



Munich Personal RePEc Archive

An introduction to multilevel analysis: Is the individual demand for health is affected by the physician of primary care?

Herrera Gómez, Marcos

Universidad de Zaragoza

2008

Online at <https://mpra.ub.uni-muenchen.de/35267/>

MPRA Paper No. 35267, posted 08 Dec 2011 17:32 UTC

Una introducción al análisis multinivel: ¿La demanda individual de salud es afectada por el médico de cabecera?

Marcos H. Herrera Gómez*

**Programa de Ciencias Económicas
Universidad de Zaragoza**

Resumen

La utilización de los modelos multinivel ha aumentado con los años, aplicándose en diferentes áreas o especialidades, pero aún no se los considera una técnica de uso habitual. Este trabajo analiza la metodología multinivel destacando sus ventajas sobre otras técnicas más tradicionales que no toman en cuenta estructuras jerárquicas.

Como ejemplo de aplicación se investiga el efecto del médico de cabecera sobre el gasto en farmacia que realiza el paciente. Se dispone de información de corte transversal, referida al año 2004, de la población que integra el sistema sanitario público español en la Comunidad Autónoma de Aragón.

Para el caso de Zaragoza, se realiza la estimación multinivel, encontrando efectos significativos del médico de cabecera en el gasto que realizan los pacientes en medicamentos.

Respecto a las otras variables explicativas, los resultados más importantes se resumen en que, efectivamente, la edad y el sexo del paciente influyen positivamente en el gasto. Siendo las mujeres las que generan un mayor gasto para la edad media de la muestra. La situación de pensionista o de ser español influye positivamente en el gasto en farmacia.

La cantidad de datos nulos en la muestra requiere una corrección por sesgo de selección. La aplicación del modelo de Heckman muestra un fuerte sesgo, teniendo que ser precavidos en las conclusiones respecto al efecto de las variables explicativas. La posibilidad de sesgo en las estimaciones multinivel abre un tema de particular interés para futuras investigaciones.

Palabras Claves: multinivel, estructura jerárquica, sesgo de selección, gasto en salud.

- Octubre 2008 -

* Mail para correspondencia: mherreragomez@gmail.com

Agradecimientos: El autor agradece al Dr. Jesús Mur y a la Dra. Ana Angulo por sus valiosos comentarios.

1. Introducción

El término *multinivel* hace referencia a estructuras de datos jerarquizados o anidados. Esto significa que la información sobre las observaciones esta contenida en varios niveles; un ejemplo pueden ser alumnos dentro de escuelas, o pacientes agrupados por médicos de cabecera. El nivel más bajo de información se refiere al alumno o paciente y el nivel superior contiene la información referida a los establecimientos o al médico de cabecera. En el caso de dos niveles, podemos hablar de variables *micro* referidas a los individuos, y de variables *macro* para las del nivel superior.

La metodología multinivel es de especial interés en economía, principalmente en microeconometría, debido a que las relaciones estudiadas en los individuos suelen estar influenciadas por el contexto social al que pertenecen.

Dos procedimientos habituales en el análisis de datos multinivel son la desagregación desde el nivel superior hacia el inferior o la agregación de las variables del nivel inferior hacia el superior, de Leew [9]. En el primer caso, se adjudica a individuos el mismo valor de la variable a nivel superior. En el segundo caso, los datos individuales son agregados en unidades del nivel superior, eliminando toda la variación intra-grupal. Este último caso comúnmente se denomina *análisis contextual* o *ecológico*.

Analizar variables de diferentes niveles en un único nivel es inadecuado. Uno de los problemas que acarrea es de tipo conceptual, ocasionando la generalmente denominada *falacia ecológica*, que consiste en analizar datos a un nivel superior y formular conclusiones a un nivel inferior al analizado, Freedman [14].¹ El problema también se produce con resultados obtenidos a partir de datos desagregados y su interpretación incorrecta a nivel agregado, denominada *falacia atomística*, Diez-Roux [11].

Otro problema que presenta el *análisis contextual* es de índole estadístico. Al agregar los datos se resume mucha información individual en un único valor y esto genera pérdida de potencia estadística, Hox [24]. Por otra parte, si los datos de pocas unidades de nivel macro se desagregan, entonces se “infla” el tamaño muestral, afectando a la significancia estadística. Es decir, los errores estándares serán más pequeños que los reales (los que efectivamente toman en cuenta la estructura multinivel) y esta subestimación causará el incremento del error tipo I, Bosker y Snijders [2].

Como alternativa, es posible plantearse la estimación por separado de las unidades del nivel 2, de Leeuw [9]. Es decir, estimar una regresión por cada grupo, evitando así la agregación o desagregación de los datos. Pero esta estrategia genera una enorme cantidad de parámetros a estimar y pasa por alto la pertenencia de los grupos a un mismo sistema.

El enfoque multinivel soluciona los problemas anteriores y puede verse como una evolución dentro del análisis de regresión, modificando algunos de los supuestos básicos, como se discute en Goldstein [19] y Hox [23]. El análisis de regresión tradicional es demasiado restrictivo en cuanto a suponer la estabilidad de los parámetros entre grupos y dentro de cada grupo. De hecho, la modificación de este supuesto permite aproximarnos a los modelos de análisis multinivel.

El objetivo de este trabajo es principalmente metodológico, explorando el concepto y utilidad del análisis multinivel mediante una aplicación a la demanda de servicios de salud. Primero, se desarrollan aspectos metodológicos del enfoque multinivel a través del modelo de dos niveles, con dos especificaciones: una de intercepto aleatorio y otra de coeficientes aleatorios. Se revisan los métodos de estimación y la bondad del ajuste de dichos modelos. En la tercera sección, se realiza una aplicación a los gastos farmacéuticos con datos provenientes del sistema de salud aragonés. Se incluye una corrección por sesgo de selección. En la cuarta y última sección, se presentan las principales conclusiones.

¹ Robinson [35], presentó datos agregados de nueve regiones geográficas para analizar la relación entre el porcentaje de gente de raza negra y el nivel de analfabetismo en 1930. La *regresión ecológica*, o sea a nivel agregado, brindaba una correlación de -0.53. Por lo contrario, el análisis a nivel individual mostraba una correlación de 0.29, concluyendo que no pueden utilizarse los resultados agregados como si fuesen analizados a nivel individual. Existe una explicación estadística de este resultado que puede consultarse en Kreft [29].

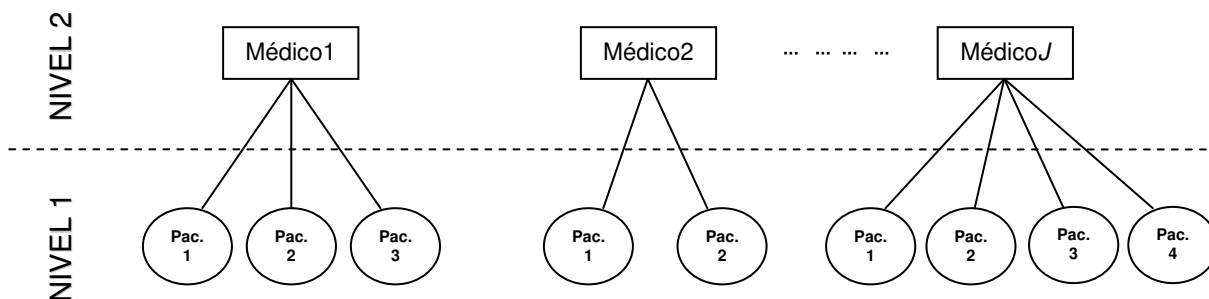
2. Aspectos metodológicos del enfoque multinivel

En estudios sociales, y principalmente en investigación educativa, se asume el supuesto básico de que los individuos interactúan dentro de estructuras o grupos sociales; lo que significa que, las decisiones individuales pueden estar influenciadas por su pertenencia a este grupo. Este contexto muchas veces no es tenido en cuenta en investigaciones de tipo empírico.

El enfoque multinivel incorpora al análisis este contexto social. Para una mejor comprensión planteamos investigar el gasto en farmacia de un grupo de individuos que pertenecen a la estructura sanitaria española. Supongamos que este gasto es íntegramente solventado por el sector público, por lo que puede pensarse que la renta del individuo no se relaciona directamente con el mismo. El gasto en farmacia se realiza bajo receta, dependiendo exclusivamente del médico que prescribe su tratamiento. Condicionando este gasto por las características propias de cada paciente, si elegimos aleatoriamente a dos pacientes, esperamos que sean más parecidos en su gasto en farmacia si pertenecen al mismo médico que si proviniesen de diferentes médicos. Pensamos que los médicos tienen cierto poder discrecional para afectar al gasto público. En otros términos, los pacientes tienen una vivencia común: los propios incentivos económicos del profesional y la opinión experta acerca de lo que el grupo debe recibir. Luego, nosotros buscamos responder a la siguiente pregunta: *¿El médico de cabecera afecta a la demanda individual del gasto en farmacia?*

Para responder a esta pregunta comenzaremos explicando la estructura jerárquica o multinivel representada en el Gráfico N° 1. Los pacientes tienen un comportamiento similar en su gasto debido a la existencia de un componente común a ellos, el médico de cabecera. En el esquema, el nivel 1 contiene la información de todos los pacientes y el nivel 2 agrupa a los pacientes en J médicos.

Gráfico N° 1: Estructura multinivel en sanidad



La estructura jerárquica establecida refleja que cada grupo de pacientes es atendido por un único médico, de modo que no hay solapamientos entre los diferentes grupos. Esta estructura genera correlación o falta de independencia en las observaciones. Tradicionalmente, el modelado estadístico toma a esta correlación como una dificultad y, por lo general, se asume que esta dependencia no existe, pero en el análisis multinivel tal correlación es esperada y explícitamente modelada.

Es importante aclarar que el análisis multinivel puede tratar con estructuras más complejas, en donde los individuos pueden pertenecer a más de un grupo. Este tipo de estructura se denomina *membresías múltiples*. Aquí nos enfocaremos en la estructura planteada sin entrar en mayor complejidad (para una revisión centrada en investigaciones educativas con estructuras complejas puede consultarse el trabajo de Fielding y Goldstein [9]).

Inspeccionando en el Gráfico N° 1, observamos que las unidades del nivel 1 no están equitativamente distribuidas entre las unidades del nivel 2. El primer médico tiene la responsabilidad de asignar el gasto en medicamentos de 3 pacientes mientras que el médico J tiene a su disposición a 4 pacientes. En tal caso diremos que los grupos no están balanceados. Los modelos multinivel pueden trabajar con grupos de diferentes tamaños.

Para contrastar el efecto medico podemos elegir el análisis de regresión tradicional e ir modificando la especificación cuando sea necesario. Partimos del modelo de regresión, sin variables explicativas, que asume que todas las observaciones son independientes:

$$y_i = \beta_0 + \varepsilon_i \quad (1)$$

Donde y_i es el gasto en farmacia del paciente i -ésimo, β_0 captura el gasto medio de toda la muestra y ε_i es un término de error individual. Este modelo no distingue los niveles.

El siguiente paso es permitir que el valor medio difiera por pertenencia a un grupo. Algebraicamente, incorporamos variables dummies indicando la pertenencia de cada paciente a un grupo. Agregamos el suscripto j , indicando que la observación i -ésima pertenece al grupo j -ésimo. Debemos tomar a un grupo como base para evitar colinealidad perfecta.

$$y_{ij} = \beta_0 + \delta_j + \varepsilon_{ij} \quad (i = 1, \dots, N ; j = 1, \dots, J - 1) \quad (2)$$

con δ_j capturando el efecto grupal en el gasto en farmacia.

A pesar de que identificamos a los pacientes con sus médicos o grupos, la ecuación (2) refleja un modelo de un solo nivel. O sea, se continúa suponiendo que la varianza residual dentro de los grupos y entre los grupos es la misma. Esta estrategia, denominada comúnmente de efectos fijos, genera inferencias restringidas a los grupos representados en la muestra. Las dificultades de su aplicación aumentan cuando el número de grupos es grande haciendo inaplicable el procedimiento.

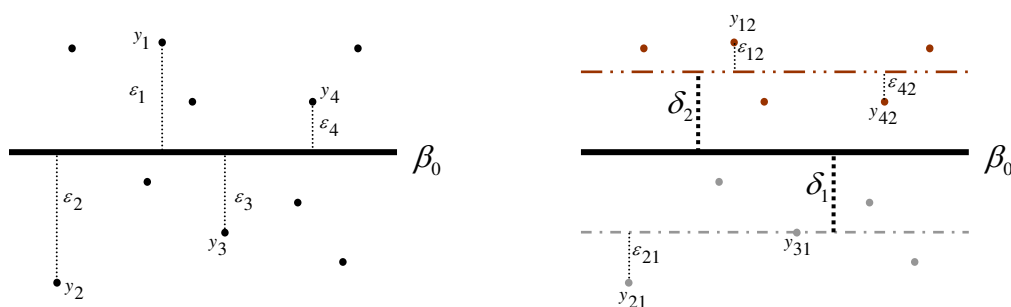
Una estrategia alternativa es la estimación multinivel, que permite realizar inferencias sobre la población, generalizando así los resultados obtenidos en la muestra, Snijders [39]. Para lograr este resultado debemos suponer que nuestros datos constituyen una muestra representativa y aleatoria de la variable jerárquica. A continuación, el análisis en dos niveles supone que los individuos y los grupos, identificados por el médico de cabecera, provienen de una muestra aleatoria con parámetros poblacionales conocidos, pero diferentes entre si. Algebraicamente, podemos representar esta modificación de la siguiente manera:

$$y_{ij} = \beta_0 + \delta_j + \varepsilon_{ij} \quad (3)$$

con δ_j capturando el efecto *aleatorio*, no observable, del grupo en el gasto en farmacia. Se supone que el término de error, ε_{ij} , se distribuye normal con media 0 y varianza σ_ε^2 . Suponemos, además, que el efecto grupal aleatorio se distribuye normal con media 0 y varianza σ_δ^2 . Este modelo es conocido como *modelo vacío* (empty model). La distribución normal es la habitualmente utilizada, aunque puede trabajarse otro tipo de distribución, Seltzer y Choi [37].

