UN ESTUDIO COMPARATIVO DEL ESTIMADOR DE MÍNIMOS CUADRADOS GENERALIZADOS PARA MODELOS DE PANEL*

Elena Casquel y Ezequiel Uriel**

WP-EC 2000-12

Correspondencia a E. Casquel: Universidad Jaume I. Dpto. Economía. Campus del Riu Sec, 12071 Castellón.

Tel.: +34 964 728608 / E-mail: casquel@eco.uji.es

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.

Primera Edición Julio 2000

Depósito Legal: V-3078-2000

---

* Este trabajo ha sido …nanciado por una ayuda a la investigación concedida por el Ivie.

** E. Casquel: Universidad Jaume I. E. Uriel: Instituto Valenciano de Investigaciones Económicas (Ivie) y Universidad de Valencia.

UN ESTUDIO COMPARATIVO DEL ESTIMADOR DE MÍNIMOS CUADRADOS GENERALIZADOS PARA MODELOS DE PANEL

Elena Casquel y Ezequiel Uriel

RESUMEN

En este artículo se realiza un estudio comparativo de los estimadores por Mínimos Cuadrados Generalizados y del contraste de Hausman para modelos de datos de panel aplicados por los programas LIMDEP, RATS y TSP. Con este objetivo se han llevado a cabo experimentos de Montecarlo en modelos en los que no existe correlación entre el componente individual y los regresores y para modelos en los que sí existe correlación. En la realización de los experimentos de Montecarlo se han utilizado variables control, con el objeto de aumentar la e…ciencia del experimento con un menor número de iteraciones.

Palabras clave: Mínimos Cuadrados Generalizados, datos de panel, experimentos de Montecarlo.

ABSTRACT

In this paper we make a comparative study of Generalized Least Squares Estimators used by the programs LIMDEP, RATS and TSP for panel data models. In adittion, we analize Hausman test computed by these programs. Monte Carlo experiments are applied in two di¤erents models: models where it is assumed that individual e¤ects and explanatory variables are independent and models where this assumption is relaxed. Futhermore, we use control variates in order to improve the precision of the experiment reducing the number of replications.

Keywords: Generalized Least Squares Estimators, Panel Data, Monte Carlo Experiments.

1 INTRODUCCIÓN.

La aplicación del método de Mínimos Cuadrados Generalizados Factibles (MCGF) a un modelo de efectos aleatorios con datos de panel requiere estimar previamente las varianzas de los dos componentes aleatorios. Como existen distintos procedimientos para la estimación de estas varianzas [Maddala and Mount (1973)], esto da lugar a estimaciones alternativas del vector de parámetros. Además, como la varianza del efecto individual es obtenida normalmente de forma residual, se presenta el problema de que ésta puede adoptar valores negativos. No existe acuerdo sobre cual sea el procedimiento más adecuado de estimación. Así, los programas LIMDEP (7.0), RATS (32) para Windows y TSP (4.4), que son los más usuales en el tratamiento de los modelos de panel, ofrecen diferentes estimaciones de este tipo de modelos.

Si en el modelo de datos de panel existe correlación entre las perturbaciones y el efecto individual, el estimador MCGF no resulta insesgado. En este caso, el estimador within (entre-grupos), obtenido a partir de modelo de efectos …jos, tiene mejores propiedades. Para contrastar la presencia de dicho tipo de correlación se suele utilizar el test de Hausman[1]. Conviene tener en cuenta a este respecto que en el trabajo Maudos y Uriel (1995) se han detectado ciertas anomalías en los programas LIMDEP y TSP en la aplicación del test deHausman.

El objetivo de este artículo es estudiar los procedimientos utilizados por los tres programas antes citados del estimador MCGF y el estimador within para modelos en los que se introduce exogeneidad en las variables explicativas, así como para modelos con correlación entre los regresores y el efecto individual. En este último caso se estudiará el test de Hausman calculado según LIMDEP y TSP para analizar su comportamiento en pequeñas muestras.

Con este …n se han realizado experimentos de Montecarlo utilizando técnicas de control de la varianza, en concreto, introduciendo variables control. Con estas técnicas, descritas en Hendry (1984), se consigue aumentar la e…ciencia del experimento con un menor número de iteraciones.

El artículo se estructura de la siguiente forma. En la segunda sección se describen los procedimientos de estimación de los modelos de efectos aleatorios y del cálculo del estadístico del contraste de Hausman que proponen los programas antes citados. En la tercera sección se presentan los resultados de los experi-

[1] Ver Hausman (1978).

mentos de Monte Carlo. En un primer apartado, tras explicar la técnica de variables control, se muestran los resultados del estimador MCGF tomando como referencia un modelo de variables explicativas exógenas. En un segundo apartado, se realizan los experimentos de Monte Carlo relajando el supuesto de exogeneidad, incluyendo el cálculo del test de Hausman. En la sección final se recogen las conclusiones.

2 ESTIMACIÓN DE MODELOS DE EFECTOS ALEATORIOS Y TEST DE HAUSMAN.

El modelo de efectos aleatorios se puede expresar de la siguiente forma[2]:

$$\begin{aligned} y_{it} &= \alpha + \beta' x_{it} + v_{it}, \\ i &= 1, \ldots, N, \\ t &= 1, \ldots, T. \end{aligned} \tag{1}$$

El valor de la variable dependiente para la i-ésima observación en el momento t, y_{it}, depende de K variables explicativas, $(x_{1it}, \ldots, x_{Kit})' = x_{it}$, que difieren entre individuos para cada t. β' es un vector de constantes $(1 \times K)$ y α es un escalar. El término de error se descompone en:

$$v_{it} = u_i + \varepsilon_{it}, \tag{2}$$

donde u_i es el efecto aleatorio específico del individuo i-ésimo.

Las hipótesis estadísticas que, usualmente, se formulan sobre las variables que integran el término de perturbación del modelo (1) son las siguientes:

$$E(\varepsilon_{it}\varepsilon_{js}) = \left\{ \begin{array}{l} \sigma_\varepsilon^2 \text{ si } i = j \;\; t = s \\ 0 \text{ en otro caso} \end{array} \right\}$$

$$E(u_i u_j) = \left\{ \begin{array}{l} \sigma_u^2 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{array} \right\}$$

$$E(\varepsilon_{it} u_j) = 0 \text{ para todo } i, j, t.$$

Bajo estos supuestos, la matriz de covarianzas del término de perturbación v_{it} es una matriz no escalar. Si se introduce el supuesto adicional de exogeneidad en las variables explicativas, el método

[2] Ver Hsiao (1986).

óptimo para estimar el modelo (1) es el de Mínimos Cuadrados Generalizados (MCG). Por otro lado, si no se cumple este último supuesto, el estimador que mejores propiedades presenta es el estimador within. El test de Hausman se utiliza para contrastar la existencia de correlación entre el efecto individual (u_i) y los regresores.

En los siguientes apartados se van a estudiar de forma sucesiva el método de estimación MCGF, bajo el supuesto de exogeneidad antes citado, los distintos enfoques existentes para obtener este estimador, así como el test de Hausman.

2.1 Estimación por Mínimos Generalizados Factibles.

El estimador de MCG aplicado al modelo (1) se puede expresar como media ponderada de $\hat{\beta}^W$ (estimador within) y $\hat{\beta}^B$ (estimador between) mediante la siguiente expresión:

$$\begin{aligned} \hat{\beta}^{MCG} &= \Delta \hat{\beta}^B + (I_K - \Delta)\hat{\beta}^W, \\ \hat{\alpha}^{MCG} &= \bar{y} - \hat{\beta}^{MCG}\bar{x}, \end{aligned} \tag{3}$$

donde :

$$\Delta = \theta T \left[\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' + \theta T \sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right]^{-1} \left[\sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right], \tag{4}$$

$$\theta = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + T\sigma_u^2}, \tag{5}$$

$$\hat{\beta}^W = \left[\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right]^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i)', \tag{6}$$

$$\hat{\beta}^B = \left[\sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right]^{-1} \sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y})'. \tag{7}$$

Si los dos componentes de la varianza fueran conocidos, la aplicación de (3) es inmediata. Ahora bien, como éste no será el caso en el mundo real, es necesario estimar previamente σ_ε^2 y σ_u^2 para obtener un estimador de θ. Cuando se utilizan los valores estimados de θ, los estimadores no son estrictamente MCG. En este caso a los estimadores obtenidos habrá que denominarlos MCGF, en concreto $\hat{\beta}^{MCGF}$. En su aplicación se pueden presentar problemas ya que parte de los diferentes procedimientos existentes para la estimación de θ obtienen la estimación de la σ_u^2 de forma residual, lo que puede provocar que ésta sea negativa. Además, cada método parte de supuestos distintos, obtiéndose diferentes estimadores de θ, y por tanto, de $\hat{\beta}^{MCGF}$. En Maudós y Uriel (1995) se estudia esta problemática.

