<u>Nota de Clase 8</u>
Introducción a modelos de Data Panel: Aplicación

Se dispone de la base de datos "ejemplo_panel.dta" que contiene data sobre el logaritmo del ingreso mínimo subjetivo: IMS (¿cuál es el ingreso mínimo que ud. considera necesario para satisfacer sus necesidades?). Se presume que el IMS aumenta en la medida que aumenta el ingreso observado. La hipótesis detrás es que personas que tienen más ingreso consideran que para satisfacer sus necesidades es necesario contar con mas recursos. Es decir, las necesidades del ser humano (o su valor) aumentan en la medida en que los ingresos (o riqueza) aumenta. Asimismo, se presume que el IMS aumenta con el tamaño de la familia. Es decir, a mayor número de miembros las presiones económicas aumentan por lo que es necesario contar con mayores ingresos para satisfacerlas. Sin embargo, en este tipo de aplicaciones existe la influencia de efectos no observables. Por ejemplo, hay gente que por naturaleza (o efectos sicológicos no observables) sentirá que necesita mas dinero para satisfacer sus necesidades. Tales efectos pueden no existir (lo que justifica una aplicación MCOC), pueden ser variables transversalmente y fijos en el tiempo (modelo MEF) o pueden ser aleatorios (MEA).

El objetivo del estudio es verificar estas dos hipótesis en un modelo de panel data para el periodo 2001 al 2006 para los hogares peruanos bajo los diferentes supuestos de los efectos no observables. Se disponen de 2,205 datos transversales para 6 años, lo que significa que en total la aplicación contiene 2,205x6=13,230 datos. Se estimará un panel corto (T>N), lineal (la variable dependiente es continua) y balanceado (no existen missing). Para empezar el análisis calculamos algunas variables descriptivas de las variables

```
. sum ly_s ly ln d_01 d_02 d_03 d_04 d_05 d_06

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        ly_s |     13230     6.53944    .7369081   2.302585   9.615806
          ly |     13230    8.715806    1.084943   4.465908   12.87422
          ln |     13230    1.383731    .5673864          0   2.772589
        d_01 |     13230    .1666667    .3726921          0          1
        d_02 |     13230    .1666667    .3726921          0          1
-------------+--------------------------------------------------------
        d_03 |     13230    .1666667    .3726921          0          1
        d_04 |     13230    .1666667    .3726921          0          1
        d_05 |     13230    .1666667    .3726921          0          1
        d_06 |     13230    .1666667    .3726921          0          1
```

Ahora procedemos a testear las hipótesis bajo diferentes supuestos sobre los efectos no observables.

A. MCO combinado

Se estima el modelo

$$\log(y\_s)_{it} = \alpha + \beta \log(y)_{it} + \theta \log(n)_{it} + \sum d_t D_t + e_{it}$$

donde las D_t corresponde a dicotómicas para cada uno de los años bajo análisis. Se propone estimar este modelo mediante MCOC. Para ello, utilice el siguiente comando

```
reg ly_s ly ln d_02 d_03 d_04 d_05 d_06

      Source |       SS       df       MS              Number of obs =   13230
-------------+------------------------------           F(  7, 13222) =  926.06
       Model |   2363.3372     7  337.619599           Prob > F      =  0.0000
    Residual |  4820.45432 13222  .364578303           R-squared     =  0.3290
-------------+------------------------------           Adj R-squared =  0.3286
       Total |  7183.79151 13229  .543033601           Root MSE      =   .6038

------------------------------------------------------------------------------
        ly_s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ly |   .4987152   .0068869    72.41   0.000     .4852159    .5122146
          ln |   .0409898   .0099222     4.13   0.000     .0215409    .0604386
        d_02 |   .2397933   .0181863    13.19   0.000     .2041455    .2754411
        d_03 |   -.572642    .020884   -27.42   0.000    -.6135777   -.5317063
        d_04 |  -.6353162   .0209953   -30.26   0.000      -.67647   -.5941623
        d_05 |  -.6231284   .0210922   -29.54   0.000    -.6644722   -.5817847
        d_06 |  -.6720618   .0214067   -31.39   0.000     -.714022   -.6301017
       _cons |   2.513242   .0513564    48.94   0.000     2.412576    2.613908
------------------------------------------------------------------------------
```