Gráficamente, podemos observar la diferencia del modelo convencional respecto al de dos niveles.

Gráfico N° 2: Modelo vacío. Un nivel versus dos niveles



El análisis multinivel plantea una descomposición del residuo. El residuo específico del grupo j , δ_j , captura la diferencia respecto a la media global. En el ejemplo, el residuo es negativo para el grupo 1 debido a que la media grupal es inferior a la global. Sucede lo contrario para el grupo 2, que tiene una media superior al valor β_0 . Los residuos ε_{ij} no capturan la diferencia respecto a la media global del individuo i , sino que reflejan el comportamiento del individuo i respecto al grupo de pertenencia j . Considerando el grupo 2 en el ejemplo, el individuo y_{42} tiene un comportamiento elevado respecto al comportamiento global pero, desde la óptica multinivel, su comportamiento es bajo respecto al grupo de pertenencia. Este comportamiento inferior a la media de su grupo es capturado por el residuo negativo ε_{42} .

Así también pueden compararse a diferentes individuos respecto a sus contextos. En el ejemplo, podemos observar que el individuo y_{42} y el individuo y_{21} tienen un comportamiento similar respecto a la media de sus respectivos grupos, ya que ambos están por debajo del valor promedio. Pero el contexto en el que estos individuos están inmersos es muy diferente, y_{42} pertenece a un grupo con un gasto muy por encima al grupo de y_{21} .

Si observamos el desarrollo realizado hasta aquí, podríamos concluir que es similar a un análisis de varianza, comúnmente denominado ANOVA, el cual puede resolverse por efectos fijos o aleatorios. La técnica de efectos aleatorios genera una descomposición de la varianza de la variable dependiente entre grupos y dentro de los grupos, realizando los mismos supuestos que utilizamos para el multinivel.

En términos del análisis de la varianza, con efectos aleatorios, estamos realizando una descomposición de la variabilidad total observada entre los niveles, suponiendo independencia entre δ_j y ε_{ij} :

$$Var(y_{ij}) = Var(\delta_j) + Var(\varepsilon_{ij}) = \sigma_\delta^2 + \sigma_\varepsilon^2 \quad (4)$$

De la descomposición surge la proporción de la varianza total explicada por el nivel grupal, concretado en el coeficiente de correlación intraclass. Este coeficiente, denominado comúnmente con la letra griega ρ , se define como:

$$\rho = \frac{\sigma_\delta^2}{\sigma_\delta^2 + \sigma_\varepsilon^2} \quad (5)$$

El valor de ρ oscila entre 0 y 1, teniendo varias interpretaciones. Puede leerse como el grado de similitud grupal de las observaciones: un valor de 0 indica que no existe diferencia *entre* los grupos y un valor de 1 indica que no hay diferencia *dentro* de los grupos. Cuanto mayor sea el valor del coeficiente, menor será la similitud entre los grupos. También puede interpretarse como índice de homogeneidad del nivel 1 dentro del nivel 2. Si las observaciones no son heterogéneas, el coeficiente debe estar cercano a 0. Por último y quizás más importante, igualmente, puede asociarse con la correlación poblacional esperada entre dos observaciones pertenecientes al mismo grupo. Por ejemplo, cuanto más parecido es el gasto en farmacia de los pacientes que acuden a un mismo médico de cabecera, es más probable que las causas del gasto tengan que ver con la influencia del médico. La ausencia de dependencia, en este caso, implica ausencia de efectos del médico de cabecera en el gasto individual.

El análisis multinivel adquiere mayor relevancia cuanto más diferentes sean las observaciones de los grupos entre si o cuanto mas alejado de 0 se encuentre ρ . Esto no quiere decir que valores pequeños del coeficiente sean siempre un argumento a favor del análisis de regresión convencional, Roberts [35].

Respecto a los residuos estimados, en un modelo multinivel el error total es $\delta_j + \varepsilon_{ij}$, el cual se estima como $e_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \hat{\beta}_0$. Para obtener los residuos grupales se estima la media

residual por grupo $\bar{e}_j = \bar{y}_j - \hat{\beta}_0$, donde \bar{y}_j es la media muestral de y_{ij} en el grupo j . Luego, el residuo del grupo j es:

$$\hat{\delta}_j = w \cdot \bar{e}_j, \text{ siendo } w = \frac{\hat{\sigma}_\delta^2}{\hat{\sigma}_\delta^2 + \frac{\hat{\sigma}_\varepsilon^2}{n_j}} \quad (6)$$

El factor w recibe la denominación de contracción residual (*shrinkage factor*). Este factor es siempre menor a 1, por lo que los residuos grupales son menores o iguales a los residuos medios. El parámetro n_j captura la cantidad de observaciones dentro del grupo j , cuanto más grande sea el grupo más cercanos estarán los residuos grupales de los grupales medios. Cuando $\hat{\sigma}_\varepsilon^2$ es notablemente mayor que $\hat{\sigma}_\delta^2$, los residuos grupales se acercarán a 0 y la media grupal $\hat{\beta}_0 + \hat{\delta}_j$ tenderá a la media general $\hat{\beta}_0$.

Los residuos del nivel 1, $\hat{\varepsilon}_{ij}$, son obtenidos por simple substracción:

$$\hat{\varepsilon}_{ij} = e_{ij} - \hat{\delta}_j = y_{ij} - \hat{\beta}_0 - \hat{\delta}_j \quad (7)$$

Estos residuos estimados reciben la denominación de *predicciones*.

La posibilidad de llegar al modelo multinivel desde el lado del análisis de varianza o por el análisis de regresión ha provocado que diferentes áreas trabajen con el modelo multinivel bajo muy diversos nombres. Para una descripción del modelo multinivel desde las ciencias experimentales consúltese a McCullagh y Nelder, [32], y para una revisión de las diversas terminologías véase los trabajos de Kreft y de Leeuw [6] y Meijer y de Leeuw [10].

2.1. Modelo de intercepto aleatorio

Hasta aquí no hemos incluido regresores de ningún tipo. Ahora extenderemos el modelo vacío incluyendo variables explicativas, discretas o continuas. Supongamos que introducimos una variable explicativa a nivel individual, x_{ij} . El modelo con esta nueva variable queda expresado de la siguiente manera:

$$y_{ij} = \beta_0 + \beta_1 \cdot x_{ij} + \delta_j + \varepsilon_{ij} \quad (8)$$

Donde y_{ij} representa la variable dependiente para el individuo i en el grupo j , β_0 es el intercepto, β_1 representa el efecto de la variable x_{ij} , δ_j es el efecto aleatorio del nivel 2, y ε_{ij} es el efecto aleatorio del nivel 1. Esta ecuación representa al comúnmente denominado *modelo de intercepto aleatorio* (random effects para datos en panel), Chamberlain [4].

Se puede observar que la ecuación (6) tiene una parte de efectos fijos y otra de efectos aleatorios:

$$y_{ij} = (\beta_0 + \beta_1 \cdot x_{ij}) + (\delta_j + \varepsilon_{ij}) \quad (9)$$

Además, la ecuación puede ser expresada sobre la base de una estructura jerárquica:

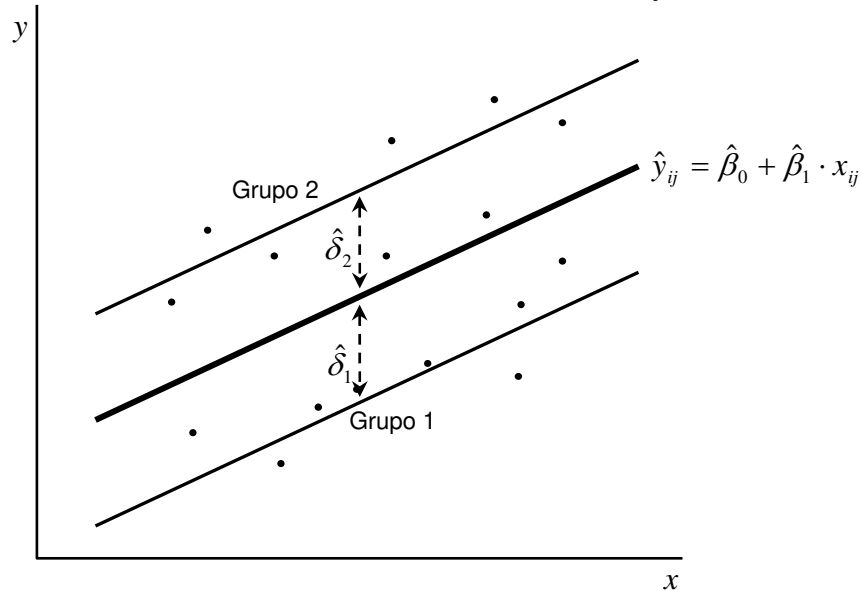
$$\text{Nivel 1: } y_{ij} = \mu_j + \beta_1 \cdot x_{ij} + \varepsilon_{ij} \quad (10)$$

$$\text{Nivel 2: } \mu_j = \beta_0 + \delta_j \quad (11)$$

Esta representación indica que existe un elemento aleatorio adicional respecto al modelo vacío. Este componente captura el efecto del nivel 2, siendo β_0 el parámetro que mide el valor medio para todos los grupos y para todas las observaciones y δ_j es la desviación del grupo j respecto de la media global. El término de error δ_j es el mismo (para todas las observaciones del mismo grupo), aunque el término de error ε_{ij} será diferente. En este modelo se asume que los grupos difieren en el intercepto y que el efecto de la variable x_{ij} es

el mismo para todas las observaciones, con independencia del grupo al que pertenezcan. Se supone además que los efectos aleatorios a nivel 2, δ_j , se distribuyen normal con media 0 y varianza σ_δ^2 .²

Gráfico N° 3: Estimación del modelo con intercepto aleatorio



El Gráfico N° 3 muestra la estimación del modelo con intercepto aleatorio. La línea de trazo grueso es la estimación global y la estimación del comportamiento esperado para los dos grupos de pacientes aparece en trazo más fino. Los dos grupos difieren en el intercepto únicamente, suponiéndose que el efecto de la variable x_{ij} es el mismo entre los individuos, sin importar el grupo de pertenencia.

Este modelo puede estimarse mediante efectos fijos, presentando las limitaciones antes comentadas. La literatura de datos panel dedica bastante espacio a la discusión de los meritos relativos entre efectos aleatorios y fijos cuando las variables explicativas están correlacionadas con la perturbación aleatoria. En tales circunstancias, la estimación por efectos aleatorios lleva a resultados inconsistentes mientras que la especificación de efectos fijos permite obtener estimaciones consistentes, Hsiao [26]. En el análisis multinivel se asume una estructura jerárquica *poblacional*, con muestreo aleatorio tanto de los grupos como de los individuos dentro de los grupos. Consecuentemente, el análisis multinivel debe incorporar necesariamente efectos aleatorios, como destacan Hox y Kreft [25].

2.2. Modelo de intercepto y pendiente aleatoria

Considerar que el efecto de la variable explicativa es constante entre los grupos puede ser demasiado restrictivo. Si modificamos el modelo de dos niveles para permitir intercepto y pendiente aleatoria, llegamos a la siguiente ecuación:

$$y_{ij} = \beta_0 + \beta_1 \cdot x_{ij} + \delta_j + \eta_j \cdot x_{ij} + \varepsilon_{ij} \quad (12)$$

Con los dos niveles identificados:

$$\text{Nivel 1: } y_{ij} = \mu_j + \beta_{1j} x_{ij} + \varepsilon_{ij} \quad (13)$$

$$\text{Nivel 2: } \mu_j = \beta_0 + \delta_j \quad (14)$$

$$\beta_{1j} = \beta_1 + \eta_j \quad (15)$$

² Este modelo es conocido en análisis longitudinal (datos panel) como de *efectos aleatorios* (random effects). Los datos longitudinales pueden considerarse como un caso multinivel, en donde los individuos constituyen el nivel 2 y las repeticiones corresponden al nivel 1, Twisk [44].

