<u>Nota de Clase 6</u>
Introducción a modelos de Data Panel: Estimación

1. Introducción

El tipo de modelos que se pretenden estimar en este curso corresponde a paneles lineales, cortos y balanceados. Este tipo de modelos, como ya se vio, presentan las ventajas de incrementar la eficiencia econométrica en la estimación, amplía el ámbito de preguntas económicas por resolverse y permite solucionar problemas econométricos como la existencia de factores no observados o sesgo por variables no incluidas. Este último beneficio está íntimamente relacionado con lo que se conoce en la literatura econométrica como heterogeneidad no observada la cual puede modelarse de diferentes formas y asumirse fija o aleatoria. Se ha optado por restringir la discusión a modelos que presentan (o no presentan) heterogeneidad en pendiente (trasversal y/o temporal) asumiendo que los factores detrás de la misma son fijos o aleatorios. En base a ello, se definen los estimador a estudiarse en esta nota de clase: MCO combinado (MCOC), MCO en primeras diferencias, MCO intra-grupo o modelo de efectos fijos(MEF), MCO inter-grupo y MEA o modelo de efectos aleatorios(MEA). Particular interés se pondrá en los estimadores MCOC. MEF y MEA, por ser los casos más generales y mayormente utilizados en la literatura

2. Modelo MCO combinado (MCOC)

En este modelo se ha impuesto la restricción que los interceptos son homogéneos, es decir, no existen efectos no observados. De este modo, se considera un modelo de la siguiente forma

$$y_{it} = \alpha + X_{it}^{'}\beta + u_{it} \tag{1a}$$

o

$$y_{it} = W_{it}B + u_{it} \tag{1b}$$

para $W_{it} = [1 \;\; X_{it}^{'}]$, $B = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ donde $u_{it} \sim IID(0, \sigma^2)$, $i = 1,2,...,N$ y $t = 1,2,...,T$.

Bajo estos supuestos (no existen efectos no observados), el modelo puede estimarse mediante MCO. Para ello las NxT observaciones se combinan en un solo vector de información. Este es un modelo bastante austero en su especificación y el hecho de combinar la variabilidad temporal y transversal de la data en una sola unidad analítica deja de lado importante información que puede ser utilizada más eficientemente en la estimación de los parámetros. Como se verá más adelante, la variabilidad temporal es la dimensión intra-grupo del set de datos y la variabilidad transversal la dimensión inter-grupo. En el MCOC se explotan ambas dimensiones, pero no de manera eficiente (cada una de las dimensiones recibe el mismo peso en la estimación, por más que alguna de ellas sea quizás más importante). Por otro lado, la consistencia de este estimador exige que $\text{cov}(u_{it}, W_{it}) = 0$. Es decir, que las variables explicativas incluidas en W_{it} no estén correlacionadas con el error y que no existan efectos no observables fijos o, en su defecto, que estos no estén correlacionados con el set variables W_{it}. El estimador toma la forma

$$\hat{B}_{MCOC} = (W^{'}W)^{-1}(W^{'}y) \qquad (2)$$

La eficiencia del estimador exige que no exista heterocedasticidad ni autocorrelación en los errores. La heterocedasticidad en este contexto debe ser testeada en la medida que puede ser un supuesto razonable la independencia de las i (por lo menos tan razonable como en un contexto de corte transversal). Más problemático es el supuesto de errores no auto-correlacionados (tan problemático como en un modelo de series de tiempo). Así, para realizar inferencia en presencia de heterocedasticidad y/o auto-correlación es necesario corregir la matriz varianza-covarianza de los estimadores. Al respecto, la corrección por White / Huber puede ser fácilmente extendida al contexto panel. El estimador quedaría expresado como (MCOC generalizado, MCOCG)

$$\hat{B}_{MCOCG} = (W^{'}\Sigma^{-1}W)^{-1}(W^{'}\Sigma^{-1}y) \qquad (3)$$

donde Σ^{-1} es la matriz que corrige por heterocedasticidad y/o autocorrelación y la varianza estimada toma la forma

