



TAREA 1

Econometría 1

Profesora: VALENTINA PAREDES

Ayudantes: HRIDAY KARNANI & MARÍA JESUS NEGRETE

Grupo: 3

Estudiantes: F. ANGUITA, N. BASTÍAS, C. CID, T. MUÑOZ O. & R. SALAZAR

26 de abril de 2024

Índice

1. Supuestos de MRL	2
2. Estimando Coeficientes Mediante MCO	3
3. Diferentes Mediciones Para el Error Estándar	3
4. Test de Hipótesis Nula β_1	5
5. Modelo con Efectos Fijos	6
6. Equivalencia entre FWL y Variables Transformadas	7
7. Simulación de Nueva Base de Datos	9
8. Aprendizajes Finales	12

1. Supuestos de MRL

El Modelo de Regresión Lineal (MRL) es, inambiguamente, el punto de partida para hacer inferencia estadística y obtener proyecciones de estimadores que tiendan al valor poblacional y capturen el efecto de relaciones lineales entre las variables. No obstante, este modelo debe cumplir ciertas condiciones de regularidad que garantizan sus propiedades. Estos supuestos son:

$$\mathbf{Y} = \beta\mathbf{X} + \mathbf{e}$$

$$\mathbb{E}[\mathbf{e}|\mathbf{X}] = 0$$

$$\mathbb{E}[y^2] < \infty$$

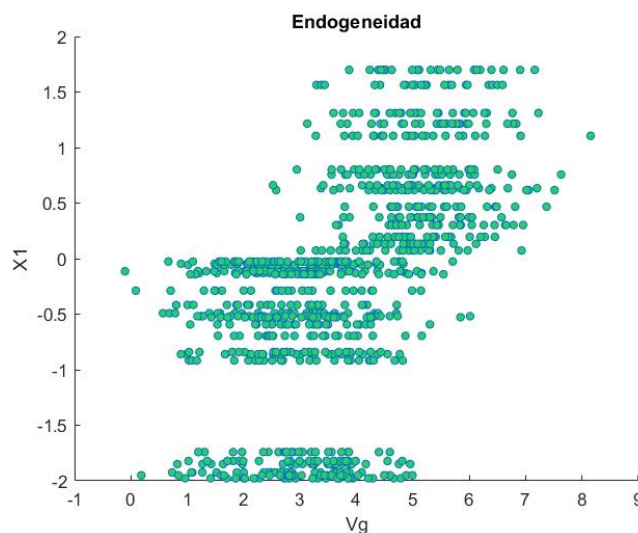
$$\mathbb{E}[\|\mathbf{X}\|^2] < \infty$$

$$\mathbf{Q}_{\mathbf{X}\mathbf{X}} = \mathbb{E}[\mathbf{X}'\mathbf{X}] > 0$$

Para el primer caso, puesto que la definición del modelo a estimar es la ecuación (2.1), se sabe que sí se cumple ya que tiene la misma forma funcional que puede ser expresado de manera matricial.

En cuanto al segundo supuesto, es importante comprender la profundidad de este, puesto que si la esperanza condicional del error en la variable explicativa es cero, significa que no se encuentran correlacionados, lo que implica una estimación sin sesgo sistemático en los errores del modelo. Sin embargo, de acuerdo a la simulación del proceso generador de datos, la distribución de la variable X_1 está en función del valor que tome el error grupal v_g , lo que establece el primer indicio de correlación. Para confirmar dicha situación, se grafican los errores grupales con la variable para verificar la premisa de correlación positiva.

Figura 1.1 Endogeneidad del Modelo de Regresión Lineal



Elaboración propia.

Es evidente, como muestra la Figura 1.1, la existencia de correlación positiva entre ambas variables, situación que causará algún grado de sesgo por endogeneidad. Entonces, se desestima el cumplimiento del supuesto de no relación entre el error y la variable X_1 . Esto se traduce en $E(v_g|X_{1ig}) \neq 0$. Adicionalmente, cabe señalar que sí existe independencia entre los grupos de

la muestra, por lo que si bien el supuesto $E(e_{ig}|X_{1ig}) = 0$ no se cumple a nivel individual, sí se estaría cumpliendo a nivel grupal.

Ahora bien, los supuestos 3 y 4 se cumplen por definición dada la forma en que han sido construidas las variables. Las esperanzas de los X ya se encuentran acotadas y por ello, sus segundos momentos también. De la misma manera se cumple para el Y , dado a que su varianza depende de la distribución de los X y e , lo cual para todos los casos está definido y se pueden estimar.

Finalmente, el hecho que la forma cuadrática de la matriz \mathbf{XX}' sea positiva implica que la matriz en sí misma sea definida positiva, supuesto que garantiza la invertibilidad de esta. Para el caso en cuestión, se cumple el supuesto puesto que no es una matriz singular y sus columnas son linealmente independientes entre sí.

