A Generalized Focused Information Criterion for GMM Model and Moment Selection

Francis J. DiTraglia

University of Pennsylvania

June 13, 2013

Generalized Focused Information Criterion (GFIC)

Purpose

Simultaneous Model and Moment Selection for GMM Estimation

Main Idea

Choose model and moment conditions to yield minimum MSE estimator of user-specified target parameter even if mis-specified.

Some Related Work

- GMM Model and Moment Selection (Andrews & Lu, 2001)
- Focused Moment Selection Criterion (DiTraglia, 2013)
- Focused Information Criterion (Claeskens & Hjort, 2003)

Key Features of GFIC

Select "Wrong" Specification on Purpose

- Choose specification to minimize MSE of associated estimator.
- Accept some bias in exchange for reduction in variance.

Focused Selection

- Select based on MSE of user-specified target parameter μ
- Different Research Goal $\Rightarrow$ Different Criterion

Local Mis-specification

- Asymptotic MSE to approximate finite sample MSE
- Local asymptotics $\Rightarrow$ bias-variance tradeoff in the limit

GFIC Model & Moment Selection Framework

Parameters

- Always estimate "protected" parameters θ
- Consider setting "nuisance" parameters γ equal to constant γ_0

Moment Conditions

- Block g is correctly specified (provided we estimate γ)
- Block h is possibly mis-specified (even if we estimate γ)

Scalar Target Parameter

- $\mu = \phi(\theta, \gamma)$

Model and Moment Selection

Which elements of γ to estimate, which MCs to use for minimum AMSE estimator of μ?

GFIC Asymptotics: Local Mis-specification

Triangular Array DGP (Only a Device!)

$$E\left[\begin{array}{c} g(Z_{ni}, \gamma_0 + \delta/\sqrt{n}, \theta_0) \\ h(Z_{ni}, \gamma_0 + \delta/\sqrt{n}, \theta_0) \end{array}\right] = \left[\begin{array}{c} 0 \\ \tau/\sqrt{n} \end{array}\right]$$

δ Controls Model Mis-specification

- Restriction $\gamma = \gamma_0$ is *false* for finite n unless $\delta = 0$
- Model mis-specification disappears in the limit

τ Controls Moment Mis-specification

- MCs in h are invalid for finite n unless $\tau = 0$
- Moment mis-specification disappears in the limit

Notation for Model and Moment Selection

Model Selection

- Full set of parameter $\beta = (\theta, \gamma)$
- Which elements of γ to estimate?
- Model Selection Vector b corresponds to γ

Moment Selection

- Full set of moment conditions $f = (g, h)$
- Which MCs to use in estimation?
- Moment Selection Vector c corresponds to f

Putting Them Together

- A particular specification (b, c)
- Set of all specifications considered $\mathcal{BC}$

Overview of GFIC Derivation

Step 1 – Limit Distribution of GMM Estimator $\widehat{\beta}(b, c)$

- Asymptotically Normal
- Biased unless γ estimated and no MCs from h used
- Smaller variance if γ set to γ_0, MCs from h used

Step 2 – Associated Target Parameter Estimator $\widehat{\mu}(b, c)$

- Asymptotically Normal, inherits bias-variance tradeoff
- AMSE $(\widehat{\mu}(b, c))$ depends on $B = \begin{bmatrix} \tau\tau' & \tau\delta' \\ \delta\tau' & \delta\delta' \end{bmatrix}$

Step 3 – GFIC is an Estimator of AMSE $(\widehat{\mu}(b, c))$

- Substitute asymptotically unbiased estimator $\widehat{B}$ of B and consistent estimators of everything else.

Estimating δ, τ – Overview

Why is this difficult?

- Local mis-specification $\Rightarrow$ no consistent estimators of δ, τ
- Can construct asymptotically unbiased estimators
- Actually need to estimate $B = \begin{bmatrix} \tau\tau' & \tau\delta' \\ \delta\tau' & \delta\delta' \end{bmatrix}$

How and when can we proceed?

