Lecture 3:
Linear Model with Panel Data:
Random Effects and Hausman Test

Bertel Schjerning

University of Copenhagen, Department of Economics

February 3, 2026

Plan for Lectures on Classical Linear Models

Lecture 1: Recap, Cross section, Least squares (W.4)

Lecture 2: Panel, Fixed effects, First differences (W.10)

Lecture 3: Panel, Random effects, Hausman test (W.10)

Next: *Non*-linear methods (W.12+13)

Empirical Strategies

Worry about whether $\mathrm{E}[\mathbf{x}_i' c_i] \neq \mathbf{0}$? Different approaches.

Fixed effects (FE) approach:

- *Eliminate* c_i from system via transformation.
- No need for model relating c_i and $\mathbf{x}_i$.

Random effects (RE) approach:

- *Model* (features of) distribution of $(c_i, \mathbf{x}_i)$.
  - E.g. specify $\mathrm{E}[c_i \mid \mathbf{x}_i]$, $\mathrm{var}(c_i \mid \mathbf{x}_i)$, entire $D(c_i \mid \mathbf{x}_i)$.
- May/may not transform data (efficiency)

Today: Restrictive but simple ("classical RE") model.

Outline

Pooled OLS

Identification and Consistency

Let $v_{it} := c_i + u_{it}$ denote composite error.

$$\textbf{POLS.1: } \mathrm{E}\left[\mathbf{x}_{it}' v_{it}\right] = \mathbf{0}, \quad t = 1, 2, \ldots, T,$$
$$\textbf{POLS.2: } \operatorname{rank} \mathrm{E}\left(\mathbf{X}_i' \mathbf{X}_i\right) = K$$

POLS.1: (Unconditional) contemporaneous exogeneity

- Suffices that $\mathrm{E}\left[u_{it} \mid \mathbf{x}_{it}\right] = 0$ all t
- ... and $\mathrm{E}\left[c_i \mid \mathbf{x}_i\right] = 0$.

POLS.2: A mild rank condition.

Under POLS.1–2, $\boldsymbol{\beta}$ identified + $\widehat{\boldsymbol{\beta}}_{POLS}$ consistent. (Check!)

Asymptotic Normality

Under POLS.1–2, $\widehat{\boldsymbol{\beta}}_{POLS}$ $\sqrt{N}$-asymptotically normal (check!),

$$\begin{aligned}\sqrt{N}(\widehat{\boldsymbol{\beta}}_{POLS} - \boldsymbol{\beta}) &\xrightarrow{d} \mathrm{N}\left(\mathbf{0}, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\right), \\ \text{where} \quad \mathbf{A} &:= \mathrm{E}\left[\mathbf{X}_i'\mathbf{X}_i\right], \\ \mathbf{B} &:= \mathrm{E}\left[\mathbf{X}_i'\mathbf{v}_i\mathbf{v}_i'\mathbf{X}_i\right], \\ \mathbf{v}_i &:= (v_{i1}, v_{i2}, \ldots, v_{iT})'.\end{aligned}$$

Note: “Meat” $\mathbf{B}$ panel version of $E(u^2\mathbf{x}'\mathbf{x})$ from cross section.

Robust variance estimation takes into account panel structure.

Variance Estimation

Under POLS.1–2, asymptotic variance is

$$\operatorname{Avar}(\widehat{\boldsymbol{\beta}}_{POLS}) = \left(\mathrm{E}\left[\mathbf{X}_i'\mathbf{X}_i\right]\right)^{-1} \mathrm{E}\left[\mathbf{X}_i'\mathbf{v}_i\mathbf{v}_i'\mathbf{X}_i\right] \left(\mathrm{E}\left[\mathbf{X}_i'\mathbf{X}_i\right]\right)^{-1} / N.$$

