

---

**Programas para la realización de Modelos Multinivel. Un análisis comparativo  
entre MLwiN, HLM, SPSS y Stata**

**Multilevel Analysis Software. A comparative study of MLwiN, HLM, SPSS and  
Stata**

Cynthia Martínez-Garrido y F. Javier Murillo

*Universidad Autónoma de Madrid*

---

**RESUMEN**

Hoy en día contamos con diferentes programas estadísticos para la estimación de los Modelos Multinivel, o Modelos Jerárquicos. En este artículo se realiza una comparación entre los cuatro programas más utilizados para ello, el MLwiN, el HLM, el SPSS y el Stata. La estrategia de análisis es hacer una ejemplificación de su uso y, a partir de la misma, describir sus ventajas y limitaciones en el desarrollo de los Modelos Multinivel. Los resultados revelan que, aunque esencialmente muestran los mismos resultados, existen diferencias en torno a la cantidad de niveles de análisis que soportan, el tratamiento de los residuos, o el método de estimación de los componentes de la varianza asignado por defecto. Con todo ello, el investigador podrá guiar sus pasos en el desarrollo de los Modelos Multinivel y tomar una decisión fundamentada en torno a qué programa utilizar en función de las necesidades propias de su estudio.

*Palabras clave:* Análisis Multinivel, Programas Estadísticos, MLwiN, HLM, SPSS, Stata.

**ABSTRACT**

Nowadays there are different statistical software to estimate Multilevel Models, or Hierarchical Models. In this paper we compare the more common software to develop Multilevel analysis: MLwiN, HLM, SPSS, and Stata. We include a brief example of their use, and describe the advantages and disadvantages of each software. Our results show there are differences between how many levels you can include, how the software works with residuals components, the estimation methods of variance components assigned by default. Researchers can see the Multilevel Analysis process, and decide which software is the best for their study.

*Keywords:* Multilevel Analysis, Statistical Software, MLwiN, HLM, SPSS, Stata.

---

Contacto:

Cynthia Martínez-Garrido.

E-mail: [Cynthia.martinez@uam.es](mailto:Cynthia.martinez@uam.es)

## 1.- Introducción

En la actualidad, cualquier desarrollo estadístico está directamente ligado al diseño del software con el que estima los estadísticos propuestos. De esta forma es habitual que, en primer lugar, se elaboren programas estadísticos sencillos, incluso gratuitos, con los que realizar las operaciones que proponen los nuevos avances. Si la propuesta resulta exitosa, un segundo paso son los grandes paquetes estadísticos los que lo incorporan en las versiones más avanzadas. De esta forma, suelen convivir diferentes software para estimar estos nuevos desarrollos, desde los específicos hasta los más generales.

Exactamente ese proceso es el que ocurre con los Modelos Multinivel (o Modelos Jerárquicos). En un primer momento, se desarrollaron dos programas para estimarlos: desde la llamada Escuela de Chicago se diseñó el programa HLM (Bryk y Raudenbush, 1992), en Europa el ML2 (Rasbash, 1989), que luego generó el MLwiN. Más adelante, paquetes estadísticos más amplios como el SPSS, el SAS o el Stata, por poner unos ejemplos, lo han incorporado en sus nuevas versiones. Pero, ¿qué diferencias hay entre ellos?

En este artículo se presenta una breve revisión sobre los pros y los contras de los principales software que existen en el mercado para desarrollar Modelos Multinivel: SPSS, Stata, Hlm, MLwiN.

## 2.- La importancia de los Modelos Multinivel y del software

La alternativa estadística para suplir las carencias de los Análisis de Regresión para datos anidados<sup>1</sup> (independencia de las observaciones y la selección de la unidad de análisis) surgió hace poco más de 25 años con los llamados Modelos Multinivel (o Modelos Jerárquicos) (Aitkin y Longford, 1986). Como luego veremos, estos modelos, en esencia, elaboran una ecuación para cada unidad de análisis, con lo que se superan las anteriores limitaciones permitiendo un mayor ajuste a la realidad. Esta estrategia metodológica es mucho más que una técnica de análisis, es un planteamiento global de la investigación, dado que exige diseñar la investigación, elaborar instrumentos y recopilar los datos y analizarlos de una forma diferente a la habitual. Aunque nació de la investigación educativa, concretamente de la investigación sobre Eficacia Escolar (Murillo, 2005), en la actualidad es ampliamente utilizado por otras ciencias tales como la Economía, la Sociología, la Epidemiología, la Zoología, la Psicología o la Botánica.

Los Modelos Multinivel consisten, en esencia, en varios modelos de regresión, uno para cada nivel de análisis (Reise y Duan, 2003; Bickel, 2007), de tal forma que se ajustan mejor a los datos. De esta manera, permiten elaborar submodelos para cada nivel y es cada uno de estos submodelos en los que se expresa la relación entre las variables dentro de un determinado nivel y nos permiten especificar cómo las variables de ese nivel influyen en las relaciones que se establecen en otros niveles (Martínez-Garrido y Murillo, 2013; Murillo, 1999, 2008).

Las principales ventajas del análisis multinivel en contraposición con métodos tradicionales son:

- Modela simultáneamente diferentes niveles de variación (por ejemplo, estudiante y aula). Esta característica permite saber qué proporción de la variación del

---

<sup>1</sup> Para profundizar en las limitaciones de los análisis de regresión: para la independencia de las observaciones se recomienda revisar los trabajos de Gelman y Hill (2007), Goldstein (2010), Heck y Thomas (2000), o Hox (2010). En cuanto a los problemas derivados de la selección de la unidad de análisis: Hill y Rowe (1996), Hox (2010) y Goldstein (2010).

rendimiento académico se debe a características propias del estudiante (nivel 1, por ejemplo, edad) y cuál a características del aula (nivel 2, por ejemplo: clima afectivo).

- Estima el efecto principal sobre el rendimiento que tiene cada variable o un conjunto de variables (factor).
- Permite que el nivel de rendimiento (intercepto “a”) y la fuerza de su relación con los factores (pendiente “b”) varíe libremente en los diferentes niveles de agregación (en el presente estudio: estudiante, aula, escuela y país).

## 2.1.- Conceptos fundamentales

Antes del breve análisis en el desarrollo formal de los modelos multinivel, se presta atención a tres conceptos fundamentales y sus implicaciones: correlación intraclase, coeficiente fijo y aleatorio e interacción internivel.

Se entiende por *Correlación Intraclase* la medida del grado de dependencia de los individuos. Se trata de un estadístico fundamental en los Modelos Multinivel que servirá para evaluar el grado de parecido entre unidades de nivel micro que pertenecen a la misma unidad macro; permite estimar el promedio de las correlaciones entre todas las posibles ordenaciones de los pares de observaciones disponibles y, por lo tanto, evitar el problema de la dependencia del orden del coeficiente de correlación. El cálculo de la Correlación Intraclase aporta un coeficiente “CCI- Coeficiente de Correlación Intraclase” (1) que constituye el índice más apropiado para cuantificar la concordancia entre diferentes mediciones de una variable numérica (Hernández Aguado *et al*, 1990; Fleiss, 1986; Prieto, Lamarca y Casado, 1998).

Una de las principales limitaciones del CCI es la dificultad de su cálculo, ya que debe ser estimado de distintas formas dependiendo del diseño del estudio (Bartko, 1966). Suele evaluarse mediante el ANOVA de efectos aleatorios conocido en el Modelo Multinivel como modelo vacío o nulo. Una correlación baja, que tome valores entre 0 y 0,39, significará que los sujetos dentro del mismo grupo son tan diferentes entre sí como los que pertenecen a otros grupos. En ese caso, los grupos no son homogéneos internamente y las observaciones son independientes (requisito necesario dentro de los modelos lineales tradicionales). Podemos decir que la correlación es alta cuando el índice toma valores superiores a 0,7. Si se ignora la presencia de esta *Correlación Intraclase*, los modelos resultantes son innecesarios y falsamente complejos, dado que aparecen relaciones significativas inexistentes.

Suponiendo que calculásemos el CCI de una estructura de dos niveles, la fórmula sería:

$$CCI = \frac{S_{nivel2}^2}{S_{nivel1}^2 + S_{nivel2}^2} \quad (1)$$

Otro concepto fundamental, y que supone la gran aportación de los modelos multinivel, es el de *coeficiente fijo* y *coeficiente aleatorio*. Mientras que los fijos son comunes a todos los sujetos, los aleatorios son variables que se distribuyen según una función de probabilidad. En una estructura multinivel los coeficientes del primer nivel (por ejemplo, estudiantes) son tratados como aleatorios en el segundo nivel (por ejemplo, aulas). Además, en los modelos multinivel se permite a los grupos desviarse de la solución central o global, tanto en el intercepto como en la pendiente. O, lo que es lo mismo, los modelos multinivel están compuestos por dos partes: una general, común a todos los contextos, que es la llamada parte fija; y otra que representa lo específico de

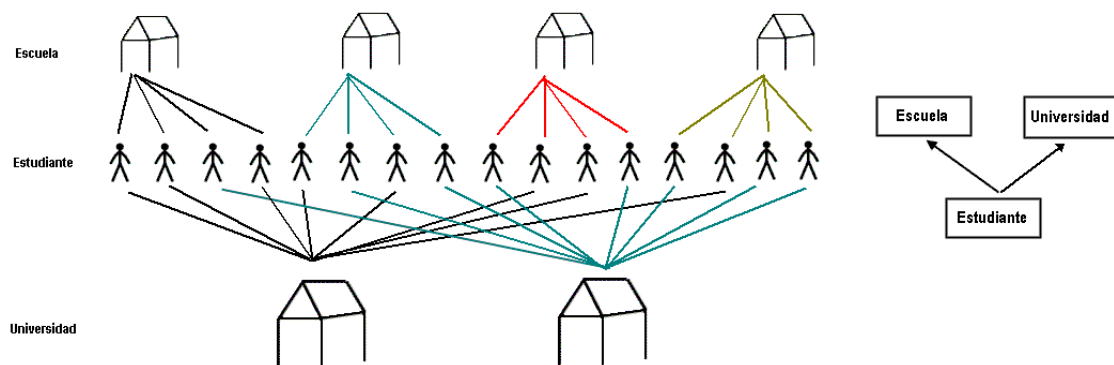
cada contexto, que varía y que se estima a través de la varianza en los distintos niveles jerárquicos.