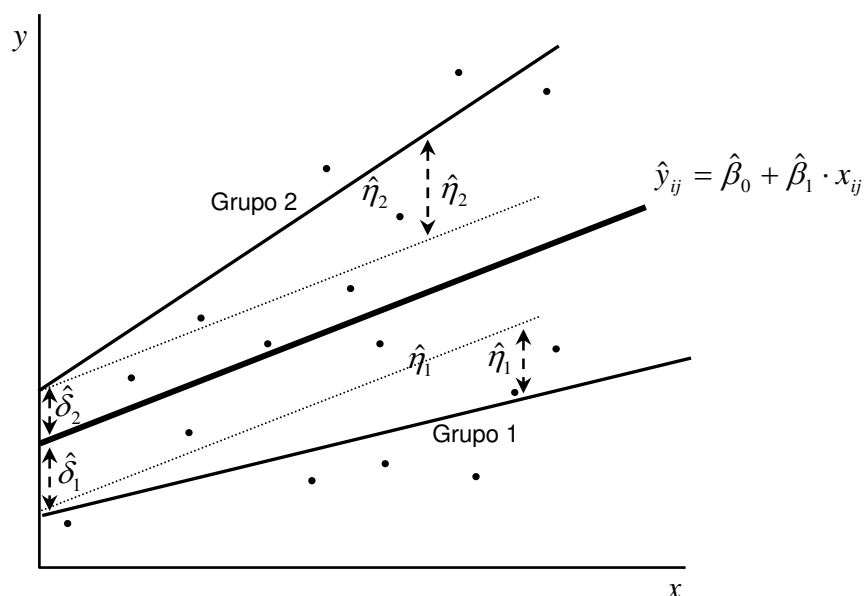
La pendiente ahora es aleatoria y puede ser descompuesta en un efecto global, común a todos los grupos, β_1 , y en un efecto aleatorio, η_j , que captura la desviación del grupo j respecto al nivel global. En este modelo, cada grupo tiene un intercepto y una pendiente diferente. El efecto aleatorio de la pendiente se supone que está distribuido con media 0 y varianza σ_η^2 . Este modelo es similar al planteado por Swamy y Tavlas [43].

Además, se asume una covarianza no necesariamente cero entre los dos efectos aleatorios, $\sigma_{\eta\delta}$. Esta covarianza captura la variación del intercepto y la pendiente entre-grupos. Un valor positivo implica que grupos con altos valores del intercepto, δ_j , tienden a tener altos valores de la pendiente, η_j .³

La ecuación (12) implica, no solo que los individuos dentro del mismo grupo están correlacionados; sino que además ésta correlación, como también la varianza de y , es dependiente de los valores de x_{ij} .⁴ Esto lo que denominamos heteroscedasticidad.

En el Gráfico N° 4, se presenta un ejemplo ilustrativo. Se observan dos líneas punteadas que son proyecciones de la estimación global, la línea de trazo más grueso, que parten del intercepto estimado para cada grupo. Los efectos aleatorios estimados del intercepto son: positivos para el grupo 2 ($\hat{\delta}_2$) y negativo para el grupo 1 ($\hat{\delta}_1$). Los efectos aleatorios estimados de la pendiente son: positivos para el grupo 2 ($\hat{\eta}_2$) y negativo para el grupo 1 ($\hat{\eta}_1$). En consecuencia, la covarianza es positiva. La estimación para los dos grupos representados diverge a medida que los valores de la variable explicativa crece, generando el efecto visual de un abanico abierto.

Gráfico N° 4: Estimación del modelo con intercepto y pendiente aleatoria



Una forma de visualizar la relación entre el intercepto y la pendiente aleatoria por grupo es mediante un diagrama de dispersión, como en el Gráfico N° 5.

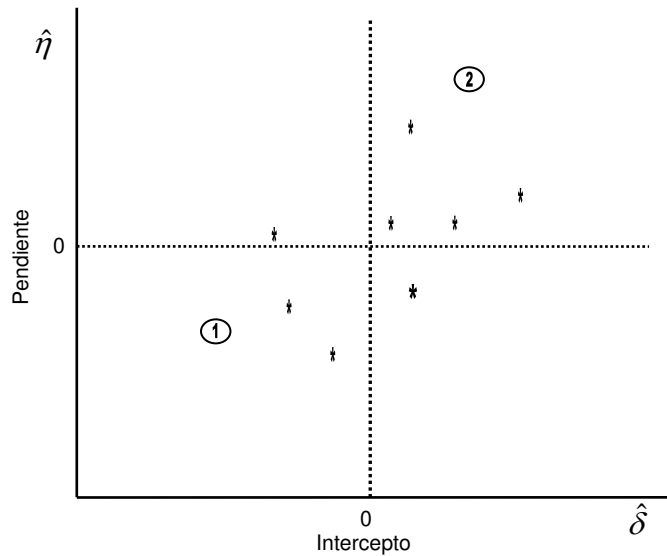
Suponiendo que tenemos más de dos grupos, por ejemplo 10, cada punto en el gráfico N° 5 representa el valor del intercepto y la pendiente aleatoria del grupo. Si dentro de estos

³ Una práctica habitual es el centrado de los datos o centrado de media global, para que la interpretación del modelo sea más simple. Pero esta transformación no es aplicable a todos los modelos multinivel y genera diversos inconvenientes, por lo que trabajaremos sin transformaciones. Para una discusión sobre el tema puede consultarse Kreft *et al.* [28] y Van Landeghem *et al.* [45].

⁴ Algebraicamente, la varianza condicional de la ecuación (12) es: $Var(y_{ij}|x_{ij}) = \sigma_\delta^2 + 2 \cdot \sigma_{\delta\eta} \cdot x_{ij} + \sigma_\eta^2 \cdot x_{ij}^2 + \sigma_\epsilon^2$, siendo la varianza entre grupos igual a $\sigma_\delta^2 + 2 \cdot \sigma_{\delta\eta} \cdot x_{ij} + \sigma_\eta^2 \cdot x_{ij}^2$.

grupos consideramos los representados en el gráfico N° 4, los podemos identificar con un número rodeado por un círculo. El grupo 1 tiene un intercepto negativo y una pendiente negativa. El grupo 2 tiene un intercepto positivo y una pendiente positiva.

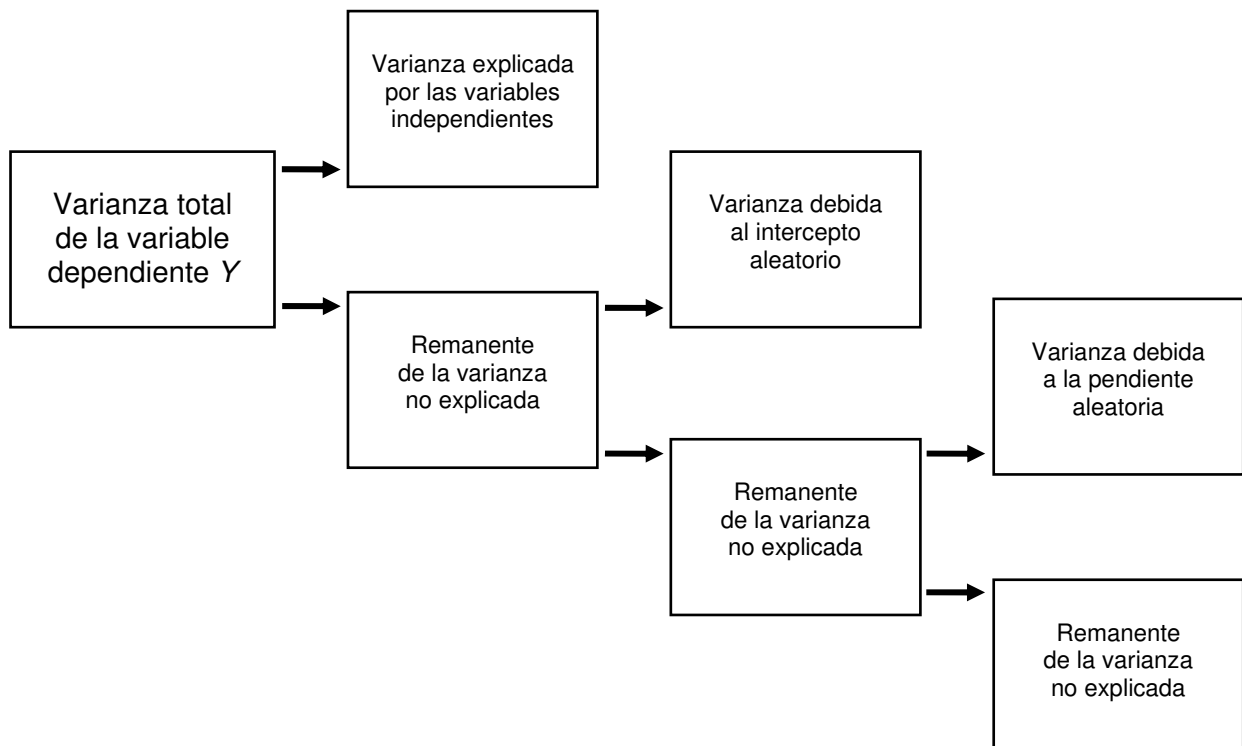
Gráfico N° 5: Relación entre los componentes de varianza



La idea general planteada por el modelo de coeficientes aleatorios es que la varianza no explicada de la variable dependiente y_{ij} se divide en dos componentes diferentes. Uno de ellos está relacionado con el intercepto aleatorio y el otro con la pendiente aleatoria. Ambos componentes, al agruparlos, resultan en la varianza a nivel de grupo.

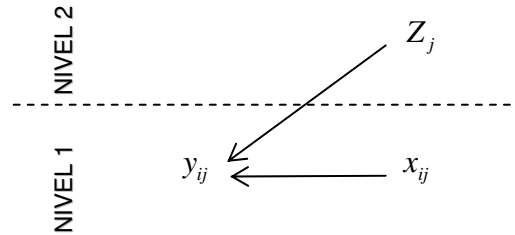
La varianza restante no explicada a nivel de grupo corresponde al nivel individual (dentro del grupo), y es denominada varianza no explicada residual. La suma de estas varianzas forma la variabilidad total. Todo lo anterior puede resumirse en el siguiente esquema:

Gráfico N° 6: Descomposición de varianza (Twisk [44], pag.81)



Una última extensión que podemos considerar se refiere a la naturaleza de las variables explicativas, hasta ahora procedentes del mismo nivel que la variable dependiente. Podemos generalizar este planteamiento al incluir en el modelo de regresión variables explicativas del nivel superior. Supongamos que tenemos variables del nivel 2, y que intentamos incorporarlas a nuestro análisis. Siguiendo a Bosker y Snijders [2], si la letra mayúscula denota a las variables de nivel 2, o *variables contextuales*, la relación es la siguiente:

Gráfico N° 7: Relación entre variables de diferentes niveles



Las variables contextuales pueden reflejar características del nivel 2, como la especialidad del médico, o pueden provenir de agregaciones del nivel 1 como, por ejemplo, la edad media de los individuos que conforman el grupo de pacientes atendidos por el mismo médico. Puede ser que la proporción de pensionistas sea un indicativo de la conducta médica o la proporción de mujeres que están bajo su prescripción. Todas estas variables, pertenecientes a la estructura jerárquica pacientes-médicos, pueden incorporarse sin temor de cometer el error de desagregación o agregación, ya que estos problemas los hemos solucionado con la descomposición del error.

Regresando al modelo de intercepto aleatorio, al incorporar la edad media de los pacientes dentro del grupo, obtenemos:

$$y_{ij} = \beta_0 + \beta_1 \cdot x_{ij} + \beta_2 \cdot Z_j + \delta_j + \varepsilon_{ij} \quad (16)$$

Siendo $Z_j = \bar{X}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij}$, para cada $j = 1, \dots, J$.

Podemos sumar y restar $\beta_1 \cdot Z_j$ para obtener una mejor interpretación del modelo:

$$y_{ij} = \beta_0 + \beta_1^* \cdot (x_{ij} - Z_j) + \beta_2^* \cdot Z_j + \delta_j + \varepsilon_{ij} \quad (17)$$

Luego, $\beta_1 = \beta_1^*$ captura el efecto intra-grupo de x y $\beta_1 + \beta_2 = \beta_2^*$ captura el efecto entre grupos de x o *efecto contextual*. La transformación realizada en la ecuación (17) es denominada *centrado grupal* marcando la diferencia con el *centrado general* comentado en la nota al pie 2.

El modelo general queda especificado de la siguiente manera:

$$y_{ij} = \beta_0 + \sum_{k=1}^K \beta_{1k} x_{ijk} + \sum_{m=1}^M \beta_{2m} Z_{jm} + \sum_{k=1}^K \eta_{jk} x_{ijk} + \eta_j + \varepsilon_{ij} \quad (18)$$

Para todo $i = 1, \dots, N$ y $j = 1, \dots, J$

Además, si es de interés, existe la posibilidad de incorporar términos de interacción entre variables de diferentes niveles.

2.3. Estimación y bondad del ajuste

El método de estimación usual es el de máxima verosimilitud, pudiendo ser completo (*Full Maximum Likelihood, FML*) o restringido (*Restricted Maximum Likelihood, REML*). La diferencia entre ambos métodos radica en que el *REML* toma en cuenta los grados de libertad perdidos por el cálculo de los efectos fijos y de esta forma genera estimaciones de varianza menos sesgadas. En muestras pequeñas la ventaja es más evidente, pero cuando la muestra es grande la diferencia entre los métodos es despreciable, Snijders y Bosker [38]. El método computacional eficiente del *REML* fue desarrollado por Goldstein y Rabash [20].