2.2 Diferentes enfoques de la estimación de θ.

A continuación se van a examinar los diferentes enfoques de estimación de θ aplicados por los programas TSP, LIMDEP y RATS.

2.2.1 Estimación en el programa TSP.

Para la estimación de σ_ε^2 y σ_u^2 programa TSP[3] utiliza como referencia la varianza del modelo de efectos aleatorios (1). Esta varianza viene dada por:

$$\sigma_v^2 = \sigma_\varepsilon^2 + \sigma_u^2. \tag{8}$$

Para estimar la varianza del primer mienbro de (8) se utiliza la suma de cuadrados de los residuos generada al estimar por MCO el modelo (1) -a los estimadores obtenidos se les denomina total en la literatura-. A esta suma se le va a llamar SCRT. Por otra parte, dado que en el modelo de efectos fijos la componente aleatoria es únicamente ε_{it}, los estimadores within permiten calcular la suma de cuadrados de los residuos -esta suma se va a denominar SCRW -que puede ser utilizada para estimar σ_ε^2. Por lo tanto, de acuerdo con lo anterior, se puede obtener la estimación de σ_u^2 de forma residual.

Los estimadores de σ_ε^2 y σ_u^2 propuestos por TSP (denominados de pequeñas muestras) vienen dados por:

$$\hat{\sigma}_\varepsilon^2 = \frac{SCRW}{N \times T - N - K}, \tag{9}$$

$$\hat{\sigma}_u^2 = \frac{SCRT}{N \times T - K - 1} - \frac{SCRW}{N \times T - N - K}. \tag{10}$$

El estimador (10) es un estimador insesgado, ya que se determina como diferencia entre los estimadores insesgados de σ_ε^2 y σ_v^2. Sin embargo, y debido precisamente a la corrección de grados de libertad para obtener estimadores insesgados, la estimación de σ_u^2 resultante puede ser negativa. En este caso, el

[3] Ver Hall (1997)

programa TSP propone los siguientes estimadores, que no son insegados pero sí consistentes. Por esta razón se denominan de muestras grandes.

El estimador de σ^2_ε para muestras grandes se define como:

$$\hat{\sigma}^2_\varepsilon = \frac{SCRW}{N \times T}; \tag{11}$$

y el estimador de σ^2_u viene dado por:

$$\hat{\sigma}^2_u = \frac{SCRT - SCRW}{N \times T}: \tag{12}$$

2.2.2 Estimación en el programa LIMDEP.

En el programa LIMDEP se utilizan tres procedimientos alternativos para estimar el efecto individual[4]. En el primer caso, puede suceder que σ^2_u sea negativa. Si esto ocurre, se pasa a una segunda forma de estimación de σ^2_u en la que también puede plantearse el mismo problema. LIMDEP propone un tercer método que por construcción la σ^2_u no puede ser negativa.

La estimación de σ^2_ε en los tres procedimientos que viene dada por:

$$\hat{\sigma}^{2*}_\varepsilon = \frac{SCRW}{N \times T - N - K}; \tag{13}$$

Para estimar el efecto individual en los dos primeros procedimientos, se toma como referencia el modelo (1) expresado en términos de medias del grupo, denominado modelo between, que es:

$$\bar{y}_i = \alpha + \beta' \bar{x}_i + u_i + \bar{\varepsilon}_i: \tag{14}$$

La varianza del modelo anterior toma el siguiente valor:

$$Var(u_i + \bar{\varepsilon}_i) = \sigma^2_u + \frac{\sigma^2_\varepsilon}{T}: \tag{15}$$

Por lo tanto, la estimación del modelo between permite obtener una suma de cuadrados de los residuos, SCRB, que se define como:

$$SCRB = \sum_{i=1}^{N} \hat{e}_i^{2*} = \sum_{i=1}^{N} \left[\bar{y}_i - (\hat{\alpha}^B + \hat{\beta}^B \bar{x}_i) \right]^2: \tag{16}$$

---

[4] Ver Greene (1995).

Obtenida esta suma de cuadrados y la estimación dada en (13) se obtiene de forma residual el primer estimador de σ_u^2 :

$$\hat{\sigma}_u^{2*} = \frac{SCRB}{N - K - 1} - \frac{\hat{\sigma}_\eta^{2*}}{T}: \tag{17}$$

En el caso de que la expresión (17) resulte negativa, el LIMDEP aplica un segundo procedimiento que consiste en sustituir SCRB por la siguiente expresión:

$$SCRB_{Tot} = \sum_{i=1}^{N} \hat{e}_i^{2**} = \sum_{i=1}^{N} \left[\bar{y}_i - (\hat{\alpha}^{Tot} + \hat{\beta}^{Tot} \bar{x}_i) \right]^2 : \tag{18}$$

Como puede observarse para el cálculo de $SCRB_{Tot}$ se utiliza el esquema (16) pero sustituyendo los estimadores between por los estimadores total. La estimación de σ_u^2 en el segundo método es, por tanto la siguiente:

$$\hat{\sigma}_u^{2**} = \frac{SCRB_{Tot}}{N - K - 1} - \frac{\hat{\sigma}_\eta^{2*}}{T}; \tag{19}$$

Dado que el estimador anterior puede ser negativo, cuando se plantea este caso el LIMDEP utiliza como tercer procedimiento el estimador propuesto Nerlove (1971), que consiste en estimar σ_u^2 directamente mediante la varianza muestral de los efectos fijos estimados en el modelo within, α_i; es decir:

$$\hat{\sigma}_u^{2***} = \frac{\sum_{i=1}^{N} (\hat{\alpha}_i - \bar{\hat{\alpha}})^2}{N}; \tag{20}$$

donde:

$$\bar{\hat{\alpha}} = \frac{\sum_{i=1}^{N} \hat{\alpha}_i}{N}: \tag{21}$$

Como es obvio este tercer estimador por construcción no puede resultar nunca negativo.

2.2.3 Estimación en el programa RATS.

El programa RATS utiliza los residuos correspondientes a la estimación total para obtener estimadores de σ_η^2 y σ_u^2[5]:

Aplicando el análisis de la varianza a estos residuos, se obtiene la siguiente descomposición:

$$SCRT = SCRb + SCRw, \tag{22}$$

[5] Ver Doan (1996).

donde SCRb es la suma de cuadrados de los residuos between-group y SCRw es la suma de los cuadrados de los residuos within-group.

Consecuentemente, las estimaciones de σ^2_ε y σ^2_u vienen dadas por:

$$\hat{\sigma}^2_\varepsilon = \frac{SCRw}{N \times T - N}; \tag{23}$$

$$\hat{\sigma}^2_u = \frac{SCRb}{N - 1}. \tag{24}$$

Aunque estos estimadores son sesgados para muestras pequeñas, la ventaja que tiene el método propuesto por el programa RATS es que siempre se obtienen estimaciones de σ^2_u no negativas.

2.3 Test de Hausman.

Como ya se ha comentado, si existe correlación entre regresores y variables explicativas el estimador MCGF sería sesgado. En ese caso, el estimador más adecuado es el estimador within.

Para contrastar la ortogonalidad entre regresores y efectos individuales aleatorios se suele utilizar el contraste de Hausman. Éste se basa en la idea de que, bajo la hipótesis nula de ortogonalidad el estimador obtenido por MCGF es consistente y eficiente. Bajo la hipótesis alternativa, este estimador es sesgado. En cambio, el estimador within es consistente bajo la hipótesis nula y la hipótesis alternativa.

Las hipótesis nula y alternativa se pueden formular de la siguiente forma:

$$\begin{aligned} H_0 &: E(u_i X_i) = 0; \\ H_1 &: E(u_i X_i) \neq 0. \end{aligned} \tag{25}$$

El estadístico propuesto por Hausman es:

$$H = (\hat{\beta}^W - \hat{\beta}^{MCGF})'[\widehat{Var}(\hat{\beta}^W) - \widehat{Var}(\hat{\beta}^{MCGF})]^{-1}(\hat{\beta}^W - \hat{\beta}^{MCGF}); \tag{26}$$

donde:

$$\widehat{Var}(\hat{\beta}^W) = \hat{\sigma}^2_\varepsilon \left[\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right]^{-1} \tag{27}$$

y

$$\widehat{Var}(\hat{\beta}^{MCGF}) = \hat{\sigma}^2_\varepsilon \left[\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' + \hat{\theta} T \sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right]^{-1}. \tag{28}$$

El cálculo de este test puede plantear problemas. Teóricamente, la diferencia entre las dos matrices de varianzas siempre es una matriz semidefinida positiva. A nivel empírico puede resultar que no se cumpla esta propiedad debido a problemas de precisión con los que trabajan los ordenadores.

2.3.1 Cálculo del test de Hausman en los programas LIMDEP y TSP.

Tanto LIMDEP como TSP utilizan (27) y (28) para estimar las varianzas de los estimadores $\hat{\beta}^W$ y $\hat{\beta}^{MCGF}$: Sin embargo, los resultados obtenidos difieren entre los dos programas.