Consistent estimator thereof:

$$\widehat{\operatorname{Avar}(\widehat{\boldsymbol{\beta}}_{POLS})} = \left(\mathbf{X}'\mathbf{X}\right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}_i'\widehat{\mathbf{v}}_i\widehat{\mathbf{v}}_i'\mathbf{X}_i\right) \left(\mathbf{X}'\mathbf{X}\right)^{-1},$$

$$\widehat{v}_{it} := y_{it} - \mathbf{x}_{it}\widehat{\boldsymbol{\beta}}_{POLS}.$$

No assumptions placed on $\mathrm{E}[\mathbf{v}_i\mathbf{v}_i' \mid \mathbf{x}_i] \Rightarrow$ robust towards

- heteroskedasticity (cond'l)
- and serial correlation (cond'l)

Discussion

Under contemporaneous exogeneity + rank condition,

$$\textbf{POLS.1: } \mathrm{E}\left[\mathbf{x}_{it}' v_{it}\right] = \mathbf{0}, \quad t = 1, 2, \ldots, T,$$
$$\textbf{POLS.2: } \operatorname{rank} \mathrm{E}\left(\mathbf{X}_i' \mathbf{X}_i\right) = K,$$

then $\widehat{\beta}_{POLS}$ $\sqrt{N}$-asymptotically normal.

... but not necessarily efficient.

Main motivation for random effects estimator (below):

Efficiency.

(Classical) Random Effects

Identifying Assumptions

(Stronger) Identifying Assumptions

Assumption RE.1:

$$\textbf{RE.1(a):}\ \mathrm{E}\left[u_{it} \mid \mathbf{x}_i, c_i\right] = 0, \quad t = 1, 2, \ldots, T,$$
$$\textbf{RE.1(b):}\ \mathrm{E}\left[c_i \mid \mathbf{x}_i\right] = \mathrm{E}\left[c_i\right] = 0.$$

RE.1(a) implies $\mathrm{E}\left[c_i u_{it}\right] = 0$ and $\mathrm{E}\left[\mathbf{x}_{it}' u_{it}\right] = \mathbf{0}$ all t

- ... also maintains *strict exogeneity*

RE.1(b) imposes c_i mean independent of $\{\mathbf{x}_{it}\}_{t=1}^{T}$

- This feature distinguishes RE from FE.

RE.1 implies $\{v_{it} = c_i + u_{it}\}_{t=1}^{T}$ uncorrelated with $\{\mathbf{x}_{it}\}_{t=1}^{T}$.

More restrictive than needed for pooled OLS.

Identifying Assumptions

Stack $\mathbf{v}_i := (v_{i1}, v_{i2}, \ldots, v_{iT})'$ $(T \times 1)$

Denote unconditional composite error variance

$$\boldsymbol{\Omega} := \mathrm{E}\left[\mathbf{v}_i \mathbf{v}_i'\right]. \qquad (T \times T)$$

Rank condition:

$$\textbf{RE.2:}\ \operatorname{rank} \mathrm{E}\left[\mathbf{X}_i' \boldsymbol{\Omega}^{-1} \mathbf{X}_i\right] = K.$$

As opposed to FE/D estimators, no time-variance restriction.

- Could (and should) include model intercept
- ... along w/ time-invariant regressors (e.g. ethnic origin)

Error Structure

Error Structure

- Could do [feasible] generalized least squares ([F]GLS).
- Consistent under mild additional conditions.
- Idea: View panel as linear system

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{v}_i, \quad \mathbf{v}_i := c_i \mathbf{j}_T + \mathbf{u}_i. \qquad (T \times 1)$$

- Apply general results for system OLS.
  - Details in W. Ch. 7.
- But wouldn't be exploiting composite error structure.
- Classical RE analysis restricts $\boldsymbol{\Omega}$ form. (Next.)