Los regresores incluidos y las dummies de tiempo incluidas explican alrededor del 33% de la varianza del IMS. Este R2 se podía también interpretar como el coeficiente de correlación ajustado entre los valores reales y los valores predichos de este modelo (esta interpretación será útil después). Antes de interpretar los coeficientes, sin embargo, resulta útil explorar si es que existe un proceso heterocedástico (lo que influirá en la signficancia) para ello desarrollamos un test Breusch-Pagan con el comando “hettest” y obtenemos los siguientes resultados

```
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of ly_s

         chi2(1)      =   126.64
         Prob > chi2  =   0.0000
```

Concluimos que el modelo estimado es caracterizado por heterocedasticidad. El valor Breusch-Pagan es 126.64 y la hipótesis nula de un proceso homocedástico del error es rechazada decisivamente por los datos para este modelo. Por ello, interpretar la significancia de los regresores estimados induciría a error. Se plantea por ello ajustar la matriz de la varianza-covarianza mediante el procedimiento de White / Huber (heteroscedasticity de forma desconocida). Para ello, estimamos el modelo usando el comando

```
. reg ly_s ly ln d_02 d_03 d_04 d_05 d_06, robust

Linear regression                                      Number of obs =   13230
                                                       F(  7, 13222) =  751.29
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.3290
                                                       Root MSE      =   .6038

------------------------------------------------------------------------------
             |               Robust
        ly_s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ly |   .4987152   .0076289    65.37   0.000     .4837615     .513669
          ln |   .0409898   .0106629     3.84   0.000      .020089    .0618905
        d_02 |   .2397933    .019581    12.25   0.000     .2014118    .2781748
        d_03 |   -.572642   .0213132   -26.87   0.000    -.6144189   -.5308651
        d_04 |  -.6353162   .0220717   -28.78   0.000    -.6785798   -.5920526
```

```
        d_05 |  -.6231284    .022681   -27.47   0.000    -.6675865   -.5786703
        d_06 |  -.6720618   .0227905   -29.49   0.000    -.7167345   -.6273892
       _cons |   2.513242   .0577557    43.52   0.000     2.400032    2.626451
------------------------------------------------------------------------------
```

Los coeficientes no cambian, pero se ajustan sus desviaciones estándar. Ahora sí, procedemos a interpretar los resultados.

- El coeficiente del ingreso esta bien determinado y la estimación puntual es 0.5. El resultado es el esperado (en signo) y sugiere que una subida del 10% en el ingreso observado genera un aumento del IMS en 5%, manteniendo el resto de condiciones constante.
- El coeficiente del tamaño del hogar es también positivo y bien determinado. Asñi si el tamaño de la familia aumenta en 10%, el IMS aumenta en 0.4%, ceteris paribus.
- El efecto de las dummies temporales (efectos temporales) también son bien determinados aunque difíciles de interpretar. En general indican que el IMS ha caído con los años.

Es importante testear la significancia conjunta de estos regresores (dummies temporales) en la medida que es la forma en la que hemos introducido la heterogeneidad temporal no observada. Así testeamos

H_o:$d_2 = d_3 = d_4 = d_5 = d_6$
H_a: H_o no es correcto

```
. test d_02 d_03 d_04 d_05 d_06

 ( 1)  d_02 = 0
 ( 2)  d_03 = 0
 ( 3)  d_04 = 0
 ( 4)  d_05 = 0
 ( 5)  d_06 = 0

       F(  5, 13222) =  466.01
            Prob > F =    0.0000
```

Con lo que se concluye que existen efectos temporales no observados.

