Cross-sectional GMM estimation under a common data shock

Serguey Khovansky   Oleksandr Zhylyevskyy

Clark University   Iowa State University

2011 Midwest Econometrics Group Meetings

October 6, 2011

Common Shocks in Cross-Sectional Data

Cross-sectional econometricians typically assume observations are **independent**

However, **independence breaks down** if population units are affected by a **common shock**

Examples:

- oil price shocks affect production costs of many firms
- interest rate shocks affect consumption of many households
- common factors affect individual stock returns

Localized and Non-Localized Shocks

Localized shock:

- dependence between observations diminishes with **distance**
- distance may be geographical, socioeconomic, time-wise, etc.

Non-localized shock:

- dependence between observations need not diminish

Consider observations $X_1, X_2, ..., X_{100}, ...$:

- localized shock: X_1, X_{100} are "less dependent" than X_1, X_2
- non-localized shock: no such relationship exists

Contribution

We propose GMM estimators for a cross-sectional model with a **non-localized common shock**

We specify conditions under which estimators are:

- consistent
- asymptotically **mixed** normal

We show that conventional Wald and OIR tests are still applicable

Data Structure

Probability space $(\Omega, \mathcal{F}, P)$

D.g.p. provides observations $X_0, X_1, X_2, \ldots$

Data structure:

- X_0 is driven by common shock
- X_i, $i = 1, 2, \ldots$, is driven by common **and** idiosyncratic shock

Examples:

- aggregate income *vs.* individual incomes
- average crop yield *vs.* individual farm crop yields
- stock market portfolio return *vs.* individual stock returns

Conditionally I.I.D. Observations

Assumption:

$X_1, X_2, \ldots$ are **conditionally i.i.d.** given σ-field $\mathcal{F}_0 \equiv \sigma(X_0)$

$\sigma(X_0)$: σ-field generated by X_0 (i.e., by common shock)

This assumption is **very mild** (Andrews, 2005):

When sample units are randomly drawn, it is compatible with:

- arbitrary dependence across population units
- different effects of common shock on population units
- heterogeneity across population units

Parameters and Moment Restrictions

Goal:

- estimate, do inference on $\underset{(p\times 1)}{\boldsymbol{\theta}_0}$: true parameter underlying d.g.p.

Parameter set is $\boldsymbol{\Theta} \subset \mathcal{R}^p$:

- $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}$
- $\boldsymbol{\Theta}$ is compact and convex

Economic model provides k **moment restrictions** $(k \geq p)$:

$$\underset{(k\times 1)}{\boldsymbol{g}}\left(X_i; \boldsymbol{\theta}, X_0\right) \text{ for } i = 1, 2, \ldots$$

For example, jth component of $\boldsymbol{g}(\cdot)$ may be:

$$\boldsymbol{g}^{(j)}\left(X_i; \boldsymbol{\theta}, X_0\right) = X_i^{\xi} - E_{\boldsymbol{\theta}}\left[X_i^{\xi} | X_0\right], \text{ where } \xi \text{ is a constant}$$

Estimators

One-step estimation using nonstochastic pos. def. $\boldsymbol{\Sigma}$:

$$Q_{1,n}(\boldsymbol{\theta}) = \left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{g}(X_i;\boldsymbol{\theta},X_0)\right)' \boldsymbol{\Sigma}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{g}(X_i;\boldsymbol{\theta},X_0)\right)$$

$$\widehat{\boldsymbol{\theta}}_{1,n} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} Q_{1,n}(\boldsymbol{\theta})$$

Two-step using $\widehat{\boldsymbol{\Sigma}}_{1,n} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{g}\left(X_i;\widehat{\boldsymbol{\theta}}_{1,n},X_0\right)\cdot\boldsymbol{g}\left(X_i;\widehat{\boldsymbol{\theta}}_{1,n},X_0\right)'$:

$$Q_{2,n}(\boldsymbol{\theta}) = \left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{g}(X_i;\boldsymbol{\theta},X_0)\right)' \widehat{\boldsymbol{\Sigma}}_{1,n}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{g}(X_i;\boldsymbol{\theta},X_0)\right)$$

$$\widehat{\boldsymbol{\theta}}_{2,n} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} Q_{2,n}(\boldsymbol{\theta})$$

