



Centro de Estudios Económicos

[www.colmex.mx](http://www.colmex.mx)

El Colegio de México, A.C.

*Serie documentos de trabajo*

**FORMACIÓN DE HABILIDADES Y LOGROS COGNITIVOS  
EN LOS NIÑOS DE MÉXICO**

Andrés Gonzalo Hincapié  
Yale University

Isidro Soloaga  
El Colegio de México

DOCUMENTO DE TRABAJO

Núm. XVII - 2010

# FORMACIÓN DE HABILIDADES Y LOGROS COGNITIVOS EN LOS NIÑOS DE MÉXICO

Andrés Gonzalo Hincapie<sup>1</sup> e Isidro Soloaga<sup>2</sup>

## Resumen

Se utilizan modelos de ecuaciones estructurales con variables latentes para estudiar la influencia de la inversión parental en la formación de la habilidad cognitiva de los niños en México, con datos de la ENNVIIH, -1 y -2. Se estudia la influencia de la habilidad parental y del rezago de habilidad del niño y se encuentra que este último es el único insumo estadísticamente significativo. Se introduce un elemento novedoso en un ejercicio más tradicional al estudiar también la tecnología de formación de logros cognitivos (no latentes) mientras la inversión sigue considerándose latente, y se encuentra que tanto el logro cognitivo rezagado como la inversión parental y el logro cognitivo parental son estadísticamente significativos en la producción de logros cognitivos. Al dividir la muestra en dos grupos de edad se encuentra evidencia de periodos sensibles para los insumos de la función de producción de logros cognitivos.

## Abstract

We use structural equation modeling, with latent variables to study the influence of parental inversion in cognitive skill formation in Mexican children. We use data coming from the ENNVIIH 1 and 2. We study the influence of parental skills and the lag of the child's skills, and we found that the latter is the only statistically significant input. We introduce a new element, in a more conventional exercise, studying the technology of formation of cognitive achievements (not latent), while we still consider inversion as latent; we found that, as the cognitive achievement lagged, parental inversion and parental cognitive achievement are statistically significant in producing cognitive achievements. When we split the sample into two age groups, we found evidence of sensitive periods for the inputs of the cognitive achievements production function.

Clasificación *JEL C31*

Palabras clave: habilidad cognitiva; variables latentes; influencia parental; análisis de factores

*Keywords: cognitive skills; latent variables; parental influence; factor analysis*

## 1. Introducción

---

<sup>1</sup> Maestro en Economía de El Colegio de México 2010-2011. Fox International Fellow at the Whitney and Betty MacMillan Center for International and Area Studies, Yale University, Agosto 2010-Julio 2011. andreshincapie@hotmail.es

<sup>2</sup> Profesor y Coordinador Académico del Centro de Estudios Económicos de El Colegio de México [isolaga@colmex.mx](mailto:isolaga@colmex.mx)

El estudio de la formación de habilidades cognitivas y no cognitivas es un tema al que se le ha prestado especial atención recientemente (ver, por ejemplo, Cunha, Heckman y Schennach (2010), Cunha y Heckman (2008)). Entre los elementos que justifican estudios de este tipo se encuentra la importancia de las habilidades en la explicación de resultados en etapas posteriores a la niñez, como la probabilidad de graduarse de la secundaria, el ingreso, la participación en actividades delictivas o los embarazos a edades tempranas. Este poder predictivo de las habilidades ha sido mostrado en algunos de los estudios citados en el compendio de la literatura hecho por Todd y Wolpin (2003), en el que los autores presentan los datos, métodos y supuestos subyacentes a las estimaciones realizadas de la tecnología de formación de logros cognitivos.

Otro argumento importante a favor de estos estudios es la reducción potencial de la desigualdad vía nivelación de las circunstancias (Roemer, 2005) en las que se desenvuelven los niños. Se ha mostrado que las brechas de habilidad cognitiva surgen a edades tempranas y se mantienen en el ciclo vital (ver figura 1 en Cunha y Heckman (2007) o figura 15 en Arias *et al.* (2010) para una ilustración en México). De esta manera, si se reconoce la importancia de las habilidades cognitivas en la predicción de resultados en la edad adulta, las acciones de política social que nivelen las circunstancias en las que se desarrollan los individuos a edades tempranas representarán, muy posiblemente, una disminución tanto de la desigualdad contemporánea como de la desigualdad en la edad adulta.

En este trabajo se abordan concretamente dos formas de acercarse al problema. En la primera parte, se considera a la habilidad cognitiva como una variable latente (que se manifiesta a través de otras variables) y se utilizan modelos de ecuaciones estructurales<sup>3</sup> con variables latentes para la estimación. El enfoque permite utilizar datos provenientes de múltiples fuentes que pueden ser más o menos informativas sobre las variables latentes.

En la segunda parte, se abandona el modelo que trata la habilidad cognitiva como variable latente y se retoma la visión más tradicional utilizando los puntajes de la prueba de Raven (logros cognitivos) mientras se introduce un elemento novedoso, al seguir considerando la inversión parental como variable latente, en lugar de

---

<sup>3</sup> SEM por sus siglas en inglés.

tener un vector de insumos parentales. Como se describirá más adelante, se considera que este modelo no utiliza toda la información disponible y la motivación para llegar a él son las posibles debilidades de las variables indicadoras de habilidad cognitiva.

Los datos provienen de un panel elaborado con la *Encuesta nacional sobre niveles de vida de los hogares*, ENNVIH, -1 y -2 para niños con edades menores o iguales a 12 años en 2005. La principal fuente de información, aunque no la única, sobre las habilidades cognitivas de los niños es el conjunto de resultados de las pruebas de matrices progresivas de Raven. Para la estimación del efecto de la inversión latente se consideran varias indicadoras, que serán descritas más adelante.

Este trabajo se diferencia claramente de los estudios relacionados con habilidades cognitivas en México, en tanto se consideran estas habilidades y la inversión como variables latentes, que se manifiestan a través de múltiples indicadores que son observados con cierto componente de error. Se diferencia, también, al utilizar datos en panel, lo que permite analizar el fenómeno de la formación de habilidades cognitivas en una perspectiva dinámica.

Los resultados principales resaltan la importancia del rezago de la habilidad cognitiva (autofomento de la habilidad cognitiva). Al utilizar el enfoque de habilidad cognitiva latente se encuentra que lo relevante es el autofomento. En el periodo tratado parece haber evidencia de lo que denominamos como la *persistencia de la historia* en la formación de la habilidad cognitiva, al menos en el rango de edad considerado; una vez que se incluye el rezago de la habilidad tanto la inversión parental como la habilidad parental pierden su poder explicativo. Al utilizar el enfoque de logros cognitivos tenemos que la inversión parental y el logro cognitivo parental juegan un papel relevante. En efecto, encontramos que los coeficientes estandarizados del puntaje rezagado, la inversión y el puntaje parental son de magnitud similar.

Como una posible explicación para las diferencias de resultados (con respecto a la inversión y habilidad parental) entre modelos se argumenta la posibilidad de que la estimación de la formación de logros cognitivos esté evitando los problemas que se enfrentan en la estimación con variables cognitivas latentes. El modelo de logros cognitivos evitaría las dificultades con las variables indicadoras de habilidad cognitiva (distintas al

puntaje en la prueba de Raven), que pueden incluir un componente de información socioeconómica de los hogares considerado también en las indicadores de la inversión latente. Dados los resultados, el argumento general no pretende ser concluyente en el tema de la formación de habilidades y logros cognitivos en México, sino ser un acercamiento al tema, sujeto a los problemas que la disponibilidad de datos genera.

Por último, utilizando los puntajes de la prueba de Raven y dos grupos de edad se estudia la existencia de una función de producción de logros cognitivos que sea específica por grupo de edad. En esta parte del análisis se rechaza la invarianza temporal de la función de formación de logros cognitivos. Se encuentra evidencia de periodos sensibles para la inversión, el autofomento y el componente parental. La función de producción de logros cognitivos es específica al grupo de edad.

Lo que resta del documento está organizado de la siguiente forma: la sección dos contiene una revisión breve de la literatura; en la tres se expone el modelo teórico; la sección cuatro trata sobre la estrategia de estimación, se da una breve introducción al SEM (por sus siglas en inglés), se citan algunos problemas que se enfrentan con las variables y con la estimación, así como algunas de sus posibles soluciones; en la sección 5 se ofrece una descripción sucinta de los datos y en la última se presentan los resultados para los modelos de formación de habilidades y logros cognitivos.

## **2. Literatura en formación de habilidades**

Hay una gama de estudios teóricos y estimaciones de la tecnología de formación de habilidades cuya existencia es reciente. Muchos de ellos utilizan puntajes en pruebas, por ejemplo, el PIAT<sup>4</sup> en Estados Unidos o la prueba de matrices progresivas de Raven en México, como variables que aproximan la capacidad cognitiva y estiman los efectos de diferentes controles: características parentales, variables de inversión y condiciones del entorno, rezagadas y contemporáneas, sobre los logros cognitivos medidos a través de estas pruebas. A las diferentes especificaciones de la función que existen en la literatura se les ha dado forma teniendo en cuenta la disponibilidad de datos históricos o contemporáneos en las bases existentes.

---

<sup>4</sup> *Peabody Individual Achievement Test.*

En este sentido, Todd y Wolpin (2004) utilizan diferentes alternativas para la estimación de la función de producción de logros cognitivos presentes en la literatura, que van desde el uso de variables instrumentales, o efectos fijos de la madre o del niño, hasta mínimos cuadrados ordinarios o funciones de valor agregado (en las que la medición rezagada de la prueba es un estadístico suficiente del historial de insumos). En su especificación preferida, variables instrumentales para los puntajes en diferencias, utilizando datos de la NLSY79 *Child Sample*, encuentran que las variables de inversión parental son determinantes, sustancialmente significativos, de los puntajes en las pruebas para los niños de Estados Unidos.

En México existen varios estudios referentes a las habilidades cognitivas. Rubalcava y Teruel (2004) relacionan las habilidades cognitivas maternas con resultados en variables de salud. Oliva, Rubalcava y Teruel (2004) relacionan habilidades cognitivas con la percepción subjetiva de salud. Altamirano y Soloaga (2009) estudian la transmisión intergeneracional de la desigualdad en habilidades cognitivas. López, Mayer y Serván (2008) tratan la transmisión intergeneracional de las habilidades cognitivas por niveles socioeconómicos.

Mayer y Serván (2009) estudian concretamente la formación de la capacidad cognitiva en México utilizando un modelo de frontera estocástica y datos de la ENNVIIH-1. En su estudio los autores encuentran que la capacidad cognitiva no depende, exclusivamente, de factores genéticos, sino también de la situación económica de la familia, de variables de política pública y de características de la localidad. Esta relación puede ser interpretada como un hallazgo en concordancia con literatura reciente que ha dejado clara la obsolescencia del debate que situaba una disyuntiva entre el componente heredado y el componente formado en vida de las habilidades, en la medida en que las variables utilizadas en su estudio pueden ser vistas como variables de inversión en la tecnología de formación de habilidades cognitivas de los niños.

