



UNIVERSIDAD
TORCUATO DI TELLA

DEPARTAMENTO DE ECONOMIA

MAESTRÍA EN ECONOMÍA APLICADA

Evidencia Econométrica de Retornos a Habilidades Cognitivas y
No-Cognitivas: Diferencias en el Mercado Laboral
Latinoamericano

Alumno: Ignacio Martín Vilieri

Tutor: *PhD* Hernán Ruffo

Fecha: 1 de junio de 2017

En memoria de *Dante Nicolás di Risio*. Mi abuelo, mi campeón, mi mejor amigo, mi héroe.

Evidencia Econométrica de Retornos a Habilidades Cognitivas y No-Cognitivas: Diferencias en el Mercado Laboral Latinoamericano

Ignacio Martín Vilieri [†]

Abstract

El estudio de los retornos asociados a los años de educación, a las habilidades cognitivas y no-cognitivas ha sido un tema abordado en numerosos trabajos en países desarrollados. Sin embargo, existe poca evidencia de dichos estudios respecto a países en desarrollo. Mediante la utilización de la base de datos *STEP* se procede a estimar el efecto marginal en salarios de las habilidades no-cognitivas y cognitivas, controlando por educación y otras variables relevantes en el mercado laboral de Colombia y Bolivia, mediante métodos multivariados y econométricos. En ambos países encontramos evidencia estadística y económica respecto a retornos positivos de las habilidades cognitivas; mientras que hallamos una relación negativa respecto a los retornos salariales y la perseverancia al esfuerzo en Colombia y la estabilidad emocional en Bolivia. No obstante, estimamos que el retorno asociado a la propensión al riesgo por parte de los individuos resulta significativa y positivamente relacionada en Colombia, al contrario de Bolivia donde dicha relación resulta negativa. Nuestras estimaciones resultan robustas ante potenciales problemas de endogeneidad, variables omitidas, como también ante correcciones de *Heckman*. Finalmente estimamos el retorno asociado a la obtención del diploma correspondiente al nivel educativo más alto alcanzado, hallando resultados significativos y robustos en ambos países para los estudios de posgrado.

Palabras clave: retornos a habilidades cognitivas, habilidades no-cognitivas, aversión al riesgo, valor del diploma, corrección por Heckman

[†] El autor desea agradecer a todas aquellas personas que siempre lo han apoyado y motivado, en especial a su familia, a su novia, a sus compañeros de cursada y a la Universidad junto a sus profesores.

1. Introducción

Siguiendo la línea de trabajos realizados en otros países respecto a la estimación de retornos a la educación, las habilidades cognitivas y a las habilidades no cognitivas, desarrollamos un análisis similar al realizado en Perú para evidenciar diferencias y similitudes entre dichos retornos y evaluar en qué medida los mercados laborales latinoamericanos se asemejan entre sí. A pesar de ser un estudio realizado en gran parte de los países desarrollados, aún no es común realizar estos en países en desarrollo. Con lo cual nuestro objetivo es poder avanzar en pos de obtener mayor conocimiento en este campo, contribuir a dilucidar cuestiones relacionadas al ámbito laboral y motivar a realizar dichos análisis en otros países de la región.

Para el presente trabajo, nos enfocamos y tomamos como punto de partida el paper realizado en Perú: “*The Returns to Cognitive and Non-Cognitive Skills in Urban Peru*”. D. Vera Tudela, JJ Díaz & O. Arias (2012). Nos proponemos obtener estimaciones de retornos a las habilidades cognitivas y no-cognitivas mediante los modelos econométricos y multivariados propuestos en el paper de referencia, mediante bases *STEP* del *World Bank*. Buscaremos corroborar si los resultados obtenidos en Perú se mantienen para Colombia y Bolivia, evidenciando las ganancias que generan la persistencia de esfuerzo y la inteligencia en países en desarrollo, como así también las demás variables que resulten (o no) significativas.