Error Variance Structure

Assumption RE.3:

$$\textbf{RE.3(a):}\ \mathrm{E}\left[\mathbf{u}_i\mathbf{u}_i' \mid \mathbf{x}_i, c_i\right] = \sigma_u^2 \mathbf{I}_T, \qquad (T \times T)$$

$$\textbf{RE.3(b):}\ \mathrm{E}\left[c_i^2 \mid \mathbf{x}_i\right] = \sigma_c^2.$$

- RE.3(a) implies $\{u_{it}\}_{t=1}^T$ (conditionally) homoskedastic + serially uncorrelated.
- RE.3(b) implies c_i conditionally homoskedastic.
- RE.1 + RE.3 imply *conditional* variance independence

$$\mathrm{E}\left[\mathbf{v}_i\mathbf{v}_i' \mid \mathbf{x}_i\right] = \mathrm{E}\left[\mathbf{v}_i\mathbf{v}_i'\right]. \qquad (= \boldsymbol{\Omega})$$

Error Variance Structure

Under RE.1 and RE.3,

$$\begin{aligned}\boldsymbol{\Omega} = \mathrm{E}\left[\mathbf{v}_i\mathbf{v}_i'\right] &= \begin{bmatrix} \sigma_c^2+\sigma_u^2 & \sigma_c^2 & \cdots & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2+\sigma_u^2 & \cdots & \vdots \\ \vdots & & \ddots & \sigma_c^2 \\ \sigma_c^2 & \cdots & \sigma_c^2 & \sigma_c^2+\sigma_u^2 \end{bmatrix}, \\ &= \sigma_u^2\mathbf{I}_T + \sigma_c^2\mathbf{j}_T\mathbf{j}_T'. \qquad (T \times T)\end{aligned}$$

We say $\boldsymbol{\Omega}$ has random effects structure.

Estimator

Estimator

Sps. (for now) access to some consistent estimators $\widehat{\sigma}_u^2$ and $\widehat{\sigma}_c^2$.

May then form

$$\widehat{\boldsymbol{\Omega}} := \widehat{\sigma}_u^2 \mathbf{I}_T + \widehat{\sigma}_c^2 \mathbf{j}_T \mathbf{j}_T'.$$

Slutsky's thm. $\Rightarrow \widehat{\boldsymbol{\Omega}} \to_p \boldsymbol{\Omega}$ (i.e., $\to_p$ coordinatewise)

Random-effects estimator = FGLS with RE structured $\widehat{\boldsymbol{\Omega}}$,

$$\widehat{\boldsymbol{\beta}}_{RE} := \left(\sum_{i=1}^{N} \mathbf{X}_i' \widehat{\boldsymbol{\Omega}}^{-1} \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}_i' \widehat{\boldsymbol{\Omega}}^{-1} \mathbf{y}_i \right).$$

Estimator

$$\widehat{\beta}_{RE} = \left(\sum_{i=1}^{N} \mathbf{X}_i' \widehat{\mathbf{\Omega}}^{-1} \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}_i' \widehat{\mathbf{\Omega}}^{-1} \mathbf{y}_i \right).$$

Special case: $\sigma_c^2 = 0$

- Follows from no c_i's [under RE.1(b)]
- Then $\mathbf{\Omega} = \sigma_u^2 \mathbf{I}_T$
- ... and asymptotically, RE collapses to POLS. (Check!)

Quasi Time-Demeaning

Quasi Time-Demeaning

Define fraction

$$\lambda := 1 - \sqrt{\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_c^2}}. \qquad (\in [0,1])$$

Linear algebra shows (check!)

$$\mathbf{\Omega}^{-1/2} = \frac{1}{\sigma_u}\left(\mathbf{I}_T - \lambda \mathbf{P}_T\right),$$

where

$$\mathbf{P}_T := \frac{1}{T}\mathbf{j}_T\mathbf{j}_T' = \begin{bmatrix} 1/T & 1/T & \cdots & 1/T \\ 1/T & 1/T & & \vdots \\ \vdots & & \ddots & \\ 1/T & \cdots & & 1/T \end{bmatrix}. \qquad (T \times T)$$

Let

$$\mathbf{C}_T := \mathbf{I}_T - \lambda \mathbf{P}_T.$$

Quasi Time-Demeaning

$$\mathbf{C}_T = \mathbf{I}_T - \lambda \mathbf{P}_T, \quad \mathbf{P}_T = \frac{1}{T}\mathbf{j}_T\mathbf{j}_T'.$$

Assume λ known (for now) $\Rightarrow$ GLS w/ $\mathbf{C}_T$ weighting possible.