$$\hat{V}[\hat{B}_{MCOCG}] = \left[\sum_{i=1}^{N}\widetilde{W}_i^{'}\widetilde{W}_i\right]^{-1}\sum_{i=1}^{N}\widetilde{W}_i^{'}\hat{u}_i\hat{u}_i^{'}\widetilde{W}_i\left[\sum_{i=1}^{N}\widetilde{W}_i^{'}\widetilde{W}_i\right]^{-1} \qquad (4)$$

considerando que $\widetilde{W}_i$ es la transformación de W_i. Similarmente, una segunda alternativa es utilizar errores estándar bootstrap. Sin embargo, el supuesto en este caso es que existe independencia en i. Según Cameron y Trivedi (2005) el estimador de la varianza por bootstrap será asintóticamente equivalente al de White / Huber[1].

3. Modelo MCO en primeras diferencias

En este modelo se considera que existen factores no observables que están correlacionados con los regresores incluidos en la ecuación por estimar. Se asume que tales factores varían transversalmente (entre individuos o grupos de individuos), pero son fijos en el tiempo. En tales condiciones y cuando se dispone de solo dos periodos en el panel, es posible plantear un modelo en diferencias como el explicado en el Ejemplo 3 de la nota de clase anterior. De este modo, se considera un modelo de la siguiente forma

$$y_{it} = \alpha_i + X_{it}^{'}\beta + u_{it} \qquad (5a)$$

donde $u_{it} \sim IID(0,\sigma^2)$, $i = 1,2,...,N$ y $t = 1,2,...,T$. Luego, tomando en cuenta el primer rezago de la relación se tiene que

$$y_{it-1} = \alpha_i + X_{it-1}^{'}\beta + u_{it} \qquad (5b)$$

[1] En Cameron y Trivedi (2005) se exploran alternativas de utilizar estructuras ARMA de los errores en el proceso de estimación

y restando (5b) de (5a) se tiene que

$$\Delta y_{it} = [\Delta X_{it}]^{'} \beta + \Delta u_{it} \qquad (6)$$

donde $\Delta y_{it} = (y_{it} - y_{it-1})$, $\Delta X_{it} = (X_{it} - X_{it-1})$ y $\Delta u_{it} = (u_{it} - u_{it-1})$. Así, se han eliminado los factores no observados y el parámetro β puede estimarse mediante MCO dando resultados consistentes. Para corregir los problemas de heterocedasticidad y autocorrelación se procede de la misma forma que en el caso anterior.

4. Modelo MCO intra-grupo o MEF (modelo de efectos fijos)

El estimador MCO intra-grupo o de Efectos Fijos es simplemente la generalización para T periodos del modelo en diferencias. De este modo, considérese un modelo como

$$y_{it} = \alpha_i + X_{it}^{'}\beta + u_{it} \qquad (7a)$$

donde u_{it} ~ $IID(0, \sigma^2)$, $i = 1,2,...,N$ y $t = 1,2,...,T$. En dicho caso, la recomendación es calcular

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it} \qquad \bar{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it} \qquad \bar{u}_i = \frac{1}{T}\sum_{t=1}^{T} u_{it}$$

y expresar el modelo en medias (ver Ejemplo 3 de la nota de clase anterior)

$$\bar{y}_i = \alpha_i + \bar{X}_i^{'}\beta + \bar{u}_i \qquad (7b)$$

para obtener la siguiente relación

$$(y_{it} - \bar{y}_i) = (X_{it} - \bar{X}_i)^{'} \beta + (u_{it} - \bar{u}_i) \qquad (8)$$

El estimador β del modelo (8) puede ser estimado por MCO y queda expresado de la siguiente forma

$$\hat{\beta}_{MEF} = \left[\sum_{i=1}^{N}\sum_{t=1}^{T} (X_{it} - \bar{X}_i)(X_{it} - \bar{X}_i)^{'} \right]^{-1} \left[\sum_{i=1}^{N}\sum_{t=1}^{T} (X_{it} - \bar{X}_i)(y_{it} - \bar{y}_i) \right] \qquad (9)$$