- $\widehat{\beta}_v = (\widehat{\theta}_v, \widehat{\gamma}_v)$ estimates *all* parameters using g *only*
- Plug $\widehat{\beta}_v$ into sample analogue of h to estimate $\tau/\sqrt{n}$
- Use $(\widehat{\gamma}_v - \gamma_0)$ to estimate $\delta/\sqrt{n}$
- Bias correction to get asymptotically unbiased estimator of B

Estimating δ, τ – Details

Limit Distribution of Bias Parameter Estimators

$$\begin{bmatrix} \widehat{\delta} \\ \widehat{\tau} \end{bmatrix} = \sqrt{n} \begin{bmatrix} (\widehat{\gamma}_v - \gamma_0) \\ h_n(\widehat{\beta}_v) \end{bmatrix} \to_d \begin{bmatrix} \delta \\ \tau \end{bmatrix} + \Psi \; N(0, \Omega)$$

- Both Ψ and Ω can be estimated consistently!

Asymptotically Unbiased Estimator of B

$$\begin{aligned} B &= \begin{bmatrix} \tau\tau' & \tau\delta' \\ \delta\tau' & \delta\delta' \end{bmatrix} \\ \widehat{B} &= \begin{bmatrix} \widehat{\tau}\widehat{\tau}' & \widehat{\tau}\widehat{\delta}' \\ \widehat{\delta}\widehat{\tau}' & \widehat{\delta}\widehat{\delta}' \end{bmatrix} - \widehat{\Psi}\widehat{\Omega}\widehat{\Psi}' \end{aligned}$$

Using the GFIC Framework

Model and Moment Selection

- Calculate $\widehat{\text{AMSE}}(\widehat{\mu}(b,c))$ for each $(b,c) \in \mathcal{BC}$
- Choose the specification with the lowest AMSE estimate.

Model and Moment Averaging

- Use AMSE estimates to construct data-dependent weights:

$$\widehat{\mu} = \sum_{(b,c)\in\mathcal{BC}} \widehat{\omega}(b,c)\widehat{\mu}(b,c)$$

- Alternatively, derive (or estimate) AMSE-optimal weights

Inference

- Correct CIs for post-selection and averaging estimators.

Simple Dynamic Panel Example – Large N, Small T

True Data Generating Process

$$y_{it} = \gamma y_{it-1} + \theta x_{it} + \eta_i + v_{it}$$

- Dynamics unless $\gamma = 0$ (assume stationary)
- Correlated individual effects $\eta_i \Rightarrow$ estimate in differences
- Regressor x_{it} predetermined but *not* strictly exogenous

Goal – Estimate θ with minimum MSE

- Model Selection Decision: set $\gamma = 0$?
- Moment Selection Decision: treat x_{it} as strictly exogenous?

Anderson & Hsiao–esque 2SLS Estimators (1982)

LW Moment Conditions:

$$\mathbb{E}\left[\begin{pmatrix} y_{i,t-2} \\ x_{i,t-1} \end{pmatrix}(\Delta y_{it} - \gamma \Delta y_{i,t-1} - \theta \Delta x_{it})\right] = 0, \text{ for } t = 3, \ldots, T$$

LS Adds the Moment Conditions:

$$\mathbb{E}\left[x_{it}\left(\Delta y_{it} - \gamma \Delta y_{i,t-1} - \theta \Delta x_{it}\right)\right] = 0, \text{ for } t = 3, \ldots, T$$

Only the LW conditions are correct

Anderson & Hsiao–esque 2SLS Estimators (1982)

W Moment Conditions:

$$\mathbb{E}\left[x_{i,t-1}\left(\Delta y_{it} - \theta \Delta x_{it}\right)\right] = 0, \text{ for } t = 2, \ldots, T$$