RE estimator arises from estimation of transformed system

$$\mathbf{C}_T\mathbf{y}_i = \mathbf{C}_T\mathbf{X}_i\boldsymbol{\beta} + \mathbf{C}_T\mathbf{v}_i.$$

New errors have variance $\mathrm{E}[(\mathbf{C}_T\mathbf{v}_i)(\mathbf{C}_T\mathbf{v}_i)'] = \sigma_u^2\mathbf{I}_T$. (Check!)

Writing out system,

$$y_{it} - \lambda\bar{y}_i = (\mathbf{x}_{it} - \lambda\bar{\mathbf{x}}_i)\,\boldsymbol{\beta} + (v_{it} - \lambda\bar{v}_i)\,, \quad t = 1, 2, \ldots, T.$$

Quasi Time-Demeaning

Let

$$\widehat{\lambda} := 1 - \sqrt{\widehat{\sigma}_u^2/(\widehat{\sigma}_u^2 + T\widehat{\sigma}_c^2)},$$
$$\widehat{\mathbf{C}}_T := \mathbf{I}_T - \widehat{\lambda}\mathbf{P}_T.$$

RE estimator takes form

$$\begin{aligned}\widehat{\beta}_{RE} &= \left(\sum_{i=1}^{N} \mathbf{X}_i'\widehat{\mathbf{C}}_T'\widehat{\mathbf{C}}_T\mathbf{X}_i\right)^{-1}\left(\sum_{i=1}^{N} \mathbf{X}_i'\widehat{\mathbf{C}}_T'\widehat{\mathbf{C}}_T\mathbf{y}_i\right)\\ &= \left(\sum_{i=1}^{N}\sum_{t=1}^{T} \breve{\mathbf{x}}_{it}'\breve{\mathbf{x}}_{it}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T} \breve{\mathbf{x}}_{it}'\breve{y}_{it}\right)\end{aligned}$$

where $\breve{\mathbf{x}}_{it} := \mathbf{x}_{it} - \widehat{\lambda}\overline{\mathbf{x}}_i, \quad \breve{y}_{it} := y_{it} - \widehat{\lambda}\overline{y}_i.$

Conclude: RE subtracts *fraction* of time-average.

Relationship: RE, FE and Pooled OLS

$$y_{it} - \widehat{\lambda}\bar{y}_i = (\mathbf{x}_{it} - \widehat{\lambda}\bar{\mathbf{x}}_i)\boldsymbol{\beta} + (v_{it} - \widehat{\lambda}\bar{v}_i),$$

$$\text{where} \quad \widehat{\lambda} := 1 - \sqrt{\frac{\widehat{\sigma}_u^2}{\widehat{\sigma}_u^2 + T\widehat{\sigma}_c^2}}.$$

RE = pooled OLS after (feasible) quasi time-demeaning

$\widehat{\lambda} \approx 1 \Rightarrow$ RE and FE tend to be close.

- Happens for large T.
- Follows from $\mathbf{C}_T \to \mathbf{I}_T - \mathbf{P}_T = \mathbf{Q}_T$ as $\lambda \to 1$.

$\widehat{\lambda} \approx 0 \Rightarrow$ RE and pooled OLS tend to be close.

- Happens with $\sigma_u^2 \gg \sigma_c^2$.

Implementation

Implementation

We need $\widehat{\sigma}_u^2$ and $\widehat{\sigma}_c^2$ consistent.

Actually, easier to consistently estimate σ_v^2 and σ_c^2.

