Datos de panel

Indice

1. Ventajas y limitaciones de la información de panel

Definición: con carácter general, llamaremos panel (datos longitudinales) a la información relativa a cortes transversales de unidades muestrales observadas a lo largo del tiempo. Las unidades maestrales podrán ser individuos, hogares, empresas, regiones, países, etc. Es decir, tenemos información para $i = 1, ..., N$ unidades muestrales, observadas a lo largo de $t = 1, ..., T$ períodos.

No obstante, podremos disponer de paneles de datos de otro tipo como por ejemplo el caso en el que observamos familias (i) y dentro de cada familia tenemos observaciones de hermanos gemelos ($t = 1, 2$). Otro caso podría ser la observación de trabajadores (i =1, …, N) en diferentes áreas regiones ($r = 1, ..., R$).

En todos los casos anteriores se ha considerado que el panel tiene dos características:

- dos dimensiones (en general, individual y temporal)
- es completo (se observan las mismas unidades muestrales a lo largo de los mismos períodos)

Sin embargo, esta no es la situación que siempre se observará y para ello, de nuevo es conveniente verlo mediante ejemplos. Supongamos que observamos trabajadores ($i = 1, ..., N$) , pertenecientes a diferentes regiones ($r = 1, ..., R$), que tienen distintas nacionalidades ($o = 1, ..., O$). O, alternativamente, observamos trabajadores ($i = 1, ..., N$) , pertenecientes a diferentes regiones ($r = 1, ..., R$), en distintos períodos de tiempo ($t = 1, ... T$). Este son casos de paneles que tienen tres dimensiones, que no tienen por qué haberse diseñado así, pero que se pueden construir de esta manera si es lo que se desea a la hora de plantear un análisis económico. (Se puede pensar de esta forma en la Encuesta Industrial o en la Encuesta de Flujos Laborales). Existe un panel europeo (PHOGUE) diseñado para representar los países de la EU-15, en el que podemos observar que sus dimensiones son individuos ($i = 1, ..., N$) u hogares ($h = 1, ..., H$), observados en diferentes momentos del tiempo ($t = 1, ... T$) en los distintos países de la UE-15 ($p = 1, ..., P$).

Por otra parte, podría darse el caso que algunas unidades muestrales fuesen observadas en unos períodos y otras unidades en otros o que tuviéramos diferente número de observaciones temporales para diferentes individuos o diferente número de individuos cada período. En estos casos, diríamos que tenemos un panel incompleto. Por ejemplo, la Encuesta Continua de Presupuestos Familiares (ECPF) se ha estado realizando trimestralmente en

España desde 1985 hasta 2005. Para evitar el sesgo por cansancio de los hogares (individuos) participantes se estableció una tasa de rotación del 12,5, de forma que un hogar (individuo) permanecía en la muestra un máximo de 8 trimestres. Dado que algunos hogares no ofrecen información durante todos los trimestres, tenemos la posibilidad de construir un panel incompleto así como un panel completo con diferentes individuos en diferentes trimestres.

El análisis (teórico y aplicado) de los datos de panel se ha constituido como uno de los campos más activos desde hace unos años por una serie de razones (ejemplo el Congreso Mundial bianual que se celebra este año en la University of Cambridge y en 2008 en la Universidad de Alicante).

1. Teóricas
2. Prácticas: Existen estudios que no se pueden llevar a cabo si no se dispone de este tipo de datos. Ejemplo 1: Consumo - Ciclo vital. Ejemplo 2. Financiación - restriciones de liquidez
3. Disponibilidad de datos
4. Disponibilidad de buenos ordenadores y programas (STATA, GAUSS, GAUSSX, LIMDEP, DPD, etc.)

Disponer de información de panel genera numerosas ventajas como:

1. Número de observaciones (normalmente N grande T fijo). Si es al revés *data field***.** Esto proporciona más grados de libertad y **más eficiencia** de los resultados.
2. **Reducción multicolinealidad**. Ejemplo 1: Ecuaciones de demanda con datos de series temporales o de corte transversal. Ejemplo 2: Modelos macroeconómicos dinámicos.
3. **Control de los efectos específicos invariantes en el tiempo** (ventaja respecto a cortes transversales y series temporales). **Control de los efectos temporales invariantes longitudinalmente** (ventaja respecto a cortes transversales).
4. Posibilidad de **ajustar modelos dinámicos (**ventaja respecto a cortes transversales) y plantear modelos más realistas que en cortes transversales y series temporales.

5. Posibilidad **de identificar efectos que son difícilmente detectables** en cortes transversales o series temporales.

6. **Evitar los sesgos de agregación** (ventaja respecto a series temporales)

La información de panel no está, sin embargo, exenta de algunas limitaciones como:

1. **Dificultades en el tratamiento de los modelos** (dinámica, dependencia de corte transversal).

