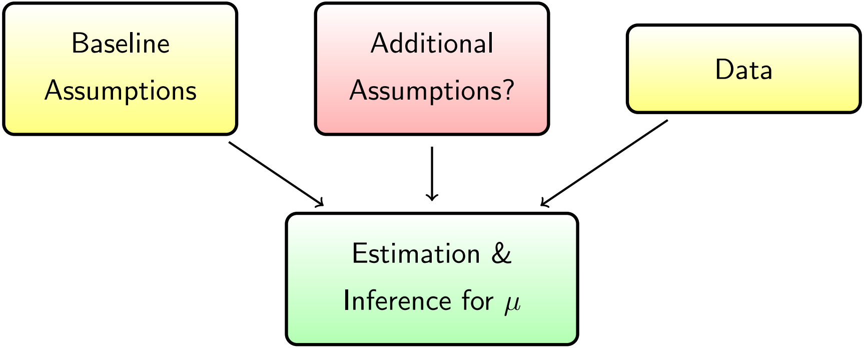
Using Invalid Instruments on Purpose: Focused Moment Selection and Averaging for GMM

Francis J. DiTraglia

University of Pennsylvania

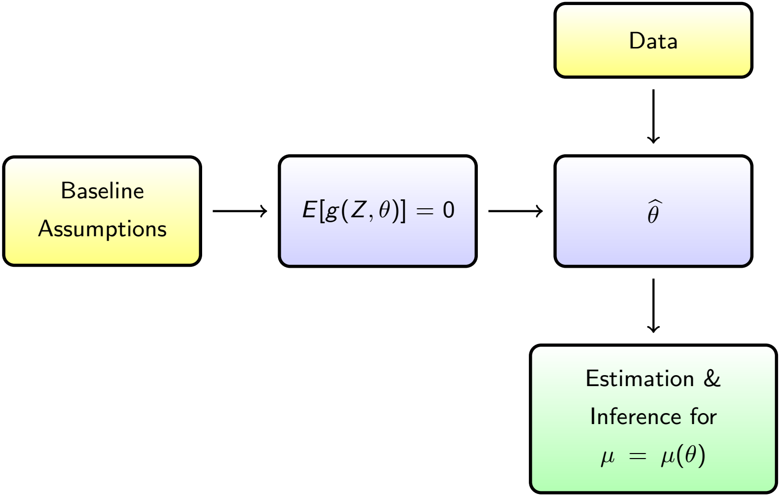
November 6, 2014

Focused Moment Selection Criterion (FMSC)

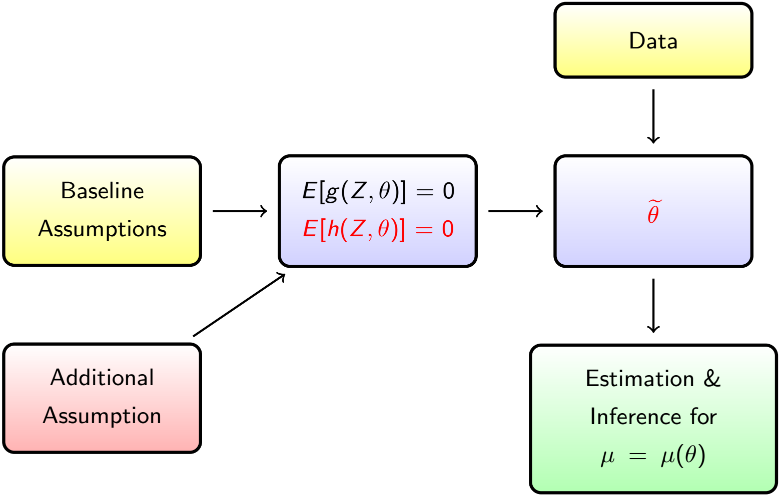


1. Choose False Assumptions on Purpose
2. Focused Choice of Assumptions
3. Local mis-specification
4. Averaging, Inference post-selection

GMM Framework



Adding Moment Conditions



Ordinary versus Two-Stage Least Squares

$$
\begin{aligned}
y_i &= \beta x_i + \epsilon_i \\
x_i &= \mathbf{z}_i' \boldsymbol{\pi} + v_i
\end{aligned}
$$