Otro enfoque en el estudio de la tecnología de formación de habilidades es el que las considera como variables latentes no observables que se manifiestan a través de otras que sí se observan. Un ejemplo, sería considerar a la capacidad cognitiva como una variable escalar no observable (aunque estos modelos no están restringidos a magnitudes escalares) y a los resultados en diferentes pruebas de capacidad intelectual como

manifestaciones de dicha variable latente a las que les cabe error; es decir, no es posible medir perfectamente la variable latente a partir de los resultados en las pruebas.

Estos modelos han buscado, especialmente, explicar algunos hallazgos en la literatura de formación de habilidades de los niños. Algunos de sus hallazgos son expuestos por Cunha *et al.* (2005). Las habilidades importan para explicar resultados en la edad adulta. Las habilidades son multidimensionales, la dicotomía entre lo genéticamente heredado y las habilidades formadas es falsa. Es posible compensar el efecto de ambientes familiares adversos. Hay periodos sensibles, las habilidades de distintos tipos parecen ser manipulables a distintas edades. El retorno a la inversión en edades tempranas es alto. Hay complementariedad de la inversión, la inversión en edades tempranas debe ser acompañada de inversión en edades subsiguientes.

Los mismos autores, junto con otros, Cunha y Heckman (2007, 2008, 2009), Cunha, Heckman y Schennach (2010) y Heckman, Stixrud y Urzua (2005) han desarrollado modelos teóricos y han hecho estimaciones de la tecnología de formación de habilidades. Cunha, Heckman y Schennach (2010) utilizan una especificación tipo CES de la tecnología de formación de habilidades y corroboran la hipótesis de existencia de periodos sensibles en edades tempranas y la de complementariedad de la inversión. En este trabajo, en el que incluyen habilidades cognitivas y no cognitivas,<sup>5</sup> los autores encuentran también que existe productividad propia y cruzada de las habilidades; es decir, un nivel mayor de habilidades no cognitivas en un periodo puede fomentar la acumulación de habilidades no cognitivas en el periodo siguiente y también la acumulación de habilidades cognitivas en el siguiente periodo.

Entre las críticas que se hacen a los modelos con variables latentes se encuentra la interpretación no directa de sus resultados. Cunha, Heckman y Schennach (2010) y Cunha y Heckman (2008) resuelven el problema anclando la escala de las variables latentes a resultados en la edad adulta -como ingreso o la probabilidad de graduarse de la secundaria- a través de ecuaciones que relacionan el nivel de la variable latente

---

<sup>5</sup> A pesar del creciente estudio de la relevancia de las habilidades no cognitivas en el conjunto de habilidades que explican resultados en la edad adulta (Postlewaite y Silverman, 2006; Heckman y Rubinstein, 2001 y Farkas, 2003) y de los hallazgos acerca de su papel no insignificante en la promoción de otras habilidades (lo que ha sido llamado productividad cruzada), en el presente trabajo se ha tenido que dejar su análisis a un lado debido a la ausencia de datos de este tipo para los niños de la muestra.

en el último periodo de formación de habilidades con los resultados. La ausencia de datos para practicar este anclaje<sup>6</sup> hace a este trabajo susceptible de dicha crítica.

### 3. El modelo

El modelo en el que se enmarcan las estimaciones realizadas en gran parte de nuestro análisis se encuentra en línea con los expuestos por Cunha *et al.* (2005) y Cunha y Heckman (2007) que integran algunos hechos estilizados sobre la formación de habilidades como son la presencia de periodos sensibles y críticos, el autofomento de las habilidades (importancia del rezago de habilidad) y la productividad cruzada (entre tipos de habilidades diferentes).

Supongamos la existencia de sólo dos tipos de habilidades, cognitivas ( $\theta^C$ ) y no cognitivas ( $\theta^N$ ), y que son magnitudes escalares.  $\theta^C$  puede ser considerada como la capacidad cognitiva del individuo, una construcción teórica que, evidentemente, no es observada de manera directa. De forma similar,  $\theta^N$  puede pensarse como la capacidad no cognitiva del individuo, también como una construcción teórica que puede recoger habilidades no cognitivas que son valoradas en el mercado (Postlewaite y Silverman, 2006) como persistencia, liderazgo y capacidad de establecer relaciones sociales.

Los insumos en la producción de estas habilidades son el acervo de habilidades acumuladas hasta el momento y la inversión realizada en formación de habilidades,  $I$ . Si suponemos efectos no inmediatos de la inversión y múltiples periodos, la forma de la función de producción de habilidades estaría dada por:

$$\theta_{t+1}^k = f_t^k(\theta_t^C, \theta_t^N, I_t) \text{ con } k \in \{C, N\} \text{ y } t \in \{1, \dots, T\} \quad (1)$$

Donde  $T$  es el periodo final de formación de habilidades.

Hay dos cosas que deben anotarse acerca de la ecuación (1). Primero, el subíndice  $t$  de la función  $f$  permite la existencia de periodos sensibles y críticos, como se mostrará más adelante. Intuitivamente, la tecnología de formación de habilidades cambia de acuerdo con el periodo, la inversión puede ser más

---

<sup>6</sup> En el modelo de habilidad cognitiva latente la unidad de medida de las variables latentes cognitivas será en puntos de la prueba de Raven. La unidad de medida para la inversión será en (logaritmo de) pesos.



productiva en unos periodos que en otros. Segundo, la inversión carece de superíndice, lo que sugiere que la inversión en habilidades no es distinguible entre los tipos de habilidad. El insumo de inversión es único y tiene efectos para ambos tipos de habilidad. Este supuesto responde a la dificultad para distinguir empíricamente entre uno y otro tipo de inversión.

*Periodos sensibles.* Son aquellos periodos  $t'$  para los cuales se cumple que

$$\frac{\partial f_t^k}{\partial I_{t'}} > \frac{\partial f_{t'}^k}{\partial I_t} \quad \forall t \text{ con } t, t' \in \{1, \dots, T\} \quad (2)$$

*Periodos críticos.* Para la formación de la habilidad  $k$  son aquellos periodos  $t'$  en los cuales se cumple que

$$\frac{\partial f_t^k}{\partial I_{t'}} \neq 0 \text{ y } \frac{\partial f_{t'}^k}{\partial I_t} = 0 \quad \forall t \neq t' \text{ con } t, t' \in \{1, \dots, T\} \quad (3)$$

*Autofomento y productividad cruzada.* Ambas sugieren que el inventario de las habilidades acumulado hasta el momento  $t$  favorece la acumulación de éstas para el periodo siguiente. Es decir

$$\frac{\partial f_t^k}{\partial \theta_t^j} > 0 \quad \text{con } k, j \in \{C, N\} \text{ y } t \in \{1, \dots, T\} \quad (4)$$

Una implicación de la tecnología de formación de habilidades en (1) es que el acervo de habilidades en  $t+1$  puede ser expresado en función de la dotación inicial de habilidades  $\theta_t^k$  y del vector intertemporal de inversión  $I_t = (I_1, I_2, \dots, I_t)$ . El resultado sugiere que las diferencias en habilidades surgidas por el *accidente de nacer* pueden ser compensadas a través de niveles apropiados de inversión. En este marco, el resultado al que hacemos referencia en la introducción como *persistencia de la historia* correspondería a periodos críticos de la inversión en edades menores a los cinco años. Un ejemplo de ello puede verse en la simulación llevada a cabo por Cunha, Heckman y Schennach (2010), en la que el planeador social maximiza los logros educativos agregados o minimiza una medida agregada del crimen.

#### 4. Estrategia de estimación

En este trabajo se asume una forma lineal<sup>7</sup> para la tecnología de formación de habilidades y se restringirá la formación de habilidades cognitivas, por las razones ya mencionadas, de manera que la tecnología estimada estará dada por una simplificación de (1):

$$\theta_{t+1}^c = \gamma_1^c \theta_{t+1}^c + \gamma_2^c I_{t+1}^c + u_{t+1}^c \text{ con } t \in \{1, \dots, T\} \quad (5)$$

Suponer una tecnología así tiene una desafortunada implicación. Supóngase una tecnología invariante en el tiempo (i.e. sin periodos sensibles) de manera que  $\gamma_1^c = \gamma_2^c \forall t, s \in \{1, \dots, T\}$ . Ello implica que el efecto ocasionado por la falta de inversión en un periodo  $t$  puede ser compensado, enteramente, por inversiones en periodos subsiguientes con un sobre costo prácticamente nulo. La intuición de dicho resultado puede ser más clara contrastando este ejemplo con uno en el que los insumos de la tecnología son complementos perfectos, de manera que, la ausencia de inversión en un periodo, lleva a la ausencia de habilidades cognitivas de ahí en adelante, independientemente del nivel de inversión.

Una forma de permitir mayor flexibilidad en la tecnología (1) es considerando una función CES que permita estudiar la elasticidad de sustitución entre los insumos de la tecnología. Cunha, Heckman y Schennach (2010) hacen esta estimación encontrando resultados consistentes con los hechos estilizados antes mencionados.

A pesar de que la tecnología en (5) puede ser restrictiva para el estudio de la formación de habilidades cognitivas, esta especificación permite estudiar la existencia de periodos sensibles y de autofomento de las habilidades (recuérdese que la productividad cruzada está fuera del alcance de nuestro estudio).

Además de la estimación de la ecuación en (5), también se estudian otras especificaciones en las que se incluyen las habilidades cognitivas maternas como variable latente, o la escolaridad o el puntaje de la madre de la prueba de Raven, como controles en la tecnología de formación de habilidades cognitivas.

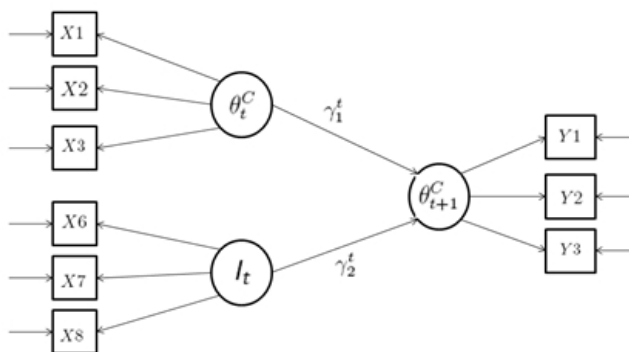
#### 4.1 Modelo de ecuaciones estructurales con variables latentes

---

<sup>7</sup> Cunha y Heckman (2007) presentan un ejemplo de cuál sería la estrategia óptima de inversión suponiendo una función de este tipo y dos periodos.

Hasta ahora se ha dicho que la capacidad cognitiva de los niños será considerada como una variable latente que no es observable directamente. Sin embargo, sí se observan manifestaciones de esta variable, en puntajes de pruebas como la de Raven, y otras variables que asimismo pueden contener información sobre la capacidad cognitiva, como sería si el niño leyó o no la semana pasada. Esta forma de entender la capacidad cognitiva se corresponde con un análisis de factores, técnica ampliamente utilizada en las ciencias sociales (ver Duncan y Goldberger, 1973) aunque no en la economía, especialmente por la dificultad para dar una interpretación clara de los coeficientes obtenidos en dicho tipo de estimaciones.

En nuestro análisis se considerará la inversión parental,  $I_t$ , también como una variable latente, de manera que para la estimación de (5) se tienen no observables del lado izquierdo y del lado derecho de la ecuación.<sup>8</sup> Un diagrama de la tecnología a estimar utilizando la notación tradicional en este tipo de estudios (Bollen, 1989) es:



**Figura 1:** Diagrama del modelo de ecuaciones estructurales con variables latentes para la estimación de la tecnología en (5).