S Adds the Moment Conditions:

$$\mathbb{E}\left[x_{it}\left(\Delta y_{it} - \theta \Delta x_{it}\right)\right] = 0, \text{ for } t = 2, \ldots, T$$

None of these moment conditions are correct

Why Use an Incorrect Specification?

$$\Delta y_{it} = \gamma \Delta y_{it-1} + \theta \Delta x_{it} + \Delta v_{it}$$

Wrong Model

- γ small $\implies$ ignore dynamics
- Adds small bias
- Large efficiency gain from from additional time period
- Further efficiency gain from one fewer parameter

Invalid MCs

- $E[x_{it} v_{it-1}]$ small $\implies$ add x_{it} as instrument for period t
- Adds small bias
- Large efficiency gain since x_{it} is a strong instrument for Δx_{it}

Simulation Setup

Similar to Andrews & Lu (2001)

- $y_{i0} = 0$
- $y_{it} = \gamma y_{it-1} + 0.5 x_{it} + \eta_i + v_{it} \qquad (t = 1, \ldots, T)$

$$\begin{bmatrix} x_i \\ \eta_i \\ v_i \end{bmatrix} \sim \text{iid } N\left(\begin{bmatrix} 0_T \\ 0 \\ 0_T \end{bmatrix}, \begin{bmatrix} I_T & 0.2\iota_T & \sigma_{xv}\Gamma \\ 0.2\iota_T' & 1 & 0_T' \\ \sigma_{xv}\Gamma' & 0_T & I_T \end{bmatrix} \right)$$

- $E[x_{it} v_{it-1}] = \sigma_{xv}$ but $E[x_{it} v_{is}] = 0, s \neq t-1$

Vary γ and σ_{xv} over a grid

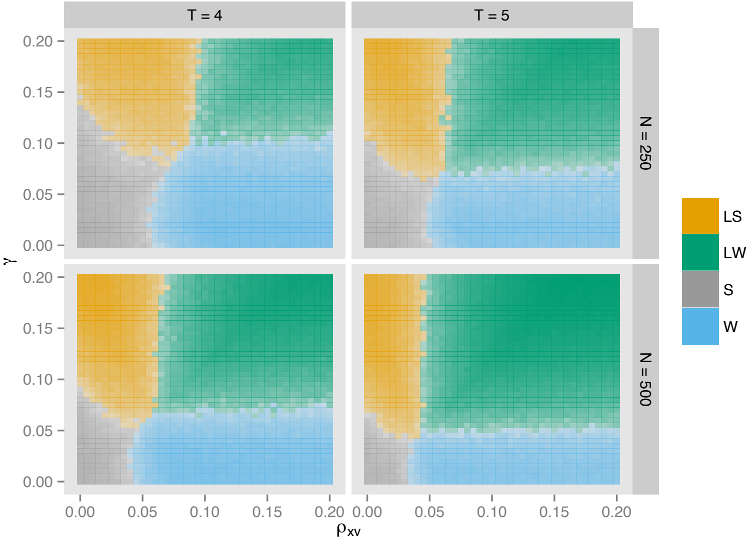

Figure: Minimum RMSE Specification at each combination of parameter values. Shading gives RMSE relative to second best specification.