B. Estimador inter-grupo

Conviene estudiar la misma hipótesis utilizando el estimador inter-grupo. Así regresionamos

$$\log(y\_s)_{it} = \alpha + \beta \log(y)_{it} + \theta \log(n)_{it} + e_{it}$$

Notese que las variables dicotomicas temporales se eliminan porque este es un estimador que analiza “solo” la variabilidad que existe entre los grupos. Así, se estima

```
. xtreg ly_s ly ln ,be i(hogar)

Between regression (regression on group means)  Number of obs      =     13230
Group variable: hogar                           Number of groups   =      2205

R-sq:  within  = 0.0056                         Obs per group: min =         6
       between = 0.5983                                        avg =       6.0
       overall = 0.1808                                        max =         6
```

```
                                                 F(2,2202)          =   1640.00
sd(u_i + avg(e_i.))=  .3667557                   Prob > F           =    0.0000

------------------------------------------------------------------------------
        ly_s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ly |   .6455503   .0120323    53.65   0.000     .6219545    .6691461
          ln |   -.024466   .0163331    -1.50   0.134    -.0564959     .007564
       _cons |   .9468025   .0989728     9.57   0.000     .7527127    1.140892
------------------------------------------------------------------------------
```

Los resultados ahora arrojan que la variable ingresos sigue siendo significativa y en la dirección correcta, pero el tamaño de la familia ya no es significativo. En alguna medida el estimador inter-grupo revela los efectos de largo plazo, por lo que la no signficancia del tamaño del hogar puede ser atribuido a que en el largo plazo esta variable no es importante. Sin embargo, hay que ser cuidadoso con esta conclusión debido a que solo se usan seis años (difícilmente revelen una relación de largo plazo) y además el modelo es sumamente austero (podrían haber otros determinantes). En todo caso, el unico regresos susceptible de ser interpretado es el ingreso cuyo estimador puntual es 0.7. Es decir, ante un aumento del 10% en el ingreso observado el IMS aumenta en 7%.

En contraste con el MCOC la interpretación del R2 es menos evidente. Es decir, no necesariamente revela que % de la varianza de la dependiente es se puede explicar con las dependientes, más bien revela el cuadrado del coeficiente de correlación entre los valores predichos y los actuales (tal como se indicó previamente). Se tienen tres tipos de estimadores

$R^2_{overall}$ = cuadrado de la correlación[$\hat{y}_{it}$, y_{it}] = 0.1808

donde $\hat{y}_{it} = x_{it}'\beta_{between}$

Es decir, la correlación ajustada entre el IMS predicho y el actual tomando en cuenta (para los valores actuales) toda la base datos. Para ello, el predicho se construye a partir de las estimaciones MCO – intergrupo y se comprara con toda la base de datos observados. Luego,

$R^2_{between}$ = cuadrado de la correlación [$\hat{y}_{i\_B}$, $\bar{y}_i$] = 0.5983

donde $\hat{y}_{i\_B} = \bar{x}_i'\beta_{between}$

Es decir, la correlación ajustada entre el IMS predicho y el actual tomando en cuenta unicamente los promedios grupales (a lo largo del tiempo). Para ello, el predicho se construye a partir de las estimaciones MCO – intergrupo y se comprara con los promedios grupales observados. Este es el R2 relevante por analizar en un contexto intergrupo. Luego,

R^2_{within} = cuadrado de la correlación [$\hat{y}_{it\_B}$ - $\hat{y}_{i\_B}$, $y_{it} - \bar{y}_i$] = 0.0056 donde $\hat{y}_{it\_B} - \hat{y}_{i\_B} = (x_{it} - \bar{x}_i)'\beta_{between}$

Es decir, la correlación ajustada entre el IMS predicho y el actual tomando en cuenta la variación temporal de la data. Es decir, se tratan de predecir las desviaciones respecto al promedio temporal utilizando el estimadot MCO – intergrupo.

Al momento de comprar los R2 se puede concluir que el estimador inter-grupo tiene un menor poder explicativo sobre la data total que el modelo MCOC (19% vs. 33%), lo cual es de suponer dada la significancia de las dummies temporales (que no aparecen en el segundo modelo). Luego, se observa que tiene un buen ajuste para explicar la variación intergrupo (60%). Esta es la medida relevante (se utiliza un modelo intergrupo para explicar variabilidad intergrupo). Asimismo, se observa que el estimador inter-grupo tiene poco poder para explicar la variabilidad intra-grupo. Esto es lógico suponer en la medida que el modelo correcto para estos efectos es el estimador inter-grupo.