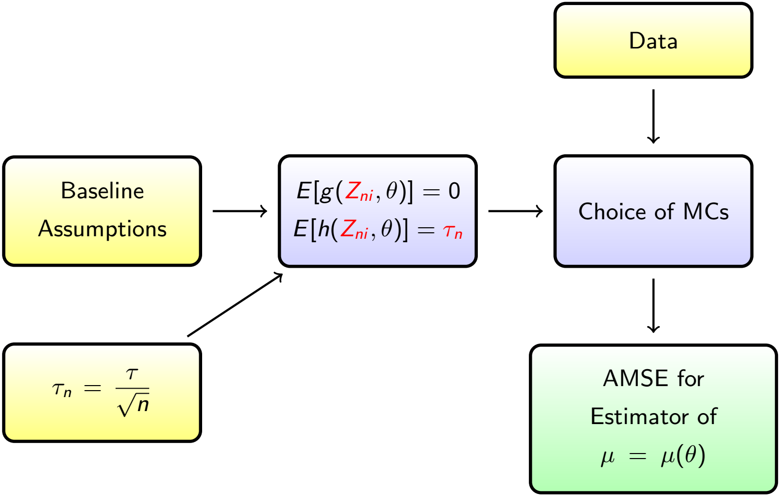
$$
\begin{aligned}
E[\mathbf{z}_i \epsilon_i] &= 0 \\
E[x_i \epsilon_i] &= ?
\end{aligned}
$$

Choosing Instrumental Variables

$$
\begin{aligned}
y_i &= \beta x_i + \epsilon_i \\
x_i &= \Pi_1' \mathbf{z}_i^{(1)} + \Pi_2' \mathbf{z}_i^{(2)} + v_i
\end{aligned}
$$

$$
\begin{aligned}
E[\mathbf{z}_i^{(1)} \epsilon_i] &= 0 \\
E[\mathbf{z}_i^{(2)} \epsilon_i] &= ?
\end{aligned}
$$

FMSC Asymptotics – Local Mis-Specification



Local Mis-Specification for OLS versus TSLS

$$
\begin{aligned}
y_i &= \beta x_i + \epsilon_i \\
x_i &= \mathbf{z}_i' \boldsymbol{\pi} + v_i
\end{aligned}
$$

$$
\begin{aligned}
E[\mathbf{z}_i \epsilon_i] &= 0 \\
E[x_i \epsilon_i] &= \tau / \sqrt{n}
\end{aligned}
$$

Local Mis-Specification for Choosing IVs

$$
\begin{aligned}
y_i &= \beta x_i + \epsilon_i \\
x_i &= \Pi_1' \mathbf{z}_i^{(1)} + \Pi_2' \mathbf{z}_i^{(2)} + v_i
\end{aligned}
$$

$$
\begin{aligned}
E[\mathbf{z}_i^{(1)} \epsilon_i] &= 0 \\
E[\mathbf{z}_i^{(1)} \epsilon_i] &= \tau / \sqrt{n}
\end{aligned}
$$

Local Mis-Specification

Triangular Array $\{Z_{ni}\colon 1 \leq i \leq n, n = 1, 2, \ldots\}$ with

(a) $E[g(Z_{ni}, \theta_0)] = 0$

(b) $E[h(Z_{ni}, \theta_0)] = n^{-1/2}\tau$

(c) $\{f(Z_{ni}, \theta_0)\colon 1 \leq i \leq n, n = 1, 2, \ldots\}$ uniformly integrable

(d) $Z_{ni} \to_d Z_i$, where the Z_i are identically distributed.

Shorthand: Write Z for Z_i

Candidate GMM Estimator

$$\widehat{\theta}_S = \underset{\theta \in \Theta}{\arg\min} \; [\Xi_S f_n(\theta)]' \, \widetilde{W}_S \; [\Xi_S f_n(\theta)]$$

$$\begin{aligned}
\Xi_S &= \text{Selection Matrix (ones and zeros)} \\
\widetilde{W}_S &= \text{Weight Matrix (p.s.d.)} \\
f_n(\theta) &= \begin{bmatrix} g_n(\theta) \\ h_n(\theta) \end{bmatrix} = \begin{bmatrix} n^{-1} \sum_{i=1}^{n} g(Z_{ni}, \theta) \\ n^{-1} \sum_{i=1}^{n} h(Z_{ni}, \theta) \end{bmatrix}
\end{aligned}$$