luego, retornando a (7b), se puede estimar cada uno de los α_i si se resuelve

$$\bar{y}_i = \hat{\alpha}_i + \bar{X}_i^{'}\hat{\beta}_{MEF} \qquad (10)$$

Sobre la base de este planteamiento se entiende porque al estimador MEF se le conoce como el intra-grupo: realiza la estimación concentrándose en la variabilidad temporal al interior de los grupos. En otras palabras, la fuente de variación utilizada para estimar el MCO son las desviaciones de cada observación respecto a su media. En este sentido, el estimador explica en que medida y_{it} difiere de y_i; y no porque y_i es diferente a y_j (para los grupos u observaciones transversales i, j). De este modo, frente al estimador MCOC, se ha desechado información, ya que solo se estaría utilizando la variabilidad temporal de la base datos (el estimador MCOC se usó la variabilidad temporal y transversal, aunque de forma ineficiente). Sin embargo, el estimador MEF es consistente bajo el supuesto de la existencia de factores no observables fijos, lo que no sucede con el MCOC que es solo consistente en la ausencia de estos factores. Por ello, ante la ausencia de efectos no observables fijos la elección óptima es el MCOC, pero sí es que estos existen la elección óptima es el MEF. Es posible testear esta proposición mediante un test F sobre la existencia de N interceptos α_i frente a la alternativa de que hay un solo intercepto α. Se plantea

$$F = \frac{(R^2_{MEF} - R^2_{MCOC})/N-1}{(1-R^2_{MEF})/NT-N-k} \sim F_{N-1,NT-N-k} \tag{11}$$

De este modo, en (11) se está comparando un modelo restringido (MCOC) frente a uno sin restringir (MEF) considerando $N-1$ restricciones y $NT-N-k$ grados de libertad del modelo sin restringir donde k es el número de parámetros estimados. Así, en base a los resultados del test

$$\begin{aligned} F &> F_{[N-1,NT-N-k]} \rightarrow no\ acepto\ H_o \quad \rightarrow\ MEF\ es\ correcto \\ F &< F_{[N-1,NT-N-k)]} \rightarrow no\ rechazo\ H_o \ \rightarrow\ MCOC\ es\ correcto \end{aligned}$$

Para derivar la varianza del estimador, conviene expresar el modelo en su forma matricial así se tiene

$$y_{i(Tx1)} = e_{Tx1}\alpha_i + X'_{i(Tx1)}\beta + u_{i(Tx1)} \tag{12a}$$

o

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ \dots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} \alpha_i + \begin{bmatrix} X'_{i1} \\ X'_{i2} \\ \dots \\ X'_{iT} \end{bmatrix} \beta + \begin{bmatrix} u_{i1} \\ u_{i2} \\ \dots \\ u_{iT} \end{bmatrix} \tag{12b}$$

Luego, se puede plantear la matriz $Q_{TxT} = I_{TxT} - \left(\frac{1}{T}\right) e_{Tx1} e'_{1xT}$ que es idempotente[2]. El rol de esta matriz es transformar la data y expresarla en términos de desviaciones de las medias eliminando el término α_i de la expresión. Así, se tiene que

$$Qy_i = Qe\alpha_i + QX'_i\beta + Qu_i \tag{13}$$

[2] La matriz idempotente Q presenta como propiedad $QQ = Q'Q = Q$ además nota que $e'e = T$.

y luego $Qe\alpha_i = e\alpha_i - (1/T)ee^{'}e\alpha_i = e\alpha_i - (1/T)eT\alpha_i = 0$, con lo que el modelo se expresa como

$$Qy_i = QX_i^{'}\beta + Qu_i \tag{14}$$

donde Qy_i , $QX_i^{'}$ y Qu_i están expresado en desviaciones de la media. Con ello, el estimador MCO toma la forma

$$\hat{\beta}_{MEF} = \left[\sum_{i=1}^{N}(X_i^{'}QX_i)\right]^{-1}\left[\sum_{i=1}^{N}(X_i^{'}Qy_i)\right]^{-1} \tag{15}$$

Nótese que la expresión (15) es igual que la expresión (9). La varianza deeste estimador, siguiendo la construcción de MCO es

$$V[\hat{\beta}_{MEF}] = \sigma^2\left[\sum_{i=1}^{N}(X_i^{'}QX_i)\right]^{-1} \tag{16a}$$

o en términos de la expresión (9)

$$V[\hat{\beta}_{MEF}] = \sigma^2\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \overline{X}_i)(X_{it} - \overline{X}_i)^{'}\right]^{-1} \tag{16b}$$

donde

$$\sigma^2 = \frac{1}{NT - N - k}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[[(y_{it} - y_i) - (X_{it} - \overline{X}_i)]\hat{\beta}_{MEF}\right]^2 \tag{16c}$$

Para corregir este estimador por heterocedasticidad y autocorrelación es posible seguir las recomendaciones ya hechas para los otros dos casos (White / Huber y bootstrap). Cameron y Trivedi (2005) argumentan que esta solución es más fácil y no necesariamente mucho menos eficiente que el diseño de un estimador específico de MCG para el caso de modelos MEF.