En el diagrama expuesto en la figura 1 se supone que se cuenta con tres indicadores por cada variable latente (exógena o endógena). La dirección a la que apuntan las flechas indica la dirección de los efectos. Movimientos en las variables latentes se ven reflejados en movimientos en sus variables indicadoras y movimientos en las variables latentes endógenas,  $\theta_t^C$  e  $I_t$ , provocan movimientos en la variable latente exógena,  $\theta_{t+1}^C$ .

<sup>8</sup> Un modelo en el que la única variable latente se encuentra del lado izquierdo de la ecuación es el MIMIC (Jöreskog y Goldberger, 1975).

Formalmente, al suponer relaciones lineales entre las variables latentes y sus variables indicadoras y entre las variables latentes endógenas y exógenas,<sup>9</sup> el modelo a estimar está dado por el siguiente sistema de ecuaciones:

$$\theta_{t+1}^c = \Gamma' \theta_t + \zeta_{t+1} \quad (6)$$

$$Y_{t+1} = \Lambda_y \theta_{t+1}^c + \epsilon_{t+1} \quad (7)$$

$$X_t = \Lambda_x \theta_t + \delta_t \quad (8)$$

Donde  $\theta_t = (\theta_t^c, I_t)'$ ,  $\Gamma' = (\gamma_1, \gamma_2)$ ,  $Y_{t+1}$  es el vector de indicadores de  $\theta_{t+1}^c$  (variable latente escalar),  $X_t$  es el vector de indicadores de  $\theta_t$ ,  $\Lambda_y$  es un vector de coeficientes y  $\Lambda_x$  es una matriz de coeficientes. Por último,  $\zeta_{t+1}$  es el error latente,  $\epsilon_{t+1}$  y  $\delta_t$  son los vectores de errores en las ecuaciones de medición. Por simplicidad se ha omitido el superíndice  $t$  de los vectores de coeficientes, lo que equivale a decir que la función es invariante en el tiempo; sin embargo, la simplificación sólo se hace para facilitar la exposición del modelo básico de estimación.

Al omitir los subíndices por simplicidad se tiene además que

$$\begin{aligned} E[\zeta] &= 0; & Var[\zeta] &= \psi \\ E[\epsilon] &= 0; & Var[\epsilon] &= \Theta_\epsilon \\ E[\delta] &= 0; & Var[\delta] &= \Theta_\delta \end{aligned} \quad (9)$$

En el modelo,  $\zeta$ ,  $\epsilon$  y  $\delta$  están mutuamente no correlacionadas, los valores esperados de las variables latentes son ceros,  $\zeta$ , no está correlacionada con  $\theta_t$ , los valores esperados de los errores son igualmente cero,  $\epsilon$  no está correlacionado con  $\theta_{t+1}^c$ , ni con  $\zeta$ , ni con  $\delta$ ;  $\delta$  no está correlacionado con  $\zeta$ , ni con  $\theta_{t+1}^c$ , ni con  $\epsilon$ , y  $E[\theta_t' \theta_t] = \Phi$ .

Si definimos el vector de datos como  $Z = [Y', X']'$  se tiene que la matriz de varianzas-covarianzas de  $Z$ ,  $Cov(Z) = \Sigma$ , queda expresada en función de los parámetros del modelo. Para la estimación de los

<sup>9</sup> Modelo LISREL, ver por ejemplo Jöreskog (2000).

parámetros del modelo se establece  $\hat{\Sigma} = \mathbf{S}$  (donde  $\mathbf{S}$  es la matriz de varianzas-covarianzas muestral) y se resuelven las ecuaciones resultantes utilizando rutinas iterativas que consideran una función criterio para reducir la discrepancia entre  $\hat{\Sigma}$  y  $\mathbf{S}$ . Dado que las ecuaciones resultantes surgen de restricciones de covarianza en función de los parámetros del modelo, no es necesario calcular los valores de las variables latentes para la estimación de los parámetros del modelo de ecuaciones estructurales con esas variables.<sup>10</sup>

Aun así, los valores de estas variables podrían ser de interés para mostrar, por ejemplo, las características de su distribución. Este ejercicio se lleva a cabo para la inversión latente en el modelo de formación de logros cognitivos (anexo B) dada su relevancia en el mismo. Se excluye la especificación de la forma en que dichos valores se obtienen ya que está fuera de las intenciones del análisis. Una forma para recuperar dichos valores<sup>11</sup> puede verse en Jöreskog (2000) y en los trabajos que el autor cita.

Las estimaciones que incluyen indicadores categóricas, como es en nuestro estudio, requieren un tratamiento un poco diferente.<sup>12</sup> En estos casos se considera la existencia de indicadores latentes con respecto a cuyo valor se determinan las variables categóricas. Por ejemplo, si la variable latente  $\theta_{t+1}^c$  se manifiesta a través de una variable dicotómica  $y_t$ , entonces el modelo de medición se transforma en un modelo de probabilidad tal que

$$y_t = \lambda_1 \theta_{t+1}^c + \epsilon_t \quad (10)$$

$$y_t = \begin{cases} 0 & \text{si } y_t \leq \alpha_1 \\ 1 & \text{si } y_t > \alpha_1 \end{cases} \quad (11)$$

En la estimación se obtiene el umbral  $\alpha_1$ . Los coeficientes de las ecuaciones de medición corresponden a la función *índice* para la variable indicadora latente, cuyo valor se evaluará con respecto al umbral estimado para dar a la variable dicotómica observada su valor (son los coeficientes de una regresión *probit* o *logit*).

<sup>10</sup> Un ejemplo puede encontrarse en el anexo A.

<sup>11</sup> El *software Mplus* permite recuperar estos valores para gran variedad de modelos.

<sup>12</sup> Ver, por ejemplo, Bollen (1989) o Everitt (1984).

La estimación de dichos modelos se hace, tradicionalmente, por máxima verosimilitud o utilizando un estimador de mínimos cuadrados ponderados. El *software Mplus*, que se utiliza en este trabajo para hacer la estimación, emplea un estimador robusto de mínimos cuadrados ponderados,<sup>13</sup> llamado WLSMV, que ha probado tener un buen desempeño estadístico y aumentar la velocidad computacional en la estimación de los modelos de variables latentes que incluyen variables categóricas (Muthén du Toit y Spisic, 1997). El método de estimación utilizado por el *Mplus* permite, además, utilizar observaciones con valores faltantes en algunas de las variables indicadoras.<sup>14</sup>

## 4.2 La formación de logros cognitivos

Consideramos que abandonar la noción de la habilidad cognitiva como variable latente es un acercamiento que no utiliza toda la información disponible sobre el problema. Si se acepta la idea de la habilidad cognitiva como una variable latente de la cual se observan algunas de sus manifestaciones, utilizar el modelo de formación de logros cognitivos (el que sólo considera los puntajes de Raven) significa desaprovechar toda la información disponible sobre la habilidad cognitiva latente. En este sentido, si se contara con un buen conjunto de datos (por ejemplo, resultados en varias pruebas de diferente tipo) que no sufriera de desventajas como las mencionadas en la sección de datos, sería preferible utilizar el SEM. Aun así, dadas las relaciones teóricas supuestas entre la habilidad latente y sus manifestaciones (como es el puntaje en la prueba de Raven), los resultados con uno u otro acercamiento no deberían diferir de manera significativa.

Una posible debilidad de la estimación del modelo, considerando a la habilidad cognitiva como variable latente, sería la pobreza informativa de los indicadores dicotómicos que se utilizan en el estudio (ver sección de datos). Recuérdese que el proceso de formación de habilidades ha sido considerado acumulativo, es decir, el inventario de habilidades crece en el tiempo, al menos en el periodo considerado en el estudio. De tal manera que la riqueza informativa de las variables dicotómicas, como indicadoras del nivel del inventario, adolecen de

---

<sup>13</sup> *Mplus* también permite hacer la estimación por máxima verosimilitud. Este tipo de estimación no ha sido utilizada por problemas de tiempo en la convergencia de las estimaciones y capacidad computacional.

<sup>14</sup> Ver Muthén y Muthén (2007).

falta de variabilidad, lo que puede impedirles capturar el fenómeno de la mejor manera.<sup>15</sup> Un agravante del problema es la alta concentración de las variables dicotómicas en el cero (ver cuadro 1). A pesar de que se observa un aumento de la masa acumulada en el valor uno, entre  $t$  y  $t + 1$ , podría pensarse que la variación es insuficiente para dar cuenta del fenómeno acumulativo que se estudia.

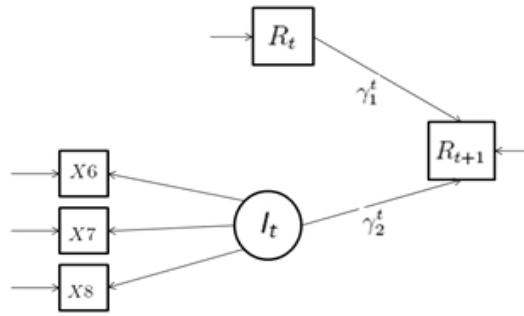
Por otra parte, dos individuos con habilidades cognitivas muy diferentes (por ejemplo, alta y baja) pueden decidir leer o utilizar *internet*. La contundencia de este problema, que no se presenta necesariamente en las pruebas de Raven, puede ser disminuido si se piensa en la existencia de un costo para los individuos, decreciente en la habilidad, de leer o utilizar *internet*. Una solución al problema sería restringirse a los resultados en la prueba de Raven y analizar la producción de logros cognitivos retornando a especificaciones más estándar, como son las mencionadas en el trabajo de Todd y Wolpin (2003).

En nuestro estudio utilizamos esta solución, aunque no se abandona la noción conceptual de la inversión parental como variable latente. Es decir, en lugar de tener un vector de variables que aproximen la inversión parental entrando directamente como variables independientes, la covarianza entre ellas es aprovechada para hacer uso del concepto de inversión parental, como variable latente, recogiendo las múltiples expresiones de la inversión y entendiéndola como una actitud latente hacia los hijos, que se manifiesta a través de las variables indicadoras utilizadas.

La solución, aunque restringida por el abandono del concepto atractivo de la habilidad cognitiva como variable latente, es novedosa en la medida en que sigue considerando la inversión parental como variable latente, en una especificación más tradicional en la que sólo se utilizan los puntajes de Raven en el análisis de la formación de logros cognitivos. El diagrama de esta especificación se muestra en la figura 2.

---

<sup>15</sup> Cunha y Heckman (2008) utilizan el puntaje en dos pruebas aprovechando la variación entre edades.



**Figura 2:** Formación de logros cognitivos. Inversión parental como variable latente.

Para la estimación del modelo se sigue usando la misma estrategia de estimación. El modelo está dado por una simplificación de las ecuaciones (6) y (8).