Consistency

Suppose:

- $\boldsymbol{g}\left(X_i;\boldsymbol{\theta},X_0\right)$ is measurable w.r.t. $\sigma\left(X_0,X_i\right)$ for all $\boldsymbol{\theta}$
- $\boldsymbol{g}\left(X_i;\boldsymbol{\theta},X_0\right)$ is a.s. differentiable in $\boldsymbol{\theta}$
- $E\left[\sup_{\boldsymbol{\theta}}\left\|\boldsymbol{g}\left(X_i;\boldsymbol{\theta},X_0\right)\right\|^2\right]<\infty$, $E\left[\sup_{\boldsymbol{\theta}}\left\|\frac{\partial\boldsymbol{g}\left(X_i;\boldsymbol{\theta},X_0\right)}{\partial\boldsymbol{\theta}}\right\|^2\right]<\infty$
- $E\left[\boldsymbol{g}\left(X_i;\boldsymbol{\theta}_0,X_0\right)|\mathcal{F}_0\right]=\mathbf{0}$ a.s.
- $E\left[\boldsymbol{g}\left(X_i;\boldsymbol{\theta},X_0\right)|\mathcal{F}_0\right]\neq\mathbf{0}$ a.s. for all $\boldsymbol{\theta}\neq\boldsymbol{\theta}_0$
- $\boldsymbol{\Sigma}_{\mathcal{F}_0}\equiv E\left[\boldsymbol{g}\left(X_i;\boldsymbol{\theta}_0,X_0\right)\cdot\boldsymbol{g}\left(X_i;\boldsymbol{\theta}_0,X_0\right)'|\mathcal{F}_0\right]$ is a.s. pos. def.

Theorem: As $n\rightarrow\infty$, $\widehat{\boldsymbol{\theta}}_{1,n}\rightarrow^p\boldsymbol{\theta}_0$ and $\widehat{\boldsymbol{\theta}}_{2,n}\rightarrow^p\boldsymbol{\theta}_0$

Asymptotic Mixed Normality

In addition, suppose:

- $\exists$ open ball $\mathcal{N}$ centered at $\boldsymbol{\theta}_0$ s.t. $\boldsymbol{g}\left(X_i;\boldsymbol{\theta},X_0\right)$ is a.s. twice differentiable in $\boldsymbol{\theta}$ on $\mathcal{N}$ and $E\left[\sup_{\boldsymbol{\theta}\in\mathcal{N}}\left\|\frac{\partial^2\boldsymbol{g}(X_i;\boldsymbol{\theta},X_0)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right\|\right]<\infty$
- $\mathbf{G}_{\mathcal{F}_0}\equiv E\left[\frac{\partial\boldsymbol{g}(X_i;\boldsymbol{\theta}_0,X_0)}{\partial\boldsymbol{\theta}'}|\mathcal{F}_0\right]$ has full column rank a.s.

Theorem: As $n\longrightarrow\infty$:

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_{1,n}-\boldsymbol{\theta}_0\right)\longrightarrow^d MN\left(\mathbf{0},\mathbf{V}_{1,\mathcal{F}_0}\right)$$
$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_{2,n}-\boldsymbol{\theta}_0\right)\longrightarrow^d MN\left(\mathbf{0},\mathbf{V}_{2,\mathcal{F}_0}\right)$$

$\mathbf{V}_{1,\mathcal{F}_0}$ and $\mathbf{V}_{2,\mathcal{F}_0}$ are a.s. pos. def. **stochastic** matrices