Una vez consideradas las diferencias entre coeficiente fijo o aleatorio, esta diferencia debe ser especificada en la construcción del Modelo Multinivel, por un lado los coeficientes fijos que conformarían la parte fija del modelo, y los coeficientes aleatorios que dan forma a la parte aleatoria del modelo, siendo:

- Parte fija: incluye los parámetros que permiten determinar una línea promedio para todos los estudiantes de todas las aulas, de todas las escuelas de cada país. Recibe ese nombre dado que supone que la intensidad de la asociación entre las variables independientes y el rendimiento es constante para todos los niveles de la muestra.
- Parte aleatoria: muestra la estimación de la variación de (a), las líneas individuales que representa a los niveles en torno de aquella línea promedio y (b) de los rendimientos observados de los estudiantes alrededor de la línea promedio de su nivel.

Un tercer concepto importante es la *interacción internivel* o la interacción entre variables que están medidas en diferentes niveles de una estructura jerárquica de datos. Hablaremos de efectos fijos cuando contemos con una variable cuyo dominio representa todos los niveles posibles que son de interés teórico, por ejemplo, condiciones de tratamiento, sexo, cultura, etnia, entre otros. Y de efectos aleatorios, cuando la variable representa una muestra de una población más amplia de valores potenciales: individuos, aulas, escuelas.

La interacción internivel puede presentarse como intervenciones basadas en grupos completos, es decir, los sujetos de un aula, aulas de una misma escuela, escuelas de una misma ciudad. O, en el caso de tratarse de datos con estructuras no jerarquizadas, las interacciones se producen de manera cruzada siendo, por ejemplo, los sujetos de una misma escuela acuden a diferentes universidades (figura 1). Para profundizar en este tema (datos con estructuras de clasificaciones cruzadas o de pertenencia múltiple) se recomienda el estudio de Goldstein (2010) y Hox (2010).



**Figura 1.** Modelo de clasificación cruzada. (Fuente: Elaboración propia)

## 2.2. – Notación

Situémonos en un modelo sencillo de dos niveles de análisis, individuo y grupo. La ecuación de regresión para dos variables independientes se expresa matemáticamente de la siguiente forma

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_j \quad (2)$$

Si permitimos que el intercepto pueda tomar diferentes valores en función de un segundo nivel, la ecuación se expresa así:

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \varepsilon_{ij} \\ \beta_{0j} &= \beta_0 + \mu_{0j} \end{aligned} \quad (3)$$

Donde:

$y_{ij}$  es la variable respuesta que tiene un alumno  $i$  en una escuela  $j$   
 $\varepsilon_{ij}$  es el error y se distribuye normalmente con una varianza constante e igual a  $\sigma_{e0}^2$ ,  
 $\beta_{0j}$  es el promedio de  $y$  de la escuela  $j$ -ésima  
 $\beta_0$  representa el “gran promedio” de  $y$  para la población cuando se anulan las variables dummy,  
 $\mu_{0j}$  es el efecto aleatorio asociado a la escuela  $j$ -ésima y se supone que tiene media cero y una varianza  $\sigma_{\mu0}^2$ .

Si, además de hacer variar el intercepto, permitimos que las pendientes sean diferentes para cada escuela, tenemos la siguiente ecuación:

Nivel 1:

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{1ij} + \beta_{2j} x_{2ij} + \varepsilon_{ij} \quad (4)$$

Nivel 2:

$$\beta_{0j} = \beta_0 + \mu_{0j}, \quad \beta_{1j} = \beta_1 + \mu_{1j}, \quad \beta_{2j} = \beta_2 + \mu_{2j}$$

Con

$$\begin{aligned} \begin{bmatrix} \mu_{0j} \\ \mu_{1j} \\ \mu_{2j} \end{bmatrix} &\sim N(0, \Omega_{\mu}) : \Omega_{\mu} = \begin{bmatrix} \sigma_{\mu0}^2 & & \\ \sigma_{\mu10} & \sigma_{\mu1}^2 & \\ \sigma_{\mu20} & \sigma_{\mu21} & \sigma_{\mu2}^2 \end{bmatrix} \\ \begin{bmatrix} \varepsilon_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} \sigma_{e0}^2 \end{bmatrix} \end{aligned} \quad (5)$$

Algunos programas, por ejemplo el HLM, marcan por defecto el intercepto igual a 1 en la ecuación de regresión. Otros, por ejemplo el MLwiN, requieren que la base de datos incluya una variable igual a uno para todos los casos, la cual ha de ser añadida explícitamente a la ecuación de regresión. Resulta, por tanto, una herramienta de gran utilidad al eliminar el intercepto de la ecuación de regresión (Hox, 2010).

### 2.3.- Procedimiento de estimación

Para estimar los Modelos Multinivel es necesario realizar las siguientes fases. Cada una de ellas dará lugar a uno o varios modelos estadísticos:

- Modelo Nulo (Modelo I).
- Modelo con las variables explicativas (Modelo II)
- Modelo Final (modelo IV).

### 1. Modelo Nulo o Modelo Vacío

El primer parámetro a estimar es la variación alrededor de la media global de la variable de producto considerando simultáneamente los dos niveles de análisis. Para llevarse a cabo se necesita estimar una media global de la variable producto, y conocer cómo varían los promedios de cada nivel de agrupación. La media global de la variable de producto y las variaciones “entre-grupos” se calculan a partir de un análisis de regresión sobre una constante que asume el valor=1 para todos los individuos. Por ello, este modelo cuenta con tan sólo dos elementos, una única variable de respuesta y la constante (o intercepto), no incluye variables explicativas en ninguno de ellos. El modelo nulo se establece como línea de base para la estimación de la varianza explicada a partir de la cual se van evaluando las aportaciones de modelos más elaborados. Así mismo, a través del cálculo del CCI para el Modelo Nulo estimamos el grado de dependencia de los individuos en los diferentes niveles de análisis.

La expresión formal del modelo nulo es:

$$Y_{ij} = \beta_{0j} + e_{ij} \quad (6)$$

$$\beta_{0j} = \beta_0 + u_{0j}$$

$$u_{0j} \sim N(0, \sigma_{u0}^2)$$

$$e_{ij} \sim N(0, \sigma_e^2)$$

Los elementos que se estiman son:

$$\beta_{0j} \quad \sigma_{u0}^2 \quad (\text{varianza del nivel 2}) \quad \text{y} \quad \sigma_e^2 \quad (\text{varianza del nivel 1}).$$

La razón de verosimilitud:  $-2\log_e$  (verosimilitud), que servirá para evaluar las aportaciones al modelo, o los criterios AIC, AICC, CAIC o BIC que aparecen al realizar los cálculos en el SPSS.

El test de razón de verosimilitud (o deviance) es un estadístico que sigue una distribución de  $\chi^2$ , por lo que su respectivo valor crítico se obtiene de su correspondiente distribución, con los grados de libertad determinados por la diferencia entre el número de parámetros que van a ser comparados. Tiene la siguiente forma:

$$D_{01} = -2\log_e \left( \frac{\lambda_0}{\lambda_1} \right) = (-2\log_e \lambda_0) - (-2\log_e \lambda_1) \quad (7)$$

Donde:

$\lambda_0$  representa la verosimilitud asociada al modelo que está subyacente en la hipótesis nula.

$\lambda_1$  representa la verosimilitud asociada al modelo de hipótesis alternativa.

### 2. Modelo II

La estimación del Modelo II incluirá las variables explicativas para cada nivel de agrupación. Este modelo se construye sobre el modelo nulo pero incorporándole, tanto en la parte fija como en la aleatoria las variables explicativas. El modelo II, para un

modelo con una única variable explicativa para cada nivel de agrupación, se expresa de la siguiente forma:

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_1 x_{1ij} + \beta_2 x_{2j} + e_{ij} \\ \beta_{0j} &= \beta_0 + u_{0j} \end{aligned} \quad (8)$$

Donde:

$\beta_1, \beta_2$  son las pendientes promedio de la regresión.

El proceso de estimación se desarrolla de manera diferente en función el software que se utilice, como veremos, el más utilizado es el de mínimos cuadrados iterativos generalizados (GLS).