Como se mencionó anteriormente, el estimador MEF es consistente bajo el supuesto de que existen factores transversales no observables correlacionados con las explicativas e invariantes en el tiempo. Con ello, se tiene una alternativa al estimador MCOC que resultaría inconsistente en estas condiciones. Sin embargo, es recomendable tomar en cuenta algunos aspectos negativos de este estimador

a) El modelo en paneles cortos puede tener una dimensión temporal bastante corta (T bajo) y gran número de observaciones transversales (N alto) con lo que se debe estimar un elevado número de interceptos. De este modo, los grados de libertad pueden verse dramáticamente reducidos en una aplicación MEF.

b) La inclusión de un número grande de parámetros por estimar puede introducir problemas de multicolinearidad derivado de la firma funcional final que adoptan los regresores lo que genera estimadores que resultan imprecisos de otra forma. Este hecho normalmente se asocia al hecho de especificar modelos austeros bajo MEF.

c) No es posible analizar factores que no cambian en el tiempo. Así, en una ecuación de salarios sería imposible incluir variables como género o condición étnica. La inclusión de factores individuales fijos tiende a eliminar su influencia en el modelo a través del control por α_i. Es decir, podemos controlar por estos factores, pero no identificar su influencia específica en el fenómeno que se intenta analizar. Si es que el interés del investigador es analizar este tipo de variables claramente el estimador MEF no es el adecuado.

d) El estimador MEF esta diseñado para corregir el problema de variables omitidas eliminando la parte de la varianza que contamina la estimación por MCOC. Entonces si es que una variable esta medida con error, el estimador de MCO en corte transversal estará sesgado hacia cero. En el contexto del estimador MEF, el sesgo por variables omitidas se exacerba si es que la variable omitida está altamente correlacionada en el tiempo. Esto sugiere la necesidad de utilizar algún estimador por VI, con la dificultad implícita que este tipo de estimadores imponen (la capacidad de hallar instrumentos adecuados).

e) El uso de estructuras dinámicas impone serios problemas econométricos en los modelos MEF. Tome en cuenta la siguiente especificación

$$y_{it} = \alpha_i + \gamma y_{it-1} + \beta X_{it} + u_{it}$$

La transformación para calcular el estimador MEF es

$$(y_{it} - y_i) = \gamma(y_{it-1} - y_{i*}) + \beta(X_{it} - \overline{X}_i) + (u_{it} - \overline{u}_i)$$

donde y_{i*} es el promedio temporal pero a partir de $t-1$. Así, claramente y_{i*} estará correlacionada con $(u_{it} - \overline{u}_i)$ debido a que $\overline{u}_i$ contiene a u_{it-1} que esta obviamente correlacionado con y_{it-1}. Como resultado el parámetro γ será sesgado y para solucionar el sesgo deberá optarse por una estimación por VI.

5. Modelo MCO inter-grupo

Alternativamente al modelo intra-grupo que solo explota la variabilidad temporal de la data, es posible plantear un estimador que explote la variación transversal de la data. Para ello, considérese un modelo del tipo

$$y_{it} = \alpha_i + X_{it}^{'}\beta + u_{it} \qquad (17a)$$

donde $u_{it} \sim IID(0,\sigma^2)$, $i = 1,2,...,N$ y $t = 1,2,...,T$. Para lo cual se calculan los siguientes términos