2. Problemas en el **diseño y recolección de la información** que pueden generar sesgos de colaboración como:

    a. **Sesgo de no respuesta** en las encuestas. Ejemplo: ESEE. Solución: datos de cohorte (no válidos para determinadas aplicaciones). Ejemplo 1: Encuesta Industrial. Ejemplo 2: ECPF.

    b. **Sesgos de autoselección**. En la ECPF acaban estando sobre-representados los jubilados e infla-representados los hogares de rentas altas.

    c. **Sesgos** (attrition). Es decir, son sesgos de no respuesta pero correspondientes a individuos que pueden desaparecer de la muestra. Por ejemplo, pueden morir o migrar.

3. Errores de medida, que suelen ser más importantes con datos de panel porque se acumulan.

4. La **dimensión temporal** de los paneles de datos suele ser **reducida**.

La principal ventaja de los datos de panel es la siguiente:

Supongamos:

$$Y_i = \alpha + \beta X_i + u_i \qquad (1)$$

$$E(X_i\, u_i) \neq 0 \qquad (2)$$

$$E(\beta) = \beta + E(X_i\, u_i)/Var(X_i) \quad (3)$$

Imposible estimar consistentemente β si disponemos de una sola observación para cada individuo. Supongamos que tenemos $T = 2$.

$$Y_{i1} = \alpha + \beta X_{i1} + u_{i1} \quad (4)$$

$$Y_{i2} = \alpha + \beta X_{i2} + u_{i2} \quad (5)$$

Si $E(u_{i1}\, u_{i2}) = 0$ se puede estimar consistentemente β en (5) instrumentando X_{i2} con X_{i1} (generalmente buen instrumento.

De otra forma: *Si* $u_{it} = \eta_y + v_{it}$ y $E(v_{it}\, X_{it}) = 0$, se puede mantener la correlación con del regresor con el error ($E(\eta_i\, X_{it}) \neq 0$) y estimar β en la siguiente ecuación:

$$Y_{i2} - Y_{i1} = \beta\,(X_{i2} - X_{i1}) + (u_{i2} - u_{i1}) \quad (6)$$

Ejemplos en los que la disponibilidad de datos de panel es fundamental para la estimación de modelos econométricos

Ejemplo 1. Funciones de producción y habilidad de los managers

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2i} + u_{it} \quad (7)$$

$i = 1, \ldots, N$ observaciones de empresas (supongamos que son explotaciones agrarias)
$t = 1, \ldots, T$ son períodos de tiempo (supongamos que tenemos datos anuales)
Y_{it} es el logaritmo del output de la empresa i en el momento t
X_{1it} es el logaritmo del trabajo (número de trabajadores, por ejemplo) de la empresa i en el momento t
X_{2i} es el logaritmo de la habilidad del manager de la empresa i, no es observable y no cambia en el tiempo
u_{it} es un término de error con $E(u_{it}) = 0$, por ejemplo, puede ser el tiempo atmosférico.

Supongamos que solamente tenemos información de un período (corte transversal) y pretendemos estimar la relación (7), es decir,

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$$

con

$$\varepsilon_i = \beta_2 X_{2i} + u_i$$

Por tanto,

$$\mathrm{E}(Y_i/X_{1i}) = \beta_0 + \beta_1 X_{1i} + E(\varepsilon_i/X_{1i})$$
$$\mathrm{E}(Y_i/X_{1i}) = \beta_0 + \beta_1 X_{1i} + \beta_2 E(X_{2i}/X_{1i})$$

Supongamos que:

$$E(X_{2i}/X_{1i}) = \lambda_1 + \lambda_2 X_{1i}$$

Y como consecuencia

$$\mathrm{E}(Y_i/X_{1i}) = (\beta_0 + \beta_2 \lambda_1) + (\beta_1 + \beta_2 \lambda_2) X_{1i} \qquad (8)$$

Por tanto, si estimamos (8), obtenemos un valor estimado para $\beta_1 + \beta_2 \lambda_2$ cuya esperanza no es es β_2. De esta forma obtenemos un estimador no consistente de la productividad del trabajo, que es la relación causal dada por (7).

- si la habilidad está positivamente correlacionada con la productividad ($\lambda_2 > 0$), entonces obtenemos una sobre-estimación del efecto de la productividad del trabajo.
- si la habilidad está negativamente correlacionada con la productividad ($\lambda_2 < 0$), entonces obtenemos una infla-estimación del efecto de la productividad del trabajo.