Además de estos métodos, se utilizan otros procedimientos que toman en cuenta la estructura jerárquica. La estimación del *modelo de medias poblacionales*, comúnmente conocido como *GEE* (Generalized Estimating Equations), se centra en la estimación de la parte fija y condensa la estructura jerárquica en una serie de parámetros que deben estimarse. La diferencia entre los enfoques es que el *GEE* es un modelo *marginal* (la variable explicada depende únicamente de las variables explicativas) y el análisis multinivel es de *efectos mixtos* (la variable explicada depende tanto de los efectos fijos como de los aleatorios). El análisis *GEE* genera estimaciones de coeficientes corregidos por la estructura jerárquica, combinando los efectos entre los grupos y dentro de los grupos, lo que produce que el coeficiente estimado tenga una doble interpretación. A diferencia de la estimación multinivel, no es posible determinar la contribución de cada nivel. El análisis *GEE* puede trabajar con diferentes estructuras de correlación dentro de los grupos: autorregresivas (generalmente, para datos longitudinales); intercambiables (la correlación es la misma para todas las observaciones dentro del grupo); no estructuradas (cada par de observaciones dentro del grupo tiene una correlación diferente). Solo para variables dependientes continuas, el análisis *GEE* con estructura de correlación intercambiable es igual al análisis multinivel con intercepto aleatorio, Twisk [44], pag. 91.

La estimación bayesiana, vía el método *MCMC* (Markov Chain Monte Carlo), es otra posibilidad que combina una distribución a priori con la información muestral, a través de la verosimilitud para obtener una distribución a posteriori. El enfoque bayesiano tiene ventajas para tratar con variables marcadamente no normales. Los detalles de este procedimiento pueden consultarse en Draper [12].

Los principales programas estadísticos tienen implementadas las rutinas de estimación por máxima verosimilitud. Esta estimación arroja el valor del logaritmo de la función de máxima verosimilitud del modelo estimado. De este valor surge un estadístico de uso habitual en la literatura multinivel, denominado *deviance*, definido como:

$$d = -2 \ln ML \quad (19)$$

Siendo $\ln ML$, el valor del logaritmo de la función de máxima verosimilitud.

En términos generales, a menor valor de *deviance* mejor ajuste. Para modelos anidados, la diferencia de *deviance* se distribuye aproximadamente (si el tamaño muestral es grande) como una chi-cuadrado con grados de libertad igual a la diferencia de los parámetros estimados en los dos modelos. Esta diferencia de *deviance* es el conocido test de razón de verosimilitud:

$$LR = -2 \ln ML_1 - (-2 \ln ML_2) = d_1 - d_2 \quad (20)$$

Siendo $\ln ML_1$, el valor del logaritmo de la función de máxima verosimilitud que excluye el efecto del parámetro bajo hipótesis nula (modelo restringido), mientras que $\ln ML_2$ lo incluye (modelo ampliado). Esto puede utilizarse para elegir entre modelos con diferentes variables explicativas o para comparar modelos que incorporan efectos aleatorios respecto a los modelos que no los incluyen.

Si los modelos no son anidados, para seleccionar entre modelos alternativos, se pueden utilizar el criterio de información de Akaike o el criterio de información bayesiana de Schwarz.

El criterio de Akaike se define como:

$$AIC = d + 2k \quad (21)$$

Siendo d la deviance y k el número de parámetros estimados.

El criterio de Schwarz es igual a:

$$BIC = d + k \cdot \ln N \quad (22)$$

Siendo d , la deviance; k , el número de parámetros estimados y; N , el número de observaciones.

A menores valores de ambos criterios, mejor ajuste a los datos, indicando que el modelo es preferido a cualquier otro alternativo.

Los procedimientos de máxima verosimilitud producen errores estándares para los parámetros estimados. Estos errores son asintóticos y, para distribuciones no normales, son válidos solo para grandes muestras. En los estudios multinivel, el principal problema es generalmente el tamaño muestral del nivel 2, debido a que casi siempre es mucho menor que el tamaño muestral a nivel 1. Entre las recomendaciones que se pueden seguir para la estimación, el tamaño mínimo que se requiere en el nivel 2 es de 50 observaciones o grupos. Por debajo de este tamaño, las estimaciones de los errores estándares de los componentes de varianza serán sesgadas, Maas y Hox [31]. Estos autores, además, revisan diversos estudios basados en simulaciones que investigan la exactitud y sesgo en los componentes de varianza y en los coeficientes fijos, destacando la tendencia a subestimar los primeros y la insesgadez de los segundos. Para una especificación multinivel simple, una variable explicativa con intercepto aleatorio y pendiente aleatoria, resaltan que es más importante un gran número de grupos que una gran cantidad de individuos por grupo. Se puede tener una base muy grande de individuos pero para la estimación multinivel lo importante es el número de grupos, según Snijders [40].

Respecto al nivel 1, un tamaño medio grupal entre 30 y 50 individuos se considera normal y apropiado para estudios de investigación educativa u organizacional. Para los estudios de investigación longitudinal, o de mediciones ocasionales, 5 de media por grupo es normal.

Una de las preguntas que puede surgirnos es: ¿Cuál es la bondad del ajuste del modelo o cuánta de la variabilidad de la variable dependiente es explicada por el modelo?. Podemos obtener tantas medidas de bondad como niveles tengamos en nuestro modelo. Respecto a la obtención de una medida de bondad del ajuste, Hox [24], pag. 63, propone una secuencia de pasos examinando la varianza residual, desde el modelo vacío hasta el que incluye todas las variables de interés. El resultado es la obtención de un pseudo R^2 , \tilde{R}^2 , dado que es una medida de mejora en la capacidad explicativa del modelo *relativa al modelo vacío*.

Para un modelo con dos niveles, el procedimiento es el siguiente:

- i. Estimar el modelo vacío para conocer la variabilidad por niveles y total, que sería igual a la suma de los componentes de varianza. Si el subíndice v identifica la estimación residual del modelo vacío, podemos escribir:

$$\text{Nivel 1: } \sigma_{\varepsilon|v}^2 \quad (23)$$

$$\text{Nivel 2: } \sigma_{\delta|v}^2 \quad (24)$$

$$\text{Variabilidad total: } \sigma_{\delta|v}^2 + \sigma_{\varepsilon|v}^2 \quad (25)$$

Donde $\sigma_{\varepsilon|v}^2$ es la varianza no explicada a nivel 1 y $\sigma_{\delta|v}^2$ es la varianza no explicada a nivel 2. Estas varianzas surgen del ANOVA con efectos aleatorios.

- ii. Estimar el modelo de intercepto aleatorio con variables explicativas a nivel individual y contextual, que denominaremos modelo alternativo 1. Los componentes de varianza de este modelo se identifican con el subíndice a . A partir de estas estimaciones podemos obtener la proporción de la varianza explicada en cada nivel:

$$\text{Nivel 1: } \tilde{R}_1^2 = \left(\frac{\sigma_{\varepsilon|v}^2 - \sigma_{\varepsilon|a}^2}{\sigma_{\varepsilon|v}^2} \right) \quad (26)$$

$$\text{Nivel 2: } \tilde{R}_2^2 = \left(\frac{\sigma_{\delta|v}^2 - \sigma_{\delta|a}^2}{\sigma_{\delta|v}^2} \right) \quad (27)$$

Donde $\sigma_{\varepsilon|a}^2$ es la varianza no explicada a nivel 1 y $\sigma_{\delta|a}^2$ es la varianza no explicada a nivel 2 del modelo alternativo 1.

- iii. Estimar el modelo de intercepto y pendiente aleatoria con las mismas variables que en el paso ii, al que nos referiremos como el modelo alternativo 2. En este caso, estimamos dos nuevos componentes de varianza, uno debido a la pendiente, $\sigma_{\eta|p}^2$, y otro es la covarianza entre la pendiente y el intercepto a nivel 2, $\sigma_{\eta\delta|p}^2$. El subíndice p identifica la varianza estimada para el modelo con pendiente e intercepto aleatorio. El subíndice η identifica la varianza perteneciente a la pendiente. Este será nuestro modelo base, similar al modelo vacío, para comparar la bondad de ajuste de modelos, con intercepto y pendiente aleatoria, que intentan explicar la variabilidad.
- iv. Estimar el modelo alternativo 3, con pendiente aleatoria incluyendo efectos cruzados entre los niveles. A continuación, podemos estimar el poder explicativo del modelo:

$$\text{Nivel 2: } \tilde{R}_2^2 = \left(\frac{\sigma_{\eta|p}^2 - \sigma_{\eta|a}^2}{\sigma_{\eta|p}^2} \right) \quad (28)$$

Esta secuencia de pasos es un procedimiento fácil de realizar, dado que todos los parámetros utilizados son brindados por los paquetes estadísticos, pero tiene algunos inconvenientes a destacar. Primero, es posible que algunos modelos arrojen \tilde{R}^2 negativos como mencionan Roberts y Hox [23], pag. 24. Segundo, en modelos con pendientes aleatorias, la varianza estimada depende de la escala de las variables explicativas. Bosker y Snijders [2], capítulo 7, proponen una corrección para el primer problema, pero el segundo queda sin resolver.

3. Aplicación: Demanda de servicios de salud. Gasto farmacéutico

La demanda individual de servicios de salud ha sido estudiada ampliamente. Las investigaciones se ha centrado en los determinantes de la misma y este objetivo ha sido abordado por diversas vías, por ejemplo, a través del número de visitas efectuadas a profesionales médicos o analizando el propio gasto sanitario.

En España, los estudios del tipo aplicado sobre el gasto sanitario se han concentrado en el gasto privado de los hogares, tomando como base de referencia los datos de la Encuesta de Presupuestos Familiares o de la Encuesta Continua de Presupuestos Familiares. Para una revisión de la literatura reciente remitimos al lector al trabajo de Clavero y González [5].

Nuestro interés, en particular, está centrado en la aplicación del modelo multinivel, y por consiguiente en la especificación econométrica, al caso de la demanda de servicios farmacéuticos. Como ponen de manifiesto Clavero y González [5], la particularidad de los datos individuales en gasto sanitario radica en la existencia de abundantes gastos nulos, lo que dirige la especificación hacia el campo de los modelos censurados. Entre estos modelos, identificamos tres alternativas:

- 1) Modelo Tobit: caracterizado por la existencia de un proceso único de generación de demanda y del nivel del gasto.
- 2) Modelo de Heckman: proceso en dos etapas, una determina la generación de demanda y otra la cantidad demandada. Es posible que las variables explicativas difieran entre las dos etapas. Heckman considera que los gastos nulos se crean en la primera etapa, de la decisión de participar, por lo que en la segunda etapa existen únicamente gastos positivos.
- 3) Modelo de Cragg: es un proceso de dos etapas, una determina la generación de demanda y otra la cantidad demandada. Las variables explicativas deben ser las mismas en ambas etapas. Los gastos nulos pueden generarse en cualquier momento, en la decisión de participar o en la de demandar, pero sin incurrir en gasto alguno.

Una revisión detallada, con aplicaciones de cada modelo, puede consultarse en Cameron y Trivedi [3]; Jones [27] realiza una descripción detallada de los diversos métodos econométricos aplicados al área de la salud. Por nuestra parte, la naturaleza misma del proceso estudiado nos guía en la elección. Nuestra investigación supone que el médico puede influenciar la demanda individual una vez que se ha realizado la demanda efectiva. El individuo recurre al médico por una determinada dolencia y, una vez realizada esta acción, el médico interviene en el proceso de demanda de medicamentos. Luego, la posibilidad de incorporar variables que no estaban en el proceso de demanda efectiva nos inclina hacia el planteamiento de Heckman [22].

En este modelo, se supone que, en la primera etapa, se realiza la generación de demanda del servicio mediante una especificación probit y, en la segunda etapa, se evalúa la cantidad demandada. Dado que esta especificación admite el uso de variables explicativas diferentes en ambas etapas, nuestra intención es resolver la segunda etapa usando un enfoque multinivel.

A continuación, pasamos a resumir la propuesta de Heckman. La primera etapa trata de calcular cual es la probabilidad de que cualquier individuo, con determinadas características socioeconómicas, genere un gasto positivo en medicamentos. Para ello, se procede a la formulación de un modelo probit:

$$y_i > 0 (h_i = 1) \quad \text{si} \quad h_i^* > 0 \quad (29)$$

$$y_i = 0 (h_i = 0) \quad \text{si} \quad h_i^* \leq 0 \quad (30)$$

$$\text{Siendo } h_i^* = x_{2i}' \cdot \beta_2 + \varepsilon_{2i} \quad i = 1, \dots, N \quad (31)$$

Donde y_i es el gasto efectuado por el individuo i en medicamentos; h_i es igual a 1 si el individuo i ha generado un gasto positivo en la misma y 0, en caso contrario.

La variable latente es h_i^* , que representa la diferencia entre las utilidades generadas por las dos *supuestas* alternativas de que dispone un individuo, acceder a un servicio sanitario o no. Las características particulares del individuo están incluidas en el vector x_{2i} .

Utilizando la ecuación (26), se puede obtener la probabilidad de observar un gasto positivo a partir del siguiente resultado:

$$P_i = P(h_i^* > 0) = P(\varepsilon_{2i} > -x_{2i}' \cdot \beta_2) = 1 - F(-x_{2i}' \cdot \beta_2) = F(x_{2i}' \cdot \beta_2) \quad (32)$$

Donde $F(\cdot)$ es la función de distribución de probabilidad del término de error ε_{2i} , determinada a partir de una función de densidad simétrica con media cero.

