The Role of Information for International Capital Flows

New Evidence From the SDDS

Konstantin M. Wacker

IMF Working Paper 12-242 (joint with Yuko Hashimoto)

FIW-wiiw-Seminar in International Economics, June 27, 2013

Motivation & Overview

- information & capital flows: why it matters
- measuring information & capital flows: the role of distance

Motivation & Overview

- information & capital flows: why it matters
- measuring information & capital flows: the role of distance
- our measure: compliance with IMF's Special Data Dissemination Standard (SDDS)
- main results: positive impact on FDI inflows, no effect on aggregate portfolio flows

Descriptive Statistic of Research Question

Previous empirical literature

- few investor diversify their portfolio internationally → 'home bias puzzle' (French and Poterba, 1991 AER)
- distance is negatively related to portfolio flows → informational friction!? (Portes et al., 2001 EER)
- Daude and Fratzscher (2008, JIE): FDI more responsive to 'information' than portfolio flows
- Harding and Javorcik (2011, EJ): 'informational' effect of IPAs for FDI

Previous empirical literature

French and Poterba (1991 AER) calculate optimal weights *w* for a fictive portfolio of US, Japanese, UK, French, German and Canadian assets and compare them to actual portfolio holdings

Table : Deviations of Optimal from Actual Correlation

	USA	**Japan**	**UK**
USA	0.9	-1.5	-0.2
Japan	-1.1	2.5	-0.3
UK	-0.7	-1.4	4.4

Previous empirical literature

- few investor diversify their portfolio internationally → 'home bias puzzle' (French and Poterba, 1991 AER)
- distance is negatively related to portfolio flows → informational friction!? (Portes et al., 2001 EER)
- Daude and Fratzscher (2008, JIE): FDI more responsive to 'information' than portfolio flows
- Harding and Javorcik (2011, EJ): 'informational' effect of IPAs for FDI

How to meassure information asymmetries?

- distance?
- bilateral migration / telephone trafic / newspaper circulation?

above-cited papers follow one or two of these approaches; but both might capture something else (especially for FDI):

- FDI agglomeration (Head et al., 1995 JIE; Kinoshita and Mody, 2001 CJE; Blonigen et al., 2005 JIE)
- FDI in space (Blonigen et al., 2007 EER; Baltagi et al., 2008 JoE)
- FDI & culture (Davies et al., 2008 RWE)

notable exception: Gelos and Wei (2005, JoF)

Measuring Information: SDDS

$$SDDS_{it} = \begin{cases} 1, & \text{if country i meets SDDS specification in year } t \\ 0, & \text{else} \end{cases}$$

- IMF board approved Special Data Dissemination Standard in 3/1996
- for interested member countries to signal economic and financial data to international capital markets
- Canada and US were first to meet specifications on February 19, 1999
- to date 70 member countries
- confirm to provide 18 categories of data of a certain quality timely

Estimable Model

for flow of capital type j to country i in year t

$$y_{it}^{j} = \Psi_{it}\theta^{j} + SDDS_{it}\lambda_{SDDS}^{j} + \eta_t + \alpha_i + \varepsilon_{it},$$

- fixed effect OLS model
- with time dummies (two-way FE)
- HAC standard errors
- extension: spatial correlation patterns in residuals

Results: Main Models (fixed effects)

Dependent Variable: ln(capital inflow)

model	FDI (3a)	portfolio (3b)
SDDS	0.4760***	-0.0545
	(0.1355)	(0.2425)
ln(GDP)	0.7116***	1.2999***
(-1)	(0.0795)	(0.2473)
growth	2.9280***	6.1644***
	(0.9967)	(2.1640)
investment rate	2.1197*	2.3474
(-1)	(1.1974)	(2.2172)
interest rate	-0.0002***	0.0007**
(-1)	(0.0001)	(0.0003)
KA open	0.1399**	-0.0552
	(0.0677)	(0.0853)
time dummies	yes	yes
observations	634	520

...to be continued...

Results: Main Models (fixed effects)

Dependent Variable: ln(capital inflow)

model	FDI (3a)	portfolio (3b)
exrate volatility	-3.3692***	-1.3807
	(0.8635)	(1.4529)
exrate volatility (-1)	-1.4820**	-22.1873**
	(0.5685)	(10.7241)
political risk	0.0207	0.1121**
	(0.0380)	(0.0512)
high-tech exports (-1)	-0.0000	0.0000
	(0.0000)	(0.0000)
export UVs (-1)	-0.0098**	-0.0055
	(0.0042)	(0.0071)
# patents (-1)	0.0000	-0.0000
	(0.0000)	(0.0000)
exrate (PPP) (-1)	-0.0003*	0.0002
	(0.0001)	(0.0002)
trade share	-0.0000	-0.0000
	(0.0000)	(0.0000)
WEO dummy / constant	n.s. / n.s.	n.s. / **