Ejemplo 2. Rendimientos de la educación y habilidad

Supongamos la siguiente relación

$$Y_{ij} = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2j} + \beta_3 X_{3j} + u_{ij} \qquad (9)$$

$i = 1, 2$ son observaciones de gemelos
$j = 1, ..., N$ son observaciones de familias a las que pertenecen los gemelos
Y_{it} es el logaritmo del salario del individuo i perteneciente a la familia j
X_{1ij} son los años de escolarización del individuo i de la familia j
X_{2j} es el logaritmo de la renta familiar de la familia j
X_{3j} es la habilidad de los individuos de la familia i por razones genéticas y culturales
u_{ij} es un término de error con $E(u_{ij}) = 0$

Supongamos que solamente disponemos de información de un corte transversal, es decir, no disponemos de datos familiares. De esta forma, el modelo que podemos estimar es:

$$Y_{ij} = \beta_0 + \beta_1 X_{1ij} + \varepsilon_{ij} \qquad (10)$$

donde $\varepsilon_{ij} = \beta_2 X_{2j} + \beta_3 X_{3j} + u_{ij}$

Si estimamos obtenemos un valor para β_1 que en media estará sesgado en la siguiente cuantía: $E(\beta_1) = \beta_1 + \beta_2 \lambda_2 + \beta_3 \delta_2$ expresión en la que hemos supuesto que

$$E(X_{2j}/X_{1ij}) = \lambda_1 + \lambda_2 X_{1ij}$$
$$E(X_{3j}/X_{1ij}) = \delta_1 + \delta_2 X_{1ij}$$

De manera que si disponemos de la renta de la familia podemos eliminar el sesgo $\beta_2 \lambda_2$ pero no podemos controlar la habilidad (diferencias genéticas y culturales – *nature and nurture* –). Para controlar estas diferencias es necesario de disponer de datos de la familia (y suponer que son constantes por familia, es decir, que los gemelos de la familia *j* tienen las mismas habilidades (porque no tienen diferencias genéticas y culturales).

2. Modelos clásicos de efectos fijos y/o aleatorios. Estimadores y contrastes

El modelo básico que se suele utilizar cuando se dispone de datos de panel en la situación clásica (T fijo, $N \to \infty$) es el siguiente:

$$Y_{it} = \beta' X_{it} + \eta_i + u_{it} \qquad (11)$$

k regresores X (no hay constante), η_i son los efectos individuales (invariantes en el tiempo). Panel: Tenemos N observaciones observadas a lo largo de T períodos (caso del panel **completo**). Todo es válido para paneles **incompletos**. (De momento obviamos los efectos temporales).

Cabría preguntarse en primer lugar si tiene sentido el modelo presentado en (11) frente al modelo sin constantes específicas salvo con una constante común como en (12):

$$Y_{it} = \beta' X_{it} + \eta + u_{it} \qquad (12)$$

Respuesta: Test de homogeneidad (F entre (11) y (12)).

Restricciones: $\eta_i = \eta$ para todo $i \Rightarrow$ MCO son consistentes y eficientes si se cumplen las hipótesis clásicas.

Si no rechazamos (12) el análisis econométrico se simplifica.

En segundo lugar podríamos preguntarnos si es suficiente el modelo (11) para recoger la heterogeneidad o es mejor plantear un modelo menos restrictivo como:

$$Y_{it} = \beta_i' X_{it} + \eta_i + u_{it} \qquad (13)$$

Respuesta: Test de homogeneidad (F entre (13) y (12)).

Restricciones: $\eta_i = \eta$ para todo i y $\beta_i = \beta$ para todo i.

Por simplicidad (y realismo) desarrollamos toda la exposición con el modelo (11). (Si no rechazamos (12) el análisis econométrico se simplifica).

Ejemplo (tomado de Greene (1983) "*Simultaneous Estimation of Factor Substitution, Economies of Scale and Non-Neutral Technical Change*" en A. Dogramaci (ed.) **Econometric Analyses of Productivity**, Kluwer-Nijoff, Boston)

Supongamos que se ajusta una función de coste (medido por los inputs consumidos que son fuel, trabajo y capital en millones de dólares) de producción de electricidad (output en millones de Kw/hora). Muestra $N = 6$, $T = 4$. Supongamos que se ajusta el modelo (12) a la siguiente ecuación:

$$log\ C_{it} = \beta\ log\ Y_{it} + \eta + u_{it} \qquad (14)$$

MCO produce los siguientes resultados:

$$log\ C_{it} = -4,17478 + 0,887987\ log\ Y_{it}$$
$$(0,2769) \quad (0,0329)$$

$$R^2 = 0,97 \qquad s^2 = 0,046$$

Modelo de efectos fijos:

Las 6 empresas parece que muestran diferencias en costes y producción que no somos capaces de captarlas por medio de la estimación de (14). Formulación habitual: las únicas diferencias se producen en la constante (trabajan con la misma productividad) y dichas diferencias se mantienen

constantes a lo largo de los 4 años (podemos achacar estas diferencias a la capacidad de gestión de los directores de estas empresas que permanecen constantes a lo largo del tiempo).