May then obtain $\widehat{\sigma}_u^2 := \widehat{\sigma}_v^2 - \widehat{\sigma}_c^2$ consistent (using RE.3).

$\widehat{\sigma}_v^2$ consistent (using RE.1–2) given by

$$\widehat{\sigma}_v^2 := \frac{1}{NT-K}\sum_{i=1}^{N}\sum_{t=1}^{T}\check{v}_{it}^2,$$

where $\check{v}_{it} := y_{it} - \mathbf{x}_{it}\widehat{\beta}_{POLS}$. (pooled OLS residuals)

POLS suitable because $\mathrm{E}[\mathbf{x}_{it}' v_{it}] = \mathbf{0}$ under RE.1.

Implementation

Under RE.3, σ_c^2 everywhere off diagonal of $\boldsymbol{\Omega}$.

Sum up $T(T-1)/2$ non-redundant variance terms,

$$\sum_{t=1}^{T-1} \sum_{s=t+1}^{T} \mathrm{E}\left[v_{it} v_{is}\right] = \frac{T\left(T-1\right)}{2} \sigma_c^2.$$

Rearrange + analogy principle yield

$$\widehat{\sigma}_c^2 := \frac{1}{NT\left(T-1\right)/2 - K} \sum_{i=1}^{N} \sum_{t=1}^{T-1} \sum_{s=t+1}^{T} \check{v}_{it} \check{v}_{is},$$

where $\quad \check{v}_{it} := y_{it} - \mathbf{x}_{it} \widehat{\beta}_{POLS}. \qquad$ (pooled OLS residuals)

Lastly, $\widehat{\sigma}_u^2 := \widehat{\sigma}_v^2 - \widehat{\sigma}_c^2$.

Alternative implementation

Alternative implementation (Within + Between)

Goal: estimate (σ_u^2, σ_c^2) and compute $\widehat{\lambda}$.

Step 1 (Within / FE): reuse FE residual variance estimator

$$\widehat{\sigma}_u^2 := \frac{1}{NT - N - K} \sum_{i=1}^{N} \sum_{t=1}^{T} \widehat{\ddot{u}}_{it}^2, \qquad \widehat{\ddot{u}}_{it} := \ddot{y}_{it} - \ddot{\mathbf{x}}_{it} \widehat{\boldsymbol{\beta}}_{FE}.$$

Step 2 (Between / BE): run OLS on individual means

$$\overline{y}_i = \overline{\mathbf{x}}_i \boldsymbol{\beta} + c_i + \overline{u}_i, \qquad \widehat{e}_{i,BE} := \overline{y}_i - \overline{\mathbf{x}}_i \widehat{\boldsymbol{\beta}}_{BE}.$$

Under RE: $\widehat{e}_{i,BE} \approx c_i + \overline{u}_i$ and $\mathrm{Var}(c_i + \overline{u}_i) = \sigma_c^2 + \sigma_u^2 / T$.

Alternative implementation (Within + Between)

Hence estimate

$$\widehat{\sigma}^2_{e,BE} := \frac{1}{N-K}\sum_{i=1}^{N}\widehat{e}^2_{i,BE}, \qquad \widehat{\sigma}^2_c := \max\left\{0,\ \widehat{\sigma}^2_{e,BE} - \widehat{\sigma}^2_u/T\right\}.$$

Then compute the quasi-demeaning parameter

$$\widehat{\lambda} := 1 - \sqrt{\frac{\widehat{\sigma}^2_u}{\widehat{\sigma}^2_u + T\widehat{\sigma}^2_c}}.$$

Implementation: run OLS on quasi-demeaned variables

$$y_{it} - \widehat{\lambda}\,\overline{y}_i = \left(\mathbf{x}_{it} - \widehat{\lambda}\,\overline{\mathbf{x}}_i\right)\boldsymbol{\beta} + \text{error}.$$