Robustness: underlying time-dependent process?

$$y_{it} = \alpha_i + \Psi_{it}\theta + \delta_{SDDS}\ t + \gamma_{non-SDDS}\ t + \varepsilon_{it} \qquad (1)$$

Table : Different time trends between SDDS subscribers and non-subscribers?

before year...	1996	1997	1998	1999
SDDS trend ($\hat{\delta}$)	-0.0764	-0.0898	-0.1173	0.0534
non-SDDS trend ($\hat{\gamma}$)	-0.0468	-0.0890	-0.1295	-0.0046
difference significant (p-val)?	0.563	0.989	0.822	0.377
observations	150	190	234	278

Spatial correlation

Idea:

- take standardized residuals

$$\eta_i = \frac{\varepsilon - \overline{\varepsilon}}{\sigma_\varepsilon}$$

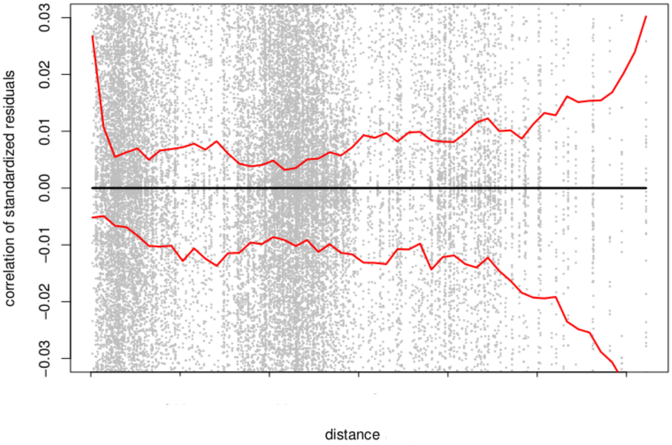
- build covariance-variance matrix $R = \eta\eta'$, where $R \in \mathbb{R}^{NT \times NT}, \ \eta \in \mathbb{R}^{NT}$
- take covariances $\rho = \{r_{ij} : \ i > j\}$ and plot them against their respective distances d_{ij}
- explore relationship between ρ and d
- local average estimator? smooting spline!

Spatial correlation

What is a significant correlation?

- asymptotic derivation of null hypothesis?
- against which alternative?
- solution: **bootstrap**:
- assign residuals (with replacement) randomly to (fixed) locations
- calculate correlation pattern
- repeat this B times and order correlations ascendingly
- take $0.05 \times B$th and $0.95 \times B$th observation to obtain 90 % confidence interval

Spatial correlation of residuals: FDI



Summary of Results

- information matters for FDI (but not aggregated portfolio) flows
- magnitude is economically relevant: about + 65 %!
- no significant pattern in spatial correlation of residuals

Economic Interpretation

- costs and economies of scale of acquiring information
- FDI not specialized in ‘investment’ but production
- shifts in portfolio investor structure may not show up on aggregate level

Policy

capital flow management

- degree of international risk diversification
- sustainability of the international investment position (cf. OECD, 2011 WEO)
- time inconsistency problem in structural reform
- marginal (private) cost of information > (expected) private opportunity cost of non-optimal investment ↔ externalities due to boom & bust cycles (cf. Bianchi, 2011 AER; Kim and Zhang, 2012 JIE; Gropp and Kadareja, 2012 JMCB)

Policy

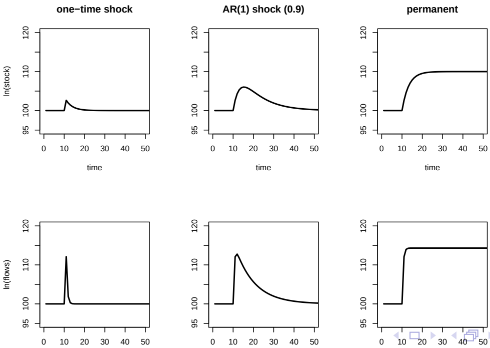
capital flow management

- degree of international risk diversification
- sustainability of the international investment position (cf. OECD, 2011 WEO)
- time inconsistency problem in structural reform
- marginal (private) cost of information $>$ (expected) private opportunity cost of non-optimal investment $\leftrightarrow$ externalities due to boom & bust cycles (cf. Bianchi, 2011 AER; Kim and Zhang, 2012 JIE; Gropp and Kadareja, 2012 JMCB)

A Simple Dynamic Model and Some FDI Evidence

$$y_{it} = \phi y_{i,t-1} + X_{it}\beta + \varepsilon_{it}$$

Figure : Response of the Dynamic Model to a Shock



FE bias in LDV Models

LDV FE: OLS on

$$\ddot{y}_{it} = \Phi \ddot{y}_{i,t-1} + \ddot{X}_{it}^{\neg y} \beta + \ddot{\varepsilon}_{it},$$

where $\ddot{x}_{it} = x_{it} - \overline{x}_i$.