2. Estimando Coeficientes Mediante MCO

Tras la simulación de los datos solicitados, mediante el método de Mínimos Cuadrados Ordinarios (MCO), se estima el siguiente modelo:

$$Y_{ig} = \beta_0 + \beta_1 X_{1ig} + \beta_2 X_{2ig} + \epsilon_{ig} + \nu_g$$

el cual se puede simplificar a lo siguiente, considerando que $e_{ig} = \epsilon_{ig} + \nu_g$:

$$Y_{ig} = \beta_0 + \beta_1 X_{1ig} + \beta_2 X_{2ig} + e_{ig} \quad (2.1)$$

La Tabla 1 a continuación presenta los resultados de las estimación:

Tabla 1 Coeficientes MCO

Coeficiente	(2.1)
β_0	-0.6588
β_1	2.4284
β_2	3.9713

Se observa que los coeficientes difieren del valor real de los parámetros, lo cual sucede debido a un sesgo presente en la estimación. Para el caso de β_0 , el parámetro real asociado es 1, no obstante, los resultados arrojan que posee un valor de -0.6588, lo cual permite ver que, para el caso de este parámetro, hay un sesgo negativo. Por otro lado, el coeficiente β_1 presenta un sesgo similar en términos de magnitud pero contrario en signo (positivo). El parámetro verdadero, en este caso, es igual a 2 y el valor estimado es mayor en 0.4. Por su parte, para el caso de β_2 , este la estimación es bastante cercana al valor verdadero del coeficiente, lo que permite intuir que la variable X_{2ig} no genera tanto sesgo en la estimación. En contraste, la variable X_{1ig} podría estar correlacionada con el término de error y por ello, producir este sesgo. Finalmente, no es de extrañar que esto ocurra, puesto que la variable X_{1ig} por construcción, correlaciona con el término de error ν_g , traducida en la explicando de este sesgo.

3. Diferentes Mediciones Para el Error Estándar

Para corroborar la eficiencia de los coeficientes estimados, se calculan 3 tipos de errores estándar: el primero asumiendo homocedasticidad y ausencia de correlación; el segundo de errores robustos, admitiendo heterocedasticidad por la matriz de *White*, y el tercero de errores clusterizados. Los resultados obtenidos se presentan en la Tabla 2.

Para estimar el error estándar bajo los supuestos de homocedasticidad y correlación igual a 0, se utilizó la estimación presentada por el libro Hansen (2022) correspondiente a la matriz de varianza y covarianza ,

$$\widehat{\text{Var}}_{\hat{\beta}} = (X'X)^{-1}s^2,$$

donde s^2 es el estimador de la varianza. El error estándar está definido como la raíz cuadrada de la diagonal de $\widehat{\text{Var}}_{\hat{\beta}}$, es decir,

$$s(\hat{\beta}) = \sqrt{\widehat{\text{Var}}_{\hat{\beta}_j}}$$

Para los demás estimadores de varianza, se considera al error estándar robusto y *cluster* como la raíz de la diagonal de las matrices varianza-convarianza correspondientes, de manera análoga a lo establecido anteriormente.

Tabla 2 Errores Estándar

Coeficiente	Errores Estándar	Errores Robusto	Errores Clusterizado
β_0	0.2281	0.2310	0.3153
β_1	0.0280	0.0270	0.0573
β_2	0.0398	0.0423	0.0388

Para el caso de los errores robustos, estos fueron calculados considerando el siguiente estimador de la varianza:

$$\widehat{\text{Var}}_{\hat{\beta}} = \left(\frac{n}{n-k} \right) (X'X)^{-1} \left(\sum_{i=1}^n X_i X_i' \hat{e}_i^2 \right) (X'X)^{-1}$$

Donde n y k son el número de observaciones y parámetros en el modelo, respectivamente.

Esta es la forma correspondiente al estimador de la varianza heterocedástico consistente, especificado como *HC1* de acuerdo a Hansen (2022). Se decide utilizar el estimador de la varianza *HC1* porque, si bien es un reescalamiento *ad-hoc*, es una práctica común.¹

Finalmente, los errores agrupados o clusterizados fueron calculados considerando el siguiente estimador:

$$\widehat{\text{Var}}_{\hat{\beta}} = a_n (X'X)^{-1} \hat{\Omega}_n (X'X)^{-1}$$

donde,

$$\hat{\Omega}_n = \sum_{g=1}^G \mathbf{X}'_g \hat{e}_g \hat{e}'_g \mathbf{X}_g, \quad a_n = \left(\frac{n-1}{n-k} \right) \cdot \left(\frac{G}{G-1} \right)$$

Aquí, $\hat{\Omega}_n$ corresponde a un término de los errores clusterizados mientras que a_n corresponde a un factor de escala, similar al presentado para el caso de errores robustos pero ajustado al número de clusters (G).²

Al comparar las 3 metodologías, se observan varias diferencias entre los valores, sin embargo, no pierden consistencia respecto a los coeficientes. De entre los 3 se observa que el error estándar correspondiente a β_0 es mayor respecto al resto, tomando el menor valor en el modelo de MCO

¹De hecho, este reescalamiento es adoptado por el *software* estadístico de Stata, lo que impulsó, aún más, su uso en este trabajo.

²Al igual que para el caso de los errores robustos, se decide utilizar este estimador de errores cluster ya que es el mismo que utiliza Stata.

asumiendo homocedasticidad y ausencia de autocorrelación. A su vez, la magnitud de los errores de β_0 para los 3 modelos permiten concluir que no es eficiente la estimación del parámetro dado a que la magnitud del error es más de un tercio de la magnitud del coeficiente.

Luego, para el caso de β_1 , se sabe con antelación que X_{1ig} está correlacionado con el término de error (ν_g). Entonces, es de esperar que la estimación del error sea más precisa admitiendo heterocedasticidad, que es justamente lo que se observa. Ahora bien, para el caso de los errores clusterizados, se observa que la magnitud del error de β_2 es menor, en comparación a las otras 2 estimaciones, situación consistente con el hecho de que las observaciones de X_{2ig} pertenecen a diferentes *clusters*. Por ello, al considerar los grupos para calcular el error, genera mayor holgura en materia de eficiencia en la estimación del parámetro, corrigiendo el error. No obstante, si bien X_{1ig} es una variable con observaciones en los mismos *clusters* que X_{2ig} , el error clusterizado de β_1 es mayor que el robusto. Esto se debe por la existencia del *trade-off* entre hacer más pequeña la varianza por las propiedades de la estimación en sí misma, y la reducción en los grados de libertad, que ponderan a la matriz de varianza-covarianza. Esto también se complementa con el hecho de que hay un sesgo en X_{1ig} .