Una primera diferencia entre ambos viene dada por la distinta estimación que hacen de γ; que ha sido descrita en el segundo apartado de esta sección.

Otra diferencia viene dada por el hecho de que LIMDEP para estimar las varianzas (27) y (28) utiliza siempre el mismo estimador de la σ^2_ε; en concreto, el estimador definido en (13). En cambio, el programa TSP no utiliza en todos los casos el mismo estimador. Cuando se utiliza (12)-el segundo procedimiento para estimar la σ^2_u-el programa TSP utiliza como estimador de σ^2_ε en (27) el estimador denominado de muestras pequeñas, (9), mientras que para estimarlo en (28) utiliza el estimador de muestras grandes, recogido en (11).

El test de Hausman obtenido tanto por TSP como por LIMDEP presenta en ocasiones resultados anómalos.

Se ha comprobado que en modelos en los que se introduce una tendencia lineal el resultado es erróneo. Al introducir una tendencia lineal, la diferencia entre las dos matrices de varianzas es una matriz singular. TSP soluciona el problema calculando la inversa generalizada, con lo que elimina las filas y columnas que son cero, mientras que LIMDEP asigna el valor cero al estadístico de contrastes del test de Hausman.

3 EXPERIMENTOS DE MONTECARLO.

En esta sección se van a comparar, mediante experimentos de Montecarlo los procedimientos de estimación y el contraste de Hausman que ofrecen los programas LIMDEP, TSP y RATS en muestras finitas.

En el primer apartado se explica en primer lugar la técnica de variables control. A continuación se obtiene el ECM y el sesgo del estimador MCGF proporcionado por los programas antes citados, así como del estimador within, tomando como referencia un modelo con variables explicativas exógenas. Estos resultados se analizan para distintos tamaños muestrales, bondad del ajuste y proporción relativa de la varianza.

En el segundo apartado se obtiene además de los resultados anteriores, el test de Hausman ofrecido por LIMDEP y TSP a partir de un modelo con correlación entre las variables explicativas y el efecto aleatorio individual.

3.1 Experimentos de Montecarlo para modelos sin correlación.

3.1.1 Método de reducción de la varianza: Variables Control.

Los experimentos de Monte Carlo[6] se usan frecuentemente para estudiar las propiedades de los estimadores en muestras …nitas. Los resultados de estos experimentos tienen inevitablemente un componente aleatorio, en el sentido de que dependen de las muestras generadas en un experimento. Para reducir la aleatoriedad a niveles aceptables es necesario llevar a cabo un gran número de iteraciones. Ahora bién, utilizando variables control se puede obtener un incremento substancial en la e…ciencia del experimento.

En la aplicación de esta técnica se eligen como variables control a variables de las que se conocen ciertas propiedades de su distribución, concretamente la media poblacional, y qué además están muy correlacionadas con el estimador o test estadístico estudiado. La divergencia entre su media poblacional y su media muestral es usada precisamente para mejorar la estimación obtenida en el experimento de Monte Carlo. Conviene señalar que las variables control sólo se usan en el contexto de los experimentos de Monte Carlo porque se calculan utilizando parámetros que no pueden ser observados en las investigaciones estadísticas. Hendry (1984) proporciona algunos ejemplos de variables control en problemas econométricos.

Regresión basada en variables control. Suponiendo que μ es un parámetro que se quiere estimar [7] usando los resultados del experimento de Monte Carlo. En cada iteración se obtiene una estimación de μ, a la que se denomina t_j y una variable control, $¿_j$:

[6] Ver Davidson and Mackinnon (1992).

[7] Ver Dadvison and Mackinnon (1993).

Sin utilizar variables control, se toma como estimador de τ; la media de las estimaciones obtenidas en las M iteraciones, es decir,

$$\bar{\theta} = \frac{1}{M}\sum_{j=1}^{M} t_j; \tag{29}$$

con varianza $V(\bar{\theta}) = \frac{V}{M}$; que puede ser estimada por:

$$\hat{V}(\bar{\theta}) = \frac{1}{M}\frac{1}{M-1}\sum_{j=1}^{M}(t_j - \bar{\theta})^2. \tag{30}$$

Si es posible obtener variables control, (29) en la mayoría de los casos no es el estimador óptimo. Entonces el estimador apropiado sería:

$$\theta^*(\beta) = \bar{\theta} - \beta(\bar{\tau} - E(\tau)); \tag{31}$$

donde $\bar{\tau}$ es la media muestral de τ_j y β es un parámetro que debe ser determinado. La elección de β es crucial. Parece natural elegir el valor de éste que minimize la varianza del estimador anterior.

La varianza de (31) es:

$$V(\theta^*(\beta)) = V(\bar{\theta}) + \beta^2 V(\bar{\tau}) - 2\beta \mathrm{Cov}(\bar{\tau}; \bar{\theta}). \tag{32}$$

Es fácil minimizar la ecuación anterior con respecto a β. El valor óptimo de β sería entonces:

$$\beta^* = \frac{\mathrm{Cov}(\bar{\tau}; \bar{\theta})}{V(\bar{\tau})}. \tag{33}$$

Sustituyendo en (32), la varianza de $\theta^*(\beta)$ se puede definir como:

$$V(\theta^*(\beta)) = V(\bar{\theta}) - \frac{\mathrm{Cov}(\bar{\tau}; \bar{\theta})^2}{V(\bar{\tau})} = (1 - \rho^2)V(\bar{\theta}); \tag{34}$$

donde ρ es el coeficiente de correlación entre τ_j y t_j. En (34), se observa que si ρ no es nulo hay un aumento de la eficiencia usando variables control.

Cuando el número de observaciones aumente, la correlación entre la variable control y el parámetro objetivo debe aumentar, ya que entonces la distribución finita se aproxima a la distribución asintótica. Como consecuencia, la eficiencia ganada con el uso de variables control es mayor a mayor tamaño muestral n. Como el coste de un experimento de Monte Carlo es proporcional a n, el incremento de eficiencia de la estimación cuando n aumenta permite reducir M al mismo tiempo.

En la mayoría de literatura de variables control [Henry(1984)], β se hace igual a 1. Esta opción parece razonable si τ_j y t_j están altamente correlacionados y tienen similares varianzas. Pero en general no es la mejor elección. Por ello, es preferible estimar λ^*. La forma más fácil de hacerlo es realizar la regresión:

$$t_j = \theta + \lambda \tau_j + u_j. \tag{35}$$

El estimador MCO de λ obtenido al regresar (35) es:

$$\hat{\lambda} = \frac{\sum_{j=1}^{M}(t_j - \bar{t})(\tau_j - \bar{\tau})}{\sum_{j=1}^{M}(\tau_j - \bar{\tau})^2} \tag{36}$$

Sustituyendo en (32) λ por $\hat{\lambda}$ se puede obtener un estimador de θ, que no es insesgado pero que verifica:

$$\sqrt{M}\left(\hat{\theta} - \hat{\lambda}(\bar{\tau} - E(\tau)) - E(\theta)\right) \xrightarrow{d} N(0; V). \tag{37}$$

3.1.2 Diseño del Experimento.

El modelo tomado como referencia en el experimento de Monte Carlo ha sido el siguiente:

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it} + u_i; \tag{38}$$

donde $\alpha = 4$, $\beta = 1$ y:

$$\begin{aligned} \varepsilon_{it} &\sim N(0; \sigma_\varepsilon^2); \\ u_i &\sim N(0; \sigma_u^2). \end{aligned} \tag{39}$$

Las x_{it} se han mantenido fijas en cada uno de los experimentos realizados.

En cada experimento se computan el ECM y el sesgo del estimador MCGF obtenido por los programas antes citados, así como del estimador within. Así mismo, se realiza el recuento en cada iteración de los procedimientos propuestos por LIMDEP y TSP para el calculo de σ_u^2.

Construcción de la variable control. Para realizar el experimento se han utilizado variables control. La variable control elegida para diseñar el experimento es:

$$\tau_j = \Delta \hat{\beta}^B + (I_K - \Delta)\hat{\beta}^W; \tag{40}$$

donde :

$$\Delta = \gamma_0 T \left[\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' + \gamma_0 T \sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right]^{-1} \left[\sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right] \tag{41}$$

y

$$\gamma_0 = \frac{\sigma^2_{0\varepsilon}}{\sigma^2_{0\varepsilon} + T\sigma^2_{0u}}; \tag{42}$$

siendo $\sigma^2_{0\varepsilon}$ y σ^2_{0u} los verdaderos valores de la varianza. La variable τ_j definida en (40) tiene una media conocida e igual a β, varianza finita y está correlacionada con los estimadores estudiados, por lo que cumple las propiedades descritas en el apartado anterior.