El proceso de análisis de los resultados del Modelo II consiste en hacer un contraste de hipótesis (9) para comprobar si  $\beta_n$  es estadísticamente significativa. Para ello planteamos las correspondientes hipótesis nula y alternativa:

$$\begin{aligned} H_0: \beta_n &= 0 \\ H_1: \beta_n &> 0 \end{aligned} \quad (9)$$

Según los presupuestos de los modelos multinivel el estimador  $\beta_n$  se distribuye de acuerdo con la Normal. La prueba estadística para contrastar la hipótesis será la *t* de Student. La prueba *t* se define como:

$$t = \frac{\hat{\beta}_n - \beta_n}{se(\hat{\beta}_n)} \quad (10)$$

En este caso, como la hipótesis nula es que  $\beta_n = 0$ , el estimador queda de la siguiente forma tras reemplazar el valor de la hipótesis:

$$t = \frac{\hat{\beta}_n - 0}{se(\hat{\beta}_n)} = \frac{\hat{\beta}_n}{se(\hat{\beta}_n)} \quad (11)$$

El valor de *t* obtenido se contrasta con la tabla de distribución de *t* de Student. Dado que en este trabajo se ha escogido un nivel de significación ( $\alpha$ ) de 0,05, el valor crítico de *t* es de 1,96. Aceptaremos la hipótesis nula para cada variable explicativa y será incorporada al modelo sólo si  $|t| < 1,96$ .

Una vez decidida la inclusión de las variables a la parte fija del modelo, es necesario discutir la inclusión de las mismas en la parte aleatoria, valorar su aportación y decidir si finalmente se deben incluir. El modelo II ajustado en la parte aleatoria se expresa de la siguiente forma:

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_{1j} x_{1ij} + \beta_{2j} x_{2j} + e_{ij} \\ \beta_{0j} &= \beta_0 + u_{0j} \end{aligned} \quad (12)$$

Y de nuevo, los contrastes de hipótesis servirán para verificar la inclusión de cada variable a la parte aleatoria del modelo.

### 3. Modelo Final

La construcción del modelo final incluirá tan sólo aquellas variables que han resultado significativas en la construcción del modelo verificando si se cumplen los supuestos de Homocedasticidad, Normalidad y Ortogonalidad propios de la regresión. Se recomienda consultar Berry (1993) para profundizar en los supuestos del análisis de regresión.

#### **3.- Análisis de los programas: descripción, ventajas, usos habituales, limitaciones**

Comencemos señalando que es completamente natural que el investigador dude a la hora de seleccionar uno u otro programa para llevar a cabo las estimaciones de los Modelos Multinivel. Mucho más natural resulta que finalmente utilice aquel que usa su colega tan sólo por que éste le apunta someras indicaciones de uso. Sabiendo que ésta no debería ser la situación más habitual, y que la elección de uno u otro programa pueden acarrear limitaciones en posteriores análisis, procedemos a explicar las ventajas e inconvenientes de cada uno.

Por un lado, los programas específicos como el MLwiN (Rabash et al., 2012) y HLM (Raudenbush, Bryk, Cheong y Congdon, 2000) aportan como gran ventaja que facilitan el ajuste de todo tipo de modelos multinivel por complejos que éstos sean y minimizan los posibles problemas de estimación. Sin embargo, acarrear limitaciones importantes en relación con la gestión de los datos o su presentación gráfica obligando al nuevo usuario a familiarizarse con el nuevo formato.

Por su parte, los programas generales (SPSS y Stata) tienen la ventaja de ser conocidos por la gran mayoría de los usuarios, pero cuentan con claras limitaciones en torno al ajuste de los Modelos Multinivel. En ambos casos el proceso de modelado para el usuario bien simula los caminos del laberinto, dado que es muy escasa la documentación disponible. A este respecto pueden consultarse varios trabajos que ayudarán en la interpretación de los Modelos Multinivel en SPSS, se trata del estudio de Norusis (2005) y el realizado por Leyland (2004). Si bien es cierto que los autores aunque no prestan demasiada atención a los elementos formales de los modelos, ni a la descripción de sus elementos, sí ofrecen una correcta y asequible interpretación, así como la sintaxis necesaria para ajustar los modelos. Para el caso del Stata, su propio blog ofrece post variados sobre la construcción, interpretación y construcción de gráficos de resultados de los modelos multinivel, con la particularidad de que permiten hacer comentarios con las dudas o sugerencias que surjan.

A continuación, se profundiza en las características de cada software y se aporta el desarrollo del modelo multinivel con cada uno. Para la estimación se ha utilizado una base de datos de carácter educativo que contiene informaciones reales de 4.248 estudiantes de educación primaria de 137 escuelas de América Latina que fueron recogidos por la UNESCO en su estudio SERCE (Segundo Estudio Regional Comparativo y Explicativo) de 2008. La base de datos no cuenta con datos perdidos, han sido previamente trabajados a través del programa SPSS. La variable de producto es:

- Rendimiento en Matemáticas: variable continua de media 500 y desviación típica de 100.

La base de datos incluye variables propias de cada nivel de análisis: los estudiantes (nivel 1) y los centros educativos (nivel 2). En concreto las variables son:



Nivel 1:

- Género, variable dummy (0 hombre, 1 mujer).
- Nivel socioeconómico de la familia, variable continua tipificada.
- Preescolarización, variable ordinal que alcanza valores de 0 a 6 en función de los años que haya estado preescolarizado el estudiante.

Nivel 2:

- Hábitat, variable dummy (0 urbano, 1 rural).
- Nivel socioeconómico de la escuela, variable continua tipificada.

El efecto que provocan estas variables de manera individual sobre la variable de producto se señala en la tabla 1.

	Rendimiento en Matemáticas
Nivel cultural familia	18,693(1,54)
NSE familia	3,355(1,49)
Género (varón-mujer)	-5,301(2,63)
Preescolarización	5,537(1,03)
NSE escuela	17,515(3,32)
Hábitat (urbano/rural)	-30,071(11,99)

**Tabla 1.** Efecto de cada variable individual sobre Rendimiento en Matemáticas

### 3.1.- MLwiN

El software MLwiN fue desarrollado esencialmente por John Rabash, a partir de su predecesor el MLn, y este del ML3. El ML3 fue creado por el *Multilevel Project*, liderado por el Institute of Education de la Universidad de Londres, proyecto ubicado desde 2005 en la Universidad de Bristol (<http://www.bristol.ac.uk/cmm/>). Frente a los software anteriores, ML3 y MLn, el MLwiN proporciona una interfaz gráfica al usuario que facilita la especificación y construcción de Modelos Multinivel, además de permitir la exploración y manipulación de los datos (tablas, gráficos, ecuaciones...).

El MLwiN ha sido diseñado sobre las aplicaciones Windows por lo que comparte características con los procesadores de texto o los paquetes estadísticos. Es a través de los estándares y convenciones de Windows que MLwiN ofrece a los usuarios la herramienta de “Ayuda” con los principales temas y contenidos propios de cada tema localizados en ventanas individuales. Además contiene un visor de variables, de datos y de ecuaciones.

La principal ventaja que ofrece utilizar el MLwiN a los usuarios es haber sido diseñado específicamente para la elaboración de Modelos Multinivel desde un centro de estudio especializado. Ello permite el ajuste perfecto a los datos, ofrece la posibilidad de albergar hasta 10.000 variables y permite la construcción de modelos multinivel de hasta cinco niveles de análisis, característica única de entre los programas que se presentan en este trabajo. Además, desde la propia página web del Centro de Modelado Multinivel el usuario puede descargarse temario sobre los Modelos Multinivel, revisar las FAQs, y estar enterado de próximos cursos y seminarios para profundizar en el manejo del software.

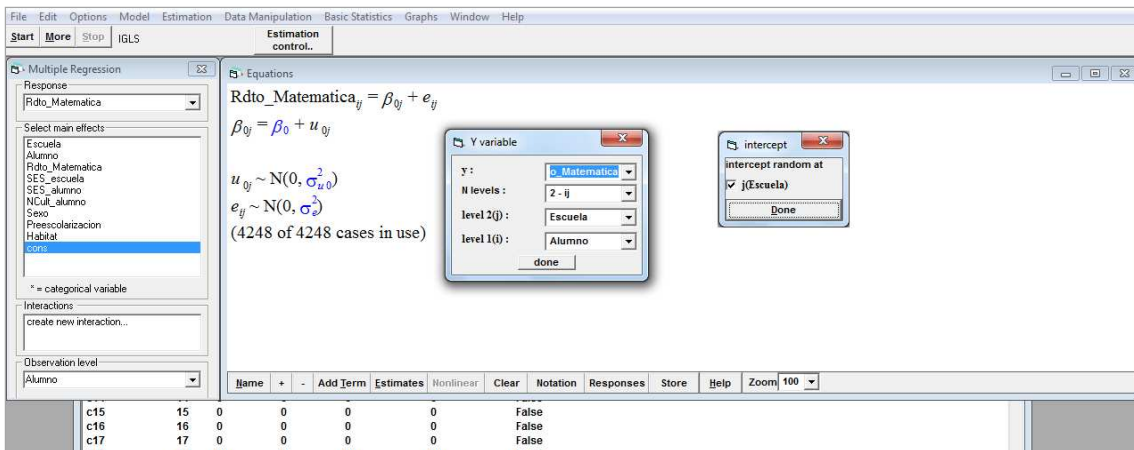
El diseño específico del MLwiN para Modelos Multinivel permite realizar Modelos con variables continuas, Modelos Logísticos con variables binomiales y Modelos Multinomiales Logísticos con variables categóricas. Así mismo, permite el tratamiento de los residuos de todos los niveles de los análisis y la modelización de datos con clasificación cruzada.