$$\begin{aligned} \overline{y}_{i,t-1} &= \frac{1}{T-1}(y_{i1} + y_{i2} + \ldots + y_{i,t-1} + y_{it} + \ldots + y_{i,T-1}) \\ \overline{\varepsilon}_{it} &= \frac{1}{T-1}(\varepsilon_{i1} + \varepsilon_{i2} + \ldots + \varepsilon_{i,t-1} + \varepsilon_{it} + \ldots + \varepsilon_{i,T-1}) \end{aligned}$$

$\Rightarrow$ Cov$(X, \varepsilon) < 0$ (Nickel, 1981) $\Rightarrow \mathbb{E}(\hat{\beta}) = \beta + (X'X)^{-1} X'\varepsilon < \beta$
bias$(\hat{\beta}_{FE})$ in panels with weak dependence: T^{-1}

Estimating FDI Models

Let y_{it} be the amount of FDI (stock) in country i at time t.

$$\begin{aligned} y_{it} &= \phi y_{it-1} + \alpha_i + x_{it}\beta + \varepsilon_{it} \\ \Rightarrow y_{it} - \phi y_{it-1} &= \alpha_i + x_{it}\beta + \varepsilon_{it} \equiv y_{it}^* \end{aligned}$$

Or, alternatively (more general):

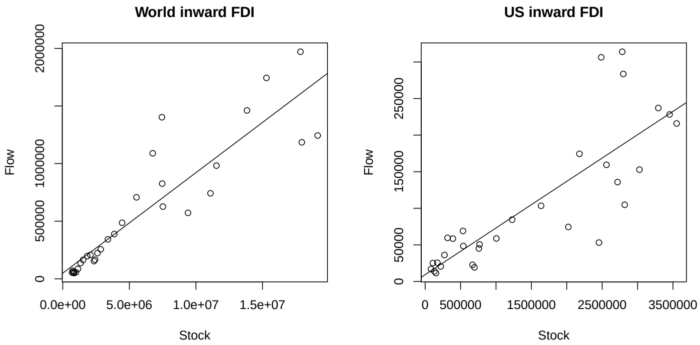
$$y_{it} = \alpha_i + x_{it}\beta_1 + x_{i,t-1}\beta_2 + x_{i,t-2}\beta_3 + x_{i,t-3}\beta_3 + \ldots + \varepsilon_{it}$$

In the steady state:

$$\begin{aligned} FDIstock_{it} &\stackrel{!}{=} FDIstock_{i,t-1} \\ FDIstock_{it} &:= FDIstock_{i,t-1} - \delta FDIstock_{i,t-1} + FDIflow_{it} \\ \delta FDIstock_{i,t} &= FDIflow_{it} \qquad \forall\ i, t. \end{aligned}$$

Estimating FDI Models

Figure : Relationship Between FDI Stock and Flow Data



Evidence

	World	US
Corellation Coefficient	0.904	0.831
$\hat{\beta}$	0.087	0.064
t-stat of $\hat{\beta}$	11.414	8.035
R-squared	0.818	0.690

Table : Regression of UNCTAD FDI Flow on Stock

Evidence

	stock	flows	assets	empl.	wages	sales	income
stock	1.00						
flows	0.77	1.00					
assets	0.93	0.67	1.00				
employment	0.71	0.51	0.69	1.00			
wages	0.80	0.54	0.83	0.89	1.00		
sales	0.87	0.62	0.84	0.88	0.95	1.00	
income	0.80	0.68	0.67	0.38	0.45	0.61	1.00
Σ	5.89	4.80	5.63	5.08	5.47	5.76	4.60

sample containing 913 observations over time (1997-2008) and host countries
Σ is the sum over *all* correlation coefficients of the measure,
not just the ones displayed.

Table : Correlation Coefficients between different BEA Measures

Estimating FDI determinants

We want to explain FDI y as a function of x:

$$y_{it} = f(x_{it}) \tag{2}$$

Effective quantity of firm's decision in t is not y but:

$$\begin{aligned} y_t^* \geq 0 &= y_t - y_{t-1} + \delta y_{t-1} && (3) \\ &= \Delta y_t + \delta y_{t-1} && (4) \\ &= y_t - (1-\delta) y_{t-1} && (5) \\ &= y_t - \phi y_{t-1}, && (6) \end{aligned}$$

hence we have to model

$$y_{it}^* = f(x_{it}) \tag{7}$$

Estimating FDI determinants

Assuming a linear relationship:

$$y^*_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it} \tag{8}$$

rearranging (8) and using (6) leads

$$\begin{aligned} y^*_{it} = y_{it} - \phi y_{it-1} &= \alpha_i + x_{it}\beta + \varepsilon_{it} \\ \Rightarrow y_{it} &= \phi y_{it-1} + \alpha_i + x_{it}\beta + \varepsilon_{it}, \end{aligned} \tag{9}$$