Utilizando la variable control, el estimador final se obtiene a partir de la siguiente formula:

$$\bar{\beta}^*(\lambda) = \bar{\beta} - \lambda(\bar{\tau} - E(\tau)); \tag{43}$$

donde:

$$\bar{\beta} = \frac{1}{M} \sum_{j=1}^{M} t_j; \tag{44}$$

La estimación de λ se ha utilizado aplicando (36).

3.1.3 Resultados del Experimento de Montecarlo.

El experimento se ha realizado para cuatro tamaños muestrales. Por una parte, con $T = 5$; se han tomado distintos tamaños de individuos $N = 20; 40; 100$: Asimismo, para $T = 10$; se ha considerado $N = 20$: El número de iteraciones aplicado a estos cuatro casos ha sido de 200, 150, 100 y 150 respectivamente. Aunque estas cifras pueden parecer no muy elevadas, hay que tener en cuenta que la utilización de variables control eleva más de 100 veces la eficiencia del experimento. Para cada uno de los tamaños muestrales se han considerado distintos coeficientes de determinación, $R^2_{Pob} = 0{:}99; 0{:}95; 0{:}9; 0{:}8$: Dado un R^2_{Pob}; y para completar la especificación, se han tomado distintas ratios entre σ^2_ε y σ^2_u; concretamente:

$$\frac{\sigma^2_\varepsilon}{\sigma^2_u} = \frac{1}{20}; \frac{1}{10}; \frac{1}{5}; 1; 5; 10; 20: \tag{45}$$

Las tablas 1 a 4 del anexo recogen el porcentaje de casos en los que se aplican cada uno de los procedimientos descritos en la sección anterior de los programas LIMDEP y TSP para la estimación de θ, para los diferentes tamaños de N y T considerados

Como se puede observar en las tablas 1 a 4, el porcentaje que utilizan los estimadores de la varianza del efecto individual según (10) en TSP (T1) y según (17) en el LIMDEP (L1), depende fundamentalmente de la ratio σ^2_μ/σ^2_u. Cuando éste es mayor o igual a 5, tanto TSP como LIMDEP obtienen porcentajes sustanciales de resultados negativos para σ^2_u. Para una proporción menor no se observan resultados negativos en ninguno de los casos estudiados. Conviene señalar que en todos los casos estudiados el segundo procedimiento propuesto por el LIMDEP (L2) se aplica a un número reducido de casos.

Al aumentar N (ver tablas $1, 2, 3$) y al aumentar T (ver Tablas $2, 4$), el porcentaje de los casos en los que (10) y (17) ofrecen resultados negativos disminuye, acentuándose este hecho cuando $\sigma^2_\mu/\sigma^2_u = 5$.

En los cuadros 1 a 4 se recogen los resultados del experimento. A los estimadores MCGF se les denomina según el procedimiento por el que han sido obtenidos que hace referencia al programa utilizado. Al estimador de efectos fijos para el que se aplica la misma formula en los tres programas se le denomina WITHIN.

En los cuadros 1 y 2 se muestra una clasificación de los estimadores, ordenados de mejores a peores resultados para $N = 20, 40, 100$ y $T = 5$, considerando $R^2_{Pob} = 0.99$. Esta ordenación se mantiene para todos los R^2_{Pob}, pues se ha observado que el cambio en la bondad del ajuste afecta al nivel de sesgo y de ECM pero no a la clasificación relativa de los estimadores. De los datos se desprende que los estimadores que proporcionan un menor sesgo y un mayor ECM son TSP y LIMDEP. El estimador proporcionado por RATS presenta un mayor sesgo y mayor ECM que los estimadores obtenidos en LIMDEP y TSP, mientrás que el estimador WITHIN es el que peores resultados ofrece.

Para tamaños pequeños de la muestra, $N = 20$, el estimador que ofrece un menor sesgo y un menor ECM es el obtenido en TSP, aunque cabe destacar que las diferencias entre los estimadores son muy pequeñas.

Conforme aumenta N, el programa que ofrece mejores resultados es el LIMDEP, sobre todo para $N = 40$.

A mayor ratio σ^2_μ/σ^2_u se acentúan las diferencias de LIMDEP y TSP con respecto a RATS y WITHIN. Asimismo, TSP ofrece mejores resultados que LIMDEP.

Cuadro 1.

Clasificación de los estimadores obtenidos

por los programas TSP, LIMDEP y RATS según su sesgo para

$T = 5$ y $N = 20; 40; 100$:

$R^2_{Pob} = 0.99$:

$T = 5$	$N = 20$	$N = 40$	$N = 100$
$\frac{\sigma_\eta^2}{\sigma_u^2} = \frac{1}{20}$	TSP; LIMDEP; $(3.49\cdot10^{-8})(3.71\cdot10^{-8})$	RATS; LIMDEP; $(4.44\cdot10^{-7})(4.71\cdot10^{-7})$	RATS; WITHIN; $(-3.4\cdot10^{-8})(-4.1\cdot10^{-8})$
	WITHIN; RATS: $(4.05\cdot10^{-8})(7.10\cdot10^{-8})$	TSP; WITHIN: $(4.72\cdot10^{-7})(4.97\cdot10^{-7})$	LIMDEP; TSP: $(-4.3\cdot10^{-8})(-4.4\cdot10^{-8})$
$\frac{\sigma_\eta^2}{\sigma_u^2} = \frac{1}{10}$	TSP; LIMDEP; $(7.79\cdot10^{-7})(7.80\cdot10^{-7})$	LIMDEP; TSP; $(7.37\cdot10^{-8})(7.6\cdot10^{-8})$	LIMDEP; TSP $(-6.1\cdot10^{-7})(-6.2\cdot10^{-7})$
	RATS; WITHIN: $(8.22\cdot10^{-7})(8.76\cdot10^{-7})$	RATS; WITHIN: $(-8.5\cdot10^{-8})(-1.1\cdot10^{-7})$	RATS; WITHIN: $(-6.5\cdot10^{-7})(-6.7\cdot10^{-7})$
$\frac{\sigma_\eta^2}{\sigma_u^2} = \frac{1}{5}$	TSP; LIMDEP; $(1.36\cdot10^{-6})(1.37\cdot10^{-6})$	LIMDEP; TSP; $(-8.9\cdot10^{-7})(-9.1\cdot10^{-7})$	LIMDEP; TSP; $(-3.2\cdot10^{-7})(-3.3\cdot10^{-7})$
	RATS; WITHIN: $(1.54\cdot10^{-6})(1.65\cdot10^{-6})$	RATS; WITHIN: $(-1.1\cdot10^{-6})(-1.1\cdot10^{-6})$	RATS; WITHIN: $(-4.0\cdot10^{-7})(-4.1\cdot10^{-7})$
$\frac{\sigma_\eta^2}{\sigma_u^2} = 1$	RATS; WITHIN; $(-1.3\cdot10^{-6})(-1.5\cdot10^{-6})$	LIMDEP; TSP; $(-4.2\cdot10^{-7})(-6.6\cdot10^{-7})$	LIMDEP; TSP; $(-3.4\cdot10^{-7})(-3.9\cdot10^{-7})$
	TSP; LIMDEP: $(-2.5\cdot10^{-6})(-2.6\cdot10^{-6})$	RATS; WITHIN: $(-6.1\cdot10^{-6})(-7.2\cdot10^{-6})$	WITHIN; RATS: $(-1.1\cdot10^{-6})(-1.2\cdot10^{-6})$
$\frac{\sigma_\eta^2}{\sigma_u^2} = 5$	TSP; LIMDEP; $(-1.3\cdot10^{-6})(-1.8\cdot10^{-6})$	LIMDEP; TSP; $(4.9\cdot10^{-7})(-2.2\cdot10^{-6})$	TSP; LIMDEP; $(-7.1\cdot10^{-7})(-1.8\cdot10^{-6})$
	RATS; WITHIN: $(5.79\cdot10^{-6})(7.70\cdot10^{-6})$	RATS; WITHIN: $(-6.3\cdot10^{-6})(-8.9\cdot10^{-6})$	RATS; WITHIN: $(-2.6\cdot10^{-6})(-2.8\cdot10^{-6})$
$\frac{\sigma_\eta^2}{\sigma_u^2} = 10$	TSP; RATS; $(-1.0\cdot10^{-6})(-1.5\cdot10^{-6})$	LIMDEP; RATS; $(-6.3\cdot10^{-9})(-1.0\cdot10^{-7})$	TSP; LIMDEP; $(-5.1\cdot10^{-7})(-8.9\cdot10^{-7})$
	WITHIN; LIMDEP: $(-2.5\cdot10^{-6})(-2.7\cdot10^{-6})$	WITHIN; TSP: $(-1.6\cdot10^{-6})(2.4\cdot10^{-6})$	RATS; WITHIN: $(5.57\cdot10^{-6})(7.3\cdot10^{-6})$
$\frac{\sigma_\eta^2}{\sigma_u^2} = 20$	LIMDEP; TSP; $(-1.2\cdot10^{-6})(-3.4\cdot10^{-6})$	WITHIN; RATS; $(1.90\cdot10^{-7})(2.9\cdot10^{-7})$	TSP; LIMDEP; $(-2.7\cdot10^{-6})(-2.9\cdot10^{-6})$
	RATS; WITHIN: $(-0.00001)(-0.00002)$	LIMDEP; TSP: $(-2.2\cdot10^{-6})(-3.2\cdot10^{-6})$	RATS; WITHIN: $(0.000012)(0.000014)$

Cuadro 2.