4. Test de Hipótesis Nula β_1

Al realizar la prueba de hipótesis, donde la hipótesis nula es $H_0 : \beta_1 = 2$, se comparan los resultados utilizando los distintos errores estándar calculados en la sección anterior, es decir, errores estándar homocedásticos, robustos y clusterizados, contra el valor crítico 1.96. Esto se traduce en, como para todos los casos el valor t es mayor al valor crítico, se rechaza la hipótesis nula con 95 % de significancia estadística³. Dicha premisa es confirmada al calcular el *p-value*, interpretado como la probabilidad de aceptar la hipótesis nula, que como se aprecia en la Tabla 3, es muy pequeño, mostrando resultados consistentes.

Tabla 3 Estadístico-T y P-Values

	Errores Estándar	Errores Robusto	Errores Clusterizados
Test-T	15.2904	15.8723	7.4749
P-Value	0.0000	0.0000	0.0000

Para el caso de los errores estándar y los errores robustos, el valor de los estadístico t es relativamente similar, por la similitud existente entre las magnitudes de los errores en ambos casos. Por otra parte, el caso de errores clusterizados, el valor del estadístico es el mínimo respecto al resto, debido al hecho de que la magnitud de los errores en β_1 es mayor. Por ello, el estadístico es menor para el *cluster* puesto que el error es más alto. Sin embargo, al margen de todo ello, un resultado general es que la hipótesis nula se rechaza para todos los casos.

Este resultado llama la atención puesto que la hipótesis nula es, en sí misma, el valor real de β_1 pero aún así, es rechazado en todos los casos. Esto se traduce en la evidente presencia del sesgo presente en la estimación del parámetro. Una razón es que el modelo de MCO captura únicamente la relación lineal que existe entre las variables explicada y explicativa. Asimismo, otro problema que genera sesgamiento en los parámetros es el efecto de endogeneidad, problema subyacente (y conocido) al proceso generador de datos.

³Para los 3 casos, se realizó un *test* de 2 colas para probar la hipótesis nula.

5. Modelo con Efectos Fijos

A la especificación con la cual se trabajó en las secciones anteriores se le incluye ahora un efecto fijo por grupo.

$$Y_{ig} = \beta_0 + \beta_1 X_{1ig} + \beta_2 X_{2ig} + e_{ig}$$

$$Y_{ig} = \beta_0 + \beta_1 X_{1ig} + \beta_2 X_{2ig} + \gamma_g + e_{ig}$$

donde el efecto fijo γ_g captura la heterogeneidad entre grupos, tal que buscaremos desarrollar un modelo que permita limpiar los efectos estimados de este efecto endógeno a los grupos.

Una forma de estimar un modelo que incluya estos efectos fijos por grupo es mediante un modelo de Mínimos Cuadrados Ordinarios donde, además de las variables dependiente y explicativas, se incluye un set de variables dummy, en particular, una dummy para cada grupo (Hansen (2022), p. 631), el cual permite incluir efecto fijo por grupo. Con ello la regresión a estimar de *dummy variable estimator* del modelo de efectos fijos tendrá la siguiente forma:

$$Y_{ig} = X\beta + D\gamma + e \quad (5.1)$$

La Tabla 4 entrega los coeficientes de interés estimados.⁴ Notamos que, a diferencia del modelo simple de MCO, el estimador β_1 es menor en magnitud (cae 0,3706), mientras que el estimador β_2 varía mínimamente. Dichas variaciones se desprenden directamente de las características del proceso generador de datos, donde X_{2ig} es un vector de números completamente aleatorios, mientras que X_{1ig} es un vector de números aleatorios que sigue una u otra distribución dependiendo del valor de ν_g para cada observación, es decir, depende de a qué grupo pertenece cada observación. Dado que la generación de X_{1ig} depende del valor que toma el efecto fijo de grupo, al limpiar la estimación de este efecto, el estimador de la variable cae. Naturalmente, dado a que la distribución de X_{2ig} no es condicional a ninguna variable, la estimación incluyendo ahora efectos fijos por grupo no afecta tanto su valor.

Tabla 4 Coeficientes FE

Coeficiente	(5.1)
β_1	2.0578
β_2	3.9738

Ahora bien, calculando las tres medidas de error⁵ utilizadas anteriormente para los estimadores del modelo con efectos fijos, cuyos resultados se presentan en la Tabla 5. Se observa que la medida de error aumenta para β_1 y disminuye para β_2 tanto en los errores estándar como en los errores robustos. Esto hace sentido ya que en el modelo simple de MCO, ν_g es una variable de heterogeneidad no observada que el modelo no alcanza a capturar, en cambio, al controlar por efectos fijos, parte de esa variación es absorbida por el efecto fijo limpiando la estimación, lo que hace variar los errores estándar estimados.

A pesar de tener una mejor estimación, tanto los errores estándar como los errores robustos tienen supuestos detrás que no se están cumpliendo (como la correlación del error con X_{1ig} distinta de 0), por lo que siguen estando sesgados, tal que el efecto anterior no es tan directo, en cambio, es más directo verlo en los errores clusterizados, los cuales disminuyen para ambos

⁴En la Tabla (14) presentada en el Anexo, se presentan los resultados de todos los coeficientes estimados, incluyendo el de los efectos fijos.