Notation: Limit Quantities

$$G = E\left[\nabla_\theta\, g(Z, \theta_0)\right], \quad H = E\left[\nabla_\theta\, h(Z, \theta_0)\right], \quad F = \begin{bmatrix} G \\ H \end{bmatrix}$$

$$\Omega = Var\left[f(Z, \theta_0)\right] = \begin{bmatrix} \Omega_{gg} & \Omega_{gh} \\ \Omega_{hg} & \Omega_{hh} \end{bmatrix}$$

$$\widetilde{W}_S \to_p W_S \text{ (p.d.)}$$

Local Mis-Specification + Standard Regularity Conditions

Every candidate estimator is consistent for θ_0 and

$$\sqrt{n}(\widehat{\theta}_S - \theta_0) \rightarrow_d -K_S \Xi_S \left(\begin{bmatrix} M_g \\ M_h \end{bmatrix} + \begin{bmatrix} 0 \\ \tau \end{bmatrix} \right)$$

$$\begin{aligned} K_S &= [F_S' W_S F_S]^{-1} F_S' W_S \\ M &= (M_g', M_h')' \\ M &\sim N(0, \Omega) \end{aligned}$$

Scalar Target Parameter μ

$$
\begin{aligned}
\mu &= \mu(\theta) && \text{Z-a.s. continuous function} \\
\mu_0 &= \mu(\theta_0) && \text{true value} \\
\widehat{\mu} &= \mu(\widehat{\theta}_S) && \text{estimator}
\end{aligned}
$$

Delta Method

$$
\sqrt{n}\left(\widehat{\mu}_S - \mu_0\right) \to_d -\nabla_\theta \mu(\theta_0)' K_S \Xi_S \left(M + \left[\begin{array}{c} 0 \\ \tau \end{array} \right] \right)
$$

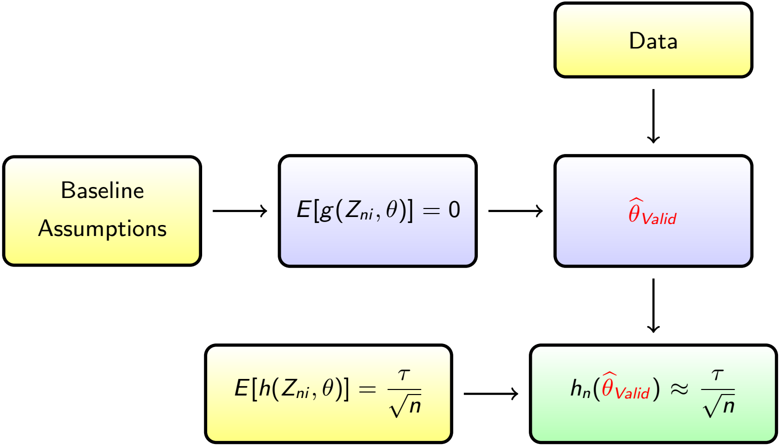
FMSC: Estimate AMSE($\widehat{\mu}_S$) and minimize over S

$$\text{AMSE}\left(\widehat{\mu}_S\right) = \nabla_\theta \mu(\theta_0)' K_S \Xi_S \left\{ \left[\begin{array}{cc} 0 & 0 \\ 0 & \tau\tau' \end{array} \right] + \Omega \right\} \Xi_S' K_S' \nabla_\theta \mu(\theta_0)$$

Estimating the unknowns

No consistent estimator of τ exists! (But everything else is easy)

A Plug-in Estimator of τ



An Asymptotically Unbiased Estimator of $\tau\tau'$

$$\sqrt{n}h_n(\widehat{\theta}_v) = \widehat{\tau} \rightarrow_d (\Psi M + \tau) \sim N_q(\tau, \Psi\Omega\Psi')$$

$$\Psi = \left[\begin{array}{cc} -HK_v & \mathbf{I}_q \end{array} \right]$$

$\widehat{\tau}\widehat{\tau}' - \widehat{\Psi}\widehat{\Omega}\widehat{\Psi}$ is an asymptotically unbiased estimator of $\tau\tau'$.

FMSC: Asymptotically Unbiased Estimator of AMSE

$$\text{FMSC}_n(S) = \nabla_\theta \mu(\widehat{\theta})' \widehat{K}_S \Xi_S \left\{ \begin{bmatrix} 0 & 0 \\ 0 & \widehat{B} \end{bmatrix} + \widehat{\Omega} \right\} \Xi_S' \widehat{K}_S' \nabla_\theta \mu(\widehat{\theta})$$

$$\widehat{B} = \widehat{\tau}\widehat{\tau}' - \widehat{\Psi}\widehat{\Omega}\widehat{\Psi}'$$

Choose S to minimize $\text{FMSC}_n(S)$ over the set of candidates $\mathscr{S}$.