### 4.3 Posibles problemas en la estimación

Los supuestos clave establecidos por Todd y Wolpin (2003) para la especificación de modelos de este tipo no son ignorados en el análisis. En efecto, se hacen dos supuestos fundamentales para la especificación. Por una parte, se supone que el rezago de la habilidad (o del puntaje según sea el caso) es un estadístico suficiente para el historial de insumos pasados. Por otra parte, dado que no se cuenta con información sobre la dotación inicial, se está suponiendo que el rezago de la habilidad (puntaje) es también un estadístico suficiente para la dotación.<sup>16</sup>

En nuestro trabajo se pone a prueba el primero de los supuestos. Todd y Wolpin (2003) sugieren hacerlo con una prueba sencilla. Proponen incluir la inversión contemporánea, al igual que sus rezagos y el rezago de habilidad (puntaje), como regresores de la habilidad (puntaje) contemporánea y observar si la inversión rezagada tiene algún poder explicativo adicional. Si el supuesto es correcto la inversión rezagada no debería tener ningún poder explicativo adicional. Como puede verse en los cuadros 7 y 8 (columnas (1) a (4)) el supuesto se mantiene para el modelo de habilidad cognitiva latente, al igual que para el modelo que considera sólo los puntajes de la prueba de Raven una vez que se incluye el puntaje parental.

<sup>16</sup> Parte del problema podría pensarse resuelto por la inclusión de la habilidad parental si se le considera como una posible aproximación del componente hereditario de la dotación inicial y se piensa en el componente idiosincrático como ortogonal a las variables exógenas incluidas en el modelo. Sin embargo, esta forma de ver el problema cambiaría la interpretación del coeficiente de la variable cognitiva parental.



## 5. Datos

En la estimación de la tecnología de formación de habilidades se utilizan datos de la ENNVIH<sup>17</sup> 2002 y 2005 para construir un panel de dos periodos con información de los niños con edades entre los cinco<sup>18</sup> y doce años, y su correspondiente información parental cognitiva y de inversión.

*Habilidad cognitiva ( $G_t$ )*. Como indicadores de la habilidad cognitiva son utilizadas las siguientes variables:

- PUNTAJE ( $R_t$ ). Puntaje de la prueba de Raven. Es una variable entera cuyo rango está entre 0 y 18, que corresponde a la asignación de un punto por cada pregunta respondida correctamente, de las 18 preguntas hechas al niño. Es la variable base del estudio. Sin embargo, la calidad de esta variable es cuestionable.

Uno de los problemas con PUNTAJE es que la correlación entre periodos es extrañamente baja (0.268) para los niños con edades menores a 12 años en el 2005.<sup>19</sup> Estas bajas medidas no corresponden a lo observado en otros estudios, como los citados en la sección de estudios de confiabilidad en Court, Raven y Raven (2006). En ellos se obtienen valores de correlación *test-retest* que van de 0.6 a poco más de 0.9 para diferentes intervalos de tiempo, tipos de individuos y rangos de edad.

Otro inconveniente es que la prueba de Raven hecha a niños y adultos de 13 o más años es distinta a la aplicada a niños de 12 años o menos. Tal característica lleva a que la comparación entre niños de edades antes y después de este umbral no sea clara, dada la característica acumulativa de las habilidades cognitivas en (5). Para eludir el problema hemos reducido el rango de comparación de edades, de manera que se consideran edades menores o iguales a 12 años en la ENNVIH-2.

Por último, las preguntas presentes en los cuestionarios de las ENNVIH 1 y 2, antes y después del umbral de edad, son una selección del total de preguntas presentes en la prueba de Raven (Raven, 2007).

---

<sup>17</sup> Ver Rubalcava y Teruel (2006, 2008).

<sup>18</sup> Edad mínima en la que los niños responden la prueba. La explicación del límite superior se encuentra más adelante.

<sup>19</sup> La baja correlación se encuentra en toda la muestra y al dividir en grupos de edad. La correlación utilizando la variable estandarizada por sexo y edad, conocida como ZRAVEN, es igualmente baja (0.239).

No se encontró ningún documento que analizara la fiabilidad de esta selección con respecto a los resultados obtenidos con la prueba completa.

- LEYO e INTERNET. Variables dicotómicas que valen uno si el niño leyó o utilizó *internet*, respectivamente, la semana pasada. Las variables están sujetas a las críticas, mencionadas anteriormente, sobre su capacidad para capturar el fenómeno acumulativo y su posible pobreza para diferenciar entre individuos de alta y baja habilidad. Una precaución más sobre el uso de estas variables es que pueden estar recogiendo, también, características socioeconómicas de los hogares de los niños y no sólo habilidades.

*Habilidad cognitiva parental* ( $R^P$ ). Se consideran las mismas variables indicadoras que se utilizan para la habilidad cognitiva de los niños más la variable RECADO (ver cuadros 1 y 2). Es una variable dicotómica que toma el valor uno si quien responde sabe leer y escribir un recado en español.

**Cuadro 1.** Estadística descriptiva. Indicadoras dicotómicas.

|                              | 2002 |      |      | 2005 |      |      |
|------------------------------|------|------|------|------|------|------|
|                              | Obs. | 0    | 1    | Obs. | 0    | 1    |
| <i>Cognitivas</i>            |      |      |      |      |      |      |
| LEYO                         | 2771 | 0.77 | 0.23 | 2773 | 0.57 | 0.43 |
| INTERNET                     | 2771 | 0.99 | 0.01 | 2773 | 0.95 | 0.05 |
| <i>Cognitivas parentales</i> |      |      |      |      |      |      |
| LEYOP                        | 2642 | 0.72 | 0.28 | 1425 | 0.70 | 0.30 |
| INTERNETP                    | 2642 | 0.98 | 0.02 | 1425 | 0.97 | 0.03 |
| RECADO                       | 2642 | 0.11 | 0.89 | 1426 | 0.10 | 0.90 |
| <i>Inversión parental</i>    |      |      |      |      |      |      |
| CULTURA                      | 2751 | 0.84 | 0.16 | 2757 | 0.86 | 0.14 |
| ACULTURA                     | 2776 | 0.85 | 0.15 | 2774 | 0.77 | 0.23 |
| CLASES                       | 2776 | 0.97 | 0.03 | 2774 | 0.95 | 0.05 |
| GRITARIA                     | -    | -    | -    | 2627 | 0.79 | 0.21 |
| PEGARIA                      | -    | -    | -    | 2619 | 0.66 | 0.34 |

LEYOP e INTERNETP son variables de la madre.

Valores en las columnas 0 y 1 son proporciones

Para el puntaje de la prueba Raven parental,  $R^P$ , se utiliza la media de los puntajes de la madre en ambos periodos, o el puntaje en el periodo disponible cuando no hay información de la madre en los dos. Para las variables RECADO, LEYO e INTERNET se usa el máximo de la observación en ambos periodos (ver cuadro 3). Con este procedimiento se considera se captura el supuesto de la invariabilidad de habilidades cognitivas parentales. Por un lado, se utiliza la media del puntaje de la prueba Raven, en un intento por capturar el valor de

largo plazo, mientras se reduce el posible error de medición y se ajusta a la baja por el efecto aprendizaje en la prueba. Por el otro, considerar el máximo de las variables dicotómicas LEYO e INTERNET pretende capturar el hecho de que la manifestación de la habilidad cognitiva, a través de estas variables indicadoras, puede quedar oculta si la madre no leyó o no utilizó *internet* por razones ajenas a su habilidad. Con respecto a la variable RECADO, la utilización del máximo de ambos periodos no parece ser la estrategia más apropiada. Sin embargo se mantiene su uso, pues al retirar las observaciones<sup>20</sup> de madres que manifiestan aprendizaje los resultados no cambian; así se evita desechar la información de sus hijos.

**Cuadro 2.** Estadística descriptiva. Indicadoras continuas.

|                              | 2002 |       |      | 2005 |       |      |
|------------------------------|------|-------|------|------|-------|------|
|                              | Obs. | Media | DE   | Obs. | Media | DE   |
| <i>Cognitivas</i>            |      |       |      |      |       |      |
| PUNTAJE                      | 2776 | 9.58  | 3.64 | 2776 | 12.37 | 3.38 |
| <i>Cognitivas parentales</i> |      |       |      |      |       |      |
| PUNTAJEP                     | 2642 | 5.25  | 2.89 | 1426 | 6.03  | 2.85 |
| <i>Inversión parental</i>    |      |       |      |      |       |      |
| CONSUMO                      | 2776 | 8.04  | 0.77 | 2774 | 8.09  | 0.78 |

PUNTAJEP corresponde al puntaje de la madre.

**Cuadro 3.** Estadística descriptiva. Indicadoras parentales.

|                    | Obs. | Media | DE   |
|--------------------|------|-------|------|
| <i>Continua</i>    |      |       |      |
| $R^P$              | 2707 | 5.48  | 2.64 |
|                    | Obs. | 0     | 1    |
| <i>Dicotómicas</i> |      |       |      |
| LEYO               | 2707 | 0.64  | 0.36 |
| INTERNET           | 2707 | 0.97  | 0.03 |
| RECADO             | 2707 | 0.10  | 0.90 |

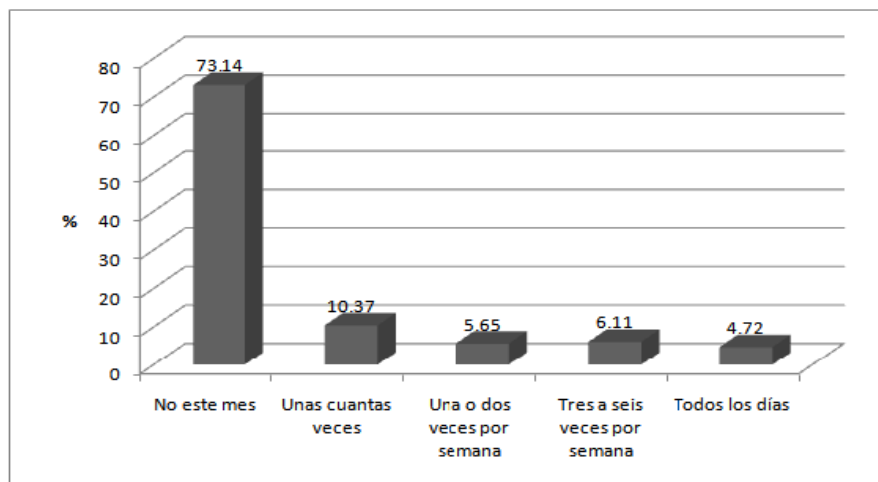
*Inversión parental, I.* Como indicadoras de la inversión parental se usan para ambos periodos las primeras cuatro variables; por disponibilidad de información las tres finales son utilizadas sólo en el último periodo.<sup>21</sup>

- CONSUMO. Logaritmo natural del consumo corriente del hogar del niño a precios de junio de 2002.
- CULTURA. Variable dicotómica que vale uno si hubo gasto en cultura y recreación en el último mes.

<sup>20</sup> Se retiran 45 observaciones en las que la madre tiene cero en la variable RECADO en la ENNVIIH -1 y uno en la ENNVIIH -2.

<sup>21</sup> Los resultados son robustos a una modificación en el tratamiento de  $I$  en la que se utilizan las mismas variables indicadoras para  $I_{t+1}$  e  $I_t$ .

- ACULTURA. Variable dicotómica que vale uno si el niño participó en actividades deportivas, culturales o de entretenimiento fuera de su vivienda durante la semana pasada.
- CLASES. Variable dicotómica que vale uno si el niño fue a algún centro educativo, por ejemplo clases particulares, sin incluir escolaridad formal durante la semana pasada.
- FLECTURA. Variable de cinco categorías creciente en la frecuencia con la que la madre que responde lee libros a sus hijos (ver figura 3).
- GRITARIA y PEGARIA. Variables dicotómicas que valen uno, respectivamente, si ante un berrinche o grito de parte del hijo, la madre que responde estaría dispuesta a gritar o golpear a su hijo.