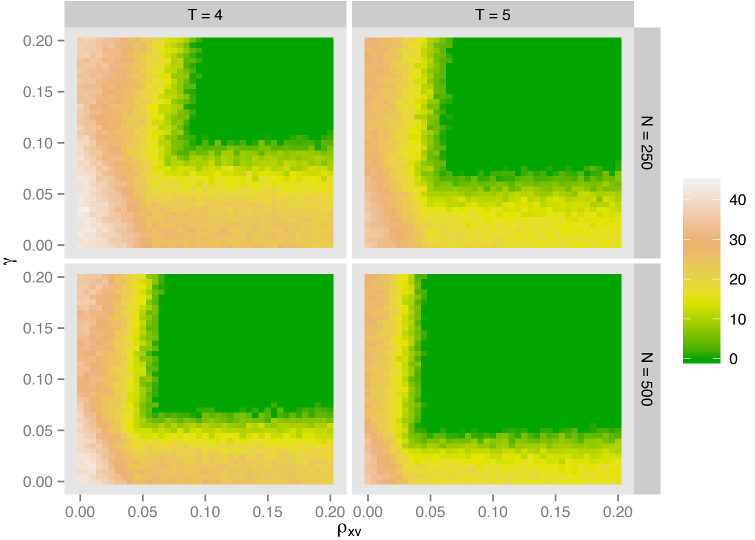

Figure: % RMSE Advantage of Best Specification (vs. LW)

	$N = 250$		$N = 500$	
	$T = 4$	$T = 5$	$T = 4$	$T = 5$
LW	19	10	13	7
LS	30	44	54	79
W	24	34	46	64
S	31	50	64	94
GFIC	17	13	15	10
J-test 10%	32	45	55	74
J-test 5%	31	47	57	79
GMM-BIC	32	48	62	87
GMM-HQ	32	46	57	77
GMM-AIC	31	39	47	57

Table: Average RMSE minus Pointwise Optimal (% points)

	$N = 250$		$N = 500$	
	$T = 4$	$T = 5$	$T = 4$	$T = 5$
LW	0	0	0	0
LS	42	81	94	154
W	49	88	105	158
S	48	92	107	171
GFIC	3	8	6	11
J-test 10%	43	78	91	140
J-test 5%	45	83	98	153
GMM-BIC	48	89	106	168
GMM-HQ	46	85	102	154
GMM-AIC	39	68	81	118

Table: Worst-case RMSE minus Minimax Optimal (% points)

Generalized Focused Information Criterion

Purpose

Simultaneous Model and Moment Selection for GMM Estimation

Key Features

- Local mis-specification framework
- Estimator of AMSE of user-specified target parameter
- Focused Selection
- Select “wrong” specification on purpose
- Works well in simulations

Points I Didn’t Emphasize Today

- Provides framework for model and moment averaging
- Correct confidence intervals

Extensions and Future Work

- More on inference/averaging
- AMSE-optimal averaging of OLS and IV estimators
- Risk functions besides MSE
- Covariate Choice in Treatment Assignment Problems (with Debopam Battacharya)

Supplementary Material

Competing Procedure: Downward J-test

1. Use S unless J-test rejects.
2. If S rejected, use W unless J-test rejects.
3. If W rejected, use LS unless J-test rejects.
4. Only use LW if all others rejected.

Competing Procedures: Andrews & Lu (2001)

J-test Statistic Minus Penalty Term

BIC-Type	$J - (\lvert c \rvert - \lvert b \rvert) \log n$
AIC-Type	$J - 2(\lvert c \rvert - \lvert b \rvert)$
HQ-Type	$J - 2.01(\lvert c \rvert - \lvert b \rvert) \log \log n$

where $|b|$ is the number of parameters estimated, and $|c|$ the number of moment conditions used. We select the specification with the lowest value of the criterion.

Average RMSE	$N = 250$		$N = 500$	
	$T = 4$	$T = 5$	$T = 4$	$T = 5$
LW	0.073	0.057	0.051	0.040
LS	0.079	0.074	0.070	0.066
W	0.075	0.069	0.066	0.061
S	0.080	0.077	0.074	0.072
GFIC	0.071	0.058	0.052	0.041
Downward J-test (10%)	0.080	0.074	0.070	0.065
Downward J-test (5%)	0.080	0.075	0.071	0.067
GMM-BIC	0.080	0.076	0.073	0.069
GMM-HQ	0.080	0.075	0.071	0.066
GMM-AIC	0.080	0.071	0.066	0.058