$$log\ C_{it} = \beta\, log\ Y_{it} + \eta_i + u_{it} \tag{15}$$

o en notación vectorial:

$$log\ C_i = \beta\, log\ Y_i + i\eta_i + u_i \tag{16}$$

$$
\begin{bmatrix} \log C_{11} \\ \log C_{12} \\ \log C_{13} \\ \log C_{14} \\ \log C_{21} \\ \log C_{22} \\ \log C_{23} \\ \log C_{24} \\ \log C_{31} \\ \log C_{32} \\ \log C_{33} \\ \log C_{34} \\ \log C_{41} \\ \log C_{42} \\ \log C_{43} \\ \log C_{44} \\ \log C_{51} \\ \log C_{52} \\ \log C_{53} \\ \log C_{54} \\ \log C_{61} \\ \log C_{62} \\ \log C_{63} \\ \log C_{64} \end{bmatrix} = \begin{bmatrix} 1,0,0,0,0,0 \\ 1,0,0,0,0,0 \\ 1,0,0,0,0,0 \\ 1,0,0,0,0,0 \\ 0,1,0,0,0,0 \\ 0,1,0,0,0,0 \\ 0,1,0,0,0,0 \\ 0,1,0,0,0,0 \\ 0,0,1,0,0,0 \\ 0,0,1,0,0,0 \\ 0,0,1,0,0,0 \\ 0,0,1,0,0,0 \\ 0,0,0,1,0,0 \\ 0,0,0,1,0,0 \\ 0,0,0,1,0,0 \\ 0,0,0,1,0,0 \\ 0,0,0,0,1,0 \\ 0,0,0,0,1,0 \\ 0,0,0,0,1,0 \\ 0,0,0,0,1,0 \\ 0,0,0,0,0,1 \\ 0,0,0,0,0,1 \\ 0,0,0,0,0,1 \\ 0,0,0,0,0,1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} + \begin{bmatrix} \log Y_{11} \\ \log Y_{12} \\ \log Y_{13} \\ \log Y_{14} \\ \log Y_{21} \\ \log Y_{22} \\ \log Y_{23} \\ \log Y_{24} \\ \log Y_{31} \\ \log Y_{32} \\ \log Y_{33} \\ \log Y_{34} \\ \log Y_{41} \\ \log Y_{42} \\ \log Y_{43} \\ \log Y_{44} \\ \log Y_{51} \\ \log Y_{52} \\ \log Y_{53} \\ \log Y_{54} \\ \log Y_{61} \\ \log Y_{62} \\ \log Y_{63} \\ \log Y_{64} \end{bmatrix} \beta + \begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \\ u_{14} \\ u_{21} \\ u_{22} \\ u_{23} \\ u_{24} \\ u_{31} \\ u_{32} \\ u_{33} \\ u_{34} \\ u_{41} \\ u_{42} \\ u_{43} \\ u_{44} \\ u_{51} \\ u_{52} \\ u_{53} \\ u_{54} \\ u_{61} \\ u_{62} \\ u_{63} \\ u_{64} \end{bmatrix} \tag{17}
$$

En notación matricial:

$$y = [d_1, d_2, d_3, d_4, d_5, d_6, X][\eta, \beta]' + u \tag{18}$$

donde d_i $i = 1, ..., 6$ son variables ficticias correspondientes a cada empresa. Si llamamos $D = [d_1, d_2, d_3, d_4, d_5, d_6]$, el modelo se puede expresar:

$$y = X\beta + D\ \eta + u \tag{19}$$

el estimador de MCO aplicado a (19) se denomina estimador Mínimo Cuadrático de variables ficticias (MCVF). **Los η son los efectos fijos**. Si N no es muy grande, MCO aplicado a (19) es muy sencillo y las propiedades son las mismas que MCO aplicados a (12). El problema es que si N es grande, la estimación de (19) implica la estimacion de $N + k$ parámetros (en nuestro ejemplo 7). (**Nota**. STATA permite estimar dicho modelo aunque N sea varios miles). Sin embargo, una simple transformación permite estimar β sin tener que estimar simultáneamente los η_i. Si aplicamos la siguiente matriz de transformación al modelo:

$$Q_d = \begin{bmatrix} Q,0,...,0 \\ 0,Q,...,0 \\ .,.,\cdots, \\ .,.,\cdots, \\ .,.,\cdots, \\ Q,0,...,Q \end{bmatrix} \tag{20}$$

$$Q = I_T - (1/T)ii` \tag{21}$$

En nuestro ejemplo,

$$Q_d = \begin{bmatrix} \begin{bmatrix}1,0,0,0\\0,1,0,0\\0,0,1,0\\0,0,0,1\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix} \\ \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}1,0,0,0\\0,1,0,0\\0,0,1,0\\0,0,0,1\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix} \\ \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}1,0,0,0\\0,1,0,0\\0,0,1,0\\0,0,0,1\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix} \\ \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}1,0,0,0\\0,1,0,0\\0,0,1,0\\0,0,0,1\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix} \\ \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}1,0,0,0\\0,1,0,0\\0,0,1,0\\0,0,0,1\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix} \\ \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}0,0,0,0\\0,0,0,0\\0,0,0,0\\0,0,0,0\end{bmatrix}, \begin{bmatrix}1,0,0,0\\0,1,0,0\\0,0,1,0\\0,0,0,1\end{bmatrix} \end{bmatrix}$$