La versión más reciente del programa MLwiN es la 2.28 (disponible desde junio de 2013). Y si bien para actualizar el software basta con acudir a la página web y hacer un simple clic, es necesario reconocer que las actualizaciones van siendo cada vez menos frecuentes, seis en el año 2010 y tan sólo dos en el año 2012. La manera en la que se adquiriera el software MLwiN puede ser considerada una ventaja o una desventaja dependiendo de dónde se resida. Para todos los residentes en UK el programa es gratuito, los demás tienen que pagar por el software<sup>2</sup>. Lo mejor de todo es que, una vez adquirida una licencia no hay que desembolsar dinero en las actualizaciones. Además, existe una versión de licencia para estudiantes de doctorado que da servicio del programa completo durante tres años de manera totalmente gratuita.

La herramienta más interesante del programa para el desarrollo de Modelos Multinivel es su ventana de ecuaciones. Se trata de un editor de ecuaciones especializado a través del cual el investigador puede definir, en forma de ecuación, el modelo haciendo "clic" en las variables y alternativas. Sin embargo, este editor de ecuaciones puede no ser tan útil si el usuario no es capaz de traducir su investigación en fórmula. Si los usuarios manejan estos conceptos, la herramienta resulta especialmente útil; en caso contrario, quizá es recomendable que usen otro programa para familiarizarse con las ecuaciones propias de un modelo multinivel antes de utilizar el MLwiN.

### *Estimación de Modelos Multinivel con MLwiN*

Una vez asignada la matriz de datos al MLwiN, el primer paso es identificar la variable dependiente (Rendimiento en Matemáticas) y asignar el número de niveles con el que se va a trabajar. Además, consideramos el intercepto aleatorio sobre los cuatro niveles de análisis.



**Figura 2.** Modelo Nulo MLwiN 2.26

<sup>2</sup> El precio por una copia única desde la web de MLwiN es de 900\$ con un 40% de descuento a los estudiantes.

Desde este momento ya es posible calcular el Modelo Nulo (tabla 1). De esta forma, el Rendimiento en Matemáticas ( $y$ ) se interpreta como el resultado de combinar el rendimiento medio de la escuela a la que pertenece ( $B_{0j}$ ) y los residuos o la variación aleatoria ( $e_{ij}$ ) en torno a esa media. Se asume que los errores o residuos se distribuyen normalmente con media cero y con igual varianza  $\sigma^2$  en todos los centros. Dicho de otra forma, se trata del tradicional modelo de ANOVA de un factor de efectos aleatorios.

La incorporación de variables explicativas al modelo y su aleatorización por los diferentes niveles permitirá la creación del modelo final (tabla 2) que posibilitará el cálculo de los efectos escolares, descontando el efecto que las variables explicativas generaban sobre la variable dependiente (figura 3).

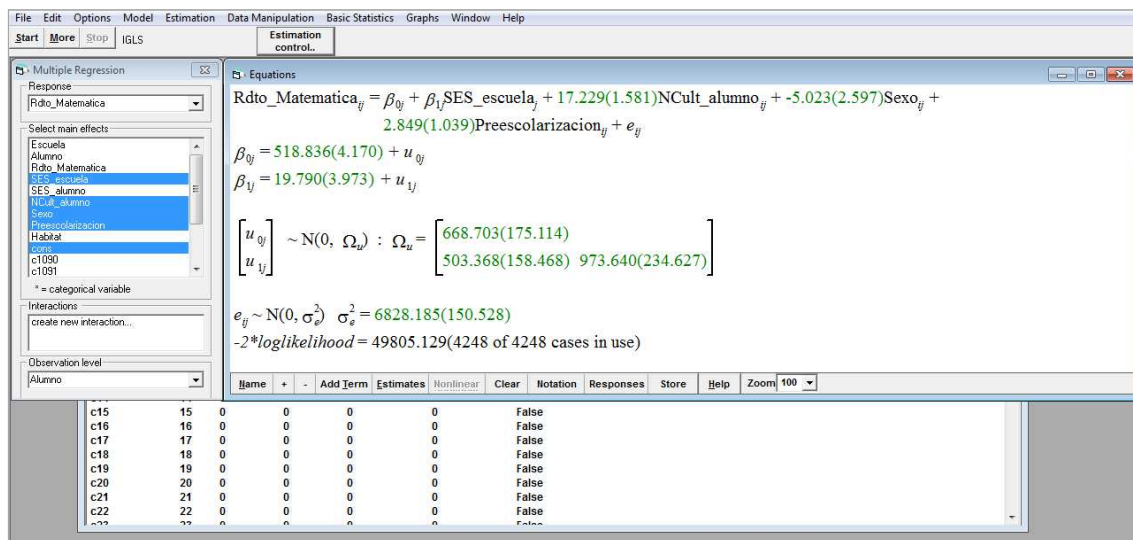


Figura 3. Modelo Final con MLwiN 2.26

MLwiN utiliza el algoritmo IGLS (mínimos cuadrados generalizados) o RIGLS (mínimos cuadrados generalizados restringidos) descritos por primera vez por Goldstein (1986) y Goldstein (1989), respectivamente. Los algoritmos consideran los niveles del modelo, los coeficientes fijos de la regresión y la matriz de componentes de varianza/covarianza. El modo en el que el software MLwiN ejecuta el algoritmo es el siguiente: el algoritmo fija los componentes de la varianza en algún valor inicial y maximiza la verosimilitud sobre los coeficientes fijos (este es justo el problema de la técnica Mínimos Cuadrados Generalizados). Entonces, fijan los coeficientes por sus actuales valores y maximizan la verosimilitud sobre los componentes de la varianza. Las dos optimaciones se alternan hasta la convergencia (puede encontrarse una descripción más detallada en Goldstein y Rasbash, 1992).

	Rendimiento en Matemáticas	
	Modelo Nulo MLwiN	Modelo Final MLwiN
Parte constante		
Intercepto	519,892(5,40)	518,836(4,17)
NSE escuela		19,790(3,97)
Nivel cultural familia		17,229(1,58)
NSE familia		NS
Género (varón-mujer)		-5,023(2,59)
Preescolarización		2,849(1,03)
Hábitat (urbano/urbano)		NS
Parte aleatoria		
Entre escuelas	3477,460(480,46)	668,703(175,11)
Entre estudiantes	7040,720(155,22)	6828,185(150,52)
-2*LogVerosimilitud	50023,308	49805,129

**Tabla 2.** Resultados del Modelo Nulo y Modelo Final con MLwiN

### 3.2.- HLM

El software HLM fue creado por Anthony Bryk en 1992, se trata de un programa diseñado específicamente para el desarrollo de Modelos Multinivel este caso en Estados Unidos. Fue a través del texto “*Hierarchical Linear Models for Social and Behavioral Research: Applications and Data Analysis Methods*” (Bryk y Raudenbush, 1992) que se ha convertido en uno de los software más utilizados. HLM se distribuye a través de SSI-Scientific Software International- (www.ssicentral.com).

El HLM ofrece al usuario una amplia gama de opciones de estimación, la imputación de ficheros desde diferentes software (SPSS, Stata...), diferentes pruebas de hipótesis de razón de verosimilitud, la creación de gráficos, y la capacidad de manejar fácilmente modelos lineales jerárquicos a través de su visor de operaciones en las que ofrece las ecuaciones de cada nivel, o su visor de modelo mixto donde se integran las ecuaciones de cada nivel en una única ecuación.

Las últimas versiones de HLM marcan claramente la diferencia frente a los paquetes estadísticos generales siendo mucho más intuitivas y cuidando el diseño la interfaz de trabajo. El HLM 7, la versión más reciente del programa puesta a la venta en marzo de 2013, incluye novedades como el cálculo de modelos de hasta cuatro niveles de análisis (anteriores versiones sólo dejaban calcular 2 y 3 niveles), o la imputación de los datos desde un único fichero de datos (versiones anteriores requerían crear tantos ficheros de datos como niveles tuviera su estudio).

Para hacerse con el HLM el usuario tan sólo tendrá que acceder a la página de la distribuidora y solicitarlo vía e-mail. La buena noticia es que para todos aquellos que ya dispongan de una versión previa del software tendrán que pagar un precio considerablemente menor por hacerse con la última versión del mismo. La página ofrece precios diferentes para los distintos tipos de licencia<sup>3</sup> e incluso ofrece una versión para el estudiante de forma totalmente gratuita<sup>4</sup>. Sin embargo, la versión para el estudiante cuenta con importantes restricciones: no incluye la herramienta Stat/Transfer para la importación de datos y limita el número de observaciones en la creación de los modelos. Por ejemplo, para un modelo de tres niveles, el número máximo de observaciones que se puede utilizar en los niveles 1, 2 y 3 es de aproximadamente 8.000, 1.700 y 60, respectivamente. Para un modelo de dos niveles el número máximo

<sup>3</sup> El precio es de 430\$ desde la web de HLM.