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it} \qquad \bar{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it} \qquad \bar{u}_i = \frac{1}{T}\sum_{t=1}^{T} u_{it}$$

y

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} \bar{y}_i \qquad \bar{X} = \frac{1}{N}\sum_{i=1}^{N} \bar{X}_i \qquad \bar{u} = \frac{1}{N}\sum_{i=1}^{N} \bar{u}_i$$

donde $\bar{y}_i$, $\bar{X}_i$ y $\bar{u}_i$ representan las medias a través del tiempo mientras que $\bar{y}$, $\bar{X}$ y $\bar{u}$ son los promedios globales a través de toda la data. Se plantea así un modelo de la forma

$$\bar{y}_i = \alpha_i + \bar{X}_i^{'}\beta + \bar{u}_i \tag{17b}$$

si se le suma y resta el término α a esta expresión se tiene que

$$\bar{y}_i = \alpha + \bar{X}_i^{'}\beta + (\alpha_i - \alpha + \bar{u}_i) \tag{17c}$$

Con esta expresión, la estimación por MCO de β determina que

$$\hat{\beta}_{INTER} = \left[\sum_{i=1}^{N}(\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})^{'}\right]^{-1}\left[\sum_{i=1}^{N}(\bar{X}_i - \bar{X})(\bar{y}_i - \bar{y})\right] \tag{18}$$

Nótese que el $\hat{\beta}_{INTER}$ (MCO inter-grupo) no es más que el estimador MCO tradicional para el caso en que $N = 1$ y representa las desviaciones del promedio temporal de la variable analizada respecto a la media total. Así, el tipo de variabilidad que explota es simplemente la que ocurre entre los grupos y no a través de ellos en el tiempo. Este estimador es consistente sólo cuando las $\bar{X}_i$ no están correlacionadas con $(\alpha_i - \alpha + \bar{u}_i)$, lo que ocurre solo cuando no hay efectos no observables ($\alpha_i = 0$) o cuando estos se asumen como variables aleatorias no correlacionados con el error (por ejemplo, en el modelo de efectos aleatorios). En todo caso, bajo los supuestos de consistencia se pueden obtener matrices de varianza covarianza robustas siguiendo las recomendaciones del modelo MCOC.

6. Modelo de Efectos Aleatorios (MEA)

En el caso del estimador MEF los factores no observables se asumieron fijos a través del tiempo, pero variables entre grupos. Así, se definió que estos factores estarían correlacionados con las variables explicativas incluidas en el modelo. En el estimador MEA asumimos un modelo donde los factores no observables son tratados como aleatorios y no correlacionados con las variables explicativas incluidas en el modelo. De manera general se asume un modelo de la forma

$$y_{it} = \alpha_{it} + \beta x_{it} + u_{it} \tag{19a}$$

donde $u_{it} \sim IID(0,\sigma^2)$, $i=1,2,...,N$ y $t=1,2,...,T$ o de la misma forma

$$y_{it} = (\lambda + \theta_i + \delta_t) + \beta x_{it} + u_{it} \tag{19b}$$

que dado el supuesto de los factores no observables θ_i y δ_t, pueden ser incorporados al término de error. Así,

$$y_{it} = \lambda + \beta x_{it} + (u_{it} + \theta_i + \delta_t) \tag{19c}$$

o

$$y_{it} = \lambda + \beta x_{it} + v_{it} \tag{19d}$$

Los siguientes supuestos caracterizan este modelo

$$E(\alpha_{it}) = E(\theta_i) = E(\delta_t) = E(u_{it}) = E(v_{it}) = 0$$

$$E(\theta_i X_{it}) = E(\delta_t X_{it}) = E(u_{it} X_{it}) = 0$$

$$V(\theta_i) = \sigma_\theta^2 \;;\; V(\delta_t) = \sigma_\delta^2 \;;\; V(u_{it}) = \sigma_u^2$$

Nótese que se deja la posibilidad de que existan ambos tipos de efectos no observables (transversales y temporales), aunque es posible restringir el modelo para que $\lambda_t = 0$ y solo existan efectos transversales o $\theta_i = 0$ y solo existan efectos temporales. Para la discusión que sigue asumamos que $\lambda_t = 0$, por lo que solo hay factores transversales que no es un supuesto muy grueso en caso de paneles cortos (en todo caso podemos retomar una especificación de T variables dummy para incorporar los factores temporales).