Different Parameter Estimates

Table : Differing Estimators and Functional Forms

model	(4)	(5)	(6)
dep. var.	log(stock)	log(flow)	log(sales)
LDV	0.7367	0.1550	0.8927
log(real GDP)	0.1093	0.5019	0.0765
a) long-run coef.	0.4150	0.5940	0.7132
b) static FE coef.	0.3573	0.6996	0.4537
c) static RE coef.	0.4158	0.5231	0.4782
d) BE coef.	0.5171	0.4929	0.5367

Note: Parameter estimates do not display any information about statistical significance.

Table : SDDS Data Coverage

Category	Component (example)
Real Sector	
National Accounts	GDP by categories
Production Indices	idustrial / commodity production
Labor market	(un)employment, wages
Price Indices	CPI, producer price index
Fiscal Sector	
General Government Operations	revenue, expenditure, financing
Central Government Operations	revenue, expenditure, financing
Central Government Debt	domestic / foreign (by currency)
Financial Sector	
Analytical Banking Accounts	money, credit
Analytical Central Bank Accounts	reserve money, domestic claims, external position
Interest Rates	government security rates
Stock Market	share price index
External Sector	
Balance of Payments	goods and services
Reserves	reserves
Merchandise Trade	exports and imports
International Investment Position	
Exchange Rates	spot rates, 3- and 6-month forward markets
External Debt	debt of government, monetary authorities, banking and other sectors
Population	

Table : List of SDDS subscribers

Country	subsc	meta	spec	Country	subsc	meta	spec
Argentina	1996	1996	1999	Jordan	2010	2010	2010
Armenia	2003	2003	2003	Kazakhstan	2003	2003	2003
Australia	1996	1998	2001	Korea	1996	1998	1999
Austria	1996	1997	2001	Kyrgyz Republic	2004	2004	2004
Belarus	2004	2004	2004	Latvia	1996	1997	1999
Belgium	1996	1997	2001	Lithuania	1996	1997	1999
Brazil	2001	2001	2001	Luxembourg	2006	2006	2006
Bulgaria	2003	2003	2003	Malaysia	1996	1996	2000
Canada	1996	1996	1999	Malta	2009	2009	2009
Chile	1996	1997	2000	Mexico	1996	1996	2000
Colombia	1996	1997	2000	Moldova	2006	2006	2006
Costa Rica	2001	2001	2001	Morocco	2005	2005	2005
Croatia	1996	1996	2001	Netherlands	1996	1996	2000
Cyprus	2009	2009	2009	Norway	1996	1996	2000
Czech Republic	1998	1998	1999	Peru	1996	1996	1999
Denmark	1996	1996	2000	Philippines	1996	1996	2001
Ecuador	1998	1998	2000	Poland	1996	1996	2000
Egypt	2005	2005	2005	Portugal	1997	1998	2000
El Salvador	1998	1998	1999	Romania	2005	2005	2005
Estonia	1998	1999	2000	Russia	2005	2005	2005
Finland	1996	1996	2000	Singapore	1996	1996	2001
France	1996	1996	2000	Slovak Republic	1996	1998	1999
Georgia	2010	2010	2010	Slovenia	1996	1996	2000
Germany	1996	1997	2000	South Africa	1996	1996	2000
Greece	2002	2002	2002	Spain	1996	1998	2000
Hong Kong SAR	1996	1997	2000	Sweden	1996	1996	2000
Hungary	1996	1997	2000	Switzerland	1996	1996	2001
Iceland	1996	1998	2004	Thailand	1996	1996	2000

Dependent Variable: Capital Inflow

- flow data (log, real values in US-$)
- either FDI/portfolio
- taken from IFS where available (> 1993)
- WEO data elsewhere (control dummy)

Investment

$$I = I(\underset{+}{Y}, \underset{-}{i}) \quad (10)$$

⇒ (expected) returns, (expected) costs

- real GDP (log, US-$, WEO)
- growth (p.a. % change in real GDP p.c., WEO)
- investment rate (gross cap form / GDP, WEO)
- capital account openness (Ito/Chinn index)
- trade share ([X+M]/GDP, IFS)
- interest rate (money market rate p.a. % over LIBOR, IFS)
- real exchange rate (implied PPP exchange rate, nat currency to US-$, WEO)
- *productivity*: high-tech exports (% or world total, WDI), export unit values (WEO), # of patents (OECD)
- *risk/uncertainty*

Risk

uncertainty and risk (can) have costs:

- sunk entry costs (Melitz, 2003 E'metrica)
- risk-averse investors
- indirect effects, e.g. productivity (Ramey and Ramey, 1995 AER; Agion et al., 2009 JME)

Risk

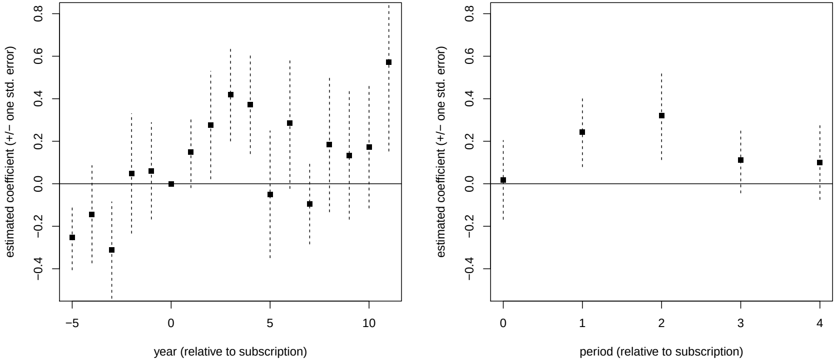
we include two types of risk:

- **political** risk indicator (ICRG; cf. Lucas, 1990 AER; Wei, 2000 RES/Brookings; Papaioannou, 2009 JDE; Kesternich and Schnitzer, 2010 JIE)
- (macro-)**economic**: exchange rate volatility (IFS, nat. currency per SDR; cf. Cushman, 1985 RES; Campa, 1993 RES)

$$\text{Exrtvol}_{it} = \sum_{m=t-1_{(7/12)}}^{t_{(6/12)}} \frac{(e_m - e_{m-1})^2}{e_{m-1}}, \tag{11}$$

where $t_{(1/12)}$ denotes the first month of year t

Model checking (FDI): Dynamics



Robustness: Identification assumptions

$$y_{it}^{j} = \Psi_{it}\theta^{j} + SDDS_{it}\lambda_{SDDS}^{j} + \eta_t + \alpha_i + \varepsilon_{it},$$

let $X^{\neg\alpha}$ be the matrix collecting $\Psi, SDDS, \eta$ and X furthermore includes α. β is the parameter vector including $\theta, \lambda, \eta, \alpha$. Then under

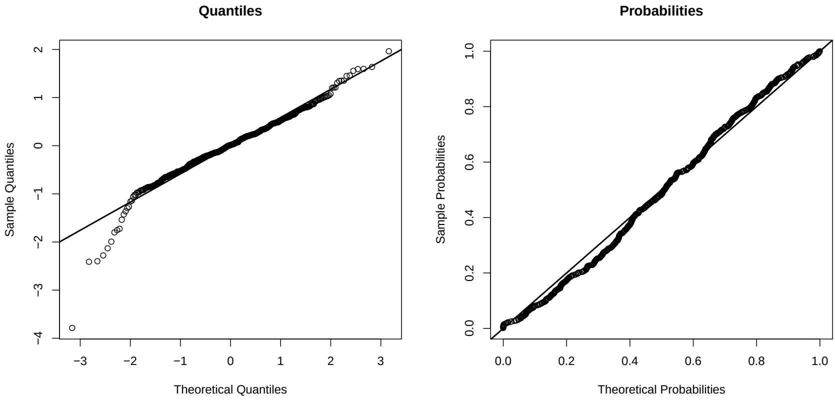
$$\mathbb{E}(X_{it}^{\neg\alpha}\alpha_i) \neq 0 \; \forall \; i, t \text{ and} \tag{12}$$

$$\mathbb{E}(X_{it}\varepsilon_{it}) = 0 \; \forall \; i, t \tag{13}$$

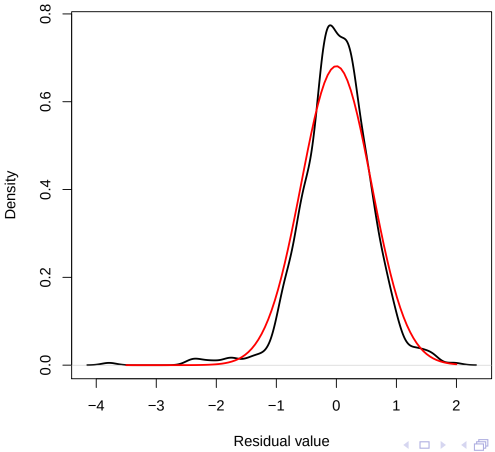
$\hat{\beta} = (X'X)^{-1}X'y$ is generally BLUE for β.

What can go wrong?