Worst-Case RMSE	$N = 250$		$N = 500$	
	$T = 4$	$T = 5$	$T = 4$	$T = 5$
LW	0.084	0.064	0.059	0.045
LS	0.120	0.116	0.115	0.113
W	0.125	0.120	0.122	0.115
S	0.125	0.123	0.122	0.121
GFIC	0.087	0.069	0.063	0.049
Downward J-test (10%)	0.120	0.114	0.113	0.107
Downward J-test (5%)	0.122	0.117	0.117	0.113
GMM-BIC	0.125	0.121	0.122	0.119
GMM-HQ	0.123	0.118	0.120	0.113
GMM-AIC	0.117	0.107	0.107	0.097

Notation

Sample Analogue of Moment Conditions

$$f_n(\beta) = \frac{1}{n}\sum_{i=1}^{n} f(Z_{ni}, \gamma, \theta) = \left[\begin{array}{c} g_n(\beta) \\ h_n(\beta) \end{array} \right] = \left[\begin{array}{c} n^{-1}\sum_{i=1}^{n} g(Z_{ni}, \gamma, \theta) \\ n^{-1}\sum_{i=1}^{n} h(Z_{ni}, \gamma, \theta) \end{array} \right]$$

PSD Weighting Matrix

$$\widetilde{W} = \left[\begin{array}{cc} \widetilde{W}_{gg} & \widetilde{W}_{gh} \\ \widetilde{W}_{hg} & \widetilde{W}_{hh} \end{array} \right]$$

Estimators

Each model/moment selection pair $(b, c) \in \mathcal{BC}$ defines a GMM estimator

$$\widehat{\beta}(b,c) = \underset{\beta^{(b)} \in \mathbf{B}^{(b)}}{\arg\min} \left[\Xi_c f_n \left(\beta^{(b)}, \gamma_0^{(-b)} \right) \right]' \left[\Xi_c \widetilde{W} \Xi_c' \right] \left[\Xi_c f_n \left(\beta^{(b)}, \gamma_0^{(-b)} \right) \right]$$

Under local mis-specification, *each* of these yields a consistent estimator of θ. Estimators based on an incorrect specification, however, show a bias in their limiting distributions.

More Notation

$$F = \left[\begin{array}{cc} \nabla_{\gamma'} g\left(Z, \gamma_0, \theta_0\right) & \nabla_{\theta'} g\left(Z, \gamma_0, \theta_0\right) \\ \nabla_{\gamma'} h\left(Z, \gamma_0, \theta_0\right) & \nabla_{\theta'} h\left(Z, \gamma_0, \theta_0\right) \end{array} \right]$$

$$F = \left[\begin{array}{cc} F_\gamma & F_\theta \end{array} \right] = \left[\begin{array}{cc} G_\gamma & G_\theta \\ H_\gamma & H_\theta \end{array} \right] = \left[\begin{array}{c} G \\ H \end{array} \right]$$

$$\Omega = Var \left[\begin{array}{c} g(Z, \gamma_0, \theta_0) \\ h(Z, \gamma_0, \theta_0) \end{array} \right] = \left[\begin{array}{cc} \Omega_{gg} & \Omega_{gh} \\ \Omega_{hg} & \Omega_{hh} \end{array} \right]$$

N.B. These expressions involve the limiting random variable Z rather than Z_{ni} so expectations are taken with respect to a distribution for which all MCs have expectation zero at (γ_0, θ_0).