En cada caso, la estimación requiere la implementación de MCG diferente al estimador MCOC encontrado anteriormente. Ello es así incluso a pesar que los efectos no observables no se encuentran correlacionados con las X_{it} debido a que esta independencia no es una condición suficiente para la aplicación de MCOC. Como se mencionó, la eficiencia del estimador MCOC se ve reducida ya que otorga igual peso a las dos fuentes de variabilidad de la data (transversal y temporal). El estimador MEA intenta corregir este problema a partir de un sistema de pesos que pondere de manera diferente cada una de las fuentes variabilidad dando un mayor a aquella que ofrezca mayor información.

El estimador por MCG cualquiera se define de la siguiente manera

$$\hat{B}_{MCG} = (x'\Omega^{-1}x)^{-1}(x'\Omega^{-1}y) \tag{20}$$

El objetivo es construir adecuadamente Ω^{-1}. Se propone para ello la expresión

$$'\Omega^{-1} = \frac{1}{\sigma_u^2}\left(I_{TxT} - \frac{\sigma_\theta^2}{\sigma_u^2 + T\sigma_\theta^2} ee' \right) \qquad (21)$$

Esta matriz de TxT pre-multiplica los vectores y_i , x_i y v_i . En forma matricial

$$y_i^* = \Omega^{-1}\begin{bmatrix} y_{i1} \\ y_{i2} \\ \dots \\ y_{iT} \end{bmatrix} \qquad x_i^* = \Omega^{-1}\begin{bmatrix} x_{i1} \\ x_{i2} \\ \dots \\ x_{iT} \end{bmatrix} \qquad v_i^* = \Omega^{-1}\begin{bmatrix} v_{i1} \\ v_{i2} \\ \dots \\ v_{iT} \end{bmatrix}$$

lo que da como resultado el modelo

$$y_{it}^* = \lambda + \beta x_{it}^* + v_{it}^* \qquad (22)$$

sobre el que se aplica MCO

$$\hat{\beta}_{MCG} = (x^{*'} x^*)^{-1}(x' y^*) = (x'\Omega^{-1}x)^{-1}(x'\Omega^{-1}y) = \hat{\beta}_{MEA} \qquad (23)$$

Nótese que si $\sigma_\theta^2 = 0$, el estimador $\hat{\beta}_{MEA}$ colapsa al estimador $\hat{\beta}_{MCOC}$. Dado que σ_θ^2 y σ_u^2 son desconocidos es necesario estimarlos por lo que utilizaremos una metodología de MCG factibles (hallamos un estimador $\hat{\Omega}^{-1}$). Para ello, utilizaremos la propiedad que el estimador MEA es una suma ponderada entre los estimadores intra-grupo e inter-grupo. Así, a partir de una estimación MEF (intra-grupo) obtenemos el valor de σ_u^2 (varianza del estimador intra-grupo) y si definimos a la expresión σ_B^2 como la varianza del estimador inter-grupo, a partir de

$$\sigma_\theta^2 = \sigma_B^2 - \frac{\sigma_u^2}{T} \qquad (24)$$

se obtiene σ_θ^2 y se tienen todos los términos necesarios para hallar $\hat{\Omega}^{-1}$. Justamente una interpretación conceptual del estimador MEA es que es una suma ponderada de los estimadores inter-grupo e intra-grupo. Se puede plantear

$$\hat{\beta}_{MEA} = \Theta_{kxk}\hat{\beta}_{INTER} + (I_{kxk} - \Theta_{kxk})\hat{\beta}_{MEF} \qquad (25)$$

donde k es el número de parámetros por estimar (el numero de variables incluidas en X_{it}). La matriz Θ_{kxk} será proporcional a la inversa de la matriz varianza-covarianza de cada estimador. De este modo, mientras más pequeña es la varianza de $\hat{\beta}_{INTER}$ (y mas preciso es el estimador) se le dará mas peso en el cálculo de $\hat{\beta}_{MEA}$ según (25). Una interpretación conceptual similar, pero para el caso de MCOC es que

$$\hat{\beta}_{MCOC} = \Phi_{kxk}\hat{\beta}_{INTER} + (I_{kxk} - \Phi_{kxk})\hat{\beta}_{MEF} \qquad (26)$$

donde la matriz Φ_{kxk} da igual peso a ambos estimadores, por lo que información no es aprovechada adecuadamente. Esta es la razón por la que el estimador MEA es más eficiente que MCOC.