Clasificación de los estimadores obtenidos por los programas TSP, LIMDEP y RATS según su ECM para $T = 5$ y $N = 20, 40, 100$.

$R^2_{Pob} = 0.99$.

$T = 5$	$N = 20$	$N = 40$	$N = 100$
$\frac{\sigma_\varepsilon^2}{\sigma_u^2} = \frac{1}{20}$	RATS, LIMDEP, $(1.50\,10^{-13})(1.71\,10^{-13})$ TSP, WITHIN. $(1.72\,10^{-13})(1.92\,10^{-13})$	RATS, LIMDEP, $(3.64\,10^{-13})(3.71\,10^{-13})$ TSP, WITHIN. $(3.72\,10^{-13})(3.89\,10^{-13})$	TSP, LIMDEP, $(5.25\,10^{-14})(5.261\,10^{-14})$ RATS, WITHIN. $(5.54\,10^{-14})(5.811\,10^{-14})$
$\frac{\sigma_\varepsilon^2}{\sigma_u^2} = \frac{1}{10}$	LIMDEP, TSP, $(1.13\,10^{-12})(1.14\,10^{-12})$ RATS, WITHIN. $(1.18\,10^{-12})(1.27\,10^{-12})$	LIMDEP, TSP, $(4.45\,10^{-13})(4.48\,10^{-13})$ RATS, WITHIN. $(4.97\,10^{-13})(5.51\,10^{-13})$	LIMDEP, TSP, $(5.75\,10^{-13})(5.76\,10^{-13})$ RATS, WITHIN. $(6.12\,10^{-13})(6.271\,10^{-13})$
$\frac{\sigma_\varepsilon^2}{\sigma_u^2} = \frac{1}{5}$	LIMDEP, TSP, $(3.20\,10^{-12})(3.23\,10^{-12})$ RATS, WITHIN. $(3.74\,10^{-12})(4.03\,10^{-12})$	LIMDEP, TSP, $(1.92\,10^{-12})(1.98\,10^{-12})$ RATS, WITHIN. $(2.47\,10^{-12})(2.65\,10^{-12})$	LIMDEP, TSP, $(5.85\,10^{-13})(5.90\,10^{-13})$ RATS, WITHIN. $(8.09\,10^{-13})(8.801\,10^{-13})$
$\frac{\sigma_\varepsilon^2}{\sigma_u^2} = 1$	TSP, LIMDEP, $(9.4\,10^{-12})(1.31\,10^{-11})$ RATS, WITHIN. $(1.9\,10^{-11})(2.81\,10^{-11})$	LIMDEP, TSP, $(1.03\,10^{-12})(1.28\,10^{-12})$ RATS, WITHIN. $(8.52\,10^{-12})(1.27\,10^{-11})$	LIMDEP, TSP, $(3.22\,10^{-13})(3.50\,10^{-13})$ RATS, WITHIN. $(5.22\,10^{-12})(8.09\,10^{-12})$
$\frac{\sigma_\varepsilon^2}{\sigma_u^2} = 5$	TSP, LIMDEP, $(1.51\,10^{-11})(2.20\,10^{-11})$ RATS, WITHIN. $(4.81\,10^{-11})(9.79\,10^{-11})$	LIMDEP, TSP, $(1.75\,10^{-11})(1.92\,10^{-11})$ RATS, WITHIN. $(4.11\,10^{-11})(7.58\,10^{-11})$	TSP, LIMDEP, $(5.87\,10^{-12})(9.29\,10^{-12})$ RATS, WITHIN. $(1.10\,10^{-11})(2.47\,10^{-11})$
$\frac{\sigma_\varepsilon^2}{\sigma_u^2} = 10$	TSP, LIMDEP, $(7.8\,10^{-12})(1.41\,10^{-11})$ RATS, WITHIN. $(7.3\,10^{-11})(1.41\,10^{-10})$	LIMDEP, TSP, $(7.23\,10^{-12})(1.28\,10^{-11})$ RATS, WITHIN. $(7.65\,10^{-11})(1.34\,10^{-10})$	TSP, LIMDEP, $(3.64\,10^{-12})(3.91\,10^{-12})$ RATS, WITHIN. $(2.61\,10^{-11})(4.78\,10^{-11})$
$\frac{\sigma_\varepsilon^2}{\sigma_u^2} = 20$	LIMDEP, TSP, $(9.7\,10^{-12})(2.4\,10^{-11})$ RATS, WITHIN. $(1.35\,10^{-10})(2.1\,10^{-10})$	LIMDEP, TSP, $(1.11\,10^{-11})(1.97\,10^{-11})$ RATS, WITHIN. $(1.02\,10^{-10})(1.706\,10^{-10})$	TSP, LIMDEP, $(1.13\,10^{-11})(1.17\,10^{-11})$ RATS, WITHIN. $(5.56\,10^{-11})(8.98\,10^{-11})$

En los Cuadros 3 y 4 los resultados para $N = 20$ y $T = 10$; considerando $R^2_{Pob} = 0.99$. Se puede observar que los estimadores presentan un menos sesgo y un menor ECM con respecto a los obtenidos para el mismo tamaño muestral, pero con $T = 5$ y $N = 40$.

Al igual que en los casos anteriores, LIMDEP y TSP ofrecen mejores resultados. Cuando la ratio σ^2_η/σ^2_u es menor que 1, LIMDEP presenta un menor ECM.

Los resultados ofrecidos por el programa RATS y el estimador WITHIN mejoran sensiblemente en términos absolutos pero no relativos con respecto a los otros dos tipos de estimadores si se comparan estos resultados con los obtenidos para $T = 5$.

Cuadro 3.

Clasificación de los estimadores obtenidos por los programas TSP, LIMDEP y RATS según su sesgo para $T = 10$ y $N = 20$.

$R^2_{Pob} = 0.99$:

	$T = 10$ $\quad$ $N = 20$
$\frac{\sigma^2_\eta}{\sigma^2_u} = \frac{1}{20}$	LIMDEP;TSP;RATS;WITHIN: $(-1.3\,10^{-8})(-1.4\,10^{-8})(-2.1\,10^{-8})(-2.2\,10^{-8})$
$\frac{\sigma^2_\eta}{\sigma^2_u} = \frac{1}{10}$	LIMDEP;TSP;RATS;WITHIN: $(-4.9\,10^{-7})(-5.0\,10^{-7})(-5.7\,10^{-7})(-5.9\,10^{-7})$
$\frac{\sigma^2_\eta}{\sigma^2_u} = \frac{1}{5}$	TSP;LIMDEP;WITHIN;RATS: $(2.2\,10^{-7})(2.3\,10^{-7})(2.7\,10^{-7})(2.9\,10^{-7})$
$\frac{\sigma^2_\eta}{\sigma^2_u} = 1$	TSP;LIMDEP;RATS;WITHIN: $(-9.0\,10^{-7})(1.0\,10^{-6})(1.1\,10^{-6})(1.3\,10^{-6})$
$\frac{\sigma^2_\eta}{\sigma^2_u} = 5$	RATS;LIMDEP;WITHIN;TSP: $(2.1\,10^{-7})(-6.5\,10^{-7})(-8.0\,10^{-7})(-9.6\,10^{-7})$
$\frac{\sigma^2_\eta}{\sigma^2_u} = 10$	TSP;LIMDEP;WITHIN;RATS: $(9.9\,10^{-7})(3.1\,10^{-6})(-4.5\,10^{-6})(-4.7\,10^{-6})$
$\frac{\sigma^2_\eta}{\sigma^2_u} = 20$	LIMDEP;TSP;RATS;WITHIN: $(-2.5\,10^{-7})(-1.9\,10^{-6})(-8.7\,10^{-6})(-9.2\,10^{-6})$

Cuadro 4.