**Figura 3:** Categorías de FLECTURA

## 6. Resultados

### 6.1 Formación de habilidades cognitivas

Los resultados del modelo que trata la habilidad cognitiva como variable latente se presentan en el cuadro 4.<sup>22</sup>

La variable  $\theta^F$  es la habilidad cognitiva de la madre como variable latente y  $R^F$  es el puntaje de la madre<sup>23</sup> de

<sup>22</sup>Este cuadro es una selección de las especificaciones presentes en el cuadro 7 del anexo C.

<sup>23</sup> Altamirano y Soloaga (2009) muestran, con datos de la ENNVIIH-2, que las características de la madre son las más relevantes en el desarrollo de logros cognitivos en los niños. Aún así, todos los modelos fueron repetidos utilizando el máximo de las variables

la prueba de Raven. Las variables se suponen invariantes en el tiempo dado el corto periodo que se considera y teniendo en cuenta que el proceso de formación de habilidades cognitivas de los padres ya ha pasado su etapa final  $T$  en (1). Intuitivamente el supuesto implica que las ganancias (o pérdidas) de habilidad entre un periodo y otro son prácticamente nulas para los padres.

En los modelos principales se utiliza la inversión contemporánea. En el cuadro 7 columnas (3) y (4) se observa que la inclusión de la inversión rezagada, en una ecuación que incluye la inversión contemporánea, sólo disminuye un poco la magnitud del efecto de esta última sin que la primera se torne significativa. El resultado sugiere que el efecto de la inversión parental no tiene retraso; nótese que, en el contexto del modelo, afirmar que la inversión tiene sólo efecto contemporáneo es equivalente a decir que lo relevante es lo invertido entre  $t$  y  $t + 1$ .

**Cuadro 4. Habilidad cognitiva como variable latente**

|                  | Variable latente dependiente $\theta_{t+1}$ |                   |                  |                    |                    |                   |       |
|------------------|---|-------------------|------------------|--------------------|--------------------|-------------------|-------|
|                  | (1)   | (2)               | (3)              | (4)                | (5)                | (6)               | (6')  |
| $I_{t+1}$        | 0.804<br>(15.843)                           |                   | 0.041<br>(0.186) |                    | 0.140<br>(0.733)   | 0.406<br>(1.077)  | 0.144 |
| $\theta_t$       |   | 1.093<br>(11.152) | 1.054<br>(4.532) | 1.360<br>(2.657)   | 1.355<br>(2.313)   | 1.084<br>(3.581)  | 0.778 |
| $\theta^P$       |   |                   |                  | -0.347<br>(-0.688) | -0.442<br>(-0.915) |                   |       |
| $R^P$            |   |                   |                  |                    |                    | 0.289<br>(13.293) | 0.525 |
| Observaciones    | 2776  | 2776              | 2776             | 2776               | 2776               | 2707              |       |
| p-valor $\chi^2$ | 0.000                                       | 0.000             | 0.000            | 0.000              | 0.000              | 0.000             | 0.000 |
| CFI              | 0.846                                       | 0.883             | 0.853            | 0.892              | 0.848              | 0.543             |       |
| TLI              | 0.821                                       | 0.810             | 0.851            | 0.889              | 0.873              | 0.499             |       |
| RMSEA            | 0.059                                       | 0.049             | 0.047            | 0.048              | 0.048              | 0.083             |       |

Valores z en paréntesis.

En las columnas (1) y (2) del cuadro 4 se observan los resultados de los coeficientes estandarizados<sup>24</sup> de la ecuación estructural (6)<sup>25</sup> (denotados con la letra  $\gamma$ ), incluyendo una sola de las variables latentes, inversión o

indicadoras para ambos padres como indicadoras de la habilidad cognitiva parental y de la inversión. Los resultados son robustos a este cambio.

<sup>24</sup> Estos coeficientes indican el cambio en la variable dependiente por un cambio de una desviación estándar de la variable independiente, medido en unidades de desviación estándar de la variable dependiente. Al igual que las elasticidades, tan conocidas en la ciencia económica, se utilizan para hacer comparaciones entre efectos de variables evitando el problema de la escala. Ver Muthén y Muthén (2007).

rezago de habilidad. Cada una de las variables por separado es relevante en la formación de habilidades. Sin embargo, cuando se les considera juntas (columna (3)) el rezago de habilidad mantiene su magnitud y significancia y la inversión pierde su significancia estadística. Algo similar ocurre al incluir la habilidad cognitiva parental como variable latente (columna (4)). Cuando las tres variables son consideradas juntas (columna (5)), únicamente el rezago de habilidad sigue siendo estadísticamente significativo. El hecho de que las variables por separado sean relevantes, pero juntas pierdan significancia, e incluso cambien de signo, sugiere que los resultados, posiblemente, estén evidenciando un problema de colinealidad entre las variables latentes, que puede surgir por el componente socioeconómico de las variables indicadoras LEYO e INTERNET, común a algunas de las variables indicadoras de la inversión, fundamentalmente a la variable CONSUMO.

Al tomar estos resultados a primera vista, la conclusión produce consternación: la historia persiste a pesar de la inversión. Lo importante en la formación de habilidades cognitivas a partir de los 5 años es el nivel del cual se parte. Sólo hay evidencia de autofomento en la producción de habilidades cognitivas. El resultado es compatible con la existencia de periodos críticos y sensibles para la inversión en edades menores a cinco años, resultado presente en la literatura (ver Arias *et al.* (2010) y los trabajos que citan). Ahora, es cierto que el resultado es radical en la medida en que sugiere que la inversión, y aun la habilidad cognitiva parental, *no importan* a partir de los 5 años. Resultado que puede ser provocado por colinealidad entre las variables latentes utilizadas,<sup>26</sup> y, por lo tanto, puede ser visto como un hallazgo que resalta la importancia del autofomento de habilidad, sin pretender descartar definitivamente un posible efecto de la inversión y la habilidad parental; que podría ser estudiado con un conjunto más rico de datos y relajando algunas de las simplificaciones hechas en este trabajo.

En la columna (6) se presentan los resultados abandonando la noción de habilidad cognitiva parental como variable latente y controlando sólo por el promedio del puntaje de la prueba de Raven de la madre  $R^P$ . Los

---

<sup>25</sup> Los coeficientes de las ecuaciones de medición (denotados con la letra  $\lambda$  en las ecuaciones (7) y (8)) son omitidos. Estos tienen siempre el signo esperado y son estadísticamente significativos. En el anexo D se presentan estos coeficientes para los modelos preferidos en el trabajo.

<sup>26</sup> Dada la especificación en (5) el rezago de habilidad del niño incluye un efecto de la habilidad parental (que se ha supuesto constante durante el periodo considerado).

coeficientes no están estandarizados. Los coeficientes estandarizados se encuentran en la columna (6').<sup>27</sup> Los resultados sugieren que el puntaje de la prueba de Raven de la madre sí juega un papel relevante en la producción de habilidades cognitivas de los niños. El hallazgo puede parecer paradójico dados los resultados de la estimación en la columna (5); sin embargo, el incluir el puntaje de la prueba de Raven de la madre mientras se trata la habilidad de los niños como variable latente puede estar evitando el problema de colinealidad y, por esto, resultar significativo y con el signo esperado.

Por último, el ajuste del modelo<sup>28</sup> no es el mejor. En efecto, la hipótesis nula de que la matriz de varianzas y covarianzas del modelo es igual a la poblacional se rechaza, los p-valores de los modelos son siempre menores a 0.05. No obstante, dada la simplificación lineal del modelo que se estima y los problemas enfrentados en la estimación, el resultado no es sorprendente. El ajuste medido por los índices incrementales CFI y TLI es mediocre, los valores están cerca de los aceptados comúnmente en la literatura, pero aún no suficientes para argumentar un buen ajuste. Finalmente, el RMSEA indica que el ajuste es bueno.

## 6.2 Formación de logros cognitivos

Se ha dicho que analizar el problema utilizando como variable indicadora sólo al puntaje en una prueba puede no estar integrando toda la información disponible sobre la formación de la habilidad cognitiva de los niños. Sin embargo, dadas las posibles restricciones que se enfrentan al estudiar la variable latente de habilidad cognitiva considerando las indicadores dicotómicas disponibles, otra aproximación al proceso se obtiene utilizando un procedimiento más tradicional en economía: hacer uso de los logros cognitivos como puntajes en una prueba específica como la de Raven y estudiar su formación. Los resultados de tal aproximación al problema de estudio se presentan en el cuadro 5.

Los coeficientes de las columnas (3) y (5) no son estandarizados. Sus contrapartes estandarizadas se encuentran en las columnas (3') y (5'). La inversión y los logros cognitivos parentales mantienen su importancia

---

<sup>27</sup> Desafortunadamente el *software Mplus* no provee los errores estándar de los coeficientes estandarizados en modelos que consideran del lado derecho de la ecuación estructural combinaciones de variables latentes y observables. Esta es la razón por la cual no son presentados dichos errores.

<sup>28</sup> Ver Bagozzi y Yi (1988), Yu (2002) y Hooper, Coughlan y Mullen (2008) para una explicación y discusión de las medidas de ajuste y Barret (2007) para una visión crítica de las medidas de ajuste y los umbrales de decisión utilizados en la literatura de modelos estructurales con variables latentes.

en la producción de logros cognitivos de los niños aun cuando se les incluye juntos en un mismo modelo con el rezago del puntaje del niño. En efecto, si tomamos a primera vista los resultados estandarizados en la columna (5'), el orden de magnitud de los efectos estandarizados de la inversión parental, el logro rezagado y el logro parental, es similar.

Estos resultados llevan a una pregunta obligada: ¿qué puede estar motivando la diferencia de resultados al estudiar la formación de habilidades y la de logros cognitivos? La respuesta podría encontrarse en los problemas de estimación mencionados y en las dificultades que se enfrentan con las variables indicadoras dicotómicas de habilidad. Para poner a prueba las posibles explicaciones se requeriría de un conjunto más rico de datos (resultados en otro tipo de pruebas) que, desafortunadamente, no se encuentra disponible.

**Cuadro 5.** Formación de logros cognitivos

|                              | Variable dependiente $R_{t+1}$ |                   |                   |       |                   |                   |       |
|------------------------------|--------------------------------|-------------------|-------------------|-------|-------------------|-------------------|-------|
|                              | (1)                            | (2)               | (3)               | (3')  | (4)               | (5)               | (5')  |
| $I_{t+1}$                    | 0.305<br>(13.856)              |                   | 1.518<br>(9.142)  | 0.241 |                   | 1.119<br>(6.511)  | 0.169 |
| $R_t$                        |                                | 0.261<br>(14.786) | 0.243<br>(14.335) | 0.262 | 0.189<br>(10.124) | 0.176<br>(10.285) | 0.190 |
| $R^P$                        |                                |                   |                   |       | 0.240<br>(12.950) | 0.307<br>(12.477) | 0.240 |
| Observaciones                | 2776                           | 2776              | 2776              |       | 2707              | 2707              |       |
| p-valor $\chi^2 \setminus F$ | 0.000                          | 0.000             | 0.000             |       | 0.000             | 0.000             |       |
| CFI $\setminus R^2$          | 0.858                          | 0.068             | 0.760             |       | 0.123             | 0.658             |       |
| TLI                          | 0.826                          |                   | 0.700             |       |                   | 0.564             |       |
| RMSEA                        | 0.064                          |                   | 0.077             |       |                   | 0.084             |       |

Valores z en paréntesis.