A (Very) Special Case of the FMSC

Under homoskedasticity, FMSC selection in the OLS versus TSLS example is *identical* to a Durbin-Hausman-Wu test with $\alpha \approx 0.16$

$$\widehat{\tau} = n^{-1/2}\mathbf{x}'(\mathbf{y} - \mathbf{x}\widetilde{\beta}_{TSLS})$$

OLS gets benefit of the doubt, but not as much as $\alpha = 0.05,\ 0.1$

Limit Distribution of FMSC

$FMSC_n(S) \rightarrow_d FMSC_S$, where

$$
\begin{aligned}
\text{FMSC}_S &= \nabla_\theta \mu(\theta_0)' K_S \Xi_S \left\{ \left[\begin{array}{cc} 0 & 0 \\ 0 & B \end{array} \right] + \Omega \right\} \Xi_S' K_S' \nabla_\theta \mu(\theta_0) \\
B &= (\Psi M + \tau)(\Psi M + \tau)' - \Psi \Omega \Psi'
\end{aligned}
$$

Conservative criterion: random even in the limit.

Moment Average Estimators

$$\widehat{\mu} = \sum_{S \in \mathscr{S}} \widehat{\omega}_S \widehat{\mu}_S$$

Additional Notation

- $\widehat{\mu}$ Moment-average Estimator
- $\widehat{\mu}_S$ Estimator of target parameter under moment set S
- $\widehat{\omega}_S$ Data-dependent weight function
- $\mathscr{S}$ Collection of moment sets under consideration

Examples of Moment-Averaging Weights

Post-Moment Selection Weights

$\widehat{\omega}_S = \mathbf{1}\left\{\text{MSC}_n(S) = \min_{S' \in \mathscr{S}} \text{MSC}_n(S')\right\}$

Exponential Weights

$$\widehat{\omega}_S = \exp\left\{-\frac{\kappa}{2}\text{MSC}(S)\right\} \Big/ \sum_{S' \in \mathscr{S}} \exp\left\{-\frac{\kappa}{2}\text{MSC}(S')\right\}$$

Minimum-AMSE Weights...

Minimum AMSE-Averaging Estimator: OLS vs. TSLS

$$\widetilde{\beta}(\omega) = \omega\widehat{\beta}_{OLS} + (1-\omega)\widetilde{\beta}_{TSLS}$$

Under homoskedasticity:

$$\omega^* = \left[1 + \frac{\text{ABIAS(OLS)}^2}{\text{AVAR(TSLS)} - \text{AVAR(OLS)}}\right]^{-1}$$

Estimate by:

$$\widehat{\omega}^* = \left[1 + \frac{\max\left\{0,\ \left(\widehat{\tau}^2 - \widehat{\sigma}_\epsilon^2\widehat{\sigma}_x^2\left(\widehat{\sigma}_x^2/\widehat{\gamma}^2 - 1\right)\right)\ /\ \widehat{\sigma}_x^4\right\}}{\widehat{\sigma}_\epsilon^2(1/\widehat{\gamma}^2 - 1/\widehat{\sigma}_x^2)}\right]^{-1}$$

Where $\widehat{\gamma}^2 = n^{-1}\mathbf{x}'Z(Z'Z)^{-1}Z'\mathbf{x}$

Limit Distribution of Moment-Average Estimators

$$\widehat{\mu} = \sum_{S \in \mathscr{S}} \widehat{\omega}_S \widehat{\mu}_S$$

(i) $\sum_{S \in \mathscr{S}} \widehat{\omega}_S = 1$ a.s.

(ii) $\widehat{\omega}(S) \to_d \varphi_S(\tau, M)$ a.s.-continuous function of τ, M and consistently-estimable constants only

$$\sqrt{n}\,(\widehat{\mu} - \mu_0) \to_d \Lambda(\tau)$$

$$\Lambda(\tau) = -\nabla_\theta \mu(\theta_0)' \left[\sum_{S \in \mathscr{S}} \varphi_S(\tau, M) K_S \Xi_S \right] \left(M + \left[\begin{array}{c} 0 \\ \tau \end{array} \right] \right)$$