C. Estimador intra-grupo

Conviene estudiar la misma hipótesis utilizando el estimador intra-grupo. Así regresionamos

$$\log(y\_s)_{it} = \alpha_i + \beta \log(y)_{it} + \theta \log(n)_{it} + \sum d_t D_t + e_{it}$$

Se estima

```
. xtreg ly_s ly ln d_02 d_03 d_04 d_05 d_06,fe i(hogar)

Fixed-effects (within) regression               Number of obs      =     13230
Group variable: hogar                           Number of groups   =      2205

R-sq:  within  = 0.0559                         Obs per group: min =         6
       between = 0.3304                                        avg =       6.0
       overall = 0.2067                                        max =         6

                                                F(7,11018)         =     93.21
corr(u_i, Xb)  = 0.2865                         Prob > F           =    0.0000

------------------------------------------------------------------------------
        ly_s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ly |   .1116671   .0105843    10.55   0.000       .09092    .1324142
          ln |   .2130612   .0189224    11.26   0.000     .1759699    .2501524
        d_02 |   .2507999   .0146505    17.12   0.000     .2220822    .2795175
        d_03 |   .0037092   .0215397     0.17   0.863    -.0385124    .0459309
        d_04 |  -.0470062   .0218016    -2.16   0.031    -.0897413   -.0042711
        d_05 |  -.0250637   .0220343    -1.14   0.255    -.0682549    .0181274
        d_06 |  -.0406546   .0227511    -1.79   0.074    -.0852508    .0039417
       _cons |    5.24772   .0780878    67.20   0.000     5.094654    5.400786
-------------+----------------------------------------------------------------
     sigma_u |  .50482081
     sigma_e |   .4862629
         rho |  .51871831   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(2204, 11018) =     4.25         Prob > F = 0.0000
```

Nótese que las variables dummy temporales ahora si se pueden incorporar. Conviene especular un poco acerca de los resultados. Así, se observa que luego de corregir por heterogeneidad (transversal) no observada, la elasticidad del ingreso ha disminuido respecto al estimado de MCOC (0.1 vs 0.6) y el efecto del tamaño de la familia ha aumentado (0.2 vs. 0.4). Del mismo modo, algunos de los efectos temporales son ahora no significativos. Sin embargo, en conjunto se observa que todavía son significativos

```
. test d_02 d_03 d_04 d_05 d_06

 ( 1)  d_02 = 0
 ( 2)  d_03 = 0
 ( 3)  d_04 = 0
 ( 4)  d_05 = 0
 ( 5)  d_06 = 0

       F(  5, 11018) =   75.51
            Prob > F =    0.0000
```

Las medidas de ajuste R2 revelan que el Rwithin (el relevante para este caso) es de apenas 0.06. Este mal resultado puede ser atribuido a la poca variabilidad temporal de la data incluida (el ingreso y el numero de miembros de la familia) lo que siguiere que necesitamos algunos regresores adicionales. Este problema de especificación puede generar efectos al momento de decidir entre los diferentes modelos (como se verá mas adelante). De momento interesa comparar los R2 para el modelo inter-grupo e intra-grupo. Nótese que en este caso los calculos se realizan de la misma forma pero considerando parámetros MEF. Así, se puede observar que tanto el modelo intra-grupo como la alternativa inter-grupo tienen similares poderes explicativos sobre la variabilidad total de la data (0.2 vs. 0.18) y el poder de ambos es menor que el modelo MCOC. Luego, cada uno de ellos lo hace mejor sobre el tipo de variabilidad que intentan explicar y se observa que el estimador MEF es mejor para explicar la variabilidad inter-grupo que el modelo inter-grupo para explicar la variabilidad intra-grupo (0.33 vs. 0.006).

Luego, la correlación ente los efectos fijos y las variables explicativas es baja (0.287), esta correlación que sea baja o alta no tiene efectos sobre las propiedades de consistencia de los estimadores MEF, pero una elevada correlación hubiera tenido un impacto mayor sobre los valores puntuales de los resultados (los cambios hubieran sido mas dramáticos). La interpretación de una correlación baja es que la heterogeneidad no observada tiene una baja correlación con las variables incluidas en el modelo. Sin embargo, el % de la varianza que proviene de los efectos no observados es alta (rho=0.52) por lo que no es de sorprender que el test F de significancia sobre estos efectos (4.25) acepte su inclusión (se favorece el uso de MEF sobre MCOC).