Otra pregunta interesante sobre el modelo es si la función de producción de logros cognitivos<sup>29</sup> cambia entre rangos de edad en la niñez. Para estudiar la posibilidad se utiliza el modelo que incluye  $R_{t+1}$  rezagado

<sup>29</sup> Para el análisis por rangos de edad, el trabajo se restringe al modelo de formación de logros cognitivos pues es el contexto en el cual la inversión resulta ser importante y, además, al hacer la estimación por grupos, tratando la habilidad cognitiva como variable latente, los resultados por grupos no son muy confiables: los coeficientes de las variables latentes no son significativos e incluso varios son



sobre  $R_t$ ,  $I_{t+1}$  y  $R^P$  (columna (5) cuadro 5). La muestra es dividida en dos grupos de edad, uno formado por los niños con edades (en 2002) en el intervalo [5, 7] y otro por los niños con edades en el intervalo [8, 9].<sup>30</sup> La elección de rangos de edad se hizo para agrupar las edades más pequeñas evitando concentraciones de las variables dicotómicas en el valor cero; las variables dicotómicas LEYO e INTERNET tienen concentraciones altas en el valor cero para los niños más pequeños, lo que no es sorprendente dada la definición de dichas variables. Por otra parte, también se buscó tener cierta similitud con los rangos de dos años utilizados por Cunha y Heckman (2008).

Para estudiar la formación de logros cognitivos por periodos se hace una prueba  $\chi^2$  de diferencias en la que se enfrenta el modelo no restringido, en donde todos los coeficientes de las ecuaciones estructurales y de medición pueden variar entre grupos de edad, contra el modelo restringido en el que los coeficientes no estandarizados son iguales entre grupos de edad. Los resultados se exponen en el cuadro 6.<sup>31</sup>

En dicho cuadro se presentan los coeficientes no estandarizados (Coef.) y estandarizados (Coef. Est.) de ambos modelos. Es de notar que los coeficientes no estandarizados del modelo restringido no son iguales a los que se observan en la columna (5) del cuadro 5, pues en el modelo del cuadro 6 la restricción es sólo sobre el valor de los coeficientes no estandarizados entre grupos. Razón por la que, los coeficientes estandarizados, siguen siendo ligeramente diferentes entre grupos en el modelo restringido. En el cuadro 5 la restricción es sobre todos los parámetros del modelo.

#### **Cuadro 6. Formación de logros cognitivos por rangos de edad**

---

negativos. El resultado puede ser explicado por la calidad de las variables indicadoras. Al dividir la muestra la ausencia de suficiente variabilidad se exagera.

<sup>30</sup> El intervalo [8,9] integra también observaciones de 53 niños (4.9% de las observaciones en el intervalo) que reportaron tener 10 años en la ENNVIIH-1 y 12 años en la ENNVIIH-2. Las observaciones se utilizan porque no sufren el problema de la incompatibilidad de resultados en la prueba de Raven luego del umbral de los 12 años; recuérdese que el conjunto de preguntas de la prueba de Raven presente en la ENNVIIH es diferente para niños de más de 12 años.

<sup>31</sup> Se omiten los resultados de los coeficientes de las variables indicadoras.

Variable dependiente  $R_{t+1}$ 

|   | Modelo no restringido* |            |                  |            | Modelo restringido** |            |            |
|---|------------------------|------------|------------------|------------|----------------------|------------|------------|
|   | 5-7                    |            | 8-9              |            | 5-7                  |            | 8-9        |
|   | Coef.                  | Coef. Est. | Coef.            | Coef. Est. | Coef.                | Coef. Est. | Coef. Est. |
| $I_{t+1}$                                 | 1.047<br>(4.743)       | 0.156      | 1.117<br>(4.578) | 0.183      | 1.089<br>(6.614)     | 0.166      | 0.172      |
| $R_t$                                     | 0.102<br>(4.581)       | 0.110      | 0.246<br>(8.709) | 0.258      | 0.157<br>(8.980)     | 0.170      | 0.166      |
| $R^P$                                     | 0.361<br>(11.096)      | 0.283      | 0.250<br>(6.668) | 0.200      | 0.313<br>(12.750)    | 0.245      | 0.253      |
| Observaciones                             | 1627                   |            | 1080             |            | 2707                 |            |            |
| $\chi^2$                                  |                        |            | 603.176          |            | 591.696              |            |            |
| Grados de libertad                        |                        |            | 58               |            | 63                   |            |            |
| p-valor $\chi^2$                          |                        |            | 0.000            |            | 0.000                |            |            |
| p-valor $\chi^2$<br>prueba de diferencias | 0.001                  |            |                  |            |                      |            |            |

\*Coeficientes varían en intervalos de edad

\*\*Restringido a que los coeficientes no estandarizados sean iguales entre grupos.

Los resultados de la prueba indican que se rechaza el modelo restringido, es decir, los datos sugieren que la función de producción es diferente entre grupos de edad. El efecto de la inversión es un poco mayor para los niños con edades mayores. No obstante, la diferencia de efectos con respecto a la inversión no es muy grande. El efecto de autofomento es más del doble para los niños en el grupo de mayores edades. El efecto del componente parental es poco menos de 50% mayor para los niños más pequeños. Los resultados difieren de los obtenidos por Cunha y Heckman (2008), quienes encuentran evidencia de periodos sensibles para la inversión en edades tempranas; en su estudio<sup>32</sup> los efectos de la inversión, la variable rezagada y el puntaje materno son mayores para los niños más pequeños.

## 7. Conclusiones

En este trabajo se desarrolló una aplicación empírica de los modelos de ecuaciones estructurales con variables latentes para estudiar la formación de habilidades y logros cognitivos en los niños mexicanos entre los 5 y 12 años. Se utilizó un modelo lineal (LISREL) en donde las variables de habilidad (de los niños y parental) y de inversión parental son consideradas como variables latentes escalares, que se manifiestan a través de otras variables indicadoras que sí pueden ser observadas. Se encuentra evidencia de lo que ha sido llamado *la*

<sup>32</sup> En este trabajo los autores utilizan una variable latente de habilidad cognitiva y no puntajes, por lo que los resultados no son directamente comparables.

*persistencia de la historia*, a partir de los cinco años. Lo que importa es el autofomento de la habilidad. Una vez que se incluye el rezago de la habilidad, la inversión y habilidad parentales pierden su poder explicativo en la formación de la habilidad cognitiva. Dados los problemas mencionados en nuestro análisis, el resultado puede entenderse como un indicador de la preponderancia de la inversión previa a los cinco años, más que como un argumento a favor de una posible productividad nula de la inversión a partir de dicha edad.

En la segunda parte del estudio se abandona el tratamiento de la habilidad cognitiva como variable latente y se usan los puntajes de la prueba de Raven para estudiar la formación de logros cognitivos en una estimación más tradicional, aunque introduciendo un elemento novedoso al seguir considerando la inversión como variable latente. Aquí, se encuentra que la inversión y el logro cognitivo parental sí son importantes en la formación de logros cognitivos de los niños aun cuando se añade el rezago del puntaje. Esta diferencia en los resultados de ambos modelos (habilidades cognitivas y logros cognitivos) puede deberse al componente socioeconómico que comparten las indicadores dicotómicas de habilidad con algunas de las indicadores de inversión, en especial con el consumo.

En el marco de la formación de logros cognitivos se pone a prueba la hipótesis de varianza temporal de la función de producción y se rechaza la invariancia temporal de la misma. Se encuentra que el efecto de la inversión parental es un poco mayor en los niños de más edad. El efecto de autofomento en los niños mayores es de más del doble del efecto encontrado para los niños más pequeños. El efecto del componente de puntaje materno es mayor para los niños en el grupo de menores edades.

Dada la ausencia de más datos sobre habilidad cognitiva (por ejemplo, puntajes en otras pruebas) no puede evaluarse enteramente el origen de los resultados disímiles sobre la inversión y el componente de habilidad parental encontrados en el estudio, según se use el modelo de variables latentes o el modelo de producción de logros cognitivos. Con todo, los resultados indican direcciones para seguir investigando cuando los datos de la tercera ronda de la ENNVIH estén disponibles.

## **Bibliografía**

- Altamirano, Aniel e I. Soloaga. 2009. Trasmisión intergeneracional de la desigualdad en habilidades cognitivas (mimeo).
- Arias, Javier, O. Azuara, P. Bernal, J. J. Heckman y C. Villarreal. 2010. Policies to Promote Growth and Economic Efficiency in Mexico, IZA, Discussion Paper, núm. 4740.
- Bagozzi, Richard y Youjae Yi. 1988. On the Evaluation of Structural Equation Models, *Journal of the Academy of Marketing Science*, 16(1): 74-94.
- Barrett, Paul. 2007. Structural Equation Modeling: Adjudging Model Fit, *Personality and Individual Differences*, 42(5): 815-824.
- Bollen, Kenneth. 1989. *Structural Equations with Latent Variables*, Wiley Series in Probability and Mathematical Statistics.
- Bureau of Labor Statistics. 2010. [en línea] s/f [consultado el 30 de abril de 2010], Estados Unidos. Disponible en <http://www.bls.gov/nls/nlsy79ch.htm>
- Court, J.H., J.C. Raven y J. Raven. 2006. *Test de matrices progresivas. Escalas coloreada, general y avanzada. Manual*, Paidós, cuarta reimpresión.
- Cunha, Flavio y J.J. Heckman. 2007. The Technology of Skill Formation, *American Economic Review*, 97(2): 31-47.
- . 2008. Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation, *Journal of Human Resources*, 43(4): 738-782.
- . 2009. The Economics and Psychology of Inequality and Human Development, UCD Geary Institute discussion paper núm. 200905.
- , L. Lochner y D. Masterov. 2005. Interpreting the Evidence on Life Cycle Skill Formation, Iza, Discussion Paper, núm. 1675.
- Cunha, Flavio, J. J. Heckman, y S. Schennach. 2010. Estimating the Technology of Cognitive and Noncognitive Skill Formation, NBER working paper, núm. 15664.