Asymptotic Inference and Specification Test

Consider testing r parametric restrictions:

$$H_0 : \underset{(r\times 1)}{\mathbf{a}(\boldsymbol{\theta}_0)} = \mathbf{0}$$

Let $\mathbf{A}(\cdot)$ be Jacobian of $\mathbf{a}(\cdot)$. Under H_0, **Wald test** statistic

$$W_n \equiv n \cdot \mathbf{a}\left(\widehat{\boldsymbol{\theta}}_{2,n}\right)' \left[\mathbf{A}\left(\widehat{\boldsymbol{\theta}}_{2,n}\right)\mathbf{V}_{2,n}\mathbf{A}\left(\widehat{\boldsymbol{\theta}}_{2,n}\right)'\right]^{-1} \mathbf{a}\left(\widehat{\boldsymbol{\theta}}_{2,n}\right) \to^d \chi^2(r)$$

If the model is correctly specified, **OIR test** statistic

$$J_n \equiv n \cdot Q_{2,n}\left(\widehat{\boldsymbol{\theta}}_{2,n}\right) \to^d \chi^2(k-p)$$

Financial Model Setup

Financial assets:

- many risky assets called **stocks**
- a diversified portfolio of stocks called **market index**
- a riskless asset

Asset prices are quoted continuously, but we eventually focus only on a **cross-section of returns** between $t = 0$ and $t = T$

Market Index Price Dynamics

Dynamics of market index:

$$\frac{dM_t}{M_t} = \mu_m dt + \sigma_m dW_t$$

where drift μ_m is

$$\mu_m = r + \delta\sigma_m$$

- r: risk-free rate
- σ_m: market volatility, $\sigma_m > 0$
- δ: Sharpe ratio of market index
- $\{W_t\}$: Brownian motion; **source of common shock**

Stock Price Dynamics

Dynamics of stock i for $i = 1, 2, ...$:

$$\frac{dS_t^i}{S_t^i} = \mu_i dt + \beta_i \sigma_m dW_t + \sigma_i dZ_t^i$$

where drift μ_i is

$$\mu_i = r + \delta \beta_i \sigma_m + \gamma \sigma_i$$

- $\beta_i \sim UNI\left[\kappa_\beta, \kappa_\beta + \lambda_\beta\right]$: beta of stock i
- $\sigma_i \sim UNI\left[0, \lambda_\sigma\right]$: idiosyncratic volatility of stock i
- γ: idiosyncratic volatility premium
- $\left\{Z_t^i\right\}$: Brownian motion; **source of idiosyncratic shock**

Dependence Among Returns

Applying Itô's lemma:

$$\frac{S_T^i}{S_0^i} = \exp\left[\left(\mu_i - 0.5\beta_i^2\sigma_m^2 - 0.5\sigma_i^2\right)T + \beta_i\sigma_m W_T + \sigma_i Z_T^i\right]$$

$$\frac{M_T}{M_0} = \exp\left[\left(\mu_m - 0.5\sigma_m^2\right)T + \sigma_m W_T\right]$$

$W_T, Z_T^i \sim i.i.d.\ N\left(0, T\right)$

W_T induces **dependence** among $\frac{S_T^1}{S_0^1}, \frac{S_T^2}{S_0^2}, \ldots$

However, $\frac{S_T^1}{S_0^1}, \frac{S_T^2}{S_0^2}, \ldots$ are **conditionally i.i.d.** given $\frac{M_T}{M_0}$

Monte Carlo Design

Inputs:

- $\sigma_m = 0.20$, $\gamma = 0.50$
- $\kappa_\beta = -0.20$, $\lambda_\beta = 3.40$; $\lambda_\sigma = 0.50$
- $\delta = 0.50$, $r = 0.01$, $T = 1/12$

Identifiable parameters are $\boldsymbol{\theta} = \left(\sigma_m, \gamma, \kappa_\beta, \lambda_\beta, \lambda_\sigma\right)'$

Moment restrictions are of the form:

$$g_i\left(\xi; \boldsymbol{\theta}\right) = \left(S_T^i / S_0^i\right)^{\xi} - E_{\boldsymbol{\theta}}\left[\left(S_T^i / S_0^i\right)^{\xi} | M_T / M_0\right]$$