Simulating from the Limit Experiment

Suppose τ Known, Consistent Estimators of Everything Else

1. `for` $j \in \{1, 2, \ldots, J\}$
   (i) $M_j \overset{iid}{\sim} N_{p+q}\left(0, \widehat{\Omega}\right)$
   (ii) $\Lambda_j(\tau) = -\nabla_\theta \mu(\widehat{\theta})' \left[\sum_{S \in \mathscr{S}} \widehat{\varphi}_S(M_j + \tau)\widehat{K}_S \Xi_S\right](M_j + \tau)$
2. Using $\{\Lambda_j(\tau)\}_{j=1}^J$ calculate $\widehat{a}(\tau)$, $\widehat{b}(\tau)$ such that
$$P\left[\widehat{a}(\tau) \leq \Lambda(\tau) \leq \widehat{b}(\tau)\right] = 1 - \alpha$$
3. $P\left[\widehat{\mu} - \widehat{b}(\tau)/\sqrt{n} \leq \mu_0 \leq \widehat{\mu} - \widehat{a}(\tau)/\sqrt{n}\right] \approx 1 - \alpha$

Two-step Procedure for Conservative Intervals

1. Construct $1-\delta$ confidence region $\mathscr{T}(\widehat{\tau},\delta)$ for τ
2. For each $\tau^* \in \mathscr{T}(\widehat{\tau},\delta)$ calculate $1-\alpha$ confidence interval $\left[\widehat{a}(\tau^*), \widehat{b}(\tau^*)\right]$ for $\Lambda(\tau^*)$ as descibed on previous slide.
3. Take the lower and upper bound over the resulting intervals: $\widehat{a}_{min}(\widehat{\tau}) = \min_{\tau^* \in \mathscr{T}} \widehat{a}(\tau^*), \quad \widehat{b}_{max}(\widehat{\tau^*}) = \max_{\tau^* \in \mathscr{T}} \widehat{b}(\tau)$
4. The interval

$$\mathrm{CI}_{sim} = \left[\widehat{\mu} - \frac{\widehat{b}_{max}(\widehat{\tau})}{\sqrt{n}}, \quad \widehat{\mu} - \frac{\widehat{a}_{min}(\widehat{\tau})}{\sqrt{n}}\right]$$