(22)

$$-\frac{1}{4}\begin{bmatrix}
\begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} \\
\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} \\
\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} \\
\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} \\
\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} \\
\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} \\
\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} & \begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix}
\end{bmatrix}$$

$$Q_d Y = Q_d X \beta + Q_d D \eta + Q_d u \tag{23}$$

$$Q_d D = 0 \tag{24}$$

$$\hat{\beta} = \left[X' Q_d X\right]^{-1} \left[X' Q_d Y\right] \tag{25}$$

La ecuación (23) supone aplicar MCO en una regresión de X^d_{it} sobre Y^d_{it} donde $X^d_{it} = X_{it} - X_{im}$, $Y^d_{it} = Y_{it} - Y_{im}$ siendo $Y_{im} = (1/T)\sum_t Y_{it}$ y $X_{im} = (1/T)\sum_t X_{it}$.

Y los efectos fijos se pueden estimar como:

$$\hat{\eta} = \left[D'D\right]^{-1} D'(Y - X\hat{\beta}) \tag{26}$$

$$\hat{\eta}_i = \ddot{Y} - \hat{\beta}\ddot{X} \tag{27}$$

Las varianzas de η y β las podemos estimar como:

$$Var(\hat{\beta}) = \hat{\sigma}^2 \left[X' Q_d X\right]^{-1} \tag{28}$$

$$Var(\hat{\eta}_i) = \frac{\hat{\sigma}^2}{T} + \overline{X}_i' Var(\hat{\beta}) \overline{X}_i \tag{29}$$

$$\hat{\sigma}^2 = \frac{\sum_i \sum_t (Y_{it} - \hat{\eta}_i - \hat{\beta}'' X_{it})}{NT - N - k} \tag{30}$$

Para saber cual de los dos modelos es mejor (MCO o MCVF) podemos realizar el test citado de homogeneidad:

$$F(N-1, NT-N-k) = \frac{(R^2_{NR} - R^2_R)/(N-1)}{(1-R^2_{NR})/(NT-N-k)} \qquad (31)$$

En el ejemplo que estamos siguiendo, introduciendo 6 variables ficticias en la ecuación de costes, los resultados son:

$$\log C_{it} = \underset{(0,06113)}{0,674279 \log Y_{it}} - \underset{(0,3828)}{2,69353\, d_1} - \underset{(0,4396)}{2,9117\, d_2} -$$

$$- \underset{(0,5287)}{2,4400\, d_3} - \underset{(0,5588)}{2,13449\, d_4} - \underset{(0,5532)}{2,3108\, d_5} - \underset{(0,6081)}{1,9035\, d_6}$$

$$R^2 = 0,9924$$

$$s^2 = 0,0155 \qquad F(5, 17) = 9,708,\ F(0,01) = 4,34$$

Incluir efectos temporales en el modelo es sencillo porque lo que se suele hacer es incluir variables ficticias.

Modelo de efectos aleatorios

Si las diferencias entre las unidades muestrales (individuos, empresas, etc) no se deben sólo a cambios en la constante, si el número de unidades muestrales es muy elevado o si creemos que la heterogeneidad no observable es puramente aleatoria, el modelo presentado en (11) no sería conveniente. En el modelo de efectos aleatorios o de **componentes del error** (o errores compuestos) se asume que los η_i son extracciones independientes de una variable aleatoria. De esta manera la especificación (11) se convierte en:

$$Y_{it} = \alpha + \beta' X_{it} + \eta_i + u_{it} \qquad (32)$$

donde como antes tenemos k regresores además de la constante. Se asume que el error $v_{it} = \eta_i + u_{it}$ está formado por la parte propiamente individual y un término mixto (componentes del error). Para dichos componentes se supone, además:

$$
\begin{aligned}
&E(u_{it}) = E(\eta_i) = 0 \\
&E(u^2_{it}) = \sigma^2_u,\ E(\eta^2_i) = \sigma^2_\eta \\
&E(u_{it}\eta_j) = 0 \text{ para todos } i,\ t \text{ y } j \\
&E(u_{it}u_{js}) = 0 \text{ si } t \neq s \text{ e } i \neq j \\
&E(\eta_i\eta_j) = 0 \text{ si } i \neq j
\end{aligned}
\tag{33}
$$

Por tanto,

$$
\begin{aligned}
&E(v_{it}) = \sigma^2_u + \sigma^2_\eta \\
&E(v_{it}v_{is}) = \sigma^2_\eta
\end{aligned}
\tag{34}
$$

Implicación importante: A pesar de las hipótesis (33), el modelo presenta correlación entre las perturbaciones de diferentes períodos (para el mismo individuo). En este contexto, lo que sabemos del análisis clásico de regresión es que MCO proporciona estimadores consistentes pero MCG proporciona estimadores consistentes y eficientes. (En contexto de panel, el análisis se complica un poco).