Clasificación de los estimadores obtenidos por los programas TSP, LIMDEP y RATS según su ECM para $T = 10$ y $N = 20$:

$R^2_{Pob} = 0.99$:

	$T = 10$ $N = 20$
$\frac{\sigma^2_\varepsilon}{\sigma^2_u} = \frac{1}{20}$	LIMDEP; TSP; RATS; WITHIN: $(3.01 \cdot 10^{-14})(3.03 \cdot 10^{-14})(3.15 \cdot 10^{-14})(3.36 \cdot 10^{-14})$
$\frac{\sigma^2_\varepsilon}{\sigma^2_u} = \frac{1}{10}$	LIMDEP; TSP; RATS; WITHIN: $(3.41 \cdot 10^{-13})(3.48 \cdot 10^{-13})(3.64 \cdot 10^{-13})(3.65 \cdot 10^{-13})$
$\frac{\sigma^2_\varepsilon}{\sigma^2_u} = \frac{1}{5}$	LIMDEP; TSP; RATS; WITHIN: $(2.80 \cdot 10^{-13})(2.83 \cdot 10^{-13})(4.10 \cdot 10^{-13})(4.37 \cdot 10^{-13})$
$\frac{\sigma^2_\varepsilon}{\sigma^2_u} = 1$	TSP; LIMDEP; RATS; WITHIN: $(1.31 \cdot 10^{-12})(1.67 \cdot 10^{-12})(3.58 \cdot 10^{-12})(4.53 \cdot 10^{-12})$
$\frac{\sigma^2_\varepsilon}{\sigma^2_u} = 5$	RATS; WITHIN; TSP; LIMDEP: $(8.26 \cdot 10^{-12})(1.27 \cdot 10^{-11})(1.40 \cdot 10^{-11})(1.62 \cdot 10^{-11})$
$\frac{\sigma^2_\varepsilon}{\sigma^2_u} = 10$	TSP; LIMDEP; RATS; WITHIN: $(7.23 \cdot 10^{-12})(1.90 \cdot 10^{-11})(2.55 \cdot 10^{-11})(4.56 \cdot 10^{-11})$
$\frac{\sigma^2_\varepsilon}{\sigma^2_u} = 20$	LIMDEP; TSP; RATS; WITHIN: $(3.48 \cdot 10^{-12})(6.65 \cdot 10^{-12})(4.91 \cdot 10^{-11})(6.21 \cdot 10^{-11})$

Concluyendo este apartado, se puede decir que el programa LIMDEP es el que mejores resultados globales ofrece. Estos resultados mejoran con respecto a los obtenidos en TSP cuando aumenta el número de observaciones individuales. TSP proporciona buenos resultados tanto en sesgo como en ECM cuando la ratio $\sigma^2_\varepsilon/\sigma^2_u$ es mayor que 1. El programa RATS parece mejorar al aumentar el tamaño muestral.

3.2 Experimentos de Montecarlo en modelos con correlación.

3.2.1 Diseño del Experimento.

La variable endógena sigue la siguiente modelización:

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it} + u_i; \quad (46)$$

$$u_i = a\overline{x}_i + b_i; \quad (47)$$

donde $\mu = 4$, $\beta = 10$, $a = 2$ y :

$$\begin{aligned} \varepsilon_{it} &\sim N(0;\sigma^2_\varepsilon); \\ b_i &\sim N(0;\sigma^2_b). \end{aligned} \tag{48}$$

Las x_{it} están correlacionadas con el efecto individual u_i; según la formulación propuesta por Mundlak (1978). El experimento se ha realizado para distintos niveles de correlación, en concreto, para $\rho = 0.1; 0.15; 0.2; 0.25; 0.3; 0.35.$

En cada experimento se han computado los valores del ECM y el sesgo del estimador $\hat{\beta}^{MCGF}$ obtenido a partir de los distintos programas antes citados, así como de $\hat{\beta}^{W}$. Adicionalmente, se recoge el resultado del test de Hausman ofrecido por LIMDEP y TSP.

El objetivo de este experimento es observar si los resultados obtenidos en el apartado anterior se mantienen para niveles de correlación relativamente bajos. Además, se recoge el test de Hausman. En estos experimentos no se ha utilizado la técnica de variables control, ya que la variable control (40) es sesgada en este caso.

3.2.2 Resultados del Experimento de Montecarlo.

El experimento se ha realizado para dos tamaños muestrales. Los resultados se presentan manteniendo fijo el número de observaciones temporales, $T = 5$ y variando el número de individuos, $N = 20; 40.$ Al no utilizar variables control se han computado 1000 y 900 iteraciones respectivamente. Para cada uno de los tamaños muestrales se han considerado diferentes niveles de correlación, $\rho = 0.1; 0.15; 0.2; 0.25; 0.3; 0.35.$ Para cada valor de ρ; se han considerado las siguientes proporciones:

$$\frac{\sigma^2_\varepsilon}{\sigma^2_u} = \frac{1}{5}; 1; 5. \tag{49}$$

En este experimento se ha calculado el test de Hausman, así como las estimaciones del modelo aplicadas en la sección anterior. El Cuadro 5 muestra el porcentaje de aceptación de H_0; definida en (25) en el test de Hausman para $N = 20$ y $T = 5$ y para $N = 40$.

Cuadro 5.

Porcentaje de aceptación de la H_0 obtenido en el test de Hausman en LIMDEP y TSP para N = 20; 40 y T = 5 para un nivel de signi…catividad de 0.05.

½ = 0; 0:1; 0:15; 0:2; 0:25; 0:3; 0:35:

(Porcentaje).

		½=0		½=0:1		½=0:15		½=0:20		½=0:25		½=0:3		½=0:35	
		HT[1]	HL[2]	HT	HL	HT	HL	HT	HL	HT	HL	HT	HL	HT	HL
N = 20	$\frac{¾^2_\eta}{¾^2_u} = \frac{1}{5}$	94:6	93:3	81:0	78:4	62:6	59:7	36:7	33:7	13:9	12:5	1:2	1:0	0:0	0:0
	$\frac{¾^2_\eta}{¾^2_u} = 1$	94:2	93:3	84:9	82:6	68:4	66:4	46:0	42:2	19:8	17:7	4:1	3:7	0:0	0:0
	$\frac{¾^2_\eta}{¾^2_u} = 5$	93:6	92:5	89:1	86:0	79:6	77:1	66:6	63:3	45:4	41:0	22:7	19:4	5:4	3:2
N = 40	$\frac{¾^2_\eta}{¾^2_u} = \frac{1}{5}$	94:5	94:0	72:4	71:1	44:3	42:1	14:4	13:8	2:0	1:7	0:0	0:0	0:0	0:0
	$\frac{¾^2_\eta}{¾^2_u} = 1$	94:5	94:0	76:0	74:5	53:1	51:5	24:1	23:1	5:2	4:7	0:1	0:1	0:0	0:0
	$\frac{¾^2_\eta}{¾^2_u} = 5$	94:6	94:4	84:5	83:3	69:6	68:1	47:5	44:5	20:4	18:8	5:2	4:7	0:0	0:0

[1] HT se re…ere a los resultados del test de Hausman según los procedimientos utilizados en el programa TSP.

[2] HL se re…ere a los resultados del test de Hausman según los procedimientos utilizados en el programa LIMDEP.

Como se puede observar, conforme aumenta ½, el porcentaje de aceptación de la H_0; es decir, de la hipótesis nula de no correlación entre el regresor y las perturbaciones, es mayor. A partir de ½ = 0:25 cuando N = 20 y ½ = 0:2 si N = 40 el test de Hausman rechaza en más de un 50 por 100 de los casos la H_0. Cabe resaltar que si $¾^2_\eta = ¾^2_u = 5$, la H_0 se acepta en más casos que para $¾^2_\eta = ¾^2_u = 1=5; 1$:

Cuando N aumenta, el porcentaje de aceptación de H_0 disminuye para cada nivel de correlación y tamaño de la varianza.

Con respecto a las estimaciones del modelo por distintos procedimientos, en los cuadros 6 y 7, se muestra el sesgo y el ECM para T = 5, N = 20 y T = 5, N = 40, considerando ½ = 0; 0:1; 0:15 y 0:2: Al igual que en el apartado anterior a los estimadores MCGF se les denomina por el procedimiento que han sido obtenidos y al estimador de efectos …jos se le denomina WITHIN. No se han recogido los resultados de los estimadores para valores de ½ superiores a 0.2, pues como ya se ha resaltado, el test de Hausman rechaza de forma abrumadora la H_0 y el mejor estimador es siempre el WITHIN.

Cuadro 6.

Clasificación de los estimadores obtenidos por los programas TSP, LIMDEP y RATS según su sesgo para $T = 5; N = 20$ y $N = 40$.

(Porcentaje).