Model checking (FDI): plots of residuals vs. $\sim N$



Model checking (FDI): residual KDE vs. $\sim N$



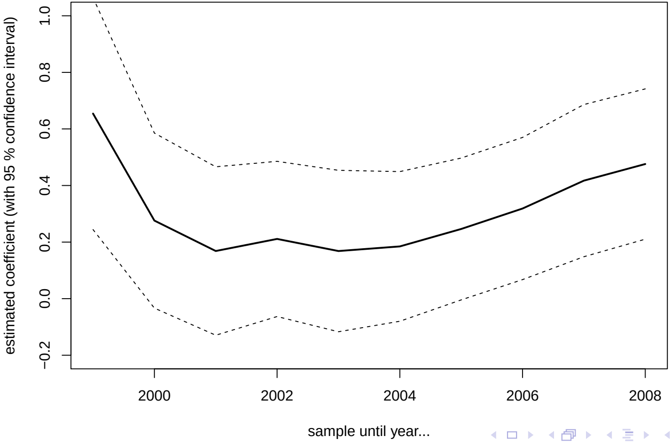
Model checking (FDI): Parameter robustness

Table : Parameter robustness

SDDS (general)	0.5016**	0.4935**	0.2056
	(0.1907)	(0.2362)	(0.1550)
SDDS (high inc)			0.6905***
			(0.2340)
Note	w/o 1996 subscribers	subscr. instead of compl.	parameter heterogeneity
observations	160	634	587

(in model 3a: 0.476***)

Model checking (FDI): Recursive regression



Spatial patterns in capital flows

- positive spill-overs / agglomeration effects (complex vertical FDI)?
- negative correlation due to portfolio risk diversification?
- relation information and distance?

Empirical evidence weak and derived under somewhat strong assumptions (Coughlin and Segev, 2000; Hernández et al., 2001 IMF; Blonigen et al., 2007 EER; Baltagi et al., 2007 JoE; Baltagi et al., 2008 JoE)

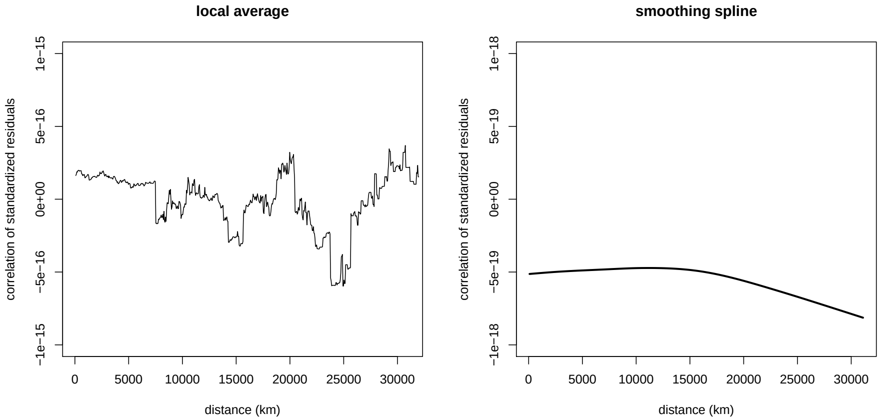
Smoothing spline

$$\hat{s}_p(x) = \underset{s(x) \in S_m(\Delta_K)}{\text{argmin}} \left[\sum_{i=1}^{n} \{y_i - s(x_i)\}^2 \ \lambda J(s) \right], \ \lambda > 0, \qquad (14)$$

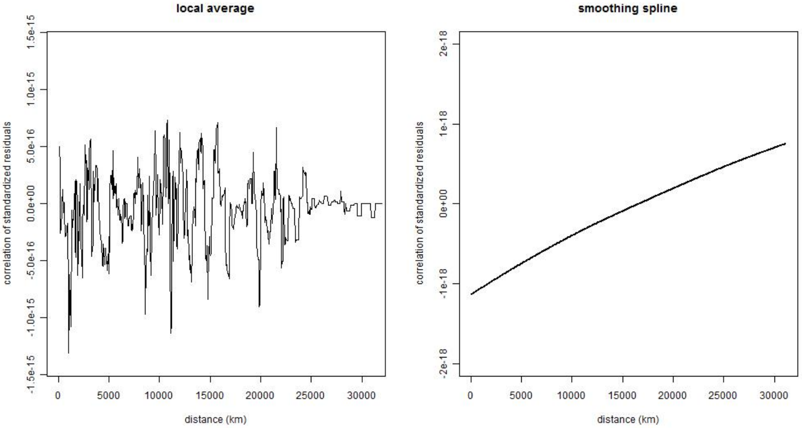
where $J(s) := \int_a^b \{s^{(m+1)}(x)\}^2 dx$ is a penalty function and $S_m(\Delta_K)$ is the spline space of degree m based on the partition Δ_K.

among all functions with $m + 1$ continuous derivatives, there is a unique function that minimizes (14)