Los efectos fijos no son interpretables en sí, pero pueden ser estimados utilizando

```
. predict fe_est, u

. list fe_est
```

No se muestran dado que son 2,205 parámetros estimados

D. Estimador MEA

Ya se concluyo a favor del MEF sobre el MCOC, a pesar que el MCOC tenía un mayor poder explicativo. Asimismo se nota que la correlación con las explicativas incluidad era bajo. Esto sugiere que tal vez un mejor supuesto sea efectos aleatorios. Para ello, se regresiona un modelo como

$$\log(y\_s)_{it} = \alpha + \beta \log(y)_{it} + \theta \log(n)_{it} + \sum d_t D_t + e_{it}$$

```
. xtreg ly_s ly ln d_02 d_03 d_04 d_05 d_06,re i(hogar)

Random-effects GLS regression                   Number of obs      =     13230
Group variable: hogar                           Number of groups   =      2205

R-sq:  within  = 0.0428                         Obs per group: min =         6
       between = 0.5689                                        avg =       6.0
       overall = 0.3220                                        max =         6

Random effects u_i ~ Gaussian                   Wald chi2(7)       =   2600.19
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
        ly_s |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ly |   .3443076   .0082733    41.62   0.000     .3280923    .3605229
          ln |   .1097441   .0128291     8.55   0.000     .0845995    .1348887
        d_02 |   .2441858   .0152509    16.01   0.000     .2142945    .2740771
        d_03 |  -.3427119   .0196154   -17.47   0.000    -.3811574   -.3042664
        d_04 |  -.4006136   .0197876   -20.25   0.000    -.4393967   -.3618305
        d_05 |   -.384532   .0199383   -19.29   0.000    -.4236104   -.3454537
        d_06 |  -.4201622   .0204186   -20.58   0.000     -.460182   -.3801424
       _cons |   3.603971   .0613006    58.79   0.000     3.483824    3.724117
-------------+----------------------------------------------------------------
     sigma_u |  .30838472
     sigma_e |   .4862629
         rho |  .28683562   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Al comprar los R2 se observa que el poder explicativo mejora. Luego, los coeficientes estimados estan bien determinados (significativos y con el signo esperado). Se observa que este modelo arroja una elasticidad del ingreso observado sobre el IMS de 0.34 (entre el 0.5 y el 0.11 estimados previamente) y una elasticidad del tamaño de familia de 0.11 (mas cercano al estimado MEF). Luego, las dummies temporales son significativas

Nótese que el supuesto implícito de la aplicación es que la correlación entre los efectos no observados y las explicativas es cero. Sin embargo, es necesario testear la prescencia de efectos aleatorios. Se procede con el comando

```
. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

        ly_s[hogar,t] = Xb + u[hogar] + e[hogar,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                    ly_s |   .5430336       .7369081
                       e |   .2364516       .4862629
                       u |   .0951011       .3083847

        Test:   Var(u) = 0
                             chi2(1) =  2454.11
                         Prob > chi2 =     0.0000
```

La hipótesis nula de que no existen efectos aleatorios se rechaza, por lo que vale la pena su inclusión (esto favorece el uso de MEA frente al MCOC).

De este modo, tenemos dos posibles modelos MEA y MEF, para testear cual es el favorito se utiliza el test de Haussman. Se procede de la siguiente manera

```
. xtreg ly_s ly ln d_02 d_03 d_04 d_05 d_06, fe i(hogar)

Fixed-effects (within) regression               Number of obs      =     13230
Group variable: hogar                           Number of groups   =      2205

R-sq:  within  = 0.0559                         Obs per group: min =         6
       between = 0.3304                                        avg =       6.0
       overall = 0.2067                                        max =         6

                                                F(7,11018)         =     93.21
corr(u_i, Xb)  = 0.2865                         Prob > F           =    0.0000