$\rho = 0; 0.1; 0.15; 0.2.$

		$\rho = 0$	$\rho = 0.1$	$\rho = 0.15$	$\rho = 0.2$
$N = 20$	$\frac{\sigma_\eta^2}{\sigma_u^2} = \frac{1}{5}$	WITHIN;RATS; (−0.0010)(−0.0011) TSP;LIMDEP: (−0.0013)(−0.0014)	RATS;WITHIN; (−0.0022)(−0.004) TSP;LIMDEP: (0.021)(0.023)	WITHIN;RATS; (−0.0026)(0.0039) TSP;LIMDEP: (0.0219)(0.0256)	WITHIN;RATS; (−0.0018)(0.0051) TSP;LIMDEP: (0.0202)(0.0226)
	$\frac{\sigma_\eta^2}{\sigma_u^2} = 1$	LIMDEP;TSP; (0.0053)(0.0055) RATS;WITHIN: (0.0062)(0.0064)	WITHIN;RATS; (0.0260)(0.004) TSP;LIMDEP: (0.1315)(0.1405)	WITHIN;RATS; (0.0161)(0.0385) TSP;LIMDEP: (0.1186)(0.1330)	WITHIN;RATS; (0.0114)(0.0332) TSP;LIMDEP: (0.1070)(0.1291)
	$\frac{\sigma_\eta^2}{\sigma_u^2} = 5$	TSP;LIMDEP; (0.0056)(0.0057) WITHIN;RATS: (0.0062)(0.0066)	WITHIN;RATS; (0.0260)(0.0857) TSP;LIMDEP: (0.3014)(0.3297)	WITHIN;RATS; (0.0160)(0.0760) TSP;LIMDEP: (0.2802)(0.3232)	WITHIN;RATS; (0.0119)(0.0697) TSP;LIMDEP: (0.2757)(0.3197)
$N = 40$	$\frac{\sigma_\eta^2}{\sigma_u^2} = \frac{1}{5}$	TSP;LIMDEP; (8.3210^{-6})(0.00002) RATS;WITHIN: (−0.0001)(−0.0002)	WITHIN;RATS; (−0.0006)(0.0035) TSP;LIMDEP: (−0.0178)(0.0196)	WITHIN;RATS; (−0.0004)(0.0037) TSP;LIMDEP: (0.0169)(0.0191)	WITHIN;RATS; (−0.0003)(0.0039) TSP;LIMDEP: (0.0156)(0.0191)
	$\frac{\sigma_\eta^2}{\sigma_u^2} = 1$	RATS;WITHIN; (−0.0059)(−0.0060) TSP;LIMDEP: (−0.0062)(−0.0063)	RATS;WITHIN; (−0.007)(−0.023) TSP;LIMDEP: (0.0522)(0.0569)	RATS;WITHIN; (0.00025)(−0.014) TSP;LIMDEP: (0.0574)(0.0656)	RATS;WITHIN; (−0.0040)(0.0120) TSP;LIMDEP: (0.0576)(0.0701)
	$\frac{\sigma_\eta^2}{\sigma_u^2} = 5$	WITHIN;RATS; (−0.007)(−0.008) TSP;LIMDEP; (−0.0102)(−0.0104)	WITHIN;RATS; (0.0298)(0.0746) TSP;LIMDEP: (−0.2517)(−0.2633)	WITHIN;RATS; (0.0192)(0.0625) TSP;LIMDEP: (0.2323)(−0.2488)	WITHIN;RATS; (−0.0136)(0.055) TSP;LIMDEP: (0.2172)(0.2414)

Cuadro 7.

Clasificación de los estimadores obtenidos por los programas TSP, LIMDEP y RATS según su ECM para $T = 5; N = 20$ y $N = 40$:

(Porcentaje).

$\rho = 0; 0.1; 0.15; 0.2$:

		$\rho = 0$	$\rho = 0.1$	$\rho = 0.15$	$\rho = 0.2$
		WITHIN;RATS;	RATS;WITHIN;	WITHIN;RATS;	WITHIN;RATS;
$N = 20$	$\frac{\sigma_\mu^2}{\sigma_u^2} = \frac{1}{5}$	$(4.00\,10^{-6})(4.00\,10^{-6})$	(0.00004)(0.00006)	(0.00002)(0.00003)	(0.00001)(0.00003)
		TSP;LIMDEP:	TSP;LIMDEP:	TSP;LIMDEP:	TSP;LIMDEP:
		$(4.93\,10^{-6})(5.02\,10^{-6})$	(0.0005)(0.0006)	(0.0004)(0.0006)	(0.0004)(0.0007)
		TSP;LIMDEP;	WITHIN;RATS;	WITHIN;RATS;	WITHIN;RATS;
	$\frac{\sigma_\mu^2}{\sigma_u^2} = 1$	(0.00004)(0.00004)	(0.0008)(0.002)	(0.0003)(0.0015)	(0.0001)(0.0011)
		RATS;WITHIN:	TSP;LIMDEP:	TSP;LIMDEP:	TSP;LIMDEP:
		(0.00005)(0.00006)	(0.0175)(0.0199)	(0.0141)(0.0177)	(0.011)(0.016)
		TSP;LIMDEP;	WITHIN;RATS;	WITHIN;RATS;	WITHIN;RATS;
	$\frac{\sigma_\mu^2}{\sigma_u^2} = 5$	(0.00009)(0.0001)	(0.0016)(0.0082)	(0.0006)(0.006)	(0.00034)(0.0049)
		RATS;WITHIN:	TSP;LIMDEP:	TSP;LIMDEP:	TSP;LIMDEP:
		(0.00010)(0.00011)	(0.091)(0.1096)	(0.083)(0.104)	(0.0762)(0.1024)
		RATS;WITHIN;	WITHIN;RATS;	WITHIN;RATS;	WITHIN;RATS;
$N = 40$	$\frac{\sigma_\mu^2}{\sigma_u^2} = \frac{1}{5}$	$(1.01\,10^{-6})(1.41\,10^{-6})$	(0.00002)(0.00003)	$(9.1\,10^{-6})$(0.0001)	$(4.0\,10^{-6})$(0.00001)
		LIMDEP;TSP;	TSP;LIMDEP:	TSP;LIMDEP:	TSP;LIMDEP:
		$(1.5\,10^{-6})(1.6\,10^{-6})$	(0.00034)(0.00039)	(0.0003)(0.0003)	(0.0002)(0.0003)
		TSP;LIMDEP;	RATS;WITHIN;	RATS;WITHIN	RATS;WITHIN
	$\frac{\sigma_\mu^2}{\sigma_u^2} = 1$	(0.00004)(0.00004)	(0.0001)(0.0006)	(0.00004)(0.0002)	(0.00003)(0.00012)
		RATS;WITHIN:	TSP;LIMDEP:	TSP;LIMDEP:	TSP;LIMDEP:
		(0.00005)(0.00005)	(0.0028)(0.0033)	(0.0033)(0.0043)	(0.0033)(0.0049)
		WITHIN;RATS;	WITHIN;RATS;	WITHIN;RATS;	WITHIN;RATS;
	$\frac{\sigma_\mu^2}{\sigma_u^2} = 5$	(0.00008)(0.0001)	(0.001)(0.0059)	(0.0005)(0.004)	(0.0002)(0.003)
		TSP;LIMDEP:	TSP;LIMDEP:	TSP;LIMDEP:	TSP;LIMDEP:
		(0.00010)(0.00011)	(0.0638)(0.0698)	(0.054)(0.0622)	(0.047)(0.058)

Como se puede observar, aunque ρ tome valores pequeños y el test de Hausman acepte la H_0, el estimador que presenta en la mayoria de los casos menor sesgo y menor ECM es el estimador WITHIN. El estimador proporcionado por el RATS es siempre más robusto que los estimadores obtenidos en LIMDEP y TSP. Además, ofrece mejores resultados que el estimador WITHIN cuando aumenta el tamaño de la muestra. Los estimadores propuestos por LIMDEP y TSP presentan siempre peores resultados.

Las diferencias de sesgo y ECM del estimador WITHIN con respecto a LIMDEP y TSP aumentan para $N = 40$; y se acentúan conforme aumenta el nivel de correlación. Este hecho también se observa en los resultados del test de Hausman, pues conforme aumenta N se obtiene un porcentaje más alto de rechazo de la H_0:

Concluyendo, se puede afirmar que cuando existe correlación entre la variable explicativa y el efecto individual los resultados son totalmente distintos a los que se obtienen en caso de que se presente dicha correlación. Si se analizan los resultados para $\rho = 0$; se observa que los mejores estimadores son los utilizados por LIMDEP y TSP, como ya se vió en el apartado anterior.

4 CONCLUSIONES.

En este artículo se ha realizado un estudio comparativo mediante utilización de experimentos de Montecarlo de los distintos estimadores MCGF propuestos por los programas más utilizados en el contexto de datos de panel, (LIMDEP, RATS y TSP), así como del estimador WITHIN. En primer lugar se han realizado experimentos de Montecarlo tomando como referencia modelos en los que el regresor es no estocástico. En segundo lugar la variable endógena se ha generado a partir de modelos con correlación entre el efecto individual y los regresores. Adicionalmente, en este segundo caso se ha calculado el test de Hausman propuesto por LIMDEP y TSP con una doble finalidad: detectar a partir de que niveles de correlación se rechaza la hipótesis nula y por lo tanto, el estimador de MCGF no sería el más adecuado y observar si estos programas ofrecen resultados anómalos del citado contraste. Los experimentos de Montecarlo se han realizado utilizando variables control, con el fín de incrementar la eficiencia del experimento.