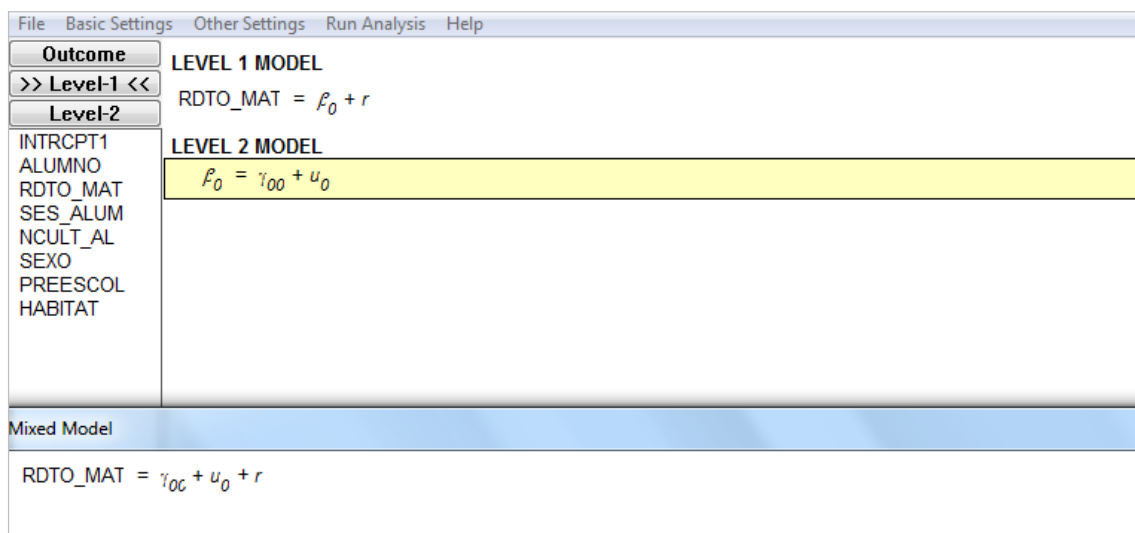
<sup>4</sup> <http://www.ssicentral.com/hlm/student.html>.

de observaciones en los dos niveles es 8.000 en el nivel-1 y de 350 en el nivel 2. Además, no podrán ser incluidos más de 5 efectos en las ecuaciones de cualquier modelo, y el total de los efectos no podrán ser más de 25.

*Estimación de Modelos Multinivel con HLM*

Para calcular el Modelo Multinivel con HLM 7, una vez seleccionado el tipo de modelo que vamos a realizar y los niveles que éste tendrá, se debe cargar el fichero de datos. Una vez cargados los datos e identificada la pertenencia de cada variable a cada nivel del modelo que se está construyendo se identifica la variable dependiente (figura 4) y se obtendría el Modelo Nulo.

Si al cargar los datos se percibe que los nombres de las variables del estudio se han convertido en ilegibles, ello es debido a que el HLM no admite más de ocho caracteres alfanuméricos en el nombre (no así en la etiqueta). Por ello, recomendamos considerar esta limitación del HLM al utilizar programas auxiliares, como SPSS o Stata.



**Figura 4.** Modelo Nulo HLM

Por defecto, el HLM de 2 niveles usa el algoritmo REML, la estimación de máxima verosimilitud restringida, mientras que para los modelos de 3 niveles el HLM usa FIML, información de máxima verosimilitud completa. Versiones más antiguas de HLM se basaban en el algoritmo EM (de máxima expectación) que es un método iterativo para la búsqueda de máxima verosimilitud o máxima *a posteriori* (MAP). Puede ampliarse información a este respecto consultando el trabajo elaborado por Leeuw (2005).

Para construir el Modelo Final (figura 5) tan sólo se deben ir incluyendo las variables predictoras en el modelo, considerando que sólo las variables que aparecen en cada nivel son aquellas que previamente el investigador ha señalado como propia para ese nivel en el momento en el que se cargaban los datos al programa. Los resultados del modelo nulo y final con HLM se presentan en la tabla 3.

File Basic Settings Other Settings Run Analysis Help	
Outcome	LEVEL 1 MODEL
>> Level-1 <<	$RDTO\_MAT = \beta_0 + \beta_1(NCULT\_AL) + \beta_2(SEXO) + \beta_3(PREESCOL) + r$
Level-2	LEVEL 2 MODEL
INTRCPT1	$\beta_0 = \gamma_{00} + \gamma_{01}(NSE\_ESC) + u_0$
ALUMNO	$\beta_1 = \gamma_{10} + u_1$
RDTO_MAT	$\beta_2 = \gamma_{20} + u_2$
SES_ALUM	$\beta_3 = \gamma_{30} + u_3$
NCULT_AL	
SEXO	
PREESCOL	
HABITAT	
Mixed Model	
$RDTO\_MAT = \gamma_{0c} + \gamma_{01}*NSE\_ESC + \gamma_{1c}*NCULT\_AL + \gamma_{2c}*SEXO + \gamma_{3c}*PREESCOL + u_0 + r$	

Figura 5. Modelo Final HLM

	Rendimiento en Matemáticas	
	Modelo Nulo HLM	Modelo Final HLM
Parte constante		
Intercepto	519,879(5,40)	524,734(5,36)
NSE escuela		19,830(9,97)
Nivel cultural familia		17,051(1,40)
NSE familia		NS
Género (varón/mujer)		-4,903(2,91)
Preescolarización		3,117(1,20)
Hábitat (urbano/rural)		NS
Parte aleatoria		
Entre escuelas	3508,032(59,22)	668,703(175,11)
Entre estudiantes	7040,633(83,90)	6845,964(82,74)
-2*LogVerosimilitud	50016,245	49836,589

Tabla 3. Resultados HLM

### 3.3.- SPSS

El software SPSS es un paquete estadístico especialmente diseñado para su uso en estudio de Ciencias Sociales. Fue creado en 1968 por Norman H. Nie, C. Hadlai Hull y Dale H. Bent. Hasta el año 1975 el *National Opinion Research Center* de la Universidad de Chicago se encargó del desarrollo, distribución y venta del programa. A partir del año 1975 estas actividades corrieron a cargo de SPSS Inc hasta que, en el mes de Junio de 2009 la empresa IBM anunció su compra adquiriendo los derechos sobre el software.

Si bien originalmente el SPSS se diseñó para utilizarse en grandes computadores, desde el año 1984 se comercializa para ordenadores personales. Las actualizaciones que regularmente se ofrecen del programa (actualmente versión 22) incluyen novedades siempre con el objetivo de hacer los análisis más intuitivos y ofrecer la gestión de los datos con mayor comodidad. A través de los diferentes comandos de los menús del SPSS (Archivo, Edición, Ver, Datos, Analizar...) y sus diferentes cuadros de diálogo se accede a toda una serie de estadísticos para aplicar sobre nuestros datos con tan sólo hacer clic (figura 6).

Al igual que el resto de los programas, el SPSS integra tutoriales en el menú “ayuda” que facilitarán la elección del estadístico apropiado y la interpretación de los resultados. Su interfaz se compone de tres "editores": Editor de Datos, con dos vistas: variables y datos donde se muestra el contenido del archivo de datos cargado. El Visor o Editor de Resultados, donde se muestran los valores de datos reales o las etiquetas de valor definidas para cada una de nuestras variables. Y la Sintaxis que permitirá programar cálculos más avanzados que los ofrecidos en las diferentes pestañas del software. Su desarrollo, ligado a los estudios de las Ciencias Sociales, han provocado que sea el programa más demandado por los usuarios.

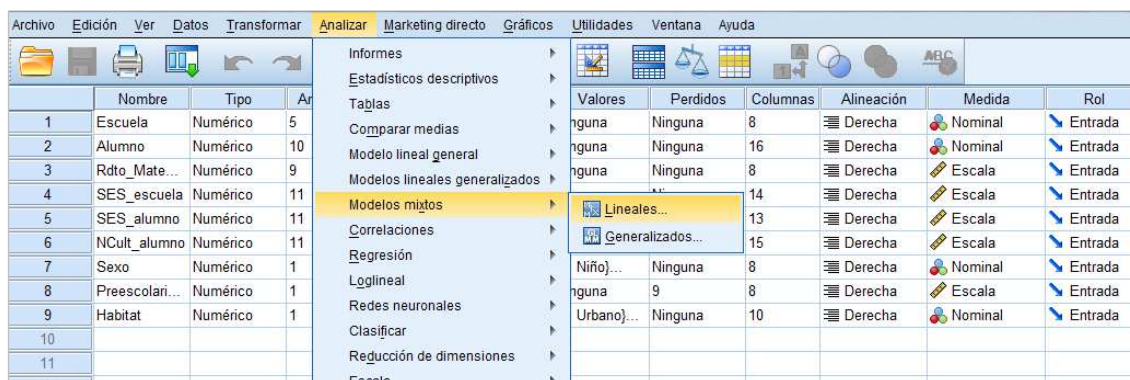


Figura 6. Interfaz SPSS

La principal ventaja del uso del SPSS frente al resto de paquetes estadísticos es que al tratarse de un paquete estadístico permite el manejo y transformación de la base de datos, elemento básico a la hora de construir las variables que formarán parte de los Modelos Multinivel que calculemos. Sin embargo, como limitación principal del software es necesario destacar que, al igual que le sucede al resto de paquetes estadísticos más generales, su uso para el desarrollo de Modelos Multinivel se ve sesgado a la luz de los resultados que puede ofrecernos. Efectivamente, el SPSS permitirá construir modelos de cuantos niveles necesitemos, pero no permitirá, por ejemplo, trabajar con los residuos de cada nivel por lo que el usuario no podrá estudiar la calidad de su modelo más allá de las diferentes pruebas sobre el estudio de la razón de verosimilitud.