Con esta información, sabemos que la matriz de varianzas - covarianzas para las T observaciones del individuo i toma la forma:

$$\Omega=\begin{bmatrix} \sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\ldots,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\ldots,\sigma_\eta^2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sigma_\eta^2,\sigma_\eta^2,\ldots,\sigma_u^2+\sigma_\eta^2 \end{bmatrix}=\sigma_u^2 I_T+\sigma_\eta^2 ii' \tag{35}$$

Y puesto en términos del ejemplo E(v’v) será:

$$V=\begin{bmatrix}
\begin{bmatrix} \sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix} \\
\begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix} \\
\begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix} \\
\begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix} \\
\begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix} \\
\begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \\ 0,0,0,0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2,\sigma_\eta^2 \\ \sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_\eta^2,\sigma_u^2+\sigma_\eta^2 \end{bmatrix}
\end{bmatrix} \tag{36}$$

MCG sólamente consiste en transformar (32) con la raíz cuadrada del inverso de $V = I_{NT} \otimes \Omega$ para conseguir que las nuevas perturbaciones cumplan todas las hipótesis. Después de esto se pueden aplicar MCO al modelo transformado.

$$
\begin{aligned}
&V^{-1/2} = I_{NT} \otimes \Omega^{-1/2} \\
&\Omega^{-1/2} = I_T - (\theta/T)\, ii' \\
&\theta = 1 - (\sigma^2_u / (T\sigma^2_\eta + \sigma^2_w)^{1/2})
\end{aligned}
\qquad (37)
$$

¿Cómo actúa esta transformación? Para las T observaciones de un individuo i cualquiera (continuemos con el ejemplo)

$$
\Omega^{-1/2} Y_i = \left[\begin{bmatrix} 1,0,0,0 \\ 0,1,0,0 \\ 0,0,1,0 \\ 0,0,0,1 \end{bmatrix} - \frac{\theta}{4} \begin{bmatrix} 1,1,1,1 \\ 1,1,1,1 \\ 1,1,1,1 \\ 1,1,1,1 \end{bmatrix} \right] \begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \end{bmatrix} = \begin{bmatrix} Y_{i1} - \theta \overline{Y}_{i.} \\ Y_{i2} - \theta \overline{Y}_{i.} \\ Y_{i3} - \theta \overline{Y}_{i.} \\ Y_{i4} - \theta \overline{Y}_{i.} \end{bmatrix} \qquad (38)
$$

y, consecuentemente,

$$
\hat{\beta}_{MCG} = (X'V^{-1}X)X'V^{-1}Y = (X'(I \otimes \Omega^{-1})X)X'(I \otimes \Omega^{-1})Y \qquad (39)
$$

Otra forma de derivar el estimador es la siguiente: Los parámetros del modelo (32) pueden ser estimados (bajo los supuestos (33)) consistentemente en cada una de las tres siguientes regresiones:

MCO sobre: $Y_{it} = \alpha + \beta' X_{it} + \eta_i + u_{it}$ (40)

MCO sobre: $Y_{it} - \overline{Y}_{i.} = \beta'(X_{it} - \overline{X}_{i.}) + u_{it} - \overline{u}_{i.}$ (41)

MCO sobre: $\overline{Y}_{i.} = \alpha + \beta' \overline{X}_{i.} + \overline{\eta}_i + \overline{u}_{i.}$ (42)

Si llamamos variación total en *X*, *Y* y covariación en *XY* a:

$$S_{XX}^t = \sum_i \sum_t (X_{it} - \overline{X})(X_{it} - \overline{X})' \quad (43)$$

$$S_{YY}^t = \sum_i \sum_t (Y_{it} - \overline{Y})(Y_{it} - \overline{Y}) \quad (44)$$

$$S_{XX}^t = \sum_i \sum_t (X_{it} - \overline{X})(Y_{it} - \overline{Y}) \quad (45)$$

Variación intragrupos:

$$S_{XX}^W = \sum_i \sum_t (X_{it} - \overline{X}_{i.})(X_{it} - \overline{X}_{i.})' \quad (46)$$

$$S_{YY}^W = \sum_i \sum_t (Y_{it} - \overline{Y}_{i.})(Y_{it} - \overline{Y}_{i.}) \quad (47)$$

$$S_{XX}^W = \sum_i \sum_t (X_{it} - \overline{X}_{i.})(Y_{it} - \overline{Y}_{i.}) \quad (48)$$