Sin embargo, es importante notar que el estimador MEA es solo consistente bajo los supuestos del modelo. Es decir, los efectos no observables no están correlacionados con los variables explicativas y son efectivamente aleatorios. Caso contrario el estimador MEA es inconsistente y más bien debería plantearse una estimación tipo MEF. La decisión no es trivial debido a que los resultados que se obtienen siguiendo cada metodología pueden tener diferencias importantes y persistentes. De ese modo la cuestión por resolver es si debemos tratar a θ_i como fijo o como aleatorio. Se plantea la siguiente discusión

a) La estimación MEF es costosa en términos de grados de libertad, problema particularmente importante en caso de paneles cortos

b) Los θ_i caracterizan la ignorancia del investigador, por lo que cabe preguntarse ¿ por qué especificar como fijo un tipo de ignorancia θ_i, mientras que otro tipo de ignorancia u_{it} como aleatoria?

c) La especificación MEF permite que el estimador haga inferencias solo "condicional a los factores fijos" en la muestra. Mientras tanto la especificación MEA permite inferencias in-condicionadas con respecto a los efectos a todo tipo de efecto poblacional.

En resumen, la principal ventaja de la estimación MEF es que es relativamente sencilla de estimar y permite eliminar factores no observables correlacionados con otras variables incluidas, un fenómeno bastante común en economía. La principal desventaja es que requiere la estimación N interceptos diferentes. Esto causa problemas porque la variación que existe en la data es utilizada para realizar esta estimación lo que puede generar imprecisión en la estimación de los parámetros de interés. Este problema es más agudo cuando exista menor variabilidad temporal en la data. Además no permite identificar efectos que se mantienen fijos en el tiempo. Todas estas variables pueden ser importancia para el diseño de políticas públicas. Mientras tanto, la principal ventaja de la estimación MEA es que se pierden menos grados de liberad en la estimación. Sin embargo, la principal desventaja es el supuesto de efectos aleatorios no correlacionados con el resto de variables explicativas. Cuando no se cumpla este supuesto (como es de suponer en economía) el estimador es incosistente.

De este modo, las ventajas de un modelo son las desventajas de otro y estas ocurren a nivel de los supuestos de estimación. Por ello, es de interés para el investigador tener una prueba que le permita discriminar entre ambos estimadores. Para ello, se utiliza un test de Haussman que compara el estimador MEA con el MEF. En principio el test es utilizado para determinar cuando la heterogeneidad omitida puede ser tratada como fija y correlacionada con las explicativas y cuando como aleatoria y no correlacionada.

Como se sabe el test de Haussman comprara un estimador que sea consistente y eficiente bajo la nula con otro que es consistente pero menos eficiente bajo la nula e incosistente bajo la alternativa. Se plantea así la siguiente prueba

H_o : *heterogeneidad es aleatoria y no correlacionada con* X_{it} (*MEA es consistente y eficiente* , *MEF es consistente y menos eficiente*)
H_1 :*heterogeneidad es aleatoria y correlacionada con* X_{it} (*MEA es inconsistente* , *MEF es consistente y eficiente*)

y el estadístico

$$H = \frac{\left[\hat{B}_{MEA} - \hat{B}_{MEF}\right]^2}{V(\hat{B}_{MEA}) - V(\hat{B}_{MEF})} \quad \sim \quad \chi_k^2 \tag{27}$$

cuando se comparan k parámetros. Así, cuando

$$H > \chi_k^2 \rightarrow No\ acepto\ H_o \rightarrow MEF$$
$$H < \chi_k^2 \rightarrow No\ rechazo\ H_o \rightarrow MEA$$

Dos cuestiones acerca del test que deben notarse. Primero, que es un test relativo que indica si el MEF o MEA se aceptan en compración con el MEA o MEF planteado. Segundo, un no rechazo de la nula y por ende aceptar MEA puede deberse a una mala especificación del MEF o pobremente determinado por las razones ya mencionadas. En todo caso, el investigador siempre debe tener presente que la elección de un modelo sobre otro dependerá también del tipo de aplicación que se esta realizando y que tipo de parámetros interesa estudiar.