Spatial correlation of residuals: FDI



Spatial correlation of residuals: portfolio



Asymptotic Variance of AR vs. MA Process

Consider the asymptotic variance $\lim_{T\to\infty} \text{Var}(\hat{\theta})$ of the limit distribution

$$\hat{\theta} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_t \to N(0, V),$$

with y following

- the *AR(1)* process $y_{t+1} = \alpha y_t + u_{t+1}$ and $u_t \sim N(0,1)$
- the *moving average* process
  $y_t = \alpha^2 u_{t-2} + \alpha u_{t-1} + u_t + \alpha u_{t+1} + \alpha^2 u_{t+2}$

Asymptotic Variance of AR(1) Process

For the *AR(1)* process

$$y_{t+1} = \alpha y_t + u_{t+1}, \quad u_t \sim N(0,1)$$

we know that $\mathbb{E}(\hat{\theta}) = 0$, $\text{Var}(y_t) = \frac{\sigma^2}{1-\alpha^2} \ \forall \ t$, $\text{Cov}(y_t, y_{t+h}) = \alpha^h \frac{\sigma^2}{1-\alpha^2} \ \forall \ t$. Hence:

$$\begin{aligned}
\lim_{T\to\infty} \text{Var}(\hat{\theta}) &= \lim_{T\to\infty} \frac{1}{T} \text{Var}\left(\sum_{t=1}^{T} y_t\right) = \\
&= \lim_{T\to\infty} \frac{1}{T} \left[\sum_{t=1}^{T} \text{Var}(y_t) + 2 \sum_{t<j} \text{Cov}(y_t, y_j)\right] = \\
&= \lim_{T\to\infty} \frac{1}{T} \left[T \frac{\sigma^2}{1-\alpha^2} + 2 \sum_{t=1}^{T} (T-t)\alpha^t \frac{\sigma^2}{1-\alpha^2}\right] = \ldots
\end{aligned}$$

Asymptotic Variance of AR(1) Process (cont'd)

$$\begin{aligned}
\lim_{T\to\infty} \text{Var}(\hat{\theta}) &= \ldots \lim_{T\to\infty} \frac{1}{T}\left[2\sum_{t=0}^{T}(T-t)\alpha^t \frac{\sigma^2}{1-\alpha^2} - T\frac{\sigma^2}{1-\alpha^2}\right] = \\
&= \lim_{T\to\infty} \frac{1}{T} T \frac{\sigma^2}{1-\alpha^2}\left[2\sum_{t=0}^{T}(1-\frac{t}{T})\alpha^t - 1\right] = \\
&= \lim_{T\to\infty} \frac{\sigma^2}{1-\alpha^2}\left[\underbrace{2\sum_{t=0}^{T}\alpha^t}_{\to\frac{2}{(1-\alpha)}} - \underbrace{2\sum_{t=0}^{T}\frac{t\alpha^t}{T}}_{\to 0} - \underbrace{1}_{=\frac{(1-\alpha)}{(1-\alpha)}}\right] = \ldots
\end{aligned}$$

Asymptotic Variance of AR(1) Process (cont'd)

$$\lim_{T\to\infty} \text{Var}(\hat{\theta}) = \ldots \frac{\sigma^2}{1-\alpha^2}\frac{1+\alpha}{1-\alpha} = \frac{\sigma^2(1+\alpha)}{(1+\alpha)(1-\alpha)^2} = \frac{\sigma^2}{(1-\alpha)^2}$$

Asymptotic Variance of MA Process

For the *moving average* process

$$y_t = \alpha^2 u_{t-2} + \alpha u_{t-1} + u_t + \alpha u_{t+1} + \alpha^2 u_{t+2},$$

again $\mathbb{E}(\hat{\theta}) = 0$ and

$$\begin{aligned}
\lim_{T\to\infty} \text{Var}(\hat{\theta}) &= \lim_{T\to\infty} \frac{1}{T} \text{Var}\left(\sum_{t=1}^{T} y_t\right) = \\
&= \lim_{T\to\infty} \frac{1}{T} \left[\sum_{t=1}^{T} \text{Var}(y_t) + 2\sum_{t<j} \text{Cov}(y_t, y_j)\right] = \\
&= \lim_{T\to\infty} \frac{1}{T} [T\sigma^2(2\alpha^4 + 2\alpha^2 + 1) + 2\sum_{i=1}^{4} (T - i)\,\text{Cov}(y_t, y_{t+i})] =
\end{aligned}$$