RE as GLS: weighted within + weighted between

Under RE.1–RE.3 (balanced panel), GLS can be written as a weighted combination of within and between fit.
Define

$$S_W(\beta) := \sum_{i=1}^{N}\sum_{t=1}^{T}\left(\ddot{y}_{it} - \ddot{\mathbf{x}}_{it}\beta\right)^2, \qquad S_B(\beta) := \sum_{i=1}^{N}\left(\bar{y}_i - \bar{\mathbf{x}}_i\beta\right)^2.$$

Then

$$\widehat{\beta}_{RE} = \operatorname*{argmin}_{\beta}\left\{\frac{1}{\sigma_u^2}S_W(\beta) + \frac{T}{\sigma_u^2 + T\sigma_c^2}S_B(\beta)\right\}.$$

- FE uses only $\ddot{y}_{it}$ and $\ddot{\mathbf{x}}_{it}$ (pure within variation).
- BE uses only $\bar{y}_i$ and $\bar{\mathbf{x}}_i$ (pure between variation),
- RE optimaliy weights **within** and **between** variation.
- The weights are driven by (σ_u^2, σ_c^2) (hence by λ).

Within and between variaation: role of λ

The RE transformation uses quasi-demeaning

$$y_{it} - \lambda \bar{y}_i = \left(\mathbf{x}_{it} - \lambda \bar{\mathbf{x}}_i\right)\boldsymbol{\beta} + \left(u_{it} - \lambda \bar{u}_i\right) + (1-\lambda)c_i.$$

$\lambda = 1 - \sqrt{\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_c^2}}$.

- λ is increasing in σ_c^2, and decreasing in σ_u^2.
- If σ_c^2 is large $\Rightarrow \lambda \approx 1$:
  - quasi-demeaning is close to *within* demeaning,
  - $(1-\lambda)c_i$ is (almost) removed,
  - RE $\approx$ FE and mainly uses within-variation.
- If σ_c^2 is small $\Rightarrow \lambda \approx 0$:
  - little demeaning,
  - RE $\approx$ pooled OLS and puts relatively more weight on between/levels variation.

Asymptotic Normality, Inference, Variance Estimation

Asymptotic Normality (Sketch Only)

- Since $\widehat{\boldsymbol{\Omega}} \to_p \boldsymbol{\Omega}$, may ignore $\boldsymbol{\Omega}$ estimation.

- Now proceed *as if*

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{RE} - \boldsymbol{\beta}) = \left(\frac{1}{N} \sum_{i=1}^{N} \mathbf{X}_i' \boldsymbol{\Omega}^{-1} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbf{X}_i' \boldsymbol{\Omega}^{-1} \mathbf{v}_i \right)$$

- Under RE.1–3, RE asymptotically normal:

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{RE} - \boldsymbol{\beta}) \xrightarrow{d} \mathrm{N}\left(\mathbf{0}, \mathbf{A}^{-1}\right),$$
$$\text{with} \quad \mathbf{A} := \mathrm{E}\left[\mathbf{X}_i' \boldsymbol{\Omega}^{-1} \mathbf{X}_i\right].$$

Efficiency

- $\text{Avar}(\widehat{\boldsymbol{\beta}}_{RE}) = \mathbf{A}^{-1}/N$.

- Thus using $\boldsymbol{\Omega}^{-1/2} = \sigma_u^{-1}\mathbf{C}_T$,

$$\begin{aligned}
\text{Avar}(\widehat{\boldsymbol{\beta}}_{RE}) &= \left(\mathrm{E}\left[\mathbf{X}_i'\boldsymbol{\Omega}^{-1}\mathbf{X}_i\right]\right)^{-1}/N \\
&= \left(\mathrm{E}\left[\mathbf{X}_i'\sigma_u^{-2}\mathbf{C}_T'\mathbf{C}_T\mathbf{X}_i\right]\right)^{-1}/N \\
&= \sigma_u^2\left\{\mathrm{E}\left[(\mathbf{C}_T\mathbf{X}_i)'(\mathbf{C}_T\mathbf{X}_i)\right]\right\}^{-1}/N.
\end{aligned}$$

- Gauss-Markov form.