Variación entre grupos:

$$S_{XX}^B = \sum_i T(\overline{X}_{i.} - \overline{X})(\overline{X}_{i.} - \overline{X})' \quad (49)$$

$$S_{YY}^{B} = \sum_{i} T(\overline{Y}_{i.} - \overline{Y})(\overline{Y}_{i.} - \overline{Y}) \tag{50}$$

$$S_{XX}^{B} = \sum_{i} (\overline{X}_{i.} - \overline{X})(\overline{Y}_{i.} - \overline{Y}) \tag{51}$$

es muy fácil comprobar que:

$$\begin{aligned} S^{t}_{xx} &= S^{B}_{xx} + S^{W}_{xx} \\ S^{t}_{xx} &= S^{B}_{xx} + S^{W}_{xx} \\ S^{t}_{xx} &= S^{B}_{xx} + S^{W}_{xx} \end{aligned} \tag{52}$$

y los tres estimadores de (40), (41) y (42)

$$\begin{aligned} \hat{\beta}_{MCO} &= (S^{t}_{XX})^{-1} S^{t}_{XY} = (S^{W}_{XX} + S^{B}_{XX})^{-1}(S^{W}_{XY} + S^{B}_{XY}) \\ \hat{\beta}_{MCVF} &= \hat{\beta}_{W} = (S^{W}_{XX})^{-1} S^{W}_{XY} \\ \hat{\beta}_{B} &= (S^{B}_{XX})^{-1} S^{B}_{XY} \end{aligned} \tag{53}$$

MCO no es más que una media ponderada de los otros dos:

Como:

$$S^{W}_{XY} = S^{W}_{XX}\hat{\beta}_{W} \qquad y \qquad S^{B}_{XY} = S^{B}_{XX}\hat{\beta}_{B} \tag{54}$$

$$\hat{\beta}_{MCO} = (S^{W}_{XX} + S^{B}_{XX})^{-1}(S^{W}_{XY} + S^{B}_{XY}) = F^{W}\hat{\beta}_{W} + F^{B}\hat{\beta}_{B} \tag{55}$$

$$F^W = (S_{XX}^W + S_{XX}^B)^{-1} S_{XX}^W \qquad F^B = I - F^W \tag{56}$$

De la misma manera, el estimador MCG es una media ponderada de los estimadores intra y entregrupos:

$$\hat{\beta}_{MCG} = \hat{F}^W \hat{\beta}_W + \hat{F}^B \hat{\beta}_B \tag{57}$$

donde ahora:

$$\hat{F}^W = (S_{XX}^W + \lambda\ S_{XX}^B)^{-1} S_{XX}^W \qquad \hat{F}^B = I - F^W \tag{58}$$

$$\lambda = \frac{\sigma_u^2}{\sigma_u^2 + T\sigma_\eta^2} = (1-\theta)^2 \tag{59}$$

Casos que se pueden dar:

$\lambda = 1 \Rightarrow$ MCG = MCO

$\lambda = 1$ si $\sigma^2_\eta = 0$, en cuyo caso, la variación intragrupos no aporta nada.

$\lambda = 0 \Rightarrow$ MCG = MCVF

$\lambda = 0$ si $\sigma^2_u = 0$ en cuyo caso toda la variación es intragrupos o si $T \to \infty$ (en esta situación los efectos no observables es como si fueran observables)

Cualquier caso intermedio, mejora la eficiencia de las estimaciones MCO y MCVF porque pondera adecuadamente todas las fuentes de variación.

El problema para la aplicación de MCG (ecuación (39)) es que desconocemos Ω, es decir, desconocemos las varianzas de los componentes del error y por tanto hemos de estimarlas. En la literatura se han propuesto varios estimadores, y en general los paquetes estadístico econométricos trabajan con los siguientes:

$$\hat{\sigma}_u^2 = \frac{\sum_i \sum_t (\hat{u}_{it} - \hat{\bar{u}}_{i.})^2}{NT - N - k} \tag{60}$$

$$\hat{\sigma}_\eta^2 = \frac{\sum_t (\hat{\bar{u}}_{i.} + \eta_i)^2}{N - k} - \frac{\hat{\sigma}_u^2}{T} \tag{61}$$

Por tanto, un procedimiento estándar para la estimación por MCG del modelo de efectos aleatorios (es que utiliza STATA) consiste en:

- utilizar una especificación de efectos fijos (intra-grupos) para estimar los parámetros (proporciona estimadores consistentes pero no eficientes)
- con los residuos del modelo de efectos fijos estimar σ_u según la expresión (60)
- de forma similar utilizar la especificación entre-grupos para estimar los parámetros
- utilizar los residuos y la estimación de σ_u para obtener una estimación de (61)
- con expresiones para estos dos parámetros se puede tener una estimación para θ que permita transformar el modelo y obtener el estimador de MCG de acuerdo a la expresión (39)

Problema: En la expresión (61) no está garantizado que la varianza sea positiva

¿Qué modelo elegir? ¿Efectos fijos o efectos aleatorios?

Esta es una decisión a tomar en base a dos informaciones:

1. Datos disponibles
2. Contrastes

Datos disponibles: Ejemplo 1. Supongamos que estamos interesados en una ecuación de crecimiento (Barro y Salas (1991)). Supongamos que disponemos de la información de 24 países de la OCDE para estimarla y contrastar cuestiones acerca de la convergencia entre ellos. Elegiríamos un modelo de efectos fijos porque estamos haciendo inferencia condicional a la muestra (muestra y población son en este caso indistinguibles).