La segunda parte en el proceso de especificación tiene por objeto explicar el gasto generado en farmacia, a partir los datos relativos a aquellos pacientes que efectivamente generan un gasto positivo:

$$y_i = y_i^* = x_{1i}' \cdot \beta_1 + \varepsilon_{1i} \quad \text{si } y_i > 0 (h_i = 1) \quad (33)$$

Para completar la propuesta tradicional de Heckman, es necesario completar las expresiones (26) y (28) mediante la especificación de la distribución de probabilidad para los errores aleatorios, ε_{1i} y ε_{2i} . Es habitual asumir una distribución normal bivariante:

$$\begin{bmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right] \quad (34)$$

La ecuación que resume el modelo formulado es:

$$y_i = x_{1i}' \cdot \beta_1 + \sigma_{12} \lambda(x_{2i}' \cdot \beta_2) + v_i \quad (35)$$

Donde $\lambda(x_{2i}' \cdot \beta_2) = \frac{\phi(x_{2i}' \cdot \beta_2)}{\Phi(x_{2i}' \cdot \beta_2)}$ es la inversa del ratio de Mills, siendo $\phi(\cdot)$ la función de densidad

y $\Phi(\cdot)$ la función acumulada de la distribución normal estándar, v_i es un término de perturbación aleatoria.

Es en esta segunda etapa donde incorporamos al modelo básico tradicional, la estructura jerárquica que subyace en los datos:

$$y_{ij} = x_{1ij}' \cdot \beta_1 + \sigma_{12} \lambda(x_{2ij}' \cdot \beta_2) + \eta_j x_{1ij}^* + \delta_j + \varepsilon_{ij} \quad (36)$$

Donde v_i es un término de error compuesto, $v_i = \eta_j x_{1ij}^* + \delta_j + \varepsilon_{ij}$, para capturar la estructura multinivel del intercepto y de la pendiente.

Esta modificación plantea un problema que no debemos obviar. El modelo de Heckman permite introducir nuevas variables explicativas, (por ejemplo, en forma de efectos fijos) pero difícilmente admite la introducción de elementos adicionales en la estructura del error, (por ejemplo, en forma de efectos aleatorios).

La revisión de la literatura no nos ayudó a encontrar una salida al problema planteado: la estimación conjunta del modelo multinivel con datos censurados. Por lo tanto, ante la imposibilidad de realizar una estimación integrada Heckman-multinivel, optamos, como estrategia alternativa, por estimar la ecuación multinivel para los gastos positivos y, posteriormente, introducir la inversa de Mills obtenida de un probit similar al de la primera etapa de Heckman, con el objetivo de disponer de una medida del sesgo de selección.

3.1. Fuente de datos y análisis descriptivo

La base empírica que utilizaremos proviene del gasto generado por toda la población cubierta por el Sistema Nacional de Salud en Aragón. Disponemos de datos individuales para cada beneficiario del Sistema de Salud en el que se detallan los costes que ha ocasionado en 2004 en hospitalización y farmacia. Además, la base de datos contiene ciertas características de los propios individuos, información procedente de la tarjeta sanitaria.

La base de datos también contiene la identificación del médico de cabecera y del médico que prescribe los medicamentos, coincidiendo en el 80% de los casos. Una limitación importante en la información radica en que no se conocen las enfermedades diagnosticadas a cada individuo. Otra limitación es la falta de datos sobre otras características personales relevantes como la renta, ocupación, formación, etc.

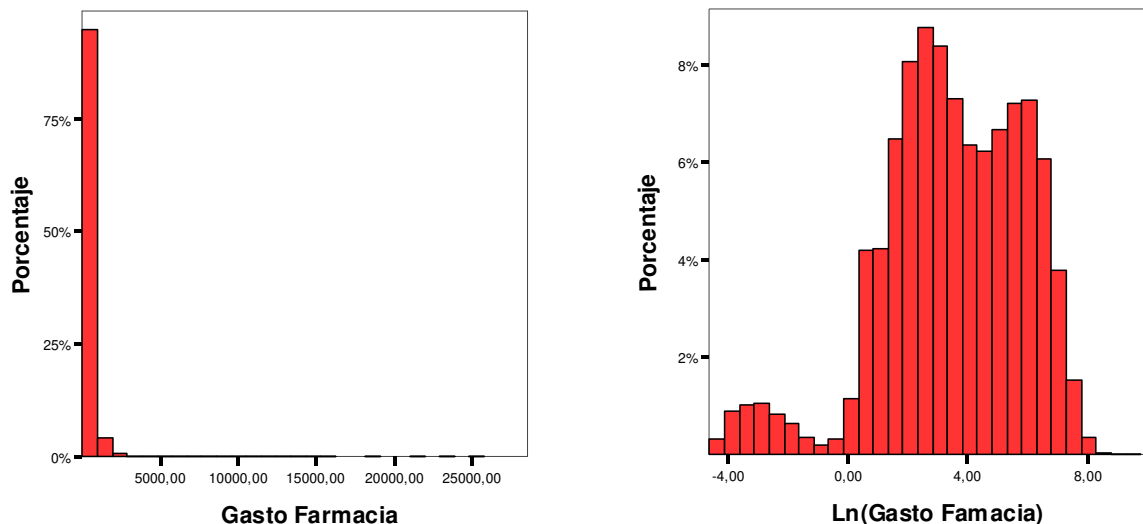
Existen 1046 de médicos de cabecera registrados en la base y asignados a diversas zonas sanitarias (divisiones geográficas de no más de 250.000 habitantes). No podemos ignorar que las zonas sanitarias constituyen un nivel de agrupación superior, implicando que los pacientes estén agrupados en médicos y los médicos en zonas sanitarias.

Para ser consecuentes con lo desarrollado en la parte metodológica, trabajaremos únicamente con dos niveles: pacientes y médicos, restringiendo nuestro marco de análisis geográfico. Por su amplitud de cobertura se eligió a la ciudad de Zaragoza como unidad geográfica homogénea (aunque la misma está dividida en 3 zonas sanitarias): el 67% de los pacientes y el 56,9% de los médicos de cabecera pertenecen a esa ciudad. Finalmente y tras depurar los registros, el tamaño de la base de datos quedó reducido a 832.699 de individuos distribuidos en 594 médicos de cabecera.

La variable dependiente, gasto en farmacia, presenta dos complicaciones: (1) El gasto positivo está muy sesgado hacia la derecha con un valor medio de 206,57 siendo muy superior a la mediana de 33,84 y (2) Existe un 34,59% de gastos nulos.

La primera complicación puede ser salvada mediante la transformación logarítmica, que permite reducir considerablemente el sesgo de la distribución: la media es de 3,48 y está cercana a la mediana de 3,52. Además, se produce una disminución de la asimetría de 7,26 hasta -0,722, mientras que la curtosis pasa de 200,92 a 3,69, también más cercana al valor esperado de 3 para una distribución normal. En el siguiente gráfico podemos observar los dos histogramas.

Gráfico Nº 6: Histograma del gasto en farmacia, en nivel y en logaritmo



El segundo problema que identificábamos es estándar en la literatura sobre sistemas de demanda y lo trataremos mediante la corrección tradicional propuesta por Heckman en la sección 3.3.

En el cuadro N° 1, se presentan las variables explicativas utilizadas en el trabajo: sexo, edad, nacionalidad, si es o no pensionista, si es o no extranjero y la especialidad del médico de cabecera.

Cuadro N° 1: Definición de las variables explicativas

Variable	Descripción
<i>Mujer</i>	Sexo del usuario, 1=mujer y 0=varón
<i>Edad</i>	Medidos en años: de 0 a 108
<i>Pensionista</i>	Actividad, 1=pensionista y 0=activo
<i>Extranjero</i>	Nacionalidad, 1=extranjero y 0=español
<i>Pediatra</i>	Especialidad del médico de cabecera, 1=Pediatra y 0=Generalista

La estadística descriptiva más relevante de estas variables se presenta en el cuadro N° 2. El cuadro se divide en dos partes en dos partes: para la muestra completa de Zaragoza y para los pacientes con gasto positivo únicamente. Los cambios más importantes que se detectan se refieren a la proporción de pensionistas y de extranjeros.

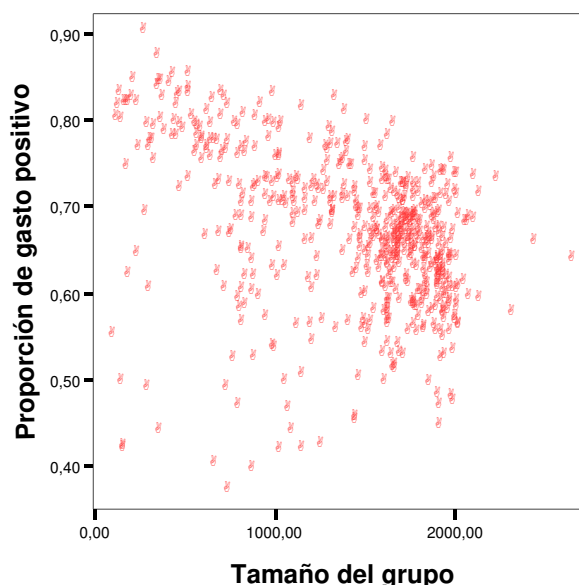
Cuadro N° 2: Definición de las variables explicativas

Variable	Muestra completa		Muestra con gasto positivo	
	Media	Desv. est.	Media	Desv. Est.
<i>Mujer</i>	0,512	0,499	0,556	0,496
<i>Edad</i>	41,92	22,639	45,71	24,109
<i>Pensionista</i>	0,262	0,440	0,355	0,478
<i>Extranjero</i>	0,068	0,252	0,045	0,207
<i>Pediatra</i>	0,113	0,317	0,112	0,315
<i>Observaciones</i>	832.699		544.609	

El número medio de pacientes por grupo es de 1402, aproximadamente, con un mínimo de 93 y un máximo de 2644. Pero, al trabajar con el gasto positivo, la cantidad media de pacientes por grupo pasa a 917, siendo el grupo más pequeño de 51 pacientes y el más grande de 1684.

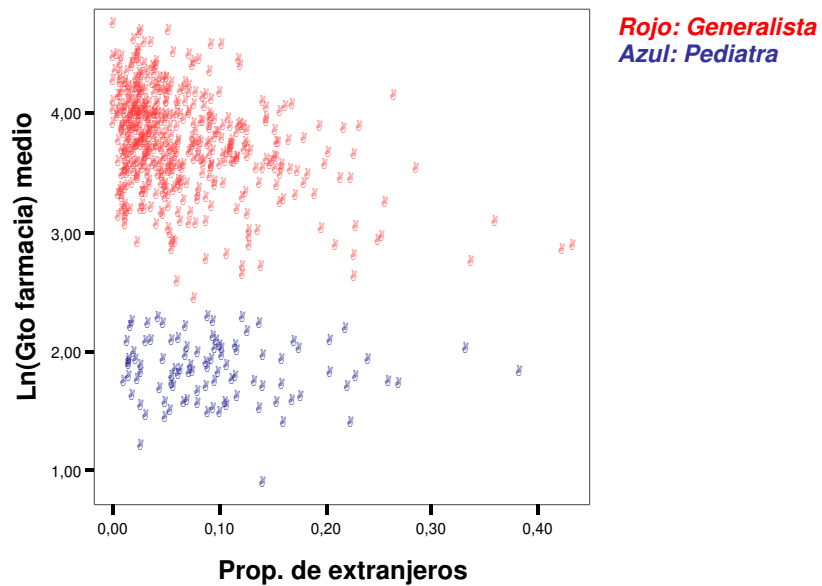
La relación entre la proporción de pacientes con gasto positivo y el tamaño grupal muestra una correlación negativa, como se puede ver en la siguiente gráfica.

Gráfico N° 7: Relación entre el tamaño del grupo y proporción de gasto positivo



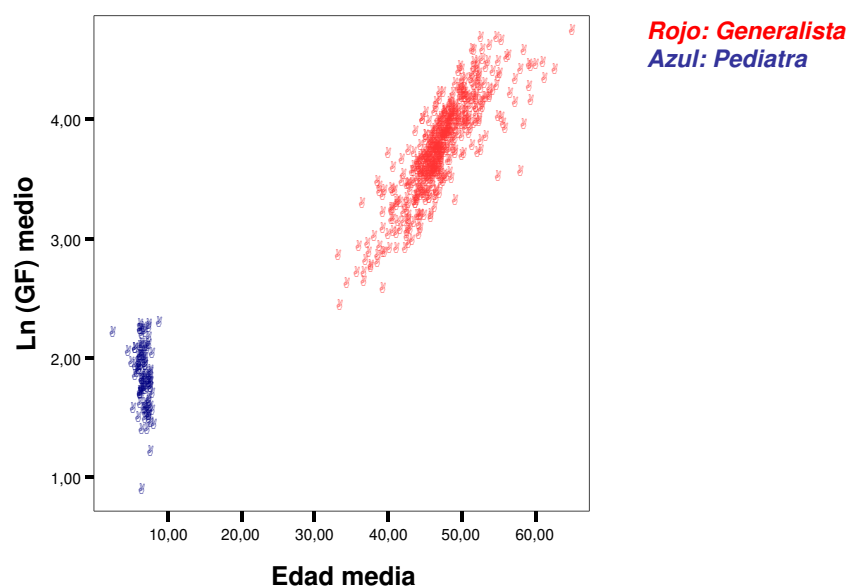
La relación entre la proporción de extranjeros en el grupo y gasto positivo medio del grupo, muestra una relación negativa para los médicos generalistas pero no para los pediatras, como se observa en el gráfico N° 8.