Theorem (Asymptotic Distribution)

$$\sqrt{n}\left(\widehat{\beta}(b,c) - \beta_0^{(b)}\right) \to_d -K(b,c)\Xi_c\left(\left[\begin{array}{c}\mathscr{N}_g\\ \mathscr{N}_h\end{array}\right] + \left[\begin{array}{c}0\\ \tau\end{array}\right] - F_\gamma\delta\right)$$

where $\beta_0^{(b)\prime} = \left(\theta_0, \gamma_0^{(b)}\right)$,

$$K(b,c) = \left[F(b,c)' W_c F(b,c)\right]^{-1} F(b,c)' W_c$$

and

$$\left[\begin{array}{c}\mathscr{N}_g\\ \mathscr{N}_h\end{array}\right] \sim N\left(\left[\begin{array}{c}0\\ 0\end{array}\right], \left[\begin{array}{cc}\Omega_{gg} & \Omega_{gh}\\ \Omega_{hg} & \Omega_{hh}\end{array}\right]\right)$$

Corollary

$\sqrt{n}\left(\widehat{\mu}(b,c)-\mu_n\right)$ converges in distribution to

$$-\nabla_\beta \varphi_0' \Xi_b' K(b,c)\Xi_c \left(\begin{bmatrix} \mathscr{N}_g \\ \mathscr{N}_h \end{bmatrix} + \begin{bmatrix} 0 \\ \tau \end{bmatrix} - F_\gamma \delta \right) - \nabla_\gamma \varphi_0' \delta$$

where $\varphi_0 = \varphi(\gamma_0, \theta_0)$, $\mu_n = \phi(\theta_0, \gamma_n)$.

- AMSE $(\widehat{\mu})$ comes as immediate consequence of this result
- Usual estimators of K, etc. consistent under local mis-spec.
- The problem is τ, δ

How and When Can We Estimate τ and δ?

No consistent estimators exist under local mis-spec. but we can construct asymptotically unbiased estimators provided:

1. There are enough moment conditions in g to identify the full parameter vector $\rightarrow$ Valid Estimator $\widehat{\beta}_v = \left(\widehat{\gamma}_v, \widehat{\theta}_v\right)'$.
2. It is possible to evaluate h_n, sample analogue of "suspect" MCs, at $\widehat{\beta}_v$. (This is usually trivial.)

Estimating δ

Corollary (Asymptotic Distribution of Valid Estimator)

$$\sqrt{n}\left(\widehat{\beta}_v - \beta_0\right) = \sqrt{n}\left(\begin{array}{c} \widehat{\gamma}_v - \gamma_0 \\ \widehat{\theta}_v - \theta_0 \end{array} \right) \to_d \left[\begin{array}{c} \delta \\ 0 \end{array} \right] - K_v \mathscr{N}_g$$

where $K_v = \left[G' W_{gg} G\right]^{-1} G' W_{gg}$ *and* $W_{gg} = plim_{N\to\infty} \widetilde{W}_{gg}$.

This immediately provides asymptotically unbiased estimator of δ, namely $\widehat{\delta} = \sqrt{n}(\widehat{\gamma}_v - \gamma_0)$ since γ_0 is known and $\mathscr{N}_g$ is mean-zero.

Estimating τ

Lemma (Asymptotically Unbiased Estimator of τ)

$$\widehat{\tau} = \sqrt{n}h_n\left(\widehat{\beta}_v\right) \to_d \tau - HK_v\mathscr{N}_g + \mathscr{N}_h$$

where $K_v = \left[G'W_{gg}G\right]^{-1}G'W_{gg}$.

This results gives asymptotically unbiased estimator of τ since $\mathscr{N}_g$ and $\mathscr{N}_h$ mean zero.

But AMSE Requires *Squared Bias*

Rewriting the Expression from Above:

$$\mathrm{BIAS}\left(\widehat{\mu}\left(b,c\right)\right)^2 = \nabla_\beta \varphi_0' M(b,c) \left[\begin{array}{cc} \tau\tau' & \tau\delta' \\ \delta\tau' & \delta\delta' \end{array} \right] M(b,c)' \nabla_\beta \varphi_0$$

where

$$M(b,c) = \Xi_b' K(b,c) \Xi_c \left[\begin{array}{cc} -G_\gamma & 0 \\ -H_\gamma & I \end{array} \right] + \left[\begin{array}{cc} I_r & 0_{r\times q} \\ 0_{p\times r} & 0_{s\times q} \end{array} \right]$$

Problem

Although $(\widehat{\delta}, \widehat{\tau})$ are asymptotically unbiased estimators of (δ, τ), $\widehat{\delta\delta'}$ is not an asymptotically unbiased estimator of $\delta\delta'$ and $(\widehat{\tau\tau'}, \widehat{\tau\delta'})$ are not asymptotically unbiased estimators of $(\tau\tau', \tau\delta')$.