Ejemplo 2. Supongamos que estamos interesados en estimar una ecuación de demanda de alimentos con datos de 6000 hogares observados durante 8 trimestres (ECPF). Efectos aleatorios porque queremos hacer inferencia incondicional, es decir, queremos que de los resultados obtenidos en la muestra se pueda inferir el comportamiento en cuanto a demanda de alimentos de los hogares españoles.

Ejemplo 3. Supongamos que estamos interesados en estimar una ecuación de número de innovaciones y disponemos de los datos de 1200 empresas observadas a lo largo de 5 años (ESEE). Efectos aleatorios porque queremos inferir de la muestra el comportamiento de las empresas industriales españolas en el tema de innovaciones tecnológicas.

Contrastes. Efectos fijos vs. efectos aleatorios se puede contrastar por medio de un test de Hausman:

H_0: No correlación efectos - variables (efectos aleatorios)

H_1: Correlación efectos variables (efectos fijos)

Bajo H_0 β_{MCG} y β_{MCVF} (y también β_{MCO}) son consistentes pero β_{MCG} es eficiente. Bajo H_1 β_{MCG} es inconsistente y β_{MCVF} (o β_{MCO}) son consistentes.

Test:

$$(\hat{\beta}_{MCVF} - \hat{\beta}_{MCG})'\left[Var(\hat{\beta}_{MCVF}) - Var(\hat{\beta}_{MCG})\right]^{-1}(\hat{\beta}_{MCVF} - \hat{\beta}_{MCG}) \tag{62}$$

que se distribuye como una χ^2 con k grados de libertad (número de parámetros excluída la constante).

Mi opinión: No es relevante si los efectos son fijos o son aleatorios sino si están o no correlacionados con las variables. Esta opinión está basada en lo que Mundlak sugería en 1978 acerca de que la distinción entre ambos es arbitraria e innecesaria. Por las siguientes razones:

1. Si el modelo está bien especificado, el estimador de efectos aleatorios es idéntico al estimador de efectos fijos.

2. Siempre podemos encontrar estimadores consistentes independientemente de como sean los efectos (porque en cualquier transformación que los elimine no tenemos que hacer referencia a ellos)

3. El test de hecho contrasta correlación entre efectos y variables y en caso de rechazar la H_0 hemos de buscar dicha transformación.

3. En general no tendrá mucho sentido un modelo en el que no haya correlación entre efectos y variables en economía. Si el modelo es estático siempre podemos obtener estimadores consistentes (β_{MCVF}). Si el modelo es dinámico, la correlación siempre está presente y se hace necesario buscar dicha transformación y emplear otros métodos para la estimación.

Ejemplo de Mundlak (1978), *Econometrica* o Chamberlain (1980) *Review of Economic Studies*

Reconciliación de los modelos de efectos fijos y aleatorios.

Supongamos el modelo básico

$$Y_{it} = \beta_i' X_{it} + \eta_i + u_{it} \qquad (63)$$

En el que sospechamos que X está correlacionado con η. Supongamos que el esquema de correlación es tan sencillo como el siguiente:

$$\eta_i = = \bar{X}_{i.}\delta + \nu_i \qquad (64)$$

y, además, supongamos que:

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

Si δ no es igual a *0*, la $E(\eta_i/X) = \bar{X}_{i.}\delta \neq 0$

Por lo que el estimador MCG aplicado a (63) producirá estimadores inconsistentes. Sin embargo podemos introducir (64) en (63) y tendremos el modelo

$$Y_{it} = \beta_i' X_{it} + \bar{X}_{i.}\delta + \nu_i + u_{it} \tag{65}$$

que será un modelo de efectos aleatorios bien especificado en el que ahora el error compuesto no está correlacionado con los regresores.

Si escribimos (65) de la siguiente forma:

$$Y_{it} = (X_{it} - \bar{X}_{i.})\beta + \bar{X}_{i.}(\beta + \delta) + \nu_i + u_{it} \tag{66}$$

$$Y_{it} = (X_{it} - \bar{X}_{i.})\phi + \bar{X}_{i.}\varphi + \nu_i + u_{it} \tag{67}$$

Estimando los parámetros de (67) tenemos

$$\hat{\phi}_{MCG} = \hat{\beta}_{WG} \Longleftrightarrow E(\hat{\beta}_{WG}) = E(\hat{\phi}_{MCG}) = \beta \tag{68}$$

$$\hat{\varphi}_{MCG} = \hat{\beta}_{BE} \Longleftrightarrow E(\hat{\beta}_{BE}) = E(\hat{\varphi}_{MCG}) = \beta' + \delta$$

por lo que una especificación correcta del modelo muestra

- que el estimador de efectos fijos y el estimador de efectos aleatorios coinciden y son consistentes.

- el estimador between es insesgado solo si no existe correlación entre regresores y efectos

3. Efectos aleatorios y regresores no exógenos