Conclude: Under RE.1–3, $\widehat{\boldsymbol{\beta}}_{RE}$ asymptotically efficient.

Variance Estimation

Under RE.1–3, consistent estimator

$$\widehat{\text{Avar}(\widehat{\boldsymbol{\beta}}_{RE})} := \left(\sum_{i=1}^{N} \mathbf{X}_i' \widehat{\Omega}^{-1} \mathbf{X}_i \right)^{-1}$$

If RE.3 in doubt, robust asymptotic variance estimator

$$\left(\sum_{i=1}^{N} \mathbf{X}_i' \widehat{\Omega}^{-1} \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}_i' \widehat{\Omega}^{-1} \widehat{\mathbf{v}}_i \widehat{\mathbf{v}}_i' \widehat{\Omega}^{-1} \mathbf{X}_i \right) \left(\sum_{i=1}^{N} \mathbf{X}_i' \widehat{\Omega}^{-1} \mathbf{X}_i \right)^{-1},$$

with $\widehat{\mathbf{v}}_i := \mathbf{y}_i - \mathbf{X}_i \widehat{\boldsymbol{\beta}}_{RE}$ being the RE residuals.

- RE residuals b/c RE.1–2 suffice for $\widehat{\boldsymbol{\beta}}_{RE}$ consistency.
- But w/o RE.3, $\widehat{\boldsymbol{\beta}}_{RE}$ need not be efficient.

Hausman Test

Hausman Test: Testing the RE Hypothesis

Under RE.1–3 and FE.2,

- Both RE and FE estimators consistent
- And RE asymptotically efficient.

But if $\mathrm{E}[\mathbf{x}_i' c_i] \neq \mathbf{0}$ [$\Rightarrow$ RE.1(b) fails],

- FE consistent
- ... while RE *in*consistent.

Compare $\widehat{\boldsymbol{\beta}}_{FE}$ and $\widehat{\boldsymbol{\beta}}_{RE}$. Significantly different?

If yes, (could) interpret as evidence against RE.1(b).

Hausman Test

Null hypothesis

$$\mathrm{H}_0: \text{ RE.1–3 and FE.2 hold.}$$

Hausman test statistic:

$$H_N := (\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{RE})' \left[\widehat{\mathrm{Avar}(\widehat{\boldsymbol{\beta}}_{FE})} - \widehat{\mathrm{Avar}(\widehat{\boldsymbol{\beta}}_{RE})}\right]^{-1} (\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{RE}).$$

- Under H_0 : $H_N \to_d \chi^2_K$.

Test: Reject at level α iff $H_N > (1-\alpha)$-quantile of χ^2_K.

- A quadratic form in $\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{RE}$.
- Under H_0, $\mathrm{Avar}(\widehat{\boldsymbol{\beta}}_{FE}) - \mathrm{Avar}(\widehat{\boldsymbol{\beta}}_{RE})$ positive semi-definite.
  - $\Rightarrow$ interpret as weighting matrix.

Discussion

Only coefficients on time-varying regressors can be contrasted.

- No within variation $\Rightarrow$ no FE estimate.
- Strictly speaking, K now # time-varying reg's

May use usual (nonrobust) Avar estimates.

But should use same σ_u^2 estimate for both $\widehat{\text{Avars}}$.

- Otherwise variance difference may be *negative* (definite).

If we drop RE.3, $\widehat{\text{Avar}(\widehat{\boldsymbol{\beta}}_{FE})} - \widehat{\text{Avar}(\widehat{\boldsymbol{\beta}}_{RE})}$ to be replaced by

$$\text{Avar}(\widehat{\widehat{\boldsymbol{\beta}}_{FE} - \widehat{\boldsymbol{\beta}}_{RE}}).$$