⁵En la Tabla (15) encontrada en Anexos, se presentan los cálculos de errores para todos los coeficientes, donde por extensión, se presentaron sólo el de los 2 coeficientes de mayor interés.

Tabla 5 Errores Estándar FE

	Errores Estándar	Errores Robusto	Errores Clusterizado
β_1	0.0327	0.0325	0.0295
β_2	0.0318	0.0323	0.0352

coeficientes. Eso indicaría que la estimación de errores clusterizados aumenta la eficiencia de los coeficientes estimados en comparación a la estimación que no incluía efecto fijo por grupo.

Finalmente, los Test-T para el modelos de efectos fijos entregan estadísticos bastante más pequeños en magnitud con *p-values* mayores a 0,05 en los tres casos (para cada tipo de error estándar). Estos resultados indican que no hay información suficiente para rechazar la hipótesis nula, tal que no podemos establecer que $\beta_1 \neq 2$. A diferencia de del modelo simple de MCO, esta vez el resultado de el test sí resulta coherente, pues conocemos que el parámetro verdadero $\beta_1 = 2$ y por ello, implica que el coeficientes ahora estimado está más cercano a su valor verdadero. Dichos resultados se presentan en la Tabla 6.

Tabla 6 Test-T modelo FE

	Errores Estándar	Errores Robusto	Errores Clusterizados
Test-T	1.8202	1.7914	1.6421
P-Value	0.0690	0.0735	0.1009

6. Equivalencia entre FWL y Variables Transformadas

Teorema de Frisch-Waugh-Lovell

Ahora, se utiliza el teorema de Frisch-Waugh-Lovell para obtener los coeficientes en el modelo con efectos fijos por grupos a partir de una regresión por variables dummy. Para esto se construye un modelo de forma matricial de la forma:

$$Y = X\beta + D\gamma + e,$$

donde X es una matriz de dimensión 1000×2 de las variables X_{1ig} y X_{2ig} tal que $X = \begin{bmatrix} X_{1ig} & X_{2ig} \end{bmatrix}$. Por otro lado, D es una matriz de dimensión 1000×40 que asigna el valor 1 para cada individuo perteneciente al grupo g (donde en cada columna tiene un vector de 25 unos correspondiente a los individuos de ello). Este modelo se puede simplificar a lo siguiente,

$$Y = \mathbf{X}\beta + e,$$

donde ahora \mathbf{X} es una matriz que se conforma por 2 particiones. Por un lado, \mathbf{X} considera la matriz D de las dummies, mientras que por otro, considera la matriz X . Lo mismo para β donde por un lado considera el γ y β . Por ello, la matriz \mathbf{X} y β corresponderían a:

$$\mathbf{X} = \begin{bmatrix} D & X \end{bmatrix}, \quad \beta' = \begin{bmatrix} \gamma & \beta \end{bmatrix}$$

De esta forma, aplicando el teorema de FWL se genera una matriz aniquilación para cada coeficiente de interés, en donde a partir de la matriz aniquilación de X , encontraremos que:

$$\hat{\gamma}_{fe} = (D' M_X D)^{-1} (D' M_X Y) = \begin{bmatrix} \widehat{\gamma}_1 \\ \widehat{\gamma}_2 \\ \vdots \\ \widehat{\gamma}_{40} \end{bmatrix}$$

Aquí, γ sería un vector de los coeficientes correspondientes a cada grupo, es decir, los coeficientes de los efectos fijos por grupo. El resultado de esta estimación se encuentra en la Tabla 16 del Anexo, donde se puede ver que el valor de estos son los mismos de la estimación realizada anteriormente ahora utilizando el modelo:

$$Y_i = \beta_0 + \beta_1 X_{1ig} + \beta_2 X_{2ig} + \gamma D_{ig} + e_{ig} \quad (6.1)$$

Por otro lado, la estimación del coeficiente correspondiente al vector X se realiza a partir de la matriz de aniquilación de D , donde:

$$\hat{\beta}_{ig} = (X' M_D X)^{-1} (X' M_D Y) = \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix},$$

Los resultados de la estimaciones se presentan en la primera columna de la Tabla 7. Aquí se puede ver que el valor de los coeficientes es consistente también con los valores anteriormente estimados en el modelo de efecto fijo de la sección anterior. Es decir, el resultado obtenido de esta estimación de regresión particionada muestra que el valor de los coeficientes es el mismo que se había estimado anteriormente mediante MCO.

Luego, de esta estimación se encuentra que el valor de los coeficientes son los mismos que los que fueron estimados con el modelo con efectos fijos de la sección anterior, correspondiente a los valores $\hat{\beta}_1 = 2.0578$ y $\hat{\beta}_2 = 3.9738$. Ambos resultados son muy cercanas al verdadero valor de los coeficientes poblacionales.