En este sentido, utilizamos la base *STEP Skills Measurement* que resulta de un programa llevado a cabo por el Banco Mundial con el objetivo de medir las habilidades en países de bajos y medios ingresos. *Skills Toward Employment and Productivity*, o *STEP*, ofrece dos tipos de instrumentos para recopilar información mediante encuestas y medir tanto la demanda como la oferta de habilidades, ellas son: la encuesta doméstica, que mide el nivel de habilidades que los individuos ofertan; y la encuesta a empleadores, que mide las habilidades que las empresas demandan.² Estos instrumentos miden entonces habilidades cognitivas y habilidades socio-emocionales mediante una serie de módulos específicamente diseñados para obtener datos concretos en dichas dimensiones, así también como habilidades relevantes del ámbito laboral. Dentro de dicha estructura, cada instrumento posee características distintas que permite obtener información respecto a la estructura de la fuerza laboral, información individual y del hogar, el comportamiento y los rasgos de la personalidad, habilidades cognitivas y no cognitivas, retribuciones y compensaciones, información educacional y de salud, entre otros. La base utilizada para el presente paper comprende información recogida entre Marzo 2012 y Junio de 2014 en Colombia y Bolivia, siendo el público comprendido adultos de entre 15 y

² Para mayores datos remitirse a *World Bank* (2014).

64 años de áreas urbanas que se encuentren empleados o no. Tomamos entonces nuestra muestra de cada país, en este caso Colombia y Bolivia, y utilizamos a la población que se encuentra empleada actualmente en zonas urbanas y donde su edad se encuentre entre 15 y 50 años. Nuestro objetivo es estimar y analizar la relación que existe entre las habilidades cognitivas y socio-emocionales respecto al salario percibido en los distintos países seleccionados para poder extender los hallazgos encontrados en países desarrollados en relación a países de América Latina y poder contribuir a generar políticas de estado y laborales más acordes a cada región.

El trabajo prosigue de la siguiente manera. En la sección 2 abordaremos de forma breve la bibliografía relacionada a habilidades cognitivas y habilidades no cognitivas, y su relación con resultados en el ámbito académico y laboral. En la sección 3 presentaremos los datos y las variables a aplicar en nuestros modelos econométricos. En la sección 4 presentamos estimaciones preliminares respecto a la relación entre las habilidades cognitivas, socio-emocionales y los ingresos laborales asociados, primero incluyendo cada grupo de habilidades por separado y luego el set completo, controlando por años de educación y sin dicha variable. En la sección 5 tratamos dos potenciales problemas de nuestro análisis: endogeneidad de la variable años de educación con las habilidades cognitivas y no cognitivas; y sesgo de selección por parte de la recolección de datos *STEP*. En la sección 6 tratamos una serie de extensiones relacionadas a estimar el ingreso marginal de poseer el diploma del nivel más alto de estudio realizado; como también a estimar la relación entre los ingresos laborales y la aversión al riesgo como proxy a la propensión de los agentes a realizar transiciones empleo-empleo. La sección 7 culmina con las conclusiones obtenidas a lo largo del presente trabajo.

2. Literatura

En nuestro paper de referencia “*The Returns to Cognitive and Non-Cognitive Skills in Urban Peru*”, los autores presentan una sección dedicada a introducir al lector en los aspectos básicos psicológicos de las habilidades cognitivas y no cognitivas, como también sus etapas de formación y la relación entre dichas habilidades y resultados en el ámbito escolar, laboral y personal. Asumimos entonces que el lector se encuentra familiarizado en general respecto a dichos aspectos y centraremos nuestro análisis en los aspectos econométricos prácticos. Para mayores referencias respecto a estudios similares en países en desarrollo y en países desarrollados se recomienda *Lindqvist & Vestman (2011)*, *Heineck & Anger (2008)*, *Nyhus & Pons (2005)*, *Nordman, Sarr & Sharma (2015)*, *Brunello & Schlotter (2011)*, *Guerra, Modecki, & Cunningham (2014)*, *Lavado, Velarde & Yamada*

(2014), Ray Miller (2015), Osborne-Groves (2005), Cueto, Muñoz & Baertl (2010) y Acosta, Muller & Sarzosa (2015), Gronau (1974) y Lewis (1974), Heckman (1976) y Heckman (1979), World Bank