Gráfico N° 8: Relación entre proporción de extranjeros y de gasto positivo por grupo



De los 594 médicos de cabecera, únicamente 98 son pediatras y los otros 496 son generalistas. Los primeros asisten a pacientes con una edad media de 7 años mientras que los segundos tienen grupos de pacientes con una media de edad cercana a 46 años. Esta marcada diferencia de edades medias tiene una extensión en lo que respecta al gasto medio por grupo, como puede visualizarse en el gráfico N° 9. El logaritmo del gasto en farmacia tiene una media de 3,7524 para los generalistas y de 1,80 para los pediatras. Además, podemos observar una relación positiva entre el gasto en farmacia y la edad media del grupo para los generalistas. En el caso de los pediatras, existe una mayor dispersión en el gasto en farmacia pero no se relaciona a la edad media grupal.

Gráfico N° 9: Relación entre edad media y gasto positivo por grupo



3.2. Resultados de los modelos alternativos

En esta sección presentamos los principales resultados obtenidos de los modelos alternativos. Siguiendo la secuencia de la sección metodológica, aplicaremos primeramente el modelo vacío para conocer si existe diferencia del gasto farmacéutico por grupo. Luego, estimamos el modelo de intercepto aleatorio y, posteriormente, el modelo de intercepto y pendiente aleatoria. Toda la información de las estimaciones se presenta en el cuadro N° 2.

Por último, realizamos un análisis de las predicciones de la parte fija y aleatoria.

Como comentábamos anteriormente, la estimación del modelo vacío puede realizarse mediante una simple descomposición de varianza o por máxima verosimilitud, nosotros aplicamos este último procedimiento. El valor del coeficiente de correlación intraclase (ρ) obtenido nos brinda una medida de la dependencia de los individuos al grupo, siendo igual a:

$$\hat{\rho} = \frac{\hat{\sigma}_{\delta}^2}{\hat{\sigma}_{\delta}^2 + \hat{\sigma}_{\varepsilon}^2} = \frac{0,64576}{0,64576 + 5,4291} = 0,1063$$

Esto significa que el 10,63% del logaritmo del gasto en farmacia de los pacientes es debido a que acuden a un mismo médico de cabecera.

Para la estimación de los dos modelos siguientes, debido a la poca cantidad de variables explicativas, hemos utilizado todas las variables disponibles con interacciones entre las mismas. La especificación econométrica que elegimos es la siguiente:

$$\begin{aligned} \ln(\text{Gasto}_{\text{Farmacia}_{ij}}) = & \beta_0 + \beta_1 \text{Mujer}_{ij} + \beta_2 \text{Edad}_{ij} + \beta_3 \text{Edad}_{ij}^2 + \beta_4 (\text{Edad}_{ij} \cdot \text{Mujer}_{ij}) + \beta_5 (\text{Edad}_{ij}^2 \cdot \text{Mujer}_{ij}) \\ & + \beta_6 \text{Pensionado}_{ij} + \beta_7 \text{Extranjero}_{ij} + \beta_8 \text{Pediatra}_j + \beta_9 (\text{Pediatra}_j \cdot \text{Edad}_{ij}) + \delta_j + \eta_j \text{Edad}_{ij} + \varepsilon_{ij} \end{aligned} \quad (37)$$

Previo a la ecuación (37), se realizaron especificaciones alternativas con el empleo de variables dummies para diferentes tramos de edades y con interacciones entre estos tramos y el sexo del paciente. Los efectos obtenidos mediante ambas especificaciones eran similares, optándose por la ecuación que presentaba mayores grados de libertad.

Los efectos aleatorios de la pendiente son referidos a la edad de los pacientes, siendo la única variable continua de que disponíamos en el modelo.

La estimación de los modelos se llevó a cabo mediante el método de Máxima Verosimilitud Restringida. Se incluyen los estadísticos de bondad del ajuste, deviance, AIC y BIC.

Todos los coeficientes tienen el signo esperado. Los valores de los coeficientes estimados de los modelos 2 y 3 son similares. Todos los coeficientes son significativos al 1%.

Si el individuo es pensionista, el gasto en farmacia es mayor. Si es extranjero, su nivel de gasto es menor. Para pacientes que no llegan al año de edad, tener como médico de cabecera a un pediatra, implica un gasto en farmacia superior, aunque este gasto va disminuyendo con la edad. Si el paciente es de sexo femenino, su gasto durante los primeros años de vida es inferior al del varón. A medida que aumenta la edad, el gasto de las mujeres supera al de los varones, aunque pasados los 80 años de edad, aproximadamente, el gasto de las mujeres vuelve a ser inferior al de los varones.

Los resultados de los efectos fijos confirman las conclusiones de otras investigaciones realizadas para España.

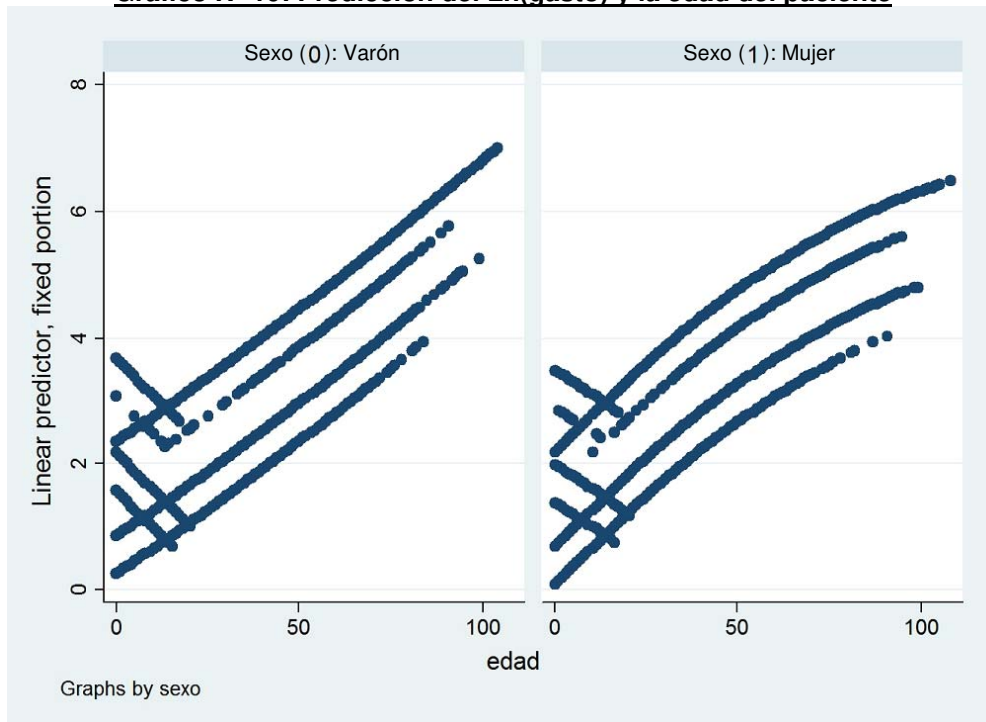
Cuadro N° 2: Estimación de modelos alternativos

Efectos Fijos	Modelo 1	Modelo 2	Modelo 3
Intercepto	3,40494 (0,3318)	0,8763 (0,0201)	0,8545 (0,0238)
Mujer		-0,1870 (0,0161)	-0,1817 (0,0161)
Edad		0,03803 (0,0008)	0,0225 (0,0008)
Edad_Mujer		0,02266 (0,0008)	0,02259 (0,0008)
Edad ²		0,00006 (0,00001)	0,00005 (0,00001)
Edad ² _Mujer		-0,00025 (0,00001)	-0,00025 (0,00001)
Pensionista		1,49399 (0,00859)	1,4941 (0,00859)
Extranjero		-0,6117 (0,0124)	-0,6033 (0,0125)
Pediatra		1,2978 (0,0190)	1,3142 (0,0429)
Pediatra_Edad		-0,09721 (0,019)	-0,0982 (0,020)
Efectos Aleatorios	Modelo 1	Modelo 2	Modelo 3
Var(Intercepto), $\hat{\sigma}_\delta^2$	0,64576 (0,0381)	0,04262 (0,0027)	0,1158 (0,0084)
Var(Edad), $\hat{\sigma}_\eta^2$			0,000028 (0,000002)
Cov(Intercepto, Edad), $\hat{\sigma}_{\delta\eta}^2$			-0,001501 (0,00013)
Var(Residuo), $\hat{\sigma}_\varepsilon^2$	5,4291 (0,0104)	3,3968 (0,0065)	3,3865 (0,0065)
Estadísticos del Modelo	Modelo 1	Modelo 2	Modelo 3
Deviance		2.214.523	2.212.154
AIC		2.214.547	2.212.182
BIC		2.214.681	2.212.339

Entre paréntesis se muestran los errores estándares.

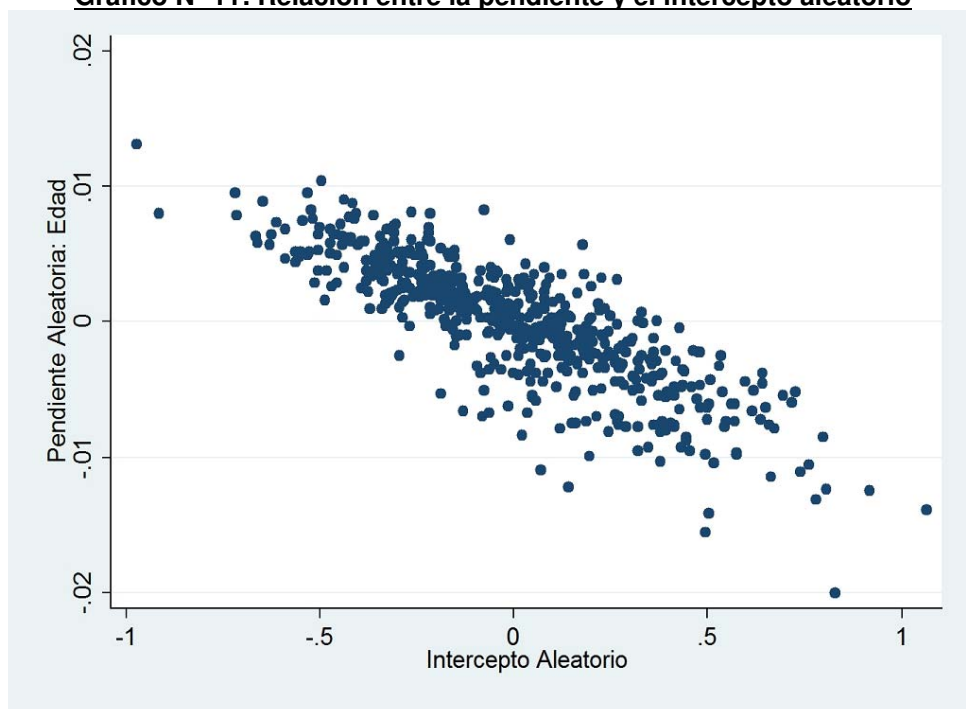
La relación entre sexo y edad del paciente y su afecto en el gasto en medicamentos puede verse más claramente en el gráfico N° 10. En el eje de las ordenadas se tienen los valores predichos según la estimación del modelo 3 y en el eje de las abscisas la edad de los pacientes. La gráfica de la derecha representa al gasto de las mujeres y la gráfica de la izquierda al gasto de los varones.

Gráfico N° 10: Predicción del Ln(gasto) y la edad del paciente



La relación entre el intercepto y la pendiente, ambas aleatorias, se muestra en el gráfico N° 11. La dispersión de puntos está alrededor de los valores esperados, cercanos a 0. La covarianza es claramente negativa: altos valores de las pendientes van acompañados de bajos valores de los interceptos para los grupos, y viceversa.

Gráfico N° 11: Relación entre la pendiente y el intercepto aleatorio



Otro aspecto a reseñar es que, cualquiera que sea el criterio que apliquemos (*deviance*, AIC o BIC), el modelo que mejor se ajusta a los datos es el 3.

Respecto a una medida de bondad de ajuste, utilizaremos las fórmulas desarrolladas en la parte 2.3. A partir de estas estimaciones podemos obtener la proporción de la varianza explicada en cada nivel del modelo 3:

$$\text{Nivel 1: } \tilde{R}_1^2 = \left(\frac{\sigma_{\varepsilon|v}^2 - \sigma_{\varepsilon|a}^2}{\sigma_{\varepsilon|v}^2} \right) = \frac{5,4291 - 3,3865}{5,4291} = 0,3762 \quad (38)$$

$$\text{Nivel 2: } \tilde{R}_2^2 = \left(\frac{\sigma_{\delta|v}^2 - \sigma_{\delta|a}^2}{\sigma_{\delta|v}^2} \right) = \frac{0,64576 - 0,1158}{0,64576} = 0,8206 \quad (39)$$

Para el nivel 1, la bondad del ajuste es baja, aunque aceptable si la comparamos con otros estudios donde se utilizan datos individuales en corte transversal. Para el nivel 2, la bondad del ajuste es muy buena.