Within Transformations

Por otro lado, se puede utilizar una regresión por variables transformadas (o *within transformations*) para estimar los coeficientes $\widehat{\beta}_1$ y $\widehat{\beta}_2$. Este consiste en construir un estimador β que es invariante en el término de error por grupos v_g . Para esto, genera una transformación de la variable (denominada por un punto arriba de ello) que corresponde a las desviaciones en torno a la media por grupos de dicha variable (dependiente o independiente). En estricto rigor, dado:

$$Y_{ig} = \beta_0 + \beta_1 X_{1ig} + \beta_2 X_{2ig} + \varepsilon_{ig} + v_g,$$

$$\bar{Y}_{ig} = \beta_0 + \beta_1 \bar{X}_{1ig} + \beta_2 \bar{X}_{2ig} + \bar{\varepsilon}_{ig} + v_g$$

De esta forma, si $\dot{Y}_{ig} = Y_{ig} - \bar{Y}_g$, $\dot{X}_{kig} = X_{kig} - \bar{X}_{kg}$,⁶ y $\dot{\varepsilon}_{ig} = \varepsilon_{ig} - \bar{\varepsilon}_g$, se tendrá el modelo de regresión por variables transformadas, el cual eliminará tanto el término constante como el error por grupos de tal forma que la ecuación a estimar será:

$$\dot{Y}_{ig} = \beta_1 \dot{X}_{ig} + \beta_2 \dot{X}_{2ig} + \dot{\varepsilon}_{ig}$$

Luego, al calcular las medias por grupos para las variables independientes y dependiente dentro de los datos y generar las variables transformadas, se encuentra que los estimadores de este modelo son iguales a $\hat{\beta}_1 = 2.0578$ y $\hat{\beta}_2 = 3.9738$, encontrando los mismos resultados que para el modelo particionado y aplicando el teorema de FWL. Estos resultados pueden verse en

Tabla 7 Equivalencia entre FWL y Within Transformations.

Coefficientes	Within	
Estimados	FWL	Transformations
$\hat{\beta}_1$	2.0578	2.0578
$\hat{\beta}_2$	3.9738	3.9738

la segunda columna de la Tabla 7. De esta forma, se demuestra que utilizar FWL es equivalente al hacer una regresión por variables transformadas.

Finalmente, ambos modelos muestran ser equivalentes puesto que de ambos enfoques observamos (implícita o explícitamente) \dot{Y} para obtener $\hat{\beta}$. En este sentido, del enfoque del teorema de FWL se sabe que $\hat{\beta}$ y \hat{e} pueden ser obtenidos al regresionar los residuos de Y sobre D en los residuos de la regresión de X sobre D , en donde el residuo de la regresión de D es igual a \dot{Y} (i.e $M_D Y = \hat{e} \equiv \dot{Y}$). Como también, sabemos que $\hat{\beta}$ y \hat{e} pueden ser obtenidos al regresionar la variable transformada \dot{Y} en la variable transformada \dot{X} . Finalmente, dado que de este último encontramos exactamente el estimador del modelo con efectos fijos $\hat{\beta}$, el enfoque a partir del teorema de FWL es equivalente.

7. Simulación de Nueva Base de Datos

En esta sección se repiten los análisis anteriores para un nuevo proceso generador de datos. La única diferencia respecto al anterior es que ahora X_{1ig} depende del valor que toma una variable aleatoria $w_i \sim U(0, 1)$ para cada observación (en lugar de depender del componente del error que varía a nivel de grupo).

Supuestos de MRL

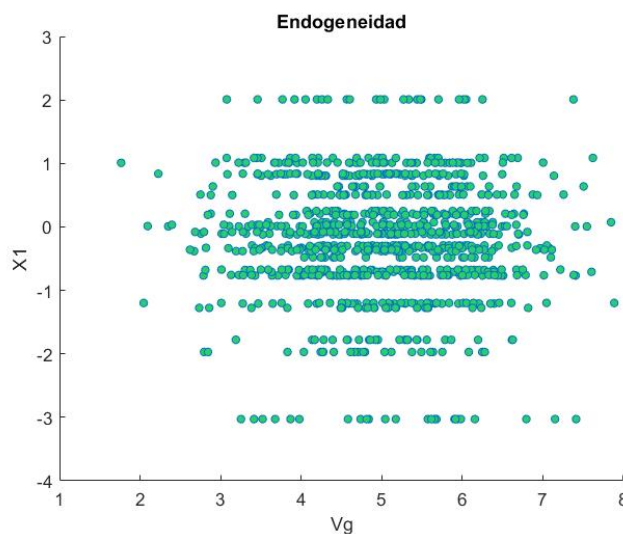
El nuevo proceso simulador de datos tiene una sola diferencia sustancial respecto al proceso anterior, tal que el cumplimiento de los supuestos de MRL antes descrito solo cambia en lo que respecta a aquellos que se ven afectados por la construcción de X_{1ig} , tal como se describe a continuación.

Igual que antes, se cumple que el modelo de la ecuación (7.1) se puede expresar en forma matricial, donde sus segundos momentos son definidos y acotados, y la matriz $\mathbf{X}\mathbf{X}'$ es definida positiva, lo que garantiza invertibilidad.

El cambio se encuentra en lo referido al supuesto $E(e|x) = 0$, que no se cumplía antes dado que había dependencia entre X_{1ig} y el error grupal v_g , pero que ahora debería cumplirse pues la construcción de X_{1ig} no depende de elementos dentro del modelo, más aun, independiente del modelo. Para confirmarlo, a continuación se grafica dicha relación en la Figura 7.1.

Se observa que no existe una correlación entre los errores grupales y X_{1ig} , evidenciando que sí se cumple el supuesto 2, junto a los demás supuestos de MRL ya mencionados.

⁶El término de \bar{X}_{kg} fue calculado como $\sum_{i \in g} \frac{X_{kig}}{N_g}$ donde N_g corresponde al tamaño del grupo.

Figura 7.1 Endogeneidad del Modelo de Regresión Lineal*Elaboración propia.*

Estimadores MCO

La primera especificación a estudiar es la siguiente:

$$Y_{ig} = \beta_0 + \beta_1 X_{1ig} + \beta_2 X_{2ig} + e_{ig} \quad (7.1)$$

donde X_{1ig} ya no correlaciona con el término de error.