Joint Distribution of $(\widehat{\delta}, \widehat{\tau})$

$$\left[\begin{array}{c} \widehat{\delta} \\ \widehat{\tau} \end{array} \right] = \sqrt{n} \left[\begin{array}{c} (\widehat{\gamma}_v - \gamma_0) \\ h_n(\widehat{\beta}_v) \end{array} \right] \rightarrow_d \left[\begin{array}{c} \delta \\ \tau \end{array} \right] + \Psi \left[\begin{array}{c} \mathcal{N}_g \\ \mathcal{N}_h \end{array} \right]$$

where

$$\Psi = \left[\begin{array}{cc} -K_v^\gamma & \mathbf{0} \\ -HK_v & I \end{array} \right]$$

Each of the quantities in the matrix pre-multiplying $(\mathcal{N}_g', \mathcal{N}_h')'$ is consistently estimable under local mis-specification, as is the variance matrix of $(\mathcal{N}_g', \mathcal{N}_h')'$.

Bias Correction

Provided that $\widehat{\Psi}$ and $\widehat{\Omega}$ are consistent estimators of Ψ and Ω,

$$\widehat{B} = \begin{bmatrix} \widehat{\tau}\widehat{\tau}' & \widehat{\tau}\widehat{\delta}' \\ \widehat{\delta}\widehat{\tau}' & \widehat{\delta}\widehat{\delta}' \end{bmatrix} - \widehat{\Psi}\widehat{\Omega}\widehat{\Psi}'$$

is an asymptotically unbiased estimator of the squared bias matrix

$$\begin{bmatrix} \tau\tau' & \tau\delta' \\ \delta\tau' & \delta\delta' \end{bmatrix}.$$

GFIC: Asymptotically Unbiased Estimator of AMSE

$$\begin{aligned}
\text{GFIC}(b,c) &= \widehat{\text{AVAR}}(b,c) + \widehat{\text{ABIAS}}^2(b,c) \\
\widehat{\text{AVAR}}(b,c) &= \nabla_\beta \widehat{\varphi}_0' \Xi_b' \widehat{K}(b,c) \widehat{\Omega}_c \widehat{K}(b,c)' \Xi_b \nabla_\beta \widehat{\varphi}_0' \\
\widehat{\text{ABIAS}}^2(b,c) &= \nabla_\beta \widehat{\varphi}_0' \widehat{M}(b,c)\ \widehat{B}\ \widehat{M}(b,c) \nabla_\beta \widehat{\varphi}_0 \\
\widehat{B} &= \begin{bmatrix} \widehat{\tau}\widehat{\tau}' & \widehat{\tau}\widehat{\delta}' \\ \widehat{\delta}\widehat{\tau}' & \widehat{\delta}\widehat{\delta}' \end{bmatrix} - \widehat{\Psi}\widehat{\Omega}\widehat{\Psi}'
\end{aligned}$$

We choose the specification (b^*, c^*) that minimizes the value of the GFIC over the candidate set $\mathcal{BC}$.