has asymptotic coverage of at least $1-(\alpha+\delta)$

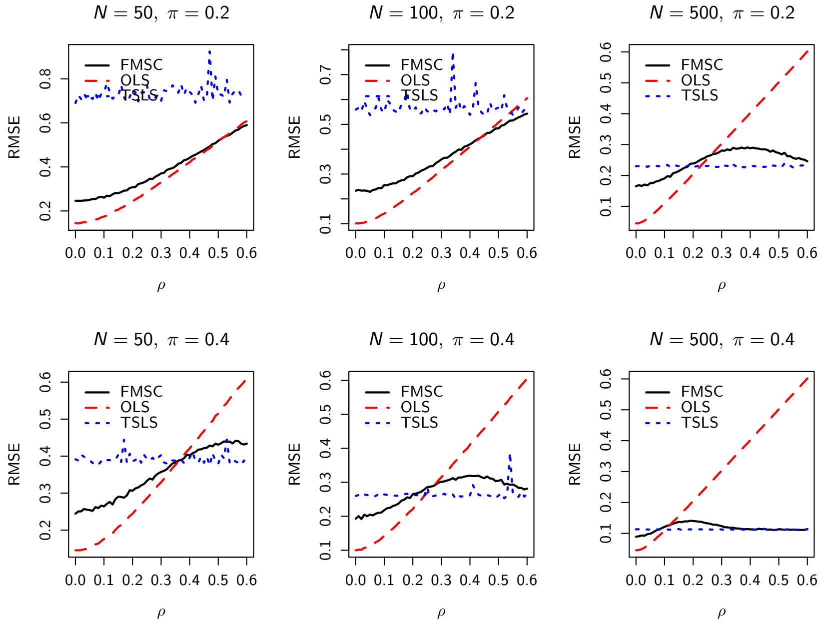
OLS versus TSLS Simulation

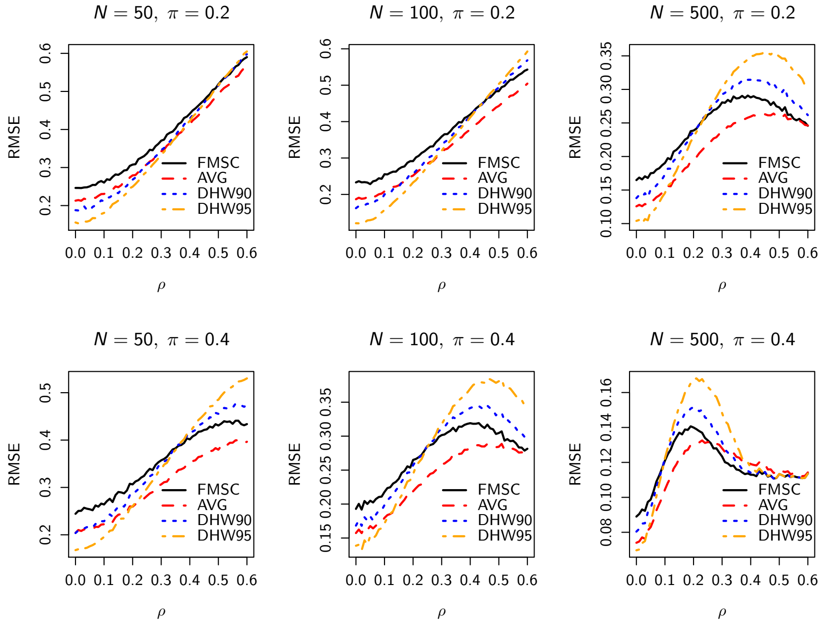
$$
\begin{aligned}
y_i &= 0.5x_i + \epsilon_i \\
x_i &= \pi(z_{1i} + z_{2i} + z_{3i}) + v_i
\end{aligned}
$$

$$(\epsilon_i, v_i, z_{1i}, z_{2i}, z_{3i}) \sim \text{ iid } N(0, \mathcal{S})$$

$$
\mathcal{S} = \begin{bmatrix}
1 & \rho & 0 & 0 & 0 \\
\rho & 1-\pi^2 & 0 & 0 & 0 \\
0 & 0 & 1/3 & 0 & 0 \\
0 & 0 & 0 & 1/3 & 0 \\
0 & 0 & 0 & 0 & 1/3
\end{bmatrix}
$$

$$Var(x) = 1, \qquad \rho = Cor(x, \epsilon), \qquad \pi^2 = \text{First-Stage } R^2$$


N = 50, π = 0.2
N = 100, π = 0.2
N = 500, π = 0.2
N = 50, π = 0.4
N = 100, π = 0.4
N = 500, π = 0.4
FMSC
OLS
TSLS
RMSE
ρ


$N = 50,\ \pi = 0.2$
$N = 100,\ \pi = 0.2$
$N = 500,\ \pi = 0.2$
$N = 50,\ \pi = 0.4$
$N = 100,\ \pi = 0.4$
$N = 500,\ \pi = 0.4$
RMSE
ρ
FMSC
AVG
DHW90
DHW95

Choosing Instrumental Variables Simulation

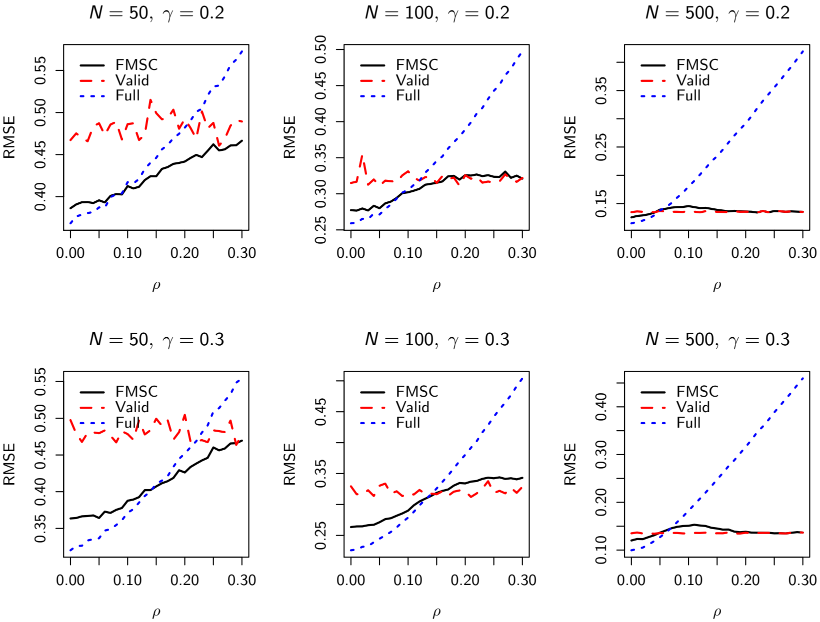
$$\begin{aligned} y_i &= 0.5x_i + \epsilon_i \\ x_i &= (z_{1i} + z_{2i} + z_{3i})/3 + \gamma w_i + v_i \end{aligned}$$