Si no existe correlación entre los regresores y el efecto individual, se observa que LIMDEP ofrece estimadores con un menor sesgo y ECM en casi todos los casos estudiados. Los resultados mejoran cuando aumenta el tamaño de la muestra, acentuándose este hecho si se incrementa el número de observaciones por individuos. El estimador propuesto por TSP es específicamente mejor, cuando la ratio σ^2_η/σ^2_u. El estimador ofrecido por RATS y el WITHIN ofrece peores resultados que los anteriores, aunque mejoran relativamente al aumentar el tamaño de la muestra.

Cuando se relaja el supuesto de exogeneidad, cabe destacar que a partir de $\rho = 0{:}25$ cuando $N = 20$

y ½ = 0:2 si N = 40 el test de Hausman rechaza en más de un 50 por 100 de los casos la H_0. Aunque el coe…ciente de correlación ½ tome valores más pequeños que los anteriormente citados y el test de Hausman acepte la H_0, el estimador que presenta en la mayoria de los casos menor sesgo y ECM es el estimador WITHIN. El estimador proporcionado por el RATS es siempre más robusto que los estimadores obtenidos en LIMDEP y TSP. Además, ofrece mejores resultados que el estimador WITHIN cuando aumenta el tamaño de la muestra. Los estimadores propuestos por LIMDEP y TSP presentan siempre peores resultados.

Cabe destacar que en ninguno de los experimentos realizados se han detectado anomalías en el cálculo del contraste de Hausman. Sin embargo se ha comprobado que la introducción de la tendencia lineal en el modelo provoca resultados anómalos en los dos programas.

ANEXO

Tabla 1.

Porcentajes de utilización en el TSP y el LIMDEP

de los procedimientos de estimación de σ_u^2 para N = 20 y T = 5:

(Porcentaje).

N = 20 T = 5							
	$\frac{\sigma_\mu^2}{\sigma_u^2}=\frac{1}{20}$	$\frac{\sigma_\mu^2}{\sigma_u^2}=\frac{1}{10}$	$\frac{\sigma_\mu^2}{\sigma_u^2}=\frac{1}{5}$	$\frac{\sigma_\mu^2}{\sigma_u^2}=1$	$\frac{\sigma_\mu^2}{\sigma_u^2}=5$	$\frac{\sigma_\mu^2}{\sigma_u^2}=10$	$\frac{\sigma_\mu^2}{\sigma_u^2}=20$
T1	100:00	100:00	100:00	100:00	67:00	51:50	46:00
T2	0:00	0:00	0:00	0:00	33:00	48:50	54:00
L1	100:00	100:00	100:00	100:00	66:50	54:00	45:50
L2	0:00	0:00	0:00	0:00	3:50	5:00	4:00
L3	0:00	0:00	0:00	0:00	30:00	41:00	50:50

[1] T1 y T2 indican el porcentaje de los experimentos que utilizan (10) y (12) respectivamente.

[2] L1;L2 y L3 indican el porcentaje de los experimentos que utilizan (17), (19) y (20).

Tabla 2.

Porcentajes de utilización en el TSP y el LIMDEP

de los procedimientos de estimación de σ_u^2 para N = 40 y T = 5:

(Porcentaje).

N = 40 T = 5							
	$\frac{\sigma_\mu^2}{\sigma_u^2}=\frac{1}{20}$	$\frac{\sigma_\mu^2}{\sigma_u^2}=\frac{1}{10}$	$\frac{\sigma_\mu^2}{\sigma_u^2}=\frac{1}{5}$	$\frac{\sigma_\mu^2}{\sigma_u^2}=1$	$\frac{\sigma_\mu^2}{\sigma_u^2}=5$	$\frac{\sigma_\mu^2}{\sigma_u^2}=10$	$\frac{\sigma_\mu^2}{\sigma_u^2}=20$
T1	100:00	100:00	100:00	100:00	70:00	58:00	48:66
T2	0:00	0:00	0:00	0:00	30:00	42:00	51:33
L1	100:00	100:00	100:00	100:00	69:33	56:66	49:33
L2	0:00	0:00	0:00	0:00	5:33	4:00	2:66
L3	0:00	0:00	0:00	0:00	25:33	39:33	48:00

Tabla 3.

Porcentajes de utilización en el TSP y el LIMDEP

de los procedimientos de estimación de N = 100 y T = 5:

(Porcentaje).

N = 100 T = 5							
	$\frac{\sigma_\eta^2}{\sigma_u^2}=\frac{1}{20}$	$\frac{\sigma_\eta^2}{\sigma_u^2}=\frac{1}{10}$	$\frac{\sigma_\eta^2}{\sigma_u^2}=\frac{1}{5}$	$\frac{\sigma_\eta^2}{\sigma_u^2}=1$	$\frac{\sigma_\eta^2}{\sigma_u^2}=5$	$\frac{\sigma_\eta^2}{\sigma_u^2}=10$	$\frac{\sigma_\eta^2}{\sigma_u^2}=20$
T1	100:00	100:00	100:00	100:00	86:00	60:00	60:00
T2	0:00	0:00	0:00	0:00	14:00	40:00	40:00
L1	100:00	100:00	100:00	100:00	87:00	59:00	59:00
L2	0:00	0:00	0:00	0:00	2:00	2:00	1:00
L3	0:00	0:00	0:00	0:00	11:00	39:00	40:00

Tabla 4.

Porcentajes de utilización en el TSP y el LIMDEP

de los procedimientos de estimación de N = 20 y T = 10:

(Porcentaje).

N = 20 T = 10							
	$\frac{\sigma_\eta^2}{\sigma_u^2}=\frac{1}{20}$	$\frac{\sigma_\eta^2}{\sigma_u^2}=\frac{1}{10}$	$\frac{\sigma_\eta^2}{\sigma_u^2}=\frac{1}{5}$	$\frac{\sigma_\eta^2}{\sigma_u^2}=1$	$\frac{\sigma_\eta^2}{\sigma_u^2}=5$	$\frac{\sigma_\eta^2}{\sigma_u^2}=10$	$\frac{\sigma_\eta^2}{\sigma_u^2}=20$
T1	100:00	100:00	100:00	100:00	82:00	52:00	45:33
T2	0:00	0:00	0:00	0:00	18:00	48:00	54:66
L1	100:00	100:00	100:00	100:00	84:00	55:33	46:00
L2	0:00	0:00	0:00	0:00	3:33	5:33	8:66
L3	0:00	0:00	0:00	0:00	12:66	39:33	45:33

4.1 REFERENCIAS.

Arellano, M. and Bond, S. (1991), "Some Test of Especi…cation for Panel Data: Monte carlo Evidence and a Application to Employment Equations". The Review of Economic Studies Limited, 58, 277-297.

Bourlange, D. and Doz, C. (1988), "Pseudo-maximum de vraisenblance: experiénces de simulations dans le cadre d'un modéle de Poisson". Annales d'Economie et Statistique, 10, 139-158.

Cornwell, C., Schmidt, P. and Wyhowski, D.(1990), "Simultaneus equations and panel data". Elservier Science Publishers B.V.

Davidson, R. and Mackinnon, J.G. (1992), "Regresion-based methods for using control variates in Monte Carlo experiments." Journal of Econometrics 54, 203-222.

Davidson, R. and Mackinnon, J.G. (1993), Estimation and Inference in Econometrics. Oxford University Press.

Doan, T. A.(1996), RATS, User's Manual Version 4.0. Estima.

Greene, W. H. (1993), Econometric Analysis. MacMillan.

Greene, W. H. (1995), LIMDEP version 7.0 User's Manual. Econometric Software Inc. Bell Port.

Hall, B. H. (1997), Time Series Processor. Version 4.4. TSP International.

Hausman, J.A. (1978), "Speci…cation Test in Econometrics". Econometrica 46, 6, 1251-1271.

Hendry, D.F (1984), Monte Carlo experimentation in econometrics, Handbook of econometrics, Vol. II., Ch.16. Hisao, C. (1986), Analysis of Panel Data. Cambrigde University Press. New York.

Maddala, G.S. and Mount ,T.D., "A comparative study of alternative estimators for variance components used in econometric applications". Journal of the American Statistical Asotiation, 68, 342, 324-328.

Maudós, J. and Uriel , E. (1995)," Análisis comparativo de las estimaciones de modelos de panel en los programas LIMDEP, TSP, RATS". Revista de Economía Aplicada, 11, 275-282.

Moulton., B.R. (1986), "Random group e¤ects and the precision of regresion estimates". Journal of Econometrics 32, 385-397.

Moulton., B.R. (1987), "Diagnostic for group e¤ects in Regresion Analysis". Journal of Business and Economic Statistics, 5, 2, 275-282.

Mundlak, Y. (1978), "On the pooling of time series and cross section data". Econometrica 46, 1, 108-123.

Nerlove (1971), "Futher Evidence on the Estimation of Dinamic Relations from a Time Series of Cross Section Data". Económetrica 39, 359-382.

Swamy,V.B. and Tavlas, G.S. (1995), "Random coe¢cient models: Theory and aplicattions".Journal of Economic Surveys 9, 2, 165-195.