Para adquirir el programa tan sólo hay que registrarse en [www.ibm.com](http://www.ibm.com) y decidir el módulo que más interesa en función de las necesidades del usuario. Cada uno de los módulos: Amos, Analytic Server, Data Collection... se adquiere con licencia independiente lo cual puede suponer un desembolso importante si se quiere disponer del paquete completo. Afortunadamente, desde la web de IBM se permite la descarga de módulos de prueba gratuitos durante 14 días que permitirán comprobar si el módulo escogido responde a las necesidades de la investigación que se esté realizando.

### *Estimación de Modelos Multinivel con SPSS*

Como se señala, estimar Modelos Multinivel desde SPSS es realmente sencillo, dado que no es necesario cambiar de software para crear, transformar o ajustar cualquiera de las variables del estudio. Su organización en forma de cuadro de diálogo facilita la puesta en marcha del proceso de modelado de una manera amable para los usuarios novatos. Puede incluso decirse que el procedimiento es “tan amigable” que ésta es la característica considerada por muchos como una limitación puesto que no permite al usuario ver la fórmula concreta que se está llevando a cabo para la estimación de los Modelos Multinivel. Como siempre, calcular el Modelo Nulo (figura 7) requerirá de

señalar la variable dependiente, la cantidad de niveles y en el apartado de Estimadores solicitar que devuelva las estimaciones de los parámetros, los contrastes sobre parámetros de covarianza y las covarianzas de los efectos aleatorios.

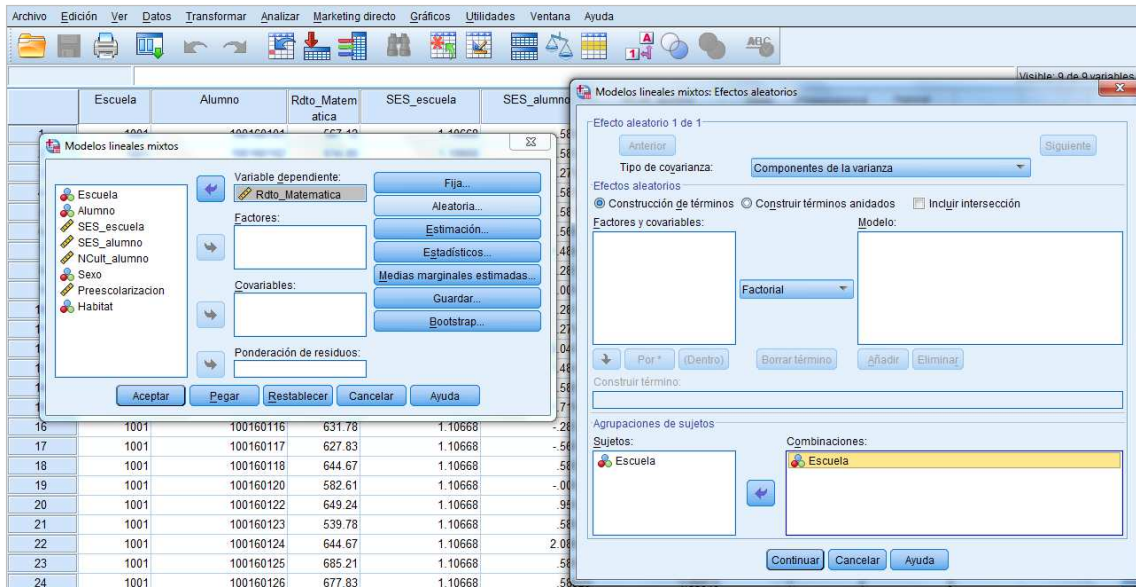


Figura 7. Cálculo del Modelo Nulo en SPSS

Los resultados aportados por el SPSS, más allá de los resultados de los coeficientes del Modelo Nulo ofrece varios estadísticos de ajuste global (13) que indican en qué medida el modelo propuesto es capaz de representar la variabilidad observada en los datos (el ajuste del modelo a los datos es tanto mejor cuanto mejor es el valor de estos estadísticos). El primero de estos estadísticos es la desviación (-2LL; véase McCullag y Nelder, 1989:33-36). El resto son modificaciones de -2LL que penalizan (aumentando) su valor mediante alguna función del número de parámetros. El segundo estadístico es AIC es el criterio de información de Akaike (Akaike, 1974); el tercero AICC es el criterio de información de Akaike corregido (Hurvich y Tsai, 1989); el cuarto CAIC es el criterio de información de Akaike consistente (Bozdogan, 1987); y el quinto BIC es el criterio de información bayesiano (Schwarz, 1978).

$$\begin{aligned}
 AIC &= -2LL + 2d \\
 AICC &= -2LL + \frac{2dn}{n-d-1} \\
 CAIC &= -2LL + d[\log(n) + 1] \\
 BIC &= -2LL + d \log(n)
 \end{aligned}
 \tag{13}$$

Donde  $LL$  se refiere al logaritmo de la verosimilitud si se utiliza el método de estimación de Máxima Verosimilitud (MV) y al logaritmo de la verosimilitud restringida si se utiliza el método (MVR). Si se utiliza MV,  $d$  es el número de parámetros asociados a los efectos fijos más el número de parámetros asociados a los efectos aleatorios, y  $n$  es el número total de casos. Si se utiliza MVR,  $d$  se trata del número total de parámetros asociados a los efectos aleatorios y  $n$  es el número total de casos menos el número de parámetros asociados a los efectos fijos.

Si bien estos estadísticos de ajuste global no tienen una interpretación directa, son muy útiles para comparar modelos alternativos siempre que uno de ellos incluya



todos los términos del otro. La diferencia entre los estadísticos  $-2LL$  correspondientes a dos modelos distintos se distribuye según chi-cuadrado con grados de libertad igual al número de parámetros en que difieren los dos modelos comparados, por tanto, la diferencia entre los estadísticos  $-2LL$  correspondientes a dos modelos distintos puede utilizarse para valorar la ganancia que se obtiene en el ajuste cuando se añaden los efectos en que difieren ambos modelos.

La figura 8 muestra cómo se incluyen los efectos principales de las variables de ajuste a la parte fija del Modelo Final en SPSS. Además, en la parte aleatoria habrá que situar el efecto aleatorio de la variable Nivel Socioeconómico de la Escuela (SES\_escuela) en el nivel escuela. Los resultados del Modelo Nulo y Final aparecen en la tabla 4.



Figura 8. Cálculo del Modelo Final en SPSS

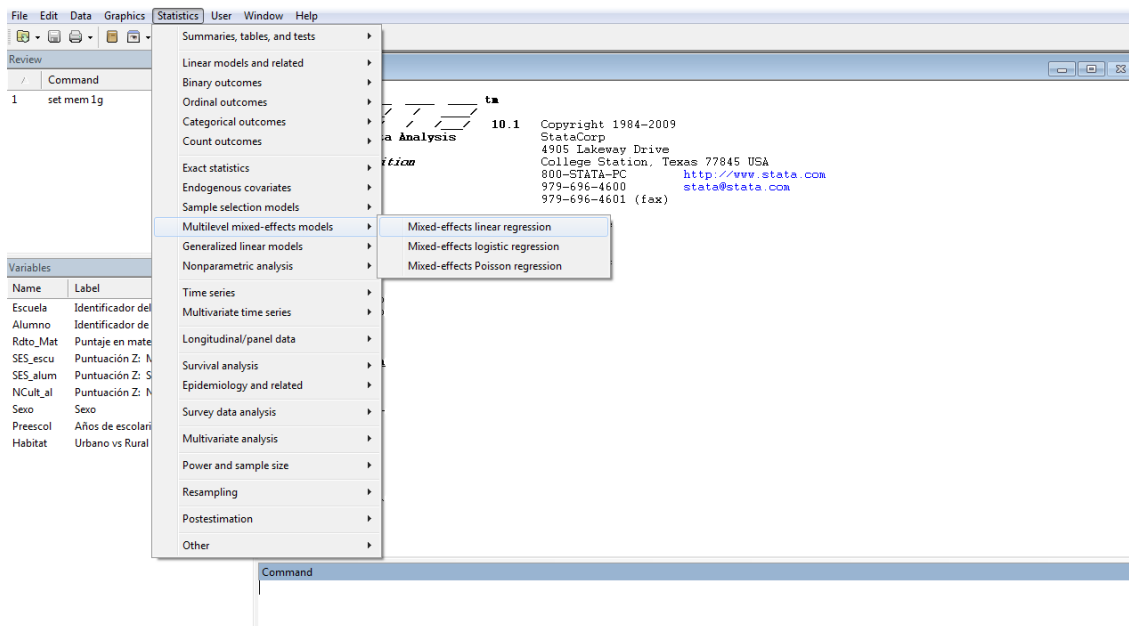
	Rendimiento en Matemáticas	
	Modelo Nulo Spss	Modelo Final Spss
<b>Parte constante</b>		
Intercepto	519,879(5,42)	518,829(4,17)
NSE escuela		19,787(3,97)
Nivel cultural familia		17,229(1,58)
NSE familia		NS
Género (varón-mujer)		-5,022(2,59)
Preescolarización		2,850(1,03)
Hábitat (rural/urbano)		NS
<b>Parte aleatoria</b>		
Entre escuelas	3508,038(505,16)	668,703(175,11)
Entre estudiantes	7040,633(155,46)	6828,185(150,52)
AIC	∞	∞
BIC	∞	∞

Tabla 4. Resultados SPSS

### 3.4.- Stata

El paquete estadístico Stata, por su parte, fue creado por StataCorp en 1985, como un programa genérico que permite a los usuarios el manejo, el análisis y la creación de gráficos con los datos. Tras dos años de espera, el pasado junio de 2013 salió al mercado la versión 13 del software.