3.3. Corrección por sesgo de selección

Hasta ahora hemos trabajado con el gasto positivo. Esta decisión ha generado que el 34,59% de las observaciones disponibles no sean tomadas en cuenta para el análisis, ocasionando posibles sesgos en las estimaciones de nuestros coeficientes.

Para tener una medida de este sesgo, procederemos a estimar el probit de la primera etapa de Heckman. Necesitamos, para ello, definir una variable que asuma valor 1 si el gasto en farmacia es positivo y valor 0 si los gastos son nulos.

La especificación econométrica elegida es la siguiente:

$$\text{Gasto_Farmacia}_i > 0 (h_i = 1) \quad \text{si } h_i^* > 0 \quad (40)$$

$$\text{Gasto_Farmacia}_i = 0 (h_i = 0) \quad \text{si } h_i^* \leq 0 \quad (41)$$

$$\text{Con } h_i^* = x_{2i}' \cdot \beta_2 + \varepsilon_{2i} \quad i = 1, \dots, N \quad (42)$$

Donde:

$$x_{2i}' \cdot \beta_2 = \left[\beta_{2,0} + \beta_{2,1} \text{Mujer}_{ij} + \beta_{2,2} \text{Edad}_{ij} + \beta_{2,3} \text{Edad}_{ij}^2 + \beta_{2,4} (\text{Edad}_{ij} \cdot \text{Mujer}_{ij}) + \beta_{2,5} (\text{Edad}_{ij}^2 \cdot \text{Mujer}_{ij}) \right. \\ \left. + \beta_{2,6} \text{Pensionado}_{ij} + \beta_{2,7} \text{Extranjero}_{ij} + \beta_{2,8} \text{Pediatra}_j + \beta_{2,9} (\text{Pediatra}_j \cdot \text{Edad}_{ij}) + \varepsilon_{2,ij} \right] \quad (43)$$

Obsérvese que en esta primera etapa no intervienen las variables relativas a la especialidad del médico de cabecera ni a la descomposición del error que captura la estructura jerárquica. Esto es debido a que consideramos que la decisión de demandar o no, servicios de salud, es propia del individuo.

Por su parte, la ecuación que completa la formulación del modelo de Heckman es la siguiente:

$$\ln(\text{Gasto_Farmacia})_{ij} = x'_{1ij} \cdot \beta_1 + \sigma_{12} \lambda(x'_{2ij} \cdot \beta_2) + \eta_j \text{Edad}_{1ij} + \delta_j + \varepsilon_{ij} \quad (44)$$

$$\text{Donde: } x'_{1ij} = \beta_{1,0} + \beta_{1,1} \text{Mujer}_{ij} + \beta_{1,2} \text{Edad}_{ij} + \beta_{1,3} \text{Edad}_{ij}^2 + \beta_{1,4} (\text{Edad}_{ij} \cdot \text{Mujer}_{ij}) + \beta_{1,5} (\text{Edad}_{ij}^2 \cdot \text{Mujer}_{ij}) \\ + \beta_{1,6} \text{Pensionado}_{ij} + \beta_{1,7} \text{Extranjero}_{ij} + \beta_{1,8} \text{Pediatra}_j + \beta_{1,9} (\text{Pediatra}_j \cdot \text{Edad}_{ij}) \quad (45)$$

En el cuadro N° 3, presentamos los resultados para la ecuación (40), denominamos a esta estimación *modelo corregido* y realizamos un paralelo con el modelo 3, *modelo sin corregir*.

Cuadro N° 3: Comparación entre el modelo con corrección y el modelo sin corrección

Efectos Fijos	Modelo con corrección	Modelo sin corrección
Intercepto	4,2197 (0,06703)	0,8545 (0,0238)
Mujer	0,2364 (0,0178)	-0,1817 (0,0161)
Edad	0,1076 (0,0015)	0,0225 (0,0008)
Edad_Mujer	-0,0589 (0,0017)	0,02259 (0,0008)
Edad ²	-0,00099 (0,00002)	0,00005 (0,00001)
Edad ² _Mujer	0,00073 (0,00002)	-0,00025 (0,00001)
Pensionista	-0,0227 (0,0296)	1,4941 (0,00859)
Extranjero	0,7267 (0,0278)	-0,6033 (0,0125)
Pediatra	0,7339 (0,0409)	1,3142 (0,0429)
Pediatra_Edad	-0,06835 (0,00205)	-0,0982 (0,020)
Inversa de Mills	-5,5580 (0,1041)	
Efectos Aleatorios	Modelo con corrección	Modelo sin corrección
Var(Intercepto), $\hat{\sigma}_\delta^2$	0,09306 (0,007)	0,1158 (0,0084)
Var(Edad), $\hat{\sigma}_\eta^2$	0,000022 (0,000002)	0,000028 (0,000002)
Cov(Intercepto, Edad), $\hat{\sigma}_{\delta\eta}^2$	-0,001108 (0,000106)	-0,001501 (0,00013)
Var(Residuo), $\hat{\sigma}_\varepsilon^2$	3,3697 (0,0064)	3,3865 (0,0065)
Estadísticos del Modelo	Modelo con corrección	Modelo sin corrección
Deviance	2.209.326	2.212.154
AIC	2.209.356	2.212.182
BIC	2.209.524	2.212.339

Entre paréntesis se muestran los errores estándares.

Comparando los coeficientes estimados, se observan diferencias apreciables en la estimación de la parte fija del modelo. El parámetro estimado para la inversa de Mills es significativo, como consecuencia del impacto del sesgo de selección que sufre el modelo 3. Para las situaciones de pensionista y extranjero, este sesgo era esperable por el cambio de representación en cada muestra, señalado en el análisis descriptivo. Los otros cambios son difíciles de explicar. Los coeficientes de la parte aleatoria son similares, el efecto médico existe en ambos modelos. El modelo corregido ajusta mejor según estadísticos del modelo. Si observamos la predicción del modelo con corrección versus sin corrección, podemos apreciar claras diferencias en relación a la edad, tanto para los varones como para las mujeres, gráfico N° 12. Las predicciones de la parte aleatoria, gráfico N° 13, de ambos modelos se asemejan, con una leve mejoría en la disminución de la dispersión del corregido.

Gráfico N° 12: Predicción del Ln(gasto) y la edad del paciente

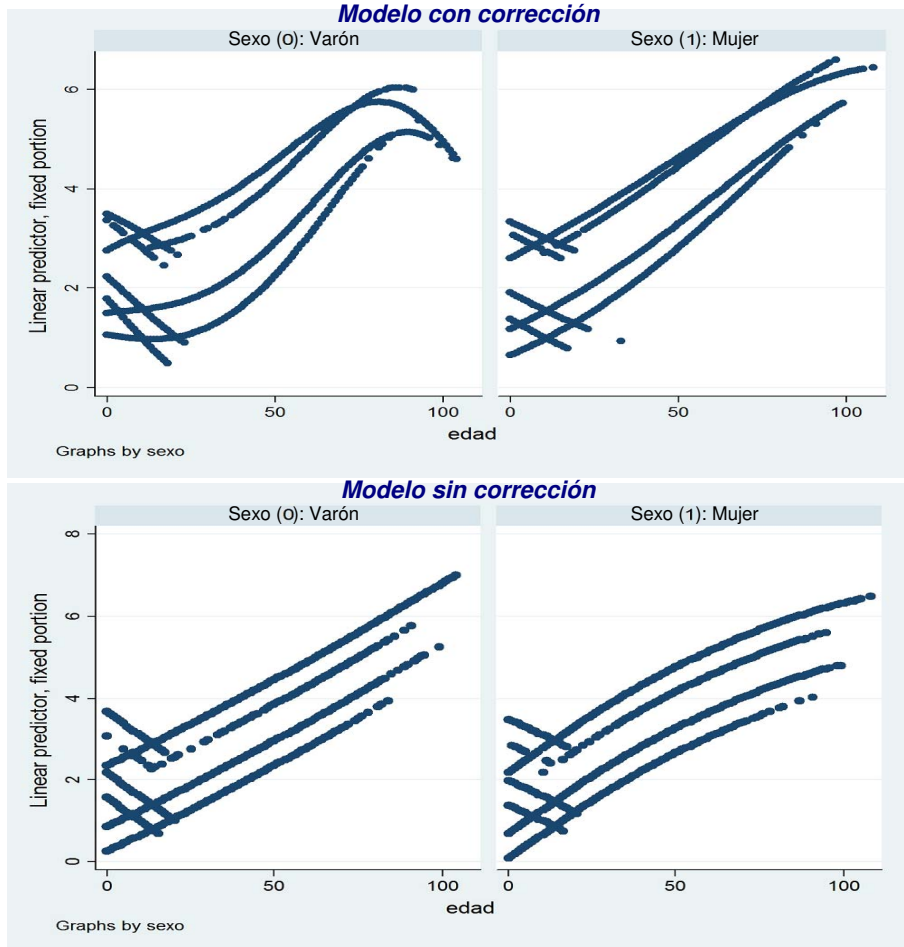
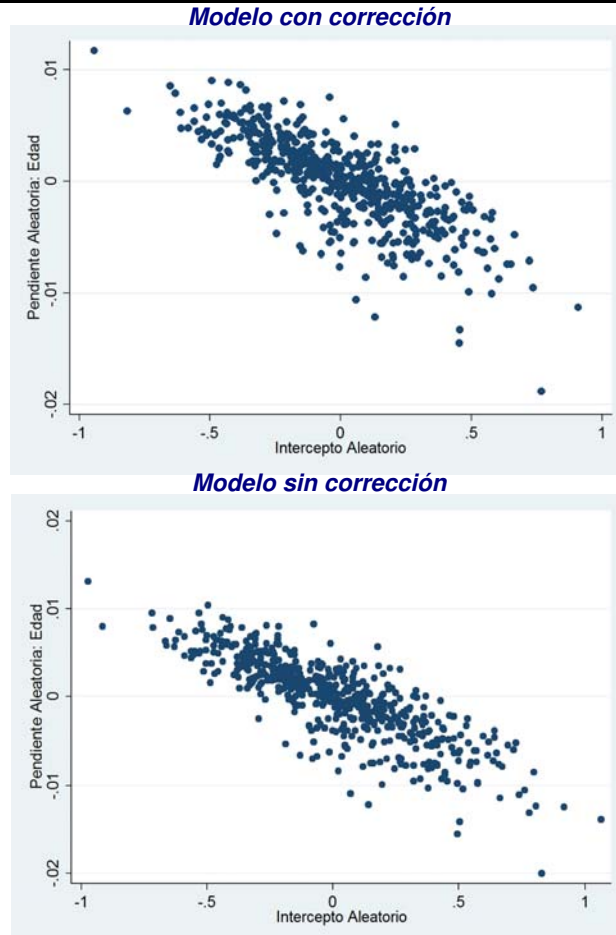


Gráfico N° 13: Relación entre la intercepto y pendiente aleatoria

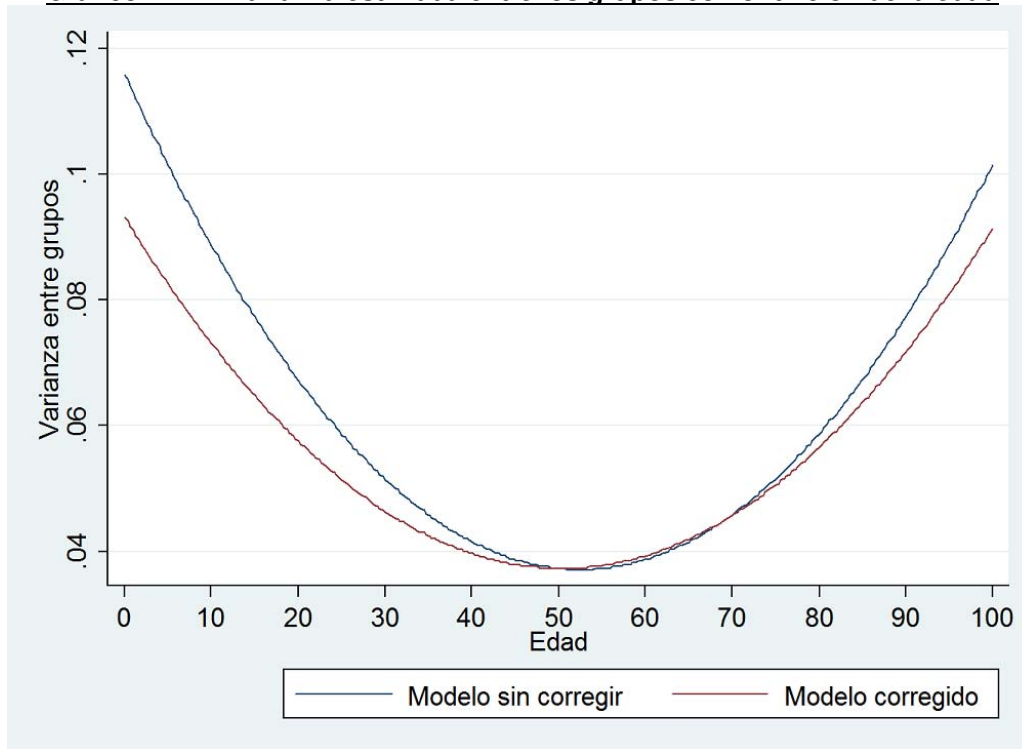


Para observar más claramente las predicciones de la parte aleatoria de ambos modelos se presenta el gráfico N° 14. Este gráfico muestra la relación de la varianza entre grupos como función de la edad, que proviene de la siguiente ecuación de varianza:

$$\text{Var}(\delta_j + \eta_j \text{Edad}_{ij}) = \sigma_\delta^2 + 2 \cdot \sigma_{\delta\eta} \text{Edad}_{ij} + \sigma_\eta^2 (\text{Edad}_{ij})^2 \quad (46)$$

Si reemplazamos en la ecuación, las varianzas y covarianzas estimadas para cada modelo, obtendremos las relaciones presentadas en el siguiente gráfico.