Post Selection Inference / Model Averaging

Consider an estimator of the form

$$\widehat{\mu} = \sum_{(b,c)\in\mathcal{BC}} \widehat{\omega}(b,c)\widehat{\mu}(b,c)$$

where $\widehat{\omega}(b,c)$ is a set of data-dependent weights

Requirements for the Weights

Let $\widehat{\omega}(b, c)$ be a function of the data $Z_{n1}, \ldots, Z_{nn}$ and (b, c) satisfying

(a) $\sum_{(b,c)\in\mathcal{BC}} \widehat{\omega}(b, c) = 1$

(b) $\widehat{\omega}(b, c) \rightarrow_d \psi(\mathscr{N}, \delta, \tau | b, c)$ jointly for all $(b, c) \in \mathcal{BC}$ where ψ is a function of the normal random vector $\mathscr{N}$, the bias parameters δ and τ, and consistently estimable quantities only.

Covers GFIC, J-test, Andrews & Lu (2001), etc.

Limit Distribution of Averaging Estimator

Since the weights sum to one:

$$\sqrt{n}(\widehat{\mu} - \mu_n) = \sum_{(b,c)\in\mathcal{BC}} \widehat{\omega}(b,c)\sqrt{n}\left(\widehat{\mu}\left(b,c\right) - \mu_n\right)$$

and $\widehat{\omega}(b,c), \widehat{\mu}(b,c)$ converge *jointly* for all $(b,c) \in \mathcal{BC}$

$$\sqrt{n}\left(\widehat{\mu} - \mu_n\right) \to_d \Lambda(\tau,\delta)$$

where

$$\boxed{\Lambda(\tau,\delta) = -\nabla_\beta \varphi_0' \sum_{(b,c)\in\mathcal{BC}} \psi(\mathscr{N},\delta,\tau|b,c)\left\{\Xi_b' K(b,c)\Xi_c \mathscr{N} + M(b,c)\left[\begin{array}{c}\delta\\ \tau\end{array}\right]\right\}}$$

Non-normal limit distribution that depends on (δ,τ)

Suppose (δ, τ) Known

(i) For each $j = 1, 2, \dots, J$, generate $\mathcal{N}_j \sim N(0, \widehat{\Omega})$

(ii) For each for $j = 1, 2, \dots, J$ set

$$\Lambda_j(\tau, \delta) = -\nabla_\beta \widehat{\varphi}_0' \sum_{(b,c) \in \mathcal{BC}} \widehat{\psi}(\mathcal{N}_j, \delta, \tau | b, c) \left\{ \Xi_b' \widehat{K}(b, c) \Xi_c \mathcal{N}_j + \widehat{M}(b, c) \begin{bmatrix} \delta \\ \tau \end{bmatrix} \right\}$$

(iii) Using $\{\Lambda_j(\delta, \tau)\}_{j=1}^J$, calculate $\widehat{a}(\delta, \tau)$, $\widehat{b}(\delta, \tau)$ such that

$$\mathbb{P}\left\{ \widehat{a}(\delta, \tau) \leq \Lambda(\delta, \tau) \leq \widehat{b}(\delta, \tau) \right\} = 1 - \alpha$$

Accounting for Estimated (δ, τ)

Let $R(\alpha_1)$ be a $(1-\alpha_1) \times 100\%$ confidence region for (δ, τ).

1. For each $(\delta, \tau) \in R(\alpha_1)$ construct a confidence interval

$$\mathbb{P}\left\{\widehat{a}(\delta,\tau) \leq \Lambda(\delta,\tau) \leq \widehat{b}(\delta,\tau)\right\} = 1-\alpha_2$$

   using the simulation procedure from the previous slide.

2. Define

$$\begin{aligned} \widehat{a}_{min}(\widehat{\delta}, \widehat{\tau}) &= \min_{(\delta,\tau)\in R(\alpha_1)} \widehat{a}(\delta,\tau) \\ \widehat{b}_{max}(\widehat{\delta}, \widehat{\tau}) &= \max_{(\delta,\tau)\in R(\alpha_1)} \widehat{b}(\delta,\tau) \end{aligned}$$

3. The following CI has asymptotic coverage of *at least* $1-(\alpha_1+\alpha_2)$

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