$$(\epsilon_i, v_i, w_i, z_{i1}, z_{2i}, z_{3i})' \sim \text{ iid } N(0, \mathcal{V})$$

$$\mathcal{V} = \begin{bmatrix} 1 & (0.5 - \gamma\rho) & \rho & 0 & 0 & 0 \\ (0.5 - \gamma\rho) & (8/9 - \gamma^2) & 0 & 0 & 0 & 0 \\ \rho & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 \end{bmatrix}$$

$$\gamma = Cor(x, w), \quad \rho = Cor(w, \epsilon), \quad \text{First-Stage } R^2 = 1/9 + \gamma^2$$

$$Var(x) = 1, \quad Cor(x, \epsilon) = 0.5$$


$N = 50,\ \gamma = 0.2$
$N = 100,\ \gamma = 0.2$
$N = 500,\ \gamma = 0.2$
$N = 50,\ \gamma = 0.3$
$N = 100,\ \gamma = 0.3$
$N = 500,\ \gamma = 0.3$
FMSC
Valid
Full
RMSE
ρ
0.00
0.10
0.20
0.30

Alternative Moment Selection Procedures

Downward J-test

Use Full instrument set unless J-test rejects.

Andrews (1999) – GMM Moment Selection Criteria

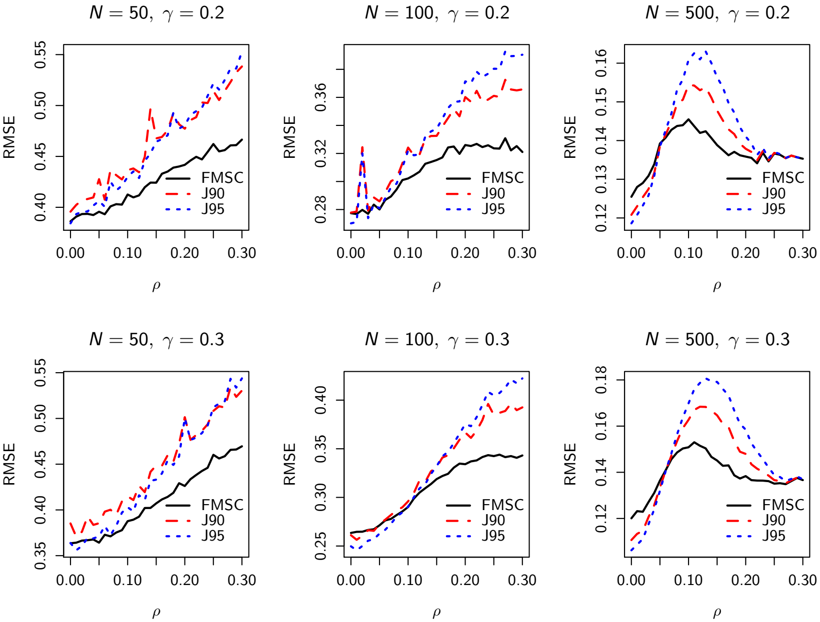
$\text{GMM-MSC}(S) = J_n(S) - \text{Bonus}$

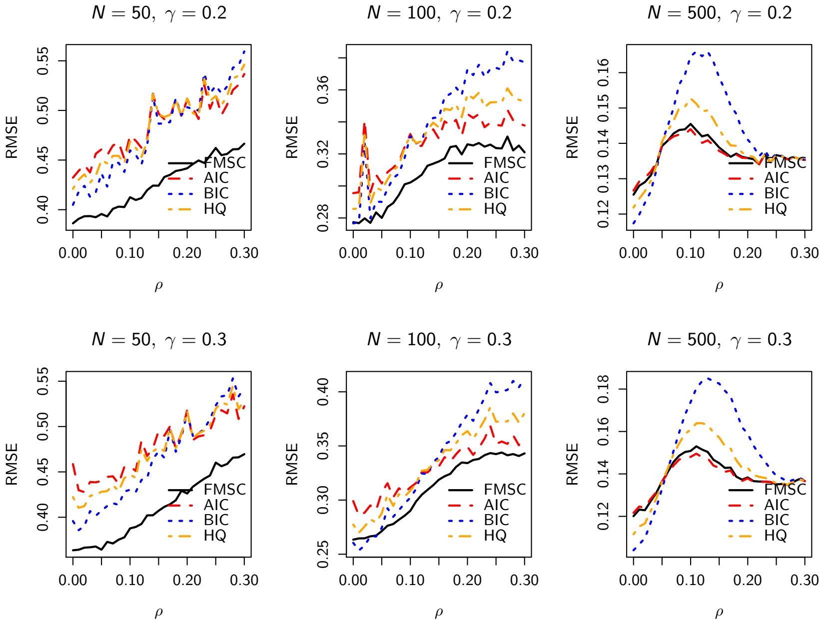
Hall & Peixe (2003) – Canonical Correlations Info. Criterion