La Tabla 8 presenta los coeficientes de la regresión por MCO de este modelo con los nuevos datos. Se puede observar que todos los coeficientes ahora son muy cercanos a los parámetros verdaderos. Al ser todas las variables del modelo generadas aleatoriamente e independientes entre sí, no hay fuente de sesgo por endogeneidad en la estimación del modelo, como sí lo había en el modelo de la primera parte, tal que los parámetros estimados logran aproximarse más al parámetro real.

Tabla 8 Coeficientes MCO

Coeficientes	(7.1)
β_0	0.9229
β_1	1.9473
β_2	4.0278

Errores Estándar

La Tabla 9 muestra las estimaciones de errores estándar, errores robustos y errores clusterezados para el modelo de MCO. De ella destaca que los errores de β_0 son de mayor magnitud que para los otros estimadores, lo que indica que su valor estimado varía más alrededor de su valor verdadero. Esto a su vez es consistente con el resultado obtenido anteriormente, donde la estimación del error para β_0 seguía siendo mayor que para el caso de los demás parámetros. También se destaca que no hay mucha diferencia entre los errores de β_1 y β_2 al calcularlos bajo el supuesto de homocedasticidad y con la medida robusta, lo que hace sentido pues el proceso

generador de datos ya no incluye dependencia entre las variables explicativas y el término de error del modelo.

Por su parte, se observa que los errores clusterizados caen para β_1 y aumentan para β_2 . Esto se puede deber al hecho de que ahora, aumenta la eficiencia de la estimación del parámetro β_1 y se puede medir mejor el error clusterizado a nivel de grupo. De la misma manera, para el caso de β_2 , pudiera estar sucediendo que en la realidad, considerando el hecho de los datos están agrupados, la varianza es mayor que lo estimado para el caso individual.

Tabla 9 Errores Estándar

	Errores Estándar	Errores Robusto	Errores Cluster
β_0	0.2973	0.3118	0.3332
β_1	0.0414	0.0433	0.0363
β_2	0.0416	0.0408	0.0534

Tests de Hipótesis

Al realizar los test de hipótesis $H_0 : \beta_1 = 2$, se observa que los estadísticos son pequeños en magnitud y los *p-value* indican que no hay suficiente evidencia para rechazar la hipótesis nula (probabilidad de rechazar H_0 mayor a 5 %). Este resultado es coherente pues sabemos que el parámetro verdadero es $\beta_1 = 2$ y como el modelo no está sesgado por construcción (no hay problema de endogeneidad), no hay elementos que distorsionen la estimación pudiendo llevar a rechazar H_0 sobre el parámetro verdadero. En línea con esto, cabe resaltar que la estimación de β_1 por MCO es muy cercano al parámetro verdadero. Dichos resultados se presentan en la Tabla 10.

Tabla 10 Test-T y P-Values

	Errores Estándar	Errores Robustos	Errores Cluster
Test t	1.2737	1.2170	1.4516
P-value	0.2031	0.2239	0.1469

Estimación con efectos fijos por grupo

En esta parte se vuelve estimar el modelo añadiendo efectos fijos por grupo como se describe en la sección 5. La Tabla 11 entrega los coeficientes de interés estimados.⁷ Se observa que con el nuevo proceso generador de datos las variaciones en los estimadores son muy pequeñas (0.03 más en β_1 , y 0.01 más bajo para β_2), aunque más cercanas al valor verdadero de los parámetros. Esto se debe a que, si bien para esta simulación de datos las variables no dependen de alguna información específica del grupo al que pertenecen, de igual manera el efecto fijo por grupo permite controlar por algunos factores quizás no observables entre grupos, el cual ayuda a limpiar más la estimación y precisar el valor de los coeficientes estimados.

Al calcular los nuevos errores con efectos fijos por grupo (Tabla 12) se puede ver que todas las medidas de errores disminuyen para los dos estimadores (β_1 , β_2) respecto a la estimación simple de MCO. Esto a su vez es consistente con el hecho de que al incluir efectos fijos, se corrige la estimación por posibles heterogeneidades que no son observadas.⁸ Por otro lado, se destaca el

⁷En la Tabla (17) en el Anexo, se presentan todos los coeficientes del modelo, incluyendo los de efecto fijo.

⁸En la Tabla (18) presentada en el Anexo contiene el detalle de los resultados de los errores para cada coeficiente.

Tabla 11 Coeficientes FE

Coeficientes	(5.1)
β_1	1.9701
β_2	4.0150

hecho de que los errores clusterizados, respecto a los del modelo MCO, caen para β_1 y aumentan para β_2 .

Tabla 12 Errores Estándar FE

	Errores Estándar	Errores Robusto	Errores Cluster
β_1	0.0305	0.0315	0.0277
β_2	0.0309	0.0307	0.0336

Finalmente, al repetir el test de hipótesis $H_0 : \beta_1 = 2$, se encuentra que los valores del Test-T son todos menores que los obtenidos con el modelo MCO con *p-values* mayores en magnitud, por lo que nuevamente implica que no hay suficiente evidencia suficiente para rechazar H_0 . Esto sigue siendo consistente con el hecho de que el parámetro estimado es cercano al valor verdadero y que la estimación mediante efecto fijo, permite tener una estimación más eficiente de los estimadores.

Tabla 13 Test-T y P-Values FE

	Errores Estándar	Errores Robustos	Errores Cluster
Test t	0.9665	0.9736	0.8880
P-value	0.3341	0.3305	0.3748

En suma, la nueva simulación evidencia que al no tener el sesgo provocado por la dependencia de la variable X_{1ig} con el error de grupo, hace que MCO entregue resultados más cercanos a los parámetros verdaderos. Por su parte, la inclusión de efectos fijos por grupo no modifica significativamente los resultados, ya que los datos no están sesgados por problema de endogeneidad o heterogeneidad no observada entre grupos.