Asymptotic Variance of MA Process (cont'd)

$$
\begin{aligned}
\lim_{T\to\infty} \text{Var}(\hat{\theta}) &= \dots \lim_{T\to\infty} \frac{1}{T}[T\sigma^2(2\alpha^4 + 2\alpha^2 + 1) + \\
&+2\sigma^2[(T-1)(2\alpha^3 + 2\alpha) + (T-2)(3\alpha^3) + \\
&+(T-3)(2\alpha^3) + (T-4)(\alpha^4)]] = \\
&= \lim_{T\to\infty} \frac{1}{T}[T\sigma^2(2\alpha^4 + 2\alpha^2 + 1) + \\
&+2\sigma^2 T(\alpha^4 + 7\alpha^3 + 2\alpha) - 2\sigma^2(2\alpha + 14\alpha^3 - 4\alpha^4)] = \\
&= \sigma^2(4\alpha^4 + 14\alpha^3 + 2\alpha^2 + 4\alpha + 1)
\end{aligned}
$$

Analytical Comparison of Asymptotic Variances of AR vs. MA Process

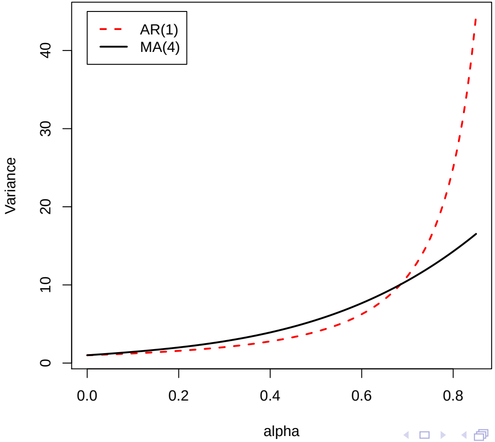
Asymptotic variance $\lim_{T\to\infty} \text{Var}(\hat{\theta})$ of

$$\hat{\theta} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_t \to N(0, V)$$

with y following the *AR(1)* process: $\frac{\sigma^2}{(1-\alpha)^2}$

with y following the *MA* process: $\sigma^2(4\alpha^4 + 14\alpha^3 + 2\alpha^2 + 4\alpha + 1)$

Graphical Comparison of Var($\hat{\theta}$)



'Standard' Panel Data Estimators

Fixed Effects Estimator: OLS on

$$y_{it} - \overline{y}_i = (x_{it} - \overline{x}_i)\beta + (u_{it} - \overline{u}_i), \tag{15}$$

where $\overline{y} = T^{-1} \sum_{t=1}^{T} y_{it}$ is the average over time; or OLS on

$$y_{it} = x_{it}\beta + \alpha_i + u_i \tag{16}$$

Random Effects Estimator:

$$\hat{\beta}_{RE} = \left(\sum_{i=1}^{N} X_i' \hat{\Omega}^{-1} X_i \right)^{-1} \left(\sum_{i=1}^{N} X_i' \hat{\Omega}^{-1} y_i \right), \tag{17}$$

where $\hat{\Omega} = \hat{\sigma}_u^2 I_T + \hat{\sigma}_\alpha^2 i_T i_T'$.

Between Estimator: OLS on

$$\overline{y}_i = \overline{x}_i\beta + \alpha_i + \overline{u}_i. \tag{18}$$

Appendix: robust VCV matrices

Robust Variance Matrix:

$$\hat{Var}(\hat{\beta}_{FE}) = (\sum_{i=1}^{N} \ddot{X}_i' \ddot{X}_i)^{-1} \sum_{i=1}^{N} \ddot{X}_i' \hat{e}_i \hat{e}_i' \ddot{X}_i (\sum_{i=1}^{N} \ddot{X}_i' \ddot{X}_i)^{-1}$$

with $\ddot{x}_{it} = x_{it} - \overline{x}_i$ and $\hat{e}_i = \hat{\ddot{\varepsilon}}_i = \ddot{y}_i - \ddot{X}_i \hat{\beta}$; controls for serial correlation and heteroskedasticity.

Cluster-Robust Variance Matrix:

$$\hat{Var}(\hat{\beta}_{FE}) = \frac{N-1}{N-k} \frac{C}{C-1} (\sum_{c=1}^{C} \ddot{X}_c' \ddot{X}_c)^{-1} \sum_{c=1}^{C} \ddot{X}_c' \hat{e}_c \hat{e}_c' \ddot{X}_c (\sum_{c=1}^{C} \ddot{X}_c' \ddot{X}_c)^{-1}$$

Appendix: Hausman Test

Note: If FE is the true model, RE is inconsistent. If RE is the true model, FE is consistent but not efficient.

General idea: compare results for an estimator θ_{eff} that is efficient (and consistent) under the null and an estimator θ_{cons} that is consistent in either case (and efficient under the alternative).

Null: Difference between θ_{eff} and θ_{cons} is not systematic $\rightarrow$ use θ_{eff}

$$H = (\hat{\theta}_{cons} - \hat{\theta}_{eff})'(\hat{Var}(\theta_{cons}) - \hat{Var}(\theta_{eff}))^{-1}(\hat{\theta}_{cons} - \hat{\theta}_{eff}) \sim \chi^2_k$$