$\text{CCIC}(S) = n \log\left[1 - R_n^2(S)\right] + \textit{Penalty}$

Penalty/Bonus Terms

Analogies to AIC, BIC, and Hannan-Quinn


$N = 50,\ \gamma = 0.2$
$N = 100,\ \gamma = 0.2$
$N = 500,\ \gamma = 0.2$
$N = 50,\ \gamma = 0.3$
$N = 100,\ \gamma = 0.3$
$N = 500,\ \gamma = 0.3$
RMSE
ρ
FMSC
J90
J95


$N = 50,\ \gamma = 0.2$
$N = 100,\ \gamma = 0.2$
$N = 500,\ \gamma = 0.2$
$N = 50,\ \gamma = 0.3$
$N = 100,\ \gamma = 0.3$
$N = 500,\ \gamma = 0.3$
RMSE
ρ
FMSC
AIC
BIC
HQ

Empirical Example: Geography or Institutions?

Institutions Rule

Acemoglu et al. (2001), Rodrik et al. (2004), Easterly & Levine (2003) – zero or negligible effects of "tropics, germs, and crops" in income per capita, controlling for institutions.

Institutions *Don't* Rule

Sachs (2003) – Large negative direct effect of malaria transmission on income.

Carstensen & Gundlach (2006)

How robust is Sachs's result?

Carstensen & Gundlach (2006)

Both Regressors Endogenous

$\ln GDPC_i = \beta_1 + \beta_2 \cdot INSTITUTIONS_i + \beta_3 \cdot MALARIA_i + \epsilon_i$

Robustness

- Various measures of *INSTITUTIONS*, *MALARIA*
- Various instrument sets
- β_3 remains large, negative and significant.

2SLS for All Results That Follow

Expand on Instrument Selection Exercise

FMSC and Corrected Confidence Intervals

1. FMSC – which instruments to estimate effect of malaria?
2. Correct CIs for Instrument Selection – effect of malaria still negative and significant?

Measures of *INSTITUTIONS* and *MALARIA*

- *rule* – Average governance indicator (Kaufmann, Kray and Mastruzzi; 2004)
- *malfal* – Proportion of population at risk of malaria transmission in 1994 (Sachs, 2001)

Instrument Sets

Baseline Instruments – Assumed Valid

- ln*mort* – Log settler mortality (per 1000), early 19th century
- *maleco* – Index of stability of malaria transmission

Further Instrument Blocks

Climate	*frost*, *humid*, *latitude*
Europe	*eurfrac*, *engfrac*
Openness	*coast*, *trade*

	μ = *malfal*			μ = *rule*		
	FMSC	posFMSC	$\widehat{\mu}$	FMSC	posFMSC	$\widehat{\mu}$
(1) Valid	3.0	3.0	-1.0	1.3	1.3	0.9
(2) Climate	3.1	3.1	-0.9	1.0	1.0	1.0
(3) Open	2.3	2.4	-1.1	1.2	1.2	0.8
(4) Eur	1.8	2.2	-1.1	0.5	0.7	0.9
(5) Climate, Eur	0.9	2.0	-1.0	0.3	0.6	0.9
(6) Climate, Open	1.9	2.3	-1.0	0.5	0.8	0.9
(7) Open, Eur	1.6	1.8	-1.2	0.8	0.8	0.8
(8) Full	0.5	1.7	-1.1	0.2	0.6	0.8
$>90\%$ CI FMSC		$(-1.6, -0.6)$			$(0.5, 1.2)$	
$>90\%$ CI posFMSC		$(-1.6, -0.6)$			$(0.6, 1.3)$	

Extensions & Future Work

- Simultaneous Model and Moment Selection – Dynamic Panel
- Welfare-Max. Treament Assignment under Budget Constraints
- Individual Heterogeneity: choosing the number of groups