Gráfico N° 14: Varianza estimada entre los grupos como función de la edad



Podemos ver como el modelo sin corrección sobreestima la varianza entre los grupos a medida que nos alejamos de los valores centrales: para edades menores a 50 años y para mayores a 70 años. La sobreestimación es más notoria para los grupos más jóvenes.

Si bien, hemos detectado sesgo en las estimaciones, las conclusiones principales no defieren. El modelo corregido, igual que el modelo con sesgo, detecta que los médicos de cabecera afectan significativamente la demanda de servicios farmacéuticos de los pacientes. La corrección del sesgo de selección realizada debe ser matizada. El modelo de Heckman supone un procedimiento integrado de las dos etapas, y la lectura debe realizarse de forma conjunta. Esta integración no ha sido desarrollada para nuestro modelo multinivel y es tema de futuros desarrollos.

4. Conclusiones

Los modelos multiniveles tienen ventajas respecto al modelo de regresión tradicional. Entre estas ventajas se destacan: i) corrección de la subestimación de los errores estándares, ii) examen de las diferencias entre los diversos grupos, iii) eliminación del problema de agregación, iv) estimación de la variabilidad de los coeficientes entre los grupos. Mientras más disímiles sean los grupos jerárquicos, mayor sesgo existirá en las estimaciones mínimo cuadráticas y más importante será la aplicación de esta herramienta.

Respecto al examen de las diferencias entre los diversos grupos, el enfoque multinivel genera inferencias no restringidas a la muestra y sin gran pérdida de grados de libertad, superando a la estimación de efectos fijos.

La técnica multinivel produce una descomposición de la varianza de la variable dependiente entre grupos y dentro de los grupos. Las diferencias son significativas con respecto al análisis de regresión tradicional, que especifica únicamente el valor medio esperado. La técnica tiene potenciales aplicaciones para corregir la heteroscedasticidad y para realizar inferencias válidas con datos no-balanceados.

El enfoque multinivel incorpora al análisis de regresión la estructura jerárquica o contexto social. Para el caso del gasto en farmacia, el análisis multinivel resulta adecuado debido a que el gasto se realiza bajo receta, dependiendo exclusivamente del médico que prescribe su otorgamiento. Condicionando este gasto por las características propias de cada paciente, encontramos evidencia significativa de que los médicos afectan el gasto farmacéutico destinado a cada paciente.

Debido a la existencia de abundantes gastos nulos individuales, se propuso una corrección por sesgo de selección al modelo multinivel. Aplicando el modelo de Heckman para nuestra especificación encontramos evidencia de sesgo en la parte fija como aleatoria del modelo. Pero los resultados principales no difieren: el médico de cabecera influye significativamente en el gasto farmacéutico de los pacientes.

El ejemplo de aplicación presentado genera algunos desafíos a tener en cuenta en futuras investigaciones como son: la extensión a más niveles, tres o más niveles jerárquicos, y la incorporación de una corrección adecuada al sesgo de selección muestral para todos los niveles, integrando ambas etapas del modelo de Heckman.

Referencias bibliográficas

- [1] Angulo, A. M., R. Barberán, P. Egea y J. Mur (2007). Los determinantes sociodemográficos del gasto sanitario. Análisis con microdatos individuales. Documentos de trabajo FUNCAS, Nº. 333. Disponible en: [//www.funcas.ceca.es/Publicaciones/InformacionArticulos/Publicaciones.asp?ID=1316](http://www.funcas.ceca.es/Publicaciones/InformacionArticulos/Publicaciones.asp?ID=1316)
- [2] Bosker R. J. y T. A. B. Snijders (1999). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. Sage. London.
- [3] Cameron A. C. y P. K. Trivedi (2005). *Microeconometrics: Methods and applications*. Cambridge University Press, Cambridge, UK.
- [4] Chamberlain G. (1984). Panel data. In Griliches Z. and M. D. Intriligator (Eds), *Handbook of econometrics*, v2, pag. 1247–1318. North-Holland, Amsterdam.
- [5] Clavero Barranquero, A. y M. L. Gonzalez Alvarez (2005). Una revisión de modelos econométricos aplicados al análisis de demanda y utilización de servicios sanitarios. Hacienda Pública Española, Nº. 173, pag. 129-162. Disponible en http://www.ief.es/Publicaciones/Revistas/Hacienda%20Publica/173_Revision.pdf
- [6] de Leeuw J. and I. G. G. Kreft (1986). Random coefficient models for multilevel analysis. *Journal of Educational Statistics*, Nº. 11, pag. 57–85. Disponible en <http://www.stat.ucla.edu/~deleeuw>
- [7] de Leeuw J. e I. G. G. Kreft (2001). Software for multilevel analysis. In A. H. Leyland and H. Goldstein, editors, *Multilevel modelling of health statistics*, pag. 187–204. Wiley, Chichester. Disponible en <http://www.stat.ucla.edu/~deleeuw>
- [8] de Leeuw J. (2005). Centering in multilevel analysis. In Everitt B. S. y D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science*, v1, pag. 247–249. Wiley, New York. Disponible en <http://www.stat.ucla.edu/~deleeuw>
- [9] de Leeuw J. (2005). Linear multilevel models. In Everitt B. S. y D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science*, v1, pag. 247–249. Wiley, New York. Disponible en <http://www.stat.ucla.edu/~deleeuw>
- [10] de Leeuw J. y E. Meijer. (2008). Introduction to multilevel analysis. In J. de Leeuw & E. Meijer (Eds.), *Handbook of multilevel analysis*. Springer, New York.
- [11] Diez-Roux, A. V. (1998). Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *American Journal Public Health*, Nº. 88, pag. 216–22.
- [12] Draper, D. (2008). Bayesian multilevel analysis and MCMC. In J. de Leeuw & E. Meijer (Eds.), *Handbook of multilevel analysis*. Springer, New York.
- [13] Fielding A. y H. Goldstein (2006). Cross-classified and membership structures in multilevel models: An introduction and review. Department for Education and Skills. DEFS Report Nº 791. Disponible en: http://www.socscistaff.bham.ac.uk/fielding/Cross_classified_review_RR791.pdf
- [14] Freedman D. (2001). Ecological inference and the ecological fallacy. In Smelser J. y P. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences*. Elsevier (2001) v.6, pag. 4027–4030. Disponible en <http://www.stat.berkeley.edu/~census/549.pdf>
- [15] Gelman A. (2006). Multilevel (hierarchical) modeling: What it can and cannot do. *Technometrics*, Nº. 48, pag. 432–435.
- [16] Gelman A., D. K. Park, S. Anselobehere, P. N. Price, and L. C. Minnete (2001). Models, assumptions and model checking in ecological regressions. *Journal of the Royal Statistical Society, Series A*, Nº. 164, pag. 101–118.
- [17] Gerdtham, U.G. y Jönsson, B. (2000). International comparisons of health expenditure: Theory, data and econometric analysis. In Culyer, A. J. and J. P. Newhouse (Eds). *Handbook of Health Economics*, pag. 12-52. Amsterdam, Elsevier Science B.V.
- [18] Goldstein H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*, Nº. 73, pag. 43–56.
- [19] Goldstein H. (2003). *Multilevel statistical models*, 3rd edition. Edward Arnold, London.

- [20] Goldstein H. y J. Rasbash (1992). Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalized least squares. *Computational Statistics & Data Analysis*, N^o. 13, pag. 63–71.
- [21] González, B., V. Ortún, A. Cabeza, A. López, J. A. Díaz y F. Álamo (2002). Evaluación del uso apropiado de medicamentos en atención primaria. ¿Cómo se puede mejorar?. *Atención Primaria*, N^o. 30 (7), pag. 467-471.
- [22] Heckman J. J. (1979). Sample selection bias as a specification error. *Econometrica*, N^o. 47, pag. 153-161.
- [23] Hox J. J. (1995). *Applied multilevel analysis*. TT-Publikaties, Amsterdam. Disponible en <http://www.geocities.com/joophox/publist/amaboek.pdf>
- [24] Hox J. J. (2002). *Multilevel analysis: Techniques and applications*. Erlbaum, Mahwah, NJ.
- [25] Hox J. J. y I. Kreft (1994). Multilevel analysis methods. *Sociological Methods & Research*, N^o. 22 (3), pag. 283-299.
- [26] Hsiao C. (2003). *Analysis of panel data*, 2nd edition. Cambridge University Press, Cambridge, UK.
- [27] Jones, A (2000). Health econometrics. In Culyer A. J. y J. P. Newhouse (Eds.), *Handbook of health economics*, Elsevier, v 1A, chapter 6, pag. 265-344.
- [28] Kreft I. G. G., J. de Leeuw, and L. S. Aiken (1994). The effect of different forms of centering in hierarchical linear models. Technical Report N^o. 30. National Institute of Statistical Sciences. Disponible en: <http://www.niss.org/technicalreports/tr30.pdf>
- [29] Kreft I. G. G. (2005). Ecological Fallacy. In Everitt B. S. y D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science*, v2, pag. 525–527. Wiley, New York.
- [30] Maas C. J. M. y T. Snijders (2003). The multilevel approach to repeated measures for complete and incomplete data. *Quality & Quantity*, N^o. 37, pag. 71–89.
- [31] Maas C. J. M. y J. J. Hox (2004). The influence of violations of assumptions on multilevel parameter estimates and their standard errors. *Computational Statistics & Data Analysis*, N^o. 46, pag. 427–440.
- [32] McCullagh P. y J. A. Nelder (1989). *Generalized linear models*, 2nd edition. Chapman & Hall, London.
- [33] Rabe-Hesketh S., A. Skrondal, y A. Pickles (2004). GLLAMM manual. Working Paper N^o. 160, U.C. Berkeley Division of Biostatistics, Berkeley, CA. Disponible en <http://www.bepress.com/ucbbiostat/paper160/>
- [34] Rasbash, J., Browne, W., Goldstein., H., Yang, M., et al. (2000). A user's guide to MLwiN (Second Edition). London, Institute of Education. Disponible en: <http://seis.bris.ac.uk/~frwjb/materials/userman.pdf>
- [35] Roberts, J. K. (2007). Group dependency in the presence of small intraclass correlation coefficients: An argument in favor of not Interpreting the ICC. Paper presented at the annual meeting of the American Educational Research Association, April 10, 2007. Disponible en <http://www.hlm-online.com/papers/depend.pdf>
- [36] Robinson, W. S. (1950). Ecological correlation and the behavior of individuals. *American Sociological Review*, N^o. 15, pag. 351-357.
- [37] Seltzer M. y K. Choi (2003). Sensitivity analysis for hierarchical models: Downweighting and identifying extreme cases using the *t* distribution. In Riese, S. y N. Duan (Eds.), *Multilevel modeling: Methodological advances, issues, and applications*. Hillsdale, N.J., Lawrence Erlbaum.
- [38] Snijders T. A. B. y R. J. Bosker (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage Publishers, London.
- [39] Snijders T. A. B. (2005). Fixed and random effects. In B. S. Everitt y D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science*, v2, pag. 664-665. Wiley.
- [40] Snijders T. A. B. (2005). Power and sample size in multilevel modeling. In B. S. Everitt y D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science*, v3, pag. 1570-1573. Wiley. Disponible en <http://stat.gamma.rug.nl/PowerSampleSizeMultilevel.pdf>

- [41] Steele, F. (2008). Module 5: Introduction to multilevel modelling concepts. LEMMA VLE, Centre for Multilevel Modelling. Disponible en: <http://www.cmm.bris.ac.uk/learning-training/course.shtml>
- [42] Subramanian, S. V. (2004). The relevance of multilevel statistical models for identifying causal neighborhood effects. *Social Science and Medicine*, N^o. 58, pag. 1961-1967.
- [43] Swamy P. A. V. B. y G. Tavlas (2001). Random coefficient models. In Baltagi B. (Ed.), *A Companion to theoretical econometrics*, pag. 310-330. Blackwell, Malden.
- [44] Twisk, Jos W. R. (2003). *Applied longitudinal data analysis for epidemiology. A practical guide*. Cambridge University Press.
- [45] Van Landeghem G., P. Onghena, y J. Van Damme (2001). The effects of different forms of centering in hierarchical linear models re-examined. Technical Report 2001-04, Universitair Centrum vor Statistiek, University of Leuven. Disponible en http://lstat.kuleuven.be/research/reports/2001/report2001_04.pdf
- [46] Verbeke G. y E. Lesaffre (1997). The effect of misspecifying the random-effects distribution in linear mixed models for longitudinal data. *Computational Statistics & Data Analysis*, N^o. 23, pag. 541–556.
- [47] Verbeke G. y G. Molenberghs (2000). *Linear mixed models for longitudinal data*. Springer, New York.