------------------------------------------------------------------------------
        ly_s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ly |   .1116671   .0105843    10.55   0.000       .09092    .1324142
          ln |   .2130612   .0189224    11.26   0.000     .1759699    .2501524
        d_02 |   .2507999   .0146505    17.12   0.000     .2220822    .2795175
        d_03 |   .0037092   .0215397     0.17   0.863    -.0385124    .0459309
        d_04 |  -.0470062   .0218016    -2.16   0.031    -.0897413   -.0042711
        d_05 |  -.0250637   .0220343    -1.14   0.255    -.0682549    .0181274
        d_06 |  -.0406546   .0227511    -1.79   0.074    -.0852508    .0039417
       _cons |    5.24772   .0780878    67.20   0.000     5.094654    5.400786
-------------+----------------------------------------------------------------
     sigma_u |  .50482081
     sigma_e |   .4862629
         rho |  .51871831   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(2204, 11018) =     4.25         Prob > F = 0.0000

. est store fixed

. xtreg ly_s ly ln d_02 d_03 d_04 d_05 d_06, re i(hogar)

Random-effects GLS regression                   Number of obs      =     13230
Group variable: hogar                           Number of groups   =      2205

R-sq:  within  = 0.0428                         Obs per group: min =         6
       between = 0.5689                                        avg =       6.0
       overall = 0.3220                                        max =         6

Random effects u_i ~ Gaussian                   Wald chi2(7)       =   2600.19
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000

------------------------------------------------------------------------------
        ly_s |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ly |   .3443076   .0082733    41.62   0.000     .3280923    .3605229
          ln |   .1097441   .0128291     8.55   0.000     .0845995    .1348887
        d_02 |   .2441858   .0152509    16.01   0.000     .2142945    .2740771
        d_03 |  -.3427119   .0196154   -17.47   0.000    -.3811574   -.3042664
        d_04 |  -.4006136   .0197876   -20.25   0.000    -.4393967   -.3618305
        d_05 |   -.384532   .0199383   -19.29   0.000    -.4236104   -.3454537
        d_06 |  -.4201622   .0204186   -20.58   0.000     -.460182   -.3801424
       _cons |   3.603971   .0613006    58.79   0.000     3.483824    3.724117
-------------+----------------------------------------------------------------
     sigma_u |  .30838472
     sigma_e |   .4862629
         rho |  .28683562   (fraction of variance due to u_i)
------------------------------------------------------------------------------

. hausman fixed

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed          .          Difference          S.E.
-------------+----------------------------------------------------------------
          ly |    .1116671     .3443076       -.2326405        .0066015
          ln |    .2130612     .1097441        .1033171        .0139094
        d_02 |    .2507999     .2441858        .0066141               .
        d_03 |    .0037092    -.3427119        .3464211        .0088991
        d_04 |   -.0470062    -.4006136        .3536074         .009152
        d_05 |   -.0250637     -.384532        .3594683        .0093794
        d_06 |   -.0406546    -.4201622        .3795076        .0100346
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg
```

```
Test:  Ho:  difference in coefficients not systematic

                  chi2(7) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =     1249.69
                Prob>chi2 =      0.0000
                (V_b-V_B is not positive definite)
```

Así, se rechaza el modelo MEA a favor del MEF y esta pasaría a ser nuestra especificación central. Sin embargo, la matriz varianza covarianza no es semidefinida positiva, por lo que el resultado de Haussman no se puede interpretar correctamente (una de las propiedad asintóticas del test ha sido violada). Por ello, cuando salga este resultado se recomienda interpretar los resultados con cuidado (quizas sea necesario dotar de mayor estructura a los modelos). Sin embargo, si obviamos este resultados en nuestra aplicación el MEF sería el modelo favorito y el MEA nuestro resultado para proveer robustez. Ojo que el el test de Breusch and Pagan para determinar la existencia de efectos aleatorios (que se acpeto) no contradice los resultados del Haussman. El primer tes indica que estos efectos existen y el segundo que tales efectos estan correlacionados con las independientes del modelo.

Un correcto reporte de resultados sería utilizar los resultados MEF como estimación central y analizar desviaciones respecto al resultado MEA (no existen en términos de signo) aunque si en valor (sobre todo para el ingeso observado). La estimación MCOC es util simplemente a nivel exploratorio.