- Duncan, Otis D. y A. Goldberger. 1973. *Structural Equation Models in the Social Sciences*, Seminar Press.
- Everitt, Brian. 1984. *An Introduction to Latent Variable Models. Monographs on Statistics and Applied Probability*, Chapman and Hall.
- Farkas, George. 2003. Cognitive Skills and Noncognitive Traits and Behaviors in Stratification Processes, *Annual Review of Sociology*, 29: 541-562.
- Heckman, James J. y J. Rubinstein. 2001. The Importance of Noncognitive Skills: Lessons from the GED Testing Program, *American Economic Review*, 91(2): 145-149.
- Heckman, James J., J. Stixrud y S. Urzua. 2005. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior, NBER Working Paper, núm. 12006.
- Hooper, D., J. Coughlan y M.R. Mullen. 2008. Structural Equation Modeling: Guidelines for Determining Model Fit, *The Electronic Journal of Business Research Methods*, 6(1): 53-60, [en línea] disponible en: [www.ejbrm.com](http://www.ejbrm.com).
- Jöreskog, Karl. 2000. Latent Variable Scores and their Uses, (mimeo).
- y A. S. Goldberger. 1975. Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable, *Journal of the American Statistical Association*, 70(351): 631-639.
- López, María Fernanda, D. Mayer Foulkes y E. E. Serván Mori. 2008. Habilidades cognitivas: transmisión intergeneracional por niveles socioeconómicos, *Estudios Económicos*, 23(1): 129-156.
- Mayer Foulkes, David y E. E. Serván Mori. 2009. Formación de la capacidad cognitiva en México: impactos económicos y de políticas públicas, *Estudios Económicos*, número extraordinario, 83-122.
- Muthén, L.K. y B. Muthén. 2007. *Mplus User's Guide*. 5a. ed., Los Angeles, Muthén y Muthén.
- Muthén, B., S. du Toit y D. Spisic. 1997. Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes, (mimeo).

- Oliva, Paulina, L. Rubalcava y G. Teruel. 2004. *Schooling. Cognitive Ability or Emotional Well Being: What Drives the Individual's Perception of Health Outcomes?* Universidad Iberoamericana.
- Postlewaite, Andrew y D. Silverman. 2006. Non-Cognitive Skills, Social Success, and Labor Market Outcomes, (mimeo).
- Pearson Education, Inc. 2010. [en línea] s/f [consultado el 10 de mayo de 2010], Estados Unidos. Disponible en <http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=PAa29060&Mode=summary>
- Raven, J.C. 2007. *Test de matrices progresivas. Escala coloreada. Cuaderno de matrices/Series A, A<sub>B</sub> y B*, Paidós, cuarta reimpresión.
- Roemer, John. 2005. Equality of Opportunity. For the New Palgrave Dictionary.
- Rubalcava, Luis y G. Teruel. 2004. The Role of Maternal Cognitive Ability on Child Health, Inter-American Development Bank, Research Network Working Paper, núm R-497.
- . 2006. *Guía de usuario de la encuesta nacional sobre niveles de vida de los hogares, primera ronda*, [en línea] disponible en: [www.ennvih-mxfls.org](http://www.ennvih-mxfls.org).
- . 2008. *Guía de usuario de la encuesta nacional sobre niveles de vida de los hogares, segunda ronda*, [en línea] disponible en: [www.ennvih-mxfls.org](http://www.ennvih-mxfls.org).
- Todd, Petra y K. Wolpin. 2003. On the Specification and Estimation of the Production Function for Cognitive Achievement, *The Economic Journal*, 113(485): F3-F33.
- . 2004. The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps, Penn Institute for Economic Research, Working Paper, núm. 04-019.
- Yu, Ching-Yun. 2002. Evaluating Cutoff Criteria of Model Fit Indices for Latent Variable Models with Binary and Continuous Outcomes, Tesis doctoral, Universidad de California, L.A.

## A. Un caso particular del modelo LISREL

Al omitir los subíndices de tiempo por simplicidad las ecuaciones que describen el modelo, (6), (7) y (8) son:

$$\theta^c = \Gamma' \theta + \zeta \quad (12)$$

$$Y = \Lambda_y \theta^c + \epsilon \quad (13)$$

$$X = \Lambda_x \theta + \delta \quad (14)$$

Dado que el vector de datos ha sido definido como  $Z = [Y', X']$  se tiene que:

$$\text{Cov}(Z) = \Sigma = \begin{pmatrix} \Sigma_{yy} & \Sigma_{yn} \\ \Sigma_{yn}' & \Sigma_{nn} \end{pmatrix} \quad (15)$$

En este trabajo,  $\theta^c$  es un escalar y  $\theta$  es un vector de 2x1 que integra la inversión parental y el rezago de la habilidad cognitiva. Al suponer 3 indicadores para cada una de las variables latentes se tiene que  $Y$  y  $X$  son vectores de 3x1 y de 6x1, respectivamente. Por lo tanto  $\Gamma$ ,  $\Lambda_y$  y  $\Lambda_x$  son vectores (matrices) de parámetros de dimensiones 2x1, 3x1 y 6x2, respectivamente.

El último ingrediente para la construcción del ejemplo del modelo LISREL son los supuestos sobre las medias, varianzas y covarianzas y están dados por:

$$\begin{aligned} E[\zeta] &= 0; & \text{Var}[\zeta] &= \psi \\ E[\epsilon] &= 0; & \text{Var}[\epsilon] &= \Theta_\epsilon \\ E[\delta] &= 0; & \text{Var}[\delta] &= \Theta_\delta \end{aligned} \quad (16)$$

En el modelo,  $\zeta$ ,  $\epsilon$  y  $\delta$  están mutuamente no correlacionadas, los valores esperados de las variables latentes son ceros,  $\zeta$ , no está correlacionada con  $\theta$ , los valores esperados de los errores son igualmente cero,  $\epsilon$  no está correlacionado con  $\theta^c$ , ni con  $\zeta$  ni con  $\delta$ ;  $\delta$  no está correlacionado con  $\zeta$ , ni con  $\theta^c$  ni con  $\epsilon$ , y  $E[\theta_t' \theta_t] = \Phi$ .

Con estos componentes pueden calcularse las matrices que conforman la matriz  $\Sigma$ , como se muestra a continuación:

$$\begin{aligned}
 \Sigma_{yy} &= E[(\Lambda_y \theta^C + \epsilon)(\Lambda_y \theta^C + \epsilon)'] \\
 &= \Lambda_y E[\theta^C \theta^C] \Lambda_y' + \Lambda_y E[\theta^C \epsilon'] + E[\epsilon \theta^C] \Lambda_y' + E[\epsilon \epsilon'] \\
 &= \Lambda_y E[(\Gamma' \theta + \zeta)(\Gamma' \theta + \zeta)'] \Lambda_y' + \Theta_\epsilon \\
 &= \Lambda_y [\Gamma' E(\theta \theta') \Gamma + E(\zeta \zeta')] \Lambda_y' + \Theta_\epsilon \\
 &= \Lambda_y [\Gamma' \Phi \Gamma + \Psi] \Lambda_y' + \Theta_\epsilon
 \end{aligned} \tag{17}$$

$$\begin{aligned}
 \Sigma_{xy} &= E[(\Lambda_y \theta^C + \epsilon)(\Lambda_x \theta + \delta)'] \\
 &= \Lambda_y E[\theta^C \theta'] \Lambda_x' + \Lambda_y E[\theta^C \delta'] + E[\epsilon \theta'] \Lambda_x' + E[\epsilon \delta'] \\
 &= \Lambda_y E[(\Gamma' \theta + \zeta) \theta'] \Lambda_x' \\
 &= \Lambda_y \Gamma' E(\theta \theta') \Lambda_x' \\
 &= \Lambda_y \Gamma' \Phi \Lambda_x'
 \end{aligned} \tag{18}$$

$$\begin{aligned}
 \Sigma_{xx} &= E[(\Lambda_x \theta + \delta)(\Lambda_x \theta + \delta)'] \\
 &= \Lambda_x E[\theta \theta'] \Lambda_x' + \Lambda_x E[\theta \delta'] + E[\delta \theta'] \Lambda_x' + E[\delta \delta'] \\
 &= \Lambda_x \Phi \Lambda_x' + \Theta_\delta
 \end{aligned} \tag{19}$$

Si reemplazamos (17), (18) y (19) en (15) se obtiene

$$\text{Cov}(\mathbf{Z}) = \Sigma = \begin{pmatrix} \Lambda_y [\Gamma' \Phi \Gamma + \Psi] \Lambda_y' + \Theta_\epsilon & \Lambda_y \Gamma' \Phi \Lambda_x' \\ (\Lambda_y \Gamma' \Phi \Lambda_x')' & \Lambda_x \Phi \Lambda_x' + \Theta_\delta \end{pmatrix} \tag{20}$$

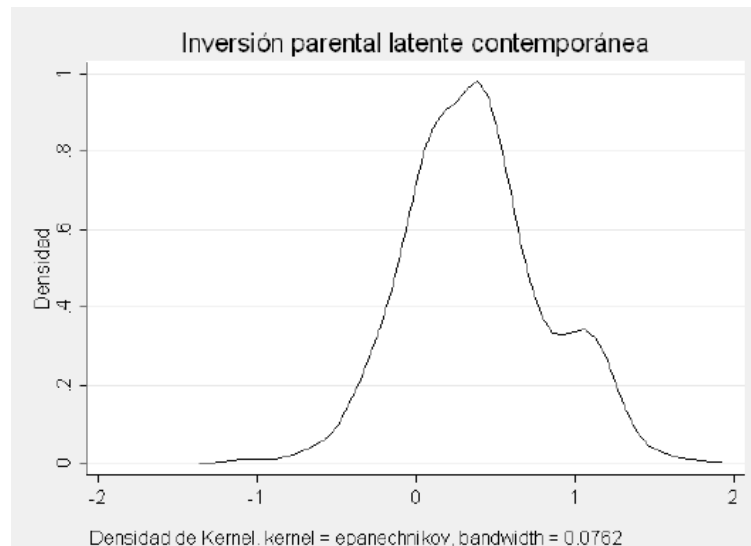
Como se observa de (20),  $\Sigma$  queda expresada **únicamente** en función de los parámetros del modelo.

Para la estimación de los parámetros del modelo se establece  $\hat{\Sigma} = \mathbf{S}$  (donde  $\hat{\Sigma}$  es la matriz estimada y  $\mathbf{S}$  es la matriz de varianzas-covarianzas muestral) y se resuelven las ecuaciones resultantes utilizando rutinas iterativas que consideran una función criterio para reducir la discrepancia entre  $\hat{\Sigma}$  y  $\mathbf{S}$ .



## B. Distribución de la inversión latente en el modelo de puntajes de la prueba de Raven

A continuación se presenta la distribución Kernel de la inversión parental latente en el modelo de puntajes de la prueba de Raven. En este modelo su efecto es significativo en la formación de logros cognitivos. La distribución corresponde a la variable  $I_{t+1}$  generada con el modelo en el cuadro 10.



**Figura 4:** Distribución Kernel de  $I_{t+1}$

La media y desviación estándar de  $I_{t+1}$  calculadas con las 2707 observaciones consideradas son, respectivamente, 0.378 y 0.442. Estos valores implican un coeficiente de variación de la inversión de 1.169, lo que indica la presencia de buena variabilidad en la inversión parental.