8. Aprendizajes Finales

A lo largo del desarrollo de la tarea, se puso en práctica de manera profunda el contenido aprendido en clases, con una perspectiva fundamental sobre modelos de regresión lineal. De lo puesto en práctica en las secciones anteriores, se hace un punteo de lo aprendido:

- Determinar cumplimiento de los supuestos de MRL para un determinado proceso generador de datos, y en consecuencia, poner de manifiesto las posibles soluciones a problemáticas inherentes a los datos en materia de inferencia estadística.
- Entender el modelo de manera transversal gracias a la estimación computacional, y con ello, comprender la naturaleza y diferencias entre diversas medidas de error estándar con el objetivo de identificar cuál es la más adecuada dada la tipología de los datos. Asimismo, entender el ajuste del modelo respecto a distintas bases de datos.

- Comprender las limitaciones de MCO (en este caso, frente a problemas de endogeneidad) y cómo lidiar con ellas mediante otros modelos, como el de efectos fijos. Se pudo verificar cómo cambia la estimación al aplicar este último logrando estimaciones más eficientes y robustas.
- Explorar distintas formas de estimación, como se vio con las distintas formas de estimar efectos fijos, las cuales son equivalentes, pero con algunas se pueden conseguir mayor eficiencia computacional, el particular el algoritmo FWL.

En resumen, esta tarea nos permitió aplicar y contrastar la teoría con un ejercicio práctico, pudiendo así comprobar si los diferentes supuestos y resultados son consistentes entre sí y aprender de la información que entregan las diferentes métricas. De esta forma, se destaca la importancia de elegir bien un modelo y los supuestos que lo sustentan para poder obtener más información de los datos y la relación entre sí.

Anexos

Tablas Resultados Modelo con Efectos Fijos

A continuación se presenta una tabla que contiene el resultado de todos los coeficientes estimados el cual incluye los coeficientes de efecto fijo.

Tabla 14 Coeficientes Modelo Efecto Fijo

Coeficiente	(5.1)
β_1	2.0578
β_2	3.9738
β_3	0.9928
β_4	0.8718
β_5	0.6733
β_6	0.9384
β_7	2.4156
β_8	0.9254
β_9	0.0274
β_{10}	0.7185
β_{11}	0.0138
β_{12}	1.0856
β_{13}	0.9429
β_{14}	1.4332
β_{15}	0.0498
β_{16}	1.0769
β_{17}	1.1627
β_{18}	-1.1377
β_{19}	0.9204
β_{20}	2.5471
β_{21}	0.3928
β_{22}	-1.0145
β_{23}	0.9218
β_{24}	1.9497
β_{25}	1.4082
β_{26}	0.4088
β_{27}	1.8067
β_{28}	-0.2048
β_{29}	-1.1250
β_{30}	0.9983
β_{31}	0.6162
β_{32}	0.8456
β_{33}	-1.0925
β_{34}	1.7582
β_{35}	-0.8610
β_{36}	1.3868
β_{37}	2.0983
β_{38}	2.2980
β_{39}	-0.9599
β_{40}	1.6282
β_{41}	0.5538
β_{42}	0.2162

Luego, aquí se presentan los resultados de la estimación de los errores para los diferentes coeficientes, incluyendo los de efecto fijo.

Tabla 15 Errores Estándar Modelo Efecto Fijo

	Errores Estándar	Errores Robusto	Errores Clusterizado
β_1	0.0327	0.0325	0.0295
β_2	0.0318	0.0323	0.0352
β_3	0.2975	0.2688	0.2111
β_4	0.2692	0.3139	0.1815
β_5	0.2628	0.2308	0.1732
β_6	0.2785	0.2441	0.1958
β_7	0.3108	0.3030	0.2262
β_8	0.2659	0.2788	0.1780
β_9	0.2766	0.2462	0.1939
β_{10}	0.2692	0.2641	0.1824
β_{11}	0.2798	0.2441	0.1974
β_{12}	0.2773	0.2792	0.1924
β_{13}	0.3071	0.2979	0.2183
β_{14}	0.2999	0.3189	0.2108
β_{15}	0.2677	0.2594	0.1795
β_{16}	0.3034	0.2907	0.2152
β_{17}	0.2960	0.3014	0.2075
β_{18}	0.2774	0.2689	0.1935
β_{19}	0.2657	0.2518	0.1747
β_{20}	0.3147	0.3467	0.2309
β_{21}	0.2846	0.2767	0.2061
β_{22}	0.2750	0.2738	0.1914
β_{23}	0.2733	0.2672	0.1890
β_{24}	0.3082	0.3334	0.2226
β_{25}	0.3152	0.3198	0.2306
β_{26}	0.2664	0.2410	0.1790
β_{27}	0.3077	0.3349	0.2223
β_{28}	0.2731	0.2527	0.1884
β_{29}	0.2749	0.2684	0.1937
β_{30}	0.2971	0.2843	0.2064
β_{31}	0.2719	0.2732	0.1850
β_{32}	0.2735	0.2827	0.1875
β_{33}	0.2687	0.2824	0.1802
β_{34}	0.3096	0.2949	0.2241
β_{35}	0.2741	0.2648	0.1888
β_{36}	0.3002	0.2845	0.2132
β_{37}	0.3050	0.2901	0.2186
β_{38}	0.3077	0.2946	0.2201
β_{39}	0.2735	0.3067	0.1878
β_{40}	0.2983	0.2976	0.2091
β_{41}	0.2636	0.2351	0.1729
β_{42}	0.2727	0.3319	0.1875