El diseño de Stata, también bajo Windows (figura 9), permite acceder a los cálculos a través de sus diferentes pestañas y añade a su interfaz un visor de comandos donde el usuario puede ejecutar directamente los comandos que más le interesen de una manera manual. Los comandos a utilizar abordan el manejo básico de los datos, los análisis descriptivos elementales, tests de hipótesis.... Incluso el software incluye comandos específicos para la investigación biomédica. En el caso de trabajar con grandes bases de datos, el comando `-set mem-` de Stata permitirá incrementar la memoria de trabajo.



**Figura9.** Interfaz Modelos Multinivel Stata 10.

Una de las principales características de Stata es que permite guardar los resultados en un archivo de texto (log file) para poder ser editados en cualquier momento. Y además, la creación de un archivo do file que guarda todos los comandos que se hayan llevado a cabo hasta el momento. La recuperación de este archivo \*.do permite ejecutar exactamente los mismos comandos guardados, o una parte de ellos para reanudar el trabajo por donde se prefiera.

Desde la página web de Stata ([www.Stata.com](http://www.Stata.com)) el usuario tiene acceso a gran cantidad de información sobre cómo utilizar Stata en diferentes sistemas operativos, visitar las FAQs, consultar el calendario de cursos y seminarios e incluso información sobre otros proveedores de software estadístico, libros, revistas especializadas en Estadística, Organizaciones de Estadística,.... Además el usuario puede consultar el listado de precios del software que varía en función del país, la duración de la licencia y el sector en el que se vaya a usar. Hacerse con una nueva versión del software requiere la adquisición de una nueva licencia.

Estimar Modelos Multinivel en Stata (incorporados a partir de su versión 10) tiene como ventaja el permitir el tratamiento de los datos sin necesidad de utilizar a otro software. Pero también sus inconvenientes, al igual que en el SPSS, no se permite guardar los archivos de errores para todos los niveles con los que se esté trabajando. Además, aportar un visor con el modelo específico que se está creando puede ser considerado como inconveniente aunque el perfecto diseño y organización de las pestañas y el visor de resultados minimizan este aspecto.

### Estimación de Modelos Multinivel con Stata

Como en los casos anteriores, una vez aumentada la memoria de trabajo del programa y cargados los datos, el usuario debe seleccionar la variable dependiente, y los niveles del estudio. Desde esta misma pantalla (figura 10) el usuario podrá elegir el método de estimación que desee, el nivel de fiabilidad y cómo presentar los resultados del modelo. Este análisis dará lugar al Modelo Nulo. La tabla 5 muestra el resultado del Modelo Nulo y Final con Stata. La estimación de los componentes de la varianza aparece por defecto a través del método de Máxima Verosimilitud Restringida, pudiendo ser cambiado por el investigador a la hora de hacer los cálculos a MV.

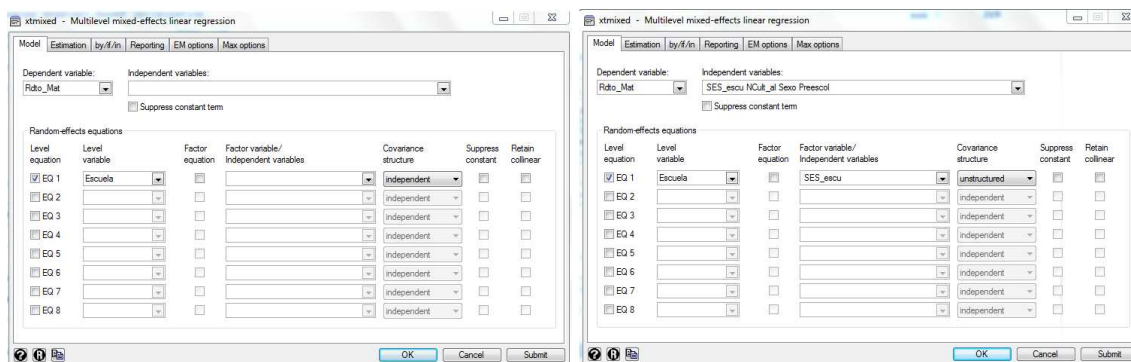


Figura 10. Construcción del Modelo Nulo y Final con Stata

	Rendimiento en Matemáticas	
	Modelo Nulo Stata	Modelo Final Stata
<b>Parte constante</b>		
Intercepto	519,892(5,40)	518,852(4,17)
NSE escuela		19,790(3,97)
Nivel cultural familia		17,230(1,58)
NSE familia		NS
Género (varón-mujer)		-5,022(2,59)
Preescolarización		2,849(1,03)
Hábitat (rural/urbano)		NS
<b>Parte aleatoria</b>		
Entre escuelas	3476,044(499,29)	668,689(168,88)
Entre estudiantes	7040,78(155,47)	6828,129(150,60)
LogVerosimilitud	-25011,649	-24902,558

Tabla 5. Resultados del Modelo Nulo y Modelo Final con Stata

### 4.- Comparación

Como hemos visto, la utilización de uno u otro software ofrece aparentemente resultados idénticos al llevar a cabo una investigación. Sin embargo, merece la pena destacar cuáles son las principales comparaciones que uno u otro programa ofrece al usuario a fin de decantarse por el uso de uno u otro para el desarrollo de Modelos Multinivel.

La clara diferencia entre los distintos programas presentados es que unos responden a un uso más general frente a un uso específico para el desarrollo de Modelos Multinivel, por tanto, en la medida que el usuario vaya a dedicarse a elaborar investigación utilizando Modelos Multinivel con mayor o menor asiduidad este será un

elemento fundamental. Sin bien HLM y MLwiN ofrecen un tratamiento específico y muchísimas más posibilidades para construir modelos tan complejos como se quiera, SPSS y Stata son mucho más comunes entre los ordenadores personales aunque cuentan con restricciones a la hora de crear modelos más complejos o profundizar en los resultados.

Muy relacionado con ser programas generales o específicos está el hecho de elaborar nuestra investigación con Modelos Multinivel con uno o al menos dos programas. Los programas específicos siempre van a necesitar el apoyo de un genérico para convertir las variables, tipificarlas, crear nuevas, o centrarlas. No se está diciendo que MLwiN o HLM no permita hacer ciertos cálculos de las variables, sino que los posibles ni son tantos, ni se desarrollan de una manera rápida y sencilla para el usuario. Por ejemplo, estandarizar una variable en Spss requiere de un par de clics, en HLM es necesario calcular una nueva variable que sea el resultado de la fórmula de coeficiente estandarizado (14)

$$Coef_{Estand} = \frac{Coef_{No:Estand} \times DesTip_{Var\_explicativa}}{DesTip_{Var\_dependiente}} \quad (14)$$

Efectivamente, el hecho de contar con limitadas operaciones en los software específicos limita radicalmente el tratamiento de los valores perdidos. Mientras por ejemplo el Spss da la opción de imputarlos a través de diferentes métodos (media, mediana, interpolación lineal, tendencia lineal al punto), el MLwiN y el HLM no permiten hacer ningún tipo de tratamiento. Además, por su parte el MLwiN forzará el procesador de la máquina y ralentizará los cálculos para cada interacción del modelo hasta llegar a un modelo conclusivo si la base de datos cuenta con valores perdidos. El HLM por su parte tampoco permite computarlos pero sí da la opción de señalar si el archivo de datos cuenta o no con valores perdidos en el primer nivel lo que ayuda a la velocidad del proceso de creación de Modelos Multinivel.

Centrados por completo en el desarrollo de Modelos Multinivel, las opciones entre las que elegir se reducen. Los programas específicos siempre permitirán un desarrollo completo y ajustado a la investigación por compleja que ésta sea. El visor de modelos, tanto del MLwiN como del HLM, es sin duda de gran utilidad a la hora de presentar los modelos construidos en el informe de resultados. Por su parte, aunque a través comandos, Stata también permite una visualización del modelo realizado.

No hay duda de que serán los datos disponibles los que guiarán la selección entre uno y otro software. La cantidad de niveles disponibles en cada programa varía y ese es un elemento a tener en cuenta y sobre lo que conviene detenerse: disponer de escasas unidades de nivel superior limita considerablemente los resultados por ello, en general, es preferible tener más unidades de nivel superior, con menos observaciones de primer nivel.

En cuanto a la selección del programa adecuado a los niveles de los análisis, si se trabaja con versiones anteriores al HLM 7 no se podrán realizar más de tres niveles de análisis. El SPSS y el Stata permiten hacer casi tantos niveles como se imagine, y el MLwiN acepta hasta cinco niveles de análisis.

