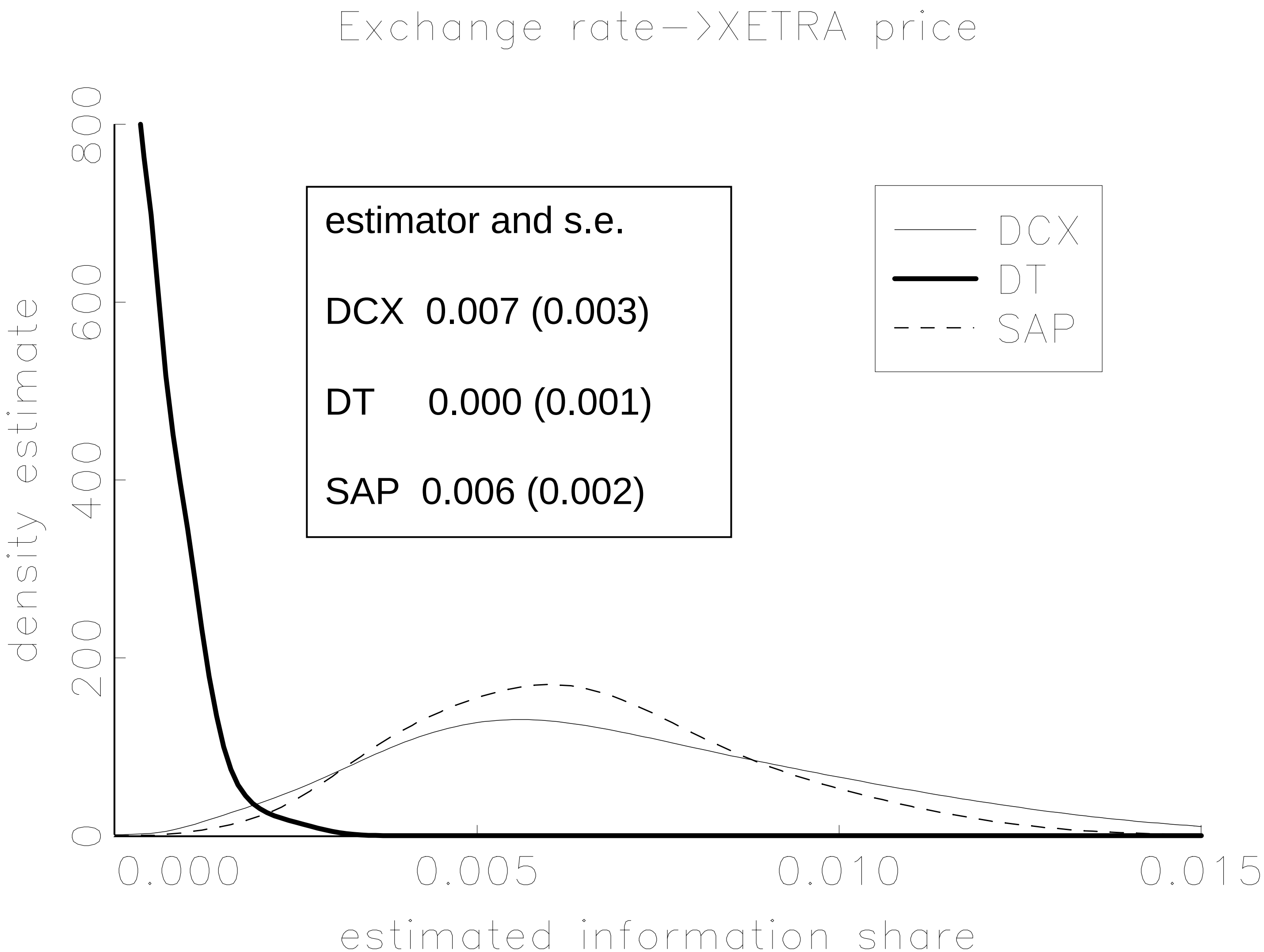
Information share of XETRA innovations w.r.t XETRA price



Information share of NYSE innovations w.r.t XETRA price



Information share of exchange rate innovations w.r.t XETRA price



Summary

⇨ One cointegrating relation between exchange rate and $ and € prices found in high frequency data.

⇨ Asymmetric price reactions in response to exchange rate shocks.

⇨ Support for “winner market takes all”-hypothesis: One market dominates price discovery.

⇨ Support for home market hypothesis.

⇨ Qualitative differences between stocks. Truly national stocks vs. stocks with larger international focus.

⇨ DaimlerChrysler: Takeover or merger among equals?

The following quote from "A Blueprint for Success", TSE, October 1998, illustrates the competitive threat from U.S. exchanges perceived by the non-U.S. exchanges.

"The TSE cannot afford to have the U.S. markets become the price discovery mechanism for Canadian interlisted stocks."

$$\begin{pmatrix} \text{permanent impact on exchange rate} \\ \text{permanent impact on -Price} \\ \text{permanent impact on \$-Price} \end{pmatrix} = \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} \\ \psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix} \times \begin{bmatrix} \varepsilon_t^e \\ \varepsilon_t^h \\ \varepsilon_t^u \end{bmatrix}$$

$$\textbf{DCX} \begin{bmatrix} 0.567\ (0.010) & 0.005\ (0.011) & 0.011\ (0.012) \\ -0.132\ (0.025) & 0.822\ (0.031) & 0.250\ (0.033) \\ 0.435\ (0.027) & 0.818\ (0.032) & 0.261\ (0.034) \end{bmatrix}$$

$$\textbf{DT} \begin{bmatrix} 0.594\ (0.006) & 0.004\ (0.007) & 0.004\ (0.008) \\ -0.046\ (0.026) & 0.879\ (0.030) & 0.081\ (0.031) \\ 0.539\ (0.027) & 0.875\ (0.030) & 0.085\ (0.031) \end{bmatrix}$$

$$\textbf{SAP} \begin{bmatrix} 0.596\ (0.007) & 0.005\ (0.008) & 0.001\ (0.008) \\ -0.149\ (0.021) & 0.689\ (0.024) & 0.287\ (0.026) \\ 0.444\ (0.023) & 0.685\ (0.025) & 0.288\ (0.026) \end{bmatrix}$$