Coefficientes Modelo FWL de Efectos Fijos

A continuación, se presentan los valores de los coeficientes estimados correspondiente a los efectos fijos mediante el modelo de FWL:

Tabla 16 Coeficientes Modelo FWL Efecto Fijo

Coeficiente	(16)
β_1	0.9928
β_2	0.8718
β_3	0.6733
β_4	0.9384
β_5	2.4156
β_6	0.9254
β_7	0.0274
β_8	0.7185
β_9	0.0138
β_{10}	1.0856
β_{11}	0.9429
β_{12}	1.4332
β_{13}	0.0498
β_{14}	1.0769
β_{15}	1.1627
β_{16}	-1.1377
β_{17}	0.9204
β_{18}	2.5471
β_{19}	0.3928
β_{20}	-1.0145
β_{21}	0.9218
β_{22}	1.9497
β_{23}	1.4082
β_{24}	0.4088
β_{25}	1.8067
β_{26}	-0.2048
β_{27}	-1.1250
β_{28}	0.9983
β_{29}	0.6162
β_{30}	0.8456
β_{31}	-1.0925
β_{32}	1.7582
β_{33}	-0.8610
β_{34}	1.3868
β_{35}	2.0983
β_{36}	2.2980
β_{37}	-0.9599
β_{38}	1.6282
β_{39}	0.5538
β_{40}	0.2162

Coeficientes Modelo Efecto Fijos Nueva Simulación

A continuación, se presentan los valores de los coeficientes estimados incluyendo los coeficientes de los efectos fijos:

Tabla 17 Coeficientes Modelo Efecto Fijo Nueva Simulación

Coeficiente	(7.1)
β_1	1.9701
β_2	4.0150
β_3	1.2093
β_4	-0.0155
β_5	1.6787
β_6	3.1903
β_7	0.4962
β_8	2.2975
β_9	0.9043
β_{10}	1.3807
β_{11}	0.7435
β_{12}	1.2410
β_{13}	-1.0405
β_{14}	0.8521
β_{15}	0.1196
β_{16}	0.9608
β_{17}	1.6109
β_{18}	-1.0053
β_{19}	0.3045
β_{20}	0.6603
β_{21}	0.7358
β_{22}	0.4691
β_{23}	0.6416
β_{24}	0.6642
β_{25}	-0.4400
β_{26}	-0.0289
β_{27}	0.9936
β_{28}	1.2912
β_{29}	2.1356
β_{30}	-1.5057
β_{31}	1.6466
β_{32}	2.0242
β_{33}	0.8522
β_{34}	0.2835
β_{35}	2.2074
β_{36}	1.8424
β_{37}	0.7905
β_{38}	0.2710
β_{39}	0.6192
β_{40}	0.9178
β_{41}	1.3985
β_{42}	1.5026

Aquí se presenta una tabla el cual contiene los resultados de los errores estándar para todos los coeficientes en el modelo de efecto fijo:

Tabla 18 Errores Estándar Modelo Efecto Fijo Nueva Simulación

	Errores Estandar	Errores Robusto	Errores Cluster
β_1	0.0305	0.0315	0.0277
β_2	0.0309	0.0307	0.0336
β_3	0.2941	0.2782	0.2264
β_4	0.2894	0.2924	0.2210
β_5	0.2865	0.2785	0.2168
β_6	0.2838	0.2533	0.2114
β_7	0.2904	0.3232	0.2226
β_8	0.2902	0.3104	0.2221
β_9	0.2965	0.2964	0.2297
β_{10}	0.2894	0.2908	0.2204
β_{11}	0.2834	0.3083	0.2115
β_{12}	0.2978	0.2914	0.2327
β_{13}	0.2871	0.3065	0.2175
β_{14}	0.2833	0.2736	0.2130
β_{15}	0.2958	0.2791	0.2304
β_{16}	0.2963	0.3086	0.2292
β_{17}	0.2990	0.3187	0.2327
β_{18}	0.2914	0.3082	0.2226
β_{19}	0.2902	0.2745	0.2216
β_{20}	0.2929	0.2988	0.2249
β_{21}	0.2846	0.2613	0.2132
β_{22}	0.2900	0.2961	0.2235
β_{23}	0.2908	0.2882	0.2232
β_{24}	0.2891	0.3160	0.2185
β_{25}	0.2913	0.2688	0.2218
β_{26}	0.2889	0.3103	0.2205
β_{27}	0.2855	0.2725	0.2157
β_{28}	0.2954	0.3299	0.2285
β_{29}	0.2979	0.3351	0.2313
β_{30}	0.2989	0.3034	0.2339
β_{31}	0.2828	0.2942	0.2124
β_{32}	0.2925	0.2964	0.2252
β_{33}	0.2879	0.2819	0.2184
β_{34}	0.2843	0.3378	0.2126
β_{35}	0.2947	0.2839	0.2313
β_{36}	0.2957	0.2957	0.2327
β_{37}	0.2893	0.2699	0.2199
β_{38}	0.2856	0.2827	0.2153
β_{39}	0.2959	0.3158	0.2292
β_{40}	0.2795	0.2838	0.2076
β_{41}	0.2863	0.2867	0.2160
β_{42}	0.2897	0.2919	0.2214

Referencias

Bruce Hansen. *Econometrics*. Princeton University Press, 2022. URL <https://users.ssc.wisc.edu/~bhansen/econometrics/>.