C. Modelos extendidos

**Cuadro 7.** Formación de habilidades cognitivas. Extendido  
Variable latente dependiente  $\theta_{t+1}$

|                  | (1)               | (2)               | (3)              | (4)               | (4')   | (5)               | (5')  | (6)              | (7)               | (8)              | (9)                | (10)               | (11)              | (11') | (12)              | (12') |
|------------------|-------------------|-------------------|------------------|-------------------|--------|-------------------|-------|------------------|-------------------|------------------|--------------------|--------------------|-------------------|-------|-------------------|-------|
| $I_t$            | 0.684<br>(14.052) |                   | 0.130<br>(1.222) | 0.271<br>(1.213)  | 0.1115 |                   |       |                  |                   |                  |                    |                    |                   |       |                   |       |
| $I_{t+1}$        |                   | 0.804<br>(15.843) | 0.666<br>(6.038) | 1.175<br>(4.959)  | 0.504  | 1.670<br>(10.389) | 0.629 | 0.479<br>(5.841) |                   | 0.041<br>(0.186) |                    | 0.140<br>(0.733)   | 0.406<br>(1.077)  | 0.144 | 0.333<br>(0.724)  | 0.112 |
| $\theta_t$       |                   |                   |                  |                   |        |                   |       |                  | 1.093<br>(11.152) | 1.054<br>(4.532) | 1.36<br>(2.657)    | 1.355<br>(2.313)   | 1.084<br>(3.581)  | 0.778 | 1.033<br>(3.187)  | 0.885 |
| $\theta^P$       |                   |                   |                  |                   |        |                   |       | 0.454<br>(5.398) |                   |                  | -0.347<br>(-0.688) | -0.442<br>(-0.915) |                   |       |                   |       |
| $R^P$            |                   |                   |                  | 0.256<br>(11.945) | 0.511  | 0.265<br>(12.368) | 0.516 |                  |                   |                  |                    |                    | 0.289<br>(13.293) | 0.525 |                   |       |
| $E^P$            |                   |                   |                  |                   |        |                   |       |                  |                   |                  |                    |                    |                   |       | 0.184<br>(11.286) | 0.490 |
| Observaciones    | 2776              | 2776              | 2776             | 2707              | 2776   | 2707              | 2776  | 2776             | 2776              | 2776             | 2776               | 2776               | 2707              | 2776  | 2483              | 2776  |
| p-valor $\chi^2$ | 0.000             | 0.000             | 0.000            | 0.000             | 0.000  | 0.000             | 0.000 | 0.000            | 0.000             | 0.000            | 0.000              | 0.000              | 0.000             | 0.000 | 0.000             | 0.000 |
| CFI              | 0.980             | 0.846             | 0.910            | 0.704             | 0.600  | 0.600             | 0.850 | 0.883            | 0.883             | 0.853            | 0.892              | 0.848              | 0.543             | 0.892 | 0.360             | 0.360 |
| TLI              | 0.973             | 0.821             | 0.923            | 0.721             | 0.526  | 0.526             | 0.864 | 0.810            | 0.810             | 0.851            | 0.889              | 0.873              | 0.499             | 0.889 | 0.280             | 0.280 |
| RMSEA            | 0.035             | 0.059             | 0.047            | 0.081             | 0.091  | 0.091             | 0.054 | 0.049            | 0.049             | 0.047            | 0.048              | 0.048              | 0.083             | 0.048 | 0.097             | 0.097 |

Valores z en paréntesis.  $E^P$  es la media de la escolaridad de la madre.

**Cuadro 8.** Formación de logros cognitivos. Extendido  
Variable dependiente  $R_{t+1}$

|                              | (1)               | (2)               | (3)              | (4)               | (4')              | (5)               | (5')  | (6)               | (6')  | (7)               | (8)               | (8')  | (9)               | (10)              | (10') | (11)              | (11') |
|------------------------------|-------------------|-------------------|------------------|-------------------|-------------------|-------------------|-------|-------------------|-------|-------------------|-------------------|-------|-------------------|-------------------|-------|-------------------|-------|
| $I_t$                        | 0.299<br>(12.209) |                   | 0.128<br>(2.101) | 0.524<br>(1.540)  | 0.087             |                   |       |                   |       |                   |                   |       |                   |                   |       |                   |       |
| $I_{t+1}$                    |                   | 0.305<br>(13.856) | 0.191<br>(3.236) | 0.685<br>(2.148)  | 0.117             | 1.277<br>(7.359)  | 0.195 | 1.285<br>(5.562)  | 0.167 |                   | 1.518<br>(9.142)  | 0.241 |                   | 1.119<br>(6.511)  | 0.197 | 1.085<br>(4.816)  | 0.139 |
| $R_t$                        |                   |                   |                  |                   |                   |                   |       |                   |       | 0.261<br>(14.786) | 0.243<br>(14.335) | 0.262 | 0.189<br>(10.124) | 0.176<br>(10.285) | 0.212 | 0.199<br>(11.216) | 0.217 |
| $R^P$                        |                   |                   |                  | 0.385<br>(15.864) | 0.301<br>(15.864) | 0.385<br>(15.864) | 0.301 |                   |       |                   |                   |       | 0.240<br>(12.950) | 0.307<br>(12.477) | 0.177 |                   |       |
| $E^P$                        |                   |                   |                  |                   |                   |                   |       | 0.216<br>(11.472) | 0.227 |                   |                   |       |                   |                   |       | 0.168<br>(8.949)  | 0.176 |
| Observaciones                | 2776              | 2776              | 2776             | 2707              | 2707              | 2707              | 2707  | 2483              | 2483  | 2707              | 2776              | 2776  | 2707              | 2707              | 2707  | 2483              |       |
| p-valor $\chi^2 \setminus F$ | 0.000             | 0.000             | 0.000            | 0.000             | 0.000             | 0.000             | 0.000 | 0.000             | 0.000 | 0.000             | 0.000             | 0.000 | 0.000             | 0.000             | 0.000 | 0.000             |       |
| CFI/ $R^2$                   | 0.990             | 0.858             | 0.926            | 0.712             | 0.613             | 0.613             | 0.324 | 0.324             | 0.324 | 0.068             | 0.760             | 0.760 | 0.123             | 0.658             | 0.658 | 0.407             |       |
| TLI                          | 0.983             | 0.826             | 0.935            | 0.717             | 0.501             | 0.501             | 0.135 | 0.135             | 0.135 |                   | 0.700             | 0.700 |                   | 0.564             | 0.564 | 0.229             |       |
| RMSEA                        | 0.037             | 0.064             | 0.048            | 0.089             | 0.103             | 0.103             | 0.139 | 0.139             | 0.139 |                   | 0.077             | 0.077 |                   | 0.084             | 0.084 | 0.116             |       |

Valores  $z(t)$  en paréntesis.  $E^P$  es la media de la escolaridad de la madre.  
Regresiones sin variables latentes incluyen estadísticos t, prueba  $F$  y  $R^2$ .

## D. Ecuaciones de medición

**Cuadro 9.** Ecuaciones de medición. Formación de habilidades cognitivas

| Coeficientes estandarizados   |          |       |           |         |
|-------------------------------|----------|-------|-----------|---------|
|                               | Estimado | E.S.  | Est./E.S. | p-valor |
| <i>Ecuaciones de medición</i> |          |       |           |         |
| $\theta_t$                    |          |       |           |         |
| PUNTAJE                       | 0.425    | 0.033 | 12.966    | 0.000   |
| LEYO                          | 0.305    | 0.035 | 8.827     | 0.000   |
| INTERNET                      | 0.746    | 0.086 | 8.709     | 0.000   |
| $\theta_{t+1}$                |          |       |           |         |
| PUNTAJE                       | 0.436    | 0.030 | 14.741    | 0.000   |
| LEYO                          | 0.310    | 0.030 | 10.426    | 0.000   |
| INTERNET                      | 0.655    | 0.049 | 13.396    | 0.000   |
| $\theta^P$                    |          |       |           |         |
| PUNTAJE                       | 0.595    | 0.020 | 30.179    | 0.000   |
| LEYO                          | 0.610    | 0.025 | 24.804    | 0.000   |
| INTERNET                      | 0.784    | 0.047 | 16.688    | 0.000   |
| RECADO                        | 0.820    | 0.028 | 29.152    | 0.000   |
| $I_{t+1}$                     |          |       |           |         |
| CONSUMO                       | 0.691    | 0.018 | 38.455    | 0.000   |
| ACULTURA                      | 0.385    | 0.030 | 12.688    | 0.000   |
| CLASES                        | 0.447    | 0.047 | 9.595     | 0.000   |
| CULTURA                       | 0.777    | 0.025 | 31.435    | 0.000   |
| PEGARIA                       | -0.278   | 0.029 | -9.446    | 0.000   |
| GRITARIA                      | -0.148   | 0.034 | -4.326    | 0.000   |
| FLECTURA                      | 0.355    | 0.027 | 13.175    | 0.000   |
| <i>Ecuación estructural</i>   |          |       |           |         |
| $\theta_t$                    | 1.355    | 0.586 | 2.313     | 0.021   |
| $I_{t+1}$                     | 0.140    | 0.191 | 0.733     | 0.464   |
| $\theta^P$                    | -0.442   | 0.483 | -0.915    | 0.360   |

Modelo (10) en la Tabla 7

En los cuadros 9 y 10 los coeficientes estandarizados, en las columnas dos y seis respectivamente, pueden ser considerados como una medición del concepto de *validez* de las variables a través de las cuales se manifiestan las variables latentes. Bollen (1989) define este concepto como la magnitud de la relación estructural directa entre las variables latentes y sus variables de medición. Si recordamos que los coeficientes de las variables dicotómicas corresponden a modelos *probit*, se observa que la validez de las variables de medición de la habilidad cognitiva se mantiene, aproximadamente, igual entre periodos. La validez de LEYO como variable de medición de la habilidad cognitiva es mayor

para las madres que para sus hijos. En inversión parental el consumo tiene validez considerable. La variable CULTURA tiene la mayor validez entre las dicotómicas, lo cual no es sorprendente dada su definición.

Como ya se mencionó, todas las variables de medición resultan ser significativas y tener los signos esperados. Son especialmente reconfortantes los signos negativos de las variables PEGARIA y GOLPEARIA.

Los coeficientes de las ecuaciones estructurales en los cuadros 9 y 10 son, respectivamente, los coeficientes de la columna (5) en el cuadro 4 y los de las columnas (5) y (5') en el cuadro 5.

**Cuadro 10.** Ecuaciones de medición. Formación de logros cognitivos

Coefficientes no estandarizados y estandarizados

|                               | Estimado | E.S.  | Est./E.S. | p-valor | Coef. Est. |
|-------------------------------|----------|-------|-----------|---------|------------|
| <i>Ecuaciones de medición</i> |          |       |           |         |            |
| $I_{t+1}$                     |          |       |           |         |            |
| CONSUMO                       | 1.000    | 0.000 | -         | -       | 0.690      |
| ACULTURA                      | 0.601    | 0.077 | 7.839     | 0.000   | 0.307      |
| CLASES                        | 0.696    | 0.109 | 6.362     | 0.000   | 0.355      |
| CULTURA                       | 1.587    | 0.154 | 10.274    | 0.000   | 0.809      |
| PEGARIA                       | -0.326   | 0.069 | -4.687    | 0.000   | -0.166     |
| GRITARIA                      | -0.275   | 0.072 | -3.799    | 0.000   | -0.140     |
| FLECTURA                      | 0.410    | 0.068 | 6.029     | 0.000   | 0.209      |
| <i>Ecuación estructural</i>   |          |       |           |         |            |
| $I_{t+1}$                     | 1.119    | 0.172 | 6.511     | 0.000   | 0.169      |
| $R_t$                         | 0.176    | 0.017 | 10.285    | 0.000   | 0.190      |
| $R^P$                         | 0.307    | 0.025 | 12.477    | 0.000   | 0.240      |

Estimado: coeficientes no estandarizados escalados en CONSUMO

Coef. Est.: coeficientes estandarizados

Modelo (10) en la Tabla 8