- vector $\boldsymbol{g}\left(S_T^i / S_0^i; \boldsymbol{\theta}, M_T / M_0\right) = \left(g_i\left(\xi_1; \boldsymbol{\theta}\right), ..., g_i\left(\xi_6; \boldsymbol{\theta}\right)\right)'$
- vector $\boldsymbol{\xi} = \left(-1.5, -1, -0.5, 0.5, 1, 1.5\right)'$

Monte Carlo Results

	Sample size n (thousands)					
	25	50	250	1,000	10,000	True value
Panel A: Means						
σ_m	0.2526	0.2382	0.2205	0.2116	0.2011	0.2000
γ	0.5560	0.5339	0.5161	0.5076	0.5020	0.5000
κ_β	−0.1316	−0.1484	−0.1476	−0.1817	−0.1978	−0.2000
λ_β	3.6166	3.5798	3.4874	3.4722	3.4303	3.4000
λ_σ	0.4989	0.4996	0.4998	0.4999	0.5000	0.5000
Panel B: RMSEs						
σ_m	0.2327	0.2102	0.1382	0.1279	0.0658	
γ	0.2105	0.1582	0.0836	0.0488	0.0182	
κ_β	0.9925	0.8817	0.7330	0.4077	0.1410	
λ_β	1.4086	1.2965	0.8896	0.8310	0.4298	
λ_σ	0.0063	0.0046	0.0020	0.0010	0.0003	
Panel C: Test sizes, H_0 : parameter = true value, %						
σ_m	15.80	13.20	8.00	7.10	5.70	5.00
γ	7.30	5.50	5.40	5.60	5.30	5.00
κ_β	8.30	6.40	5.70	5.40	4.60	5.00
λ_β	10.60	9.60	5.60	5.50	4.70	5.00
λ_σ	3.80	3.10	4.60	3.80	4.50	5.00
Panel D: OIR test size, H_0 : correct specification, %						
	19.50	15.50	11.30	8.70	8.50	5.00

Thank you!

Questions?

Econometric Literature

Localized common shock:

- general approach: Conley (1999)
- spatial effects: e.g., Kelejian & Prucha (1999)
- group effects: e.g., Lee (2007)
- social effects: e.g., Bramoullé et al. (2009)

Non-localized common shock:

- Andrews (2003)
- Andrews (2005)

Consistency: Proof Sketch

We adapt argument due to Andrews (2003) but clarify several details

Sketch:

- infer existence and measurability of estimator from standard theorem
- show pointwise convergence of objective
- show stochastic equicontinuity of objective
- establish uniform convergence of objective
- establish unique minimum of objective in the limit at $\boldsymbol{\theta}_0$ a.s.
- use the above results to prove convergence of estimator to $\boldsymbol{\theta}_0$

Stochastic Variance Terms

$\mathbf{V}_{1,\mathcal{F}_0}$ and $\mathbf{V}_{2,\mathcal{F}_0}$ are a.s. pos. def. **stochastic** matrices:

$$\mathbf{V}_{1,\mathcal{F}_0} = \left[\mathbf{G}'_{\mathcal{F}_0}\mathbf{\Sigma}^{-1}\mathbf{G}_{\mathcal{F}_0}\right]^{-1}\mathbf{G}'_{\mathcal{F}_0}\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_{\mathcal{F}_0}\mathbf{\Sigma}^{-1}\mathbf{G}_{\mathcal{F}_0}\left[\mathbf{G}'_{\mathcal{F}_0}\mathbf{\Sigma}^{-1}\mathbf{G}_{\mathcal{F}_0}\right]^{-1}$$

$$\mathbf{V}_{2,\mathcal{F}_0} = \left[\mathbf{G}'_{\mathcal{F}_0}\mathbf{\Sigma}^{-1}_{\mathcal{F}_0}\mathbf{G}_{\mathcal{F}_0}\right]^{-1}$$

Asymptotic Mixed Normality: Proof Sketch

Proof utilizes conventional techniques:

- show that $\boldsymbol{g}\left(X_1; \boldsymbol{\theta}_0, X_0\right), \boldsymbol{g}\left(X_2; \boldsymbol{\theta}_0, X_0\right), \ldots$ is m.d.s.
- mean-value expand $\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{g}\left(X_i; \widehat{\boldsymbol{\theta}}_{1,n}, X_0\right)$ around $\boldsymbol{\theta}_0$
- show that $\mathbf{G}_n\left(\widehat{\boldsymbol{\theta}}_{1,n}\right) \equiv \frac{1}{n}\sum_{i=1}^{n} \frac{\partial \boldsymbol{g}\left(X_i; \widehat{\boldsymbol{\theta}}_{1,n}, X_0\right)}{\partial \boldsymbol{\theta}'} \rightarrow^{p} \mathbf{G}_{\mathcal{F}_0}$
- invoke c.l.t. for m.d.s. to show that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \boldsymbol{g}\left(X_i; \boldsymbol{\theta}_0, X_0\right) \rightarrow^{d} \left[\boldsymbol{\Sigma}_{\mathcal{F}_0}\right]^{\frac{1}{2}} \mathbf{Z}_k$$

- invoke standard arguments to establish final result with $\mathbf{V}_{1,\mathcal{F}_0}$
- repeat steps for $\widehat{\boldsymbol{\theta}}_{2,n}$ and simplify to obtain $\mathbf{V}_{2,\mathcal{F}_0}$

Inference: Formulas

$$W_n \equiv n\mathbf{a}\left(\widehat{\boldsymbol{\theta}}_{2,n}\right)'\left[\mathbf{A}\left(\widehat{\boldsymbol{\theta}}_{2,n}\right)\mathbf{V}_{2,n}\mathbf{A}\left(\widehat{\boldsymbol{\theta}}_{2,n}\right)'\right]^{-1}\mathbf{a}\left(\widehat{\boldsymbol{\theta}}_{2,n}\right)$$

$$\mathbf{V}_{2,n} = \left[\mathbf{G}'_{2,n}\widehat{\boldsymbol{\Sigma}}_{2,n}^{-1}\mathbf{G}_{2,n}\right]^{-1}$$

$$\mathbf{G}_{2,n} = n^{-1}\sum_i \partial \boldsymbol{g}\left(X_i; \widehat{\boldsymbol{\theta}}_{2,n}, X_0\right)/\partial\boldsymbol{\theta}'$$

$$\widehat{\boldsymbol{\Sigma}}_{2,n} = n^{-1}\sum_i \boldsymbol{g}\left(X_i; \widehat{\boldsymbol{\theta}}_{2,n}, X_0\right)\cdot \boldsymbol{g}\left(X_i; \widehat{\boldsymbol{\theta}}_{2,n}, X_0\right)'$$

Finance Literature

Recall:

$$\mu_i = r + \delta\beta_i\sigma_m + \gamma\sigma_i$$

If $\gamma = 0$, our price dynamics are in line with:

- ICAPM with constant invest. opportunity set: Merton (1973)
- APT with a single market factor: Ross (1976)

But idiosyncratic volatility may be priced:

- Merton (1987), Malkiel & Xu (2006): incomplete diversification
- Epstein & Schneider (2008): ambiguity premium
- Bhootra & Hur (2011): risk-seeking in capital loss domain

Ang et al. (2006, 2009), Fu (2009):
idiosyncratic premium $\neq 0$, but **no consensus about sign**

Martingale Difference Sequence

Sequence of random variables $\{Y_i\}$ on probability space $(\Omega, \mathcal{F}, P)$ is **martingale difference sequence** (m.d.s.) with respect to filtration $\{\mathcal{F}_i\}$ if:

(i) Y_i is measurable with respect to $\mathcal{F}_i$ for all i

(ii) $E\left[|Y_i|\right] < \infty$ for all i

(iii) $E[Y_j|F_i] = 0$ a.s. for all $j > i$

Mixed Normal Distribution

Random variable Y has a **mixed normal distribution**

$$Y \sim MN\left(0, \eta^2\right)$$

if characteristic function of Y is

$$\phi_Y(t) \equiv E\left[\exp\left(itY\right)\right] = E\left[\exp\left(-\frac{1}{2}\eta^2 t^2\right)\right]$$

where η is a random variable

Y can be represented as

$$Y = \eta Z$$

where $Z \sim N(0, 1)$ and Z is **independent** of η

Law of Large Numbers for Conditionally I.I.D. R.V.'s

Let random variables $X_1, X_2, \ldots$ be defined on probability space $(\Omega, \mathcal{F}, P)$. Suppose there exists σ-field $\mathcal{F}_0 \subset \mathcal{F}$ such that, **conditional on** $\mathcal{F}_0$, $X_1, X_2, \ldots$ are i.i.d. Let $h(\cdot)$ be vector-valued function that satisfies $E\left\|h(X_i)\right\| < \infty$, where $\|\cdot\|$ is Euclidean norm. Then:

$$\frac{1}{n}\sum_{i=1}^{n} h(X_i) \to^p E\left(h(X_i)\,|\mathcal{F}_0\right) \text{ as } n \to \infty$$

Remark:

$E\left(h(X_i)\,|\mathcal{F}_0\right)$ is a random variable

See Andrews (2005, p. 1557), Hall & Heyde (1980, p. 202)

Central Limit Theorem for M.D.S.

Let $\{S_{ni}, \mathcal{F}_{ni}, 1 \le i \le k_n, n \ge 1\}$ be zero-mean, square-integrable martingale array with differences X_{ni}, and let η^2 be a.s. finite r.v. Suppose that:

(i) $\max_i |X_{ni}| \to^p 0$

(ii) $\sum_i X_{ni}^2 \to^p \eta^2$

(iii) $E\left(\max_i X_{ni}^2\right)$ is bounded in n

and σ-fields are nested: $\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$. Then:

$$S_{nk_n} = \sum_i X_{ni} \to^d Z,$$

where r.v. Z has characteristic function $E\left[\exp\left(-\frac{1}{2}\eta^2 t^2\right)\right]$

Remark: Z has a **mixed normal** distribution

See Hall & Heyde (1980, pp. 58-59)

Stochastic Equicontinuity (I)

Let $B(\theta, \delta)$ denote closed ball of radius $\delta > 0$ centered at θ. Sequence of functions $\{G_n(\theta)\}$ is **stochastically equicontinuous** on Θ if for any $\epsilon > 0$ there exists $\delta > 0$ such that

$$\limsup_{n\to\infty} P\left(\sup_{\theta\in\Theta}\sup_{\theta'\in B(\theta,\delta)}\left|G_n(\theta') - G_n(\theta)\right| > \epsilon\right) < \epsilon$$

Assumption SE-1 of Andrews (1992, p. 246):

(a) $G_n(\theta) = \hat{Q}_n(\theta) - Q_n(\theta)$, where $Q_n(\cdot)$ is nonrandom function that is continuous in θ uniformly over Θ

(b) $\left|\hat{Q}_n(\theta') - \hat{Q}_n(\theta)\right| \leq B_n h\left(d\left(\theta', \theta\right)\right)$ for any θ',$\theta \in \Theta$ a.s. for some random variable B_n and some nonrandom function h such that $h(y) \downarrow 0$ as $y \downarrow 0$, where d is metric on Θ

(c) $B_n = O_p(1)$

Stochastic Equicontinuity (II)

Lemma 1 of Andrews (1992, p. 246). If $\{G_n(\theta)\}$ satisfies Assumption SE-1, then $\{G_n(\theta)\}$ is stochastically equicontinuous on Θ

Theorem 1 of Andrews (1992, p. 244). Suppose that:

(i) Θ is totally bounded metric space

(ii) $G_n(\theta) \rightarrow^p 0$ for all $\theta \in \Theta$ (pointwise)

(iii) $\{G_n(\theta)\}$ is stochastically equicontinuous on Θ

then $G_n(\theta)$ converges **uniformly** in probability to 0:

$$\sup_{\theta \in \Theta} |G_n(\theta)| \rightarrow^p 0$$

Remark: total boundedness is weaker than compactness