Continuando con los requerimientos de los datos de la investigación, si es preciso ponderar los datos del estudio las opciones del software se debatirán principalmente entre HLM y MLwiN que sí dan esa opción, y además HLM permite que el cálculo sea aleatorio en los niveles haciendo "clic" en los errores. Por su parte los paquetes generales no posibilitan ponderar datos más allá de los de primer nivel.

Si bien el principio de parsimonia marca la construcción de los Modelos Multinivel, las interacciones entre variables aportan rica información sobre cómo la combinación de los efectos de diferentes variables impactan sobre la variable dependiente. A este respecto es importante señalar que contar con predictores correlacionados (multicolinealidad) es más grave a la hora de calcular Modelos Multinivel que en otro tipo de análisis como la regresión múltiple, dado que, efectivamente, en Modelos Multinivel se incluyen las interacciones entre los niveles. Por ello, si consideramos el cálculo de las interacciones entre las variables de diferentes niveles, el HLM y MLwiN permiten calcular si la interacción entre variables de diferentes niveles impacta sobre la variable dependiente. Sin embargo, no es posible introducir interacciones entre variables de un mismo nivel. Para este caso, será imprescindible acudir a los software generales que facilitan este tipo de operaciones.

La selección del método de estimación de los componentes de la varianza puede ser otro elemento que genere sus dudas. Desde la década de los años 70, los métodos de estimación de máxima verosimilitud y máxima verosimilitud restringida (MV Y MVR, respectivamente) han sido los más utilizados. MV presenta varias ventajas, incluyendo la habilidad de manejar algunas de las problemáticas de los métodos ANOVA (por ejemplo, la falta de unicidad, o las estimaciones de la varianza negativas). Ambos, MV y MVR, producen estimadores idénticos estimadores de efectos fijos. Por un lado MV toma en consideración los grados de libertad desde los efectos fijos y esto produce estimaciones de los componentes de la varianza menos equilibradas. Por otro lado, una desventaja del MVR es que el test ratio de verosimilitud no puede ser usado para comparar dos modelos con diferentes especificaciones en sus componentes fijos. Para muestras pequeñas, MVR es el método idóneo frente a MV, sin embargo, en muestras grandes las diferencias entre utilizar uno y otro son insignificantes (Snijders y Bosker, 1999). Puede decirse, por tanto, que “la utilización de uno u otro método de estimación responde más a una cuestión de gusto personal del investigador” (StataCorp, 2005:188).

El tratamiento de los residuos es un elemento fundamental en los estudios dentro del campo de la Investigación en Ciencias Sociales. Por ejemplo, los estudios de valor añadido permiten -o al menos intentan- aislar la contribución de la escuela y otras variables al desarrollo del aprendizaje de los alumnos. De este modo se responsabiliza “lo justo” a la escuela en torno a los resultados académicos de sus estudiantes. Así pues, para hacer estudios de Valor Añadido tan sólo será posible utilizar los programas específicos: HLM y MLwiN. El HLM permite guardar los archivos de residuos de cada nivel de análisis en archivos diferentes, y a través de, por ejemplo el SPSS, realizar las operaciones necesarias. El MLwiN, por su parte, permite además de guardar, operar con los archivos de residuos de los diferentes niveles por lo que los cálculos se realizan sin necesidad de apoyarse en ningún otro software. Como se ha dicho, ni SPSS ni Stata dan opción a guardar y operar con residuos más allá de aquellos de primer nivel.

Dos últimos elementos importantes a considerar, no ya para el desarrollo de Modelos Multinivel sino más bien para decantarse a la hora de comenzar a trabajar con uno o con otro, son la frecuencia o necesidad de actualización del software y su coste. Desde las páginas web de cada producto el usuario tiene acceso a solicitar la licencia del mismo a través de un simple e-mail. ¿Pero qué producto comprar? Para todos aquellos interesados en realizar Modelos Multinivel que residan en Reino Unido no parece que haya opción a dudas, MLwiN es gratuito; aunque habría que sopesar si las actualizaciones del software son realmente las suficientes. Para el resto de los países la decisión por uno u otro programa se agudiza. Si lo que se busca es un programa que sirva para todo, considerando las limitaciones en la creación de Modelos Multinivel, tanto el SPSS como el Stata requieren licencias una por cada módulo, otra por cada

versión, si bien es cierto que depende de la institución donde se trabaje, es muy probable que uno u otro, o incluso los dos estén ya instalados. Aprovecharse de las versiones de prueba, para estudiante, o el *renting*, es sin duda alguna una buena opción para decidirse.

## 5.- Referencias

- Aitkin, M. y Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society, Ser A*, 149, 1-43.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE. Transaction on Automatic Control*, 19, 716-723.
- Bartko, J.J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychol Rep*, 19, 3-11.
- Berry, W. (1993). *Understanding regression assumptions*. Londres: Sage.
- Bickel, R. (2007). *Multilevel analysis for applied reseach: it's just regression!*. Nueva York: The Guilford press.
- Bozdogan, H. (1987). Model selection and Akaike's selection criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52, 345-370.
- Bryk, A., y Raudenbush, S. W. (1992). *Hierarchical linear models for social and behavioral research: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Fleiss, J.L. (1986). *The design and analysis of clinical experiments*. Nueva York: Wiley.
- Gelman, A. y Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Nueva York: Cambridge University Press.
- Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika*, 73, 43-56.
- Goldstein, H. (1989). Restricted (unbiased) iterative generalized least squares estimation. *Biometrika*, 76, 622-623.
- Goldstein, H. (2010). *Multilevel statistical models* (4ª ed.). Londres: Arnold.

- Goldstein, H. y Rasbash, J. (1992). Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalized least squares. *Computational Statistics and Data Analysis*, 13, 63–71.
- Heck, R.H. y Thomas, S.L. (2000). *An Introduction to Multilevel Modeling Techniques*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Hernández Aguado, I., Porta Serra, M., Miralles, M., García Benavides, F. y Bolívar, F. (1990). La cuantificación de la variabilidad en las observaciones clínicas. *Medicina Clínica*, 95(1), 424-429.
- Hill, P.W. y Rowe, K.J. (1996). Multilevel modelling in school effectiveness research. *School Effectiveness and School Improvement*, 7, 1-34.
- HLM (1999). *HLM computer software*. Chicago, IL: Scientific Software International.
- Hox, J. (2010). *Multilevel Analysis. Techniques and Applications*. (2<sup>o</sup> ed). Nueva York: Routledge.
- Hurvich, C.M. y Tsai, C.L. (1989). A corrected Akaike information criterion for vector autoregressive model selection. *Journal of Time Series Analysis*, 14(3), 271-279.
- Leeuw, J. (2005). *Linear Multilevel Models*. En *Encyclopedia of Statistics in Behavioral Science*. Los Ángeles, CA: John Wiley & Sons, Ltd.
- Leyland, A.H. (2004). *A review of multilevel modelling in SPSS*. Bristol: Centre of Multilevel Modelling.
- Martínez-Garrido, C. y Murillo, F.J. (2013). El uso de los Modelos Multinivel en la Investigación Educativa. Estadísticas avanzadas para conocer y cambiar la educación en América Latina. En A. Salcedo (ed.), *Estadística en la Investigación: competencia transversal en la formación universitaria* (pp. 47-70). Caracas: UCV.
- McCullagh, P., y Nelder, J.A. (1989). *Generalized linear models* (2<sup>a</sup> ed.). Boca Ratón, FL: Chapman and Hall.
- MLwinN (1999). *MLwiN computer software*. Londres: Institute of Education.
- Murillo, F.J. (1999). Los modelos jerárquicos lineales aplicados a la investigación sobre eficacia escolar. *Revista de Investigación Educativa*, 17(2), 453-460.
- Murillo, F.J. (2005). *La investigación sobre eficacia escolar*. Barcelona: Octaedro.

- Murillo, F.J. (2008). Los Modelos Multinivel como herramienta para la investigación educativa. *Magis. Revista Internacional de Investigación Educativa*, 1(1), 17-34.
- Norusis, M.J. (2005). *SPSS 13.0 Advanced statistical procedures companion*. Upper Saddle River, NJ: Prentice Hall.
- Prieto, L., Lamarca, R. y Casado, A. (1998). La evaluación de la fiabilidad en las observaciones clínicas: el coeficiente de correlación intraclase. *Medicina Clínica*, 110(4), 142-145.
- Rabash, J. (1989). *ML2 software package*. Londres: Institute of Education.
- Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., Draper, D., Langford, I. y Lewis, T. (2012). *A user's guide to MlwiN (versión 2.6)*. Londres: University of Bristol.
- Raudenbush, S.W., Bryk, A.S., Cheong, Y.F., y Congdon, R. (2000). *HLM 5: Hierarchical linear and nonlinear modelling*. Lincolnwood, IL: Scientific Software International.
- Reise, P. y Duan, N. (2003). *Multilevel Modeling: Methodological Advances, Issues, and Applications*. Mahwah, NJ: Taylor & Francis.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
- Snijders, T. y Bosker, R.J. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. Londres: Sage.
- SPSS (1999). *SPSS computer software*. Chicago: SPSS, Inc.
- Stata (2001). *Stata computer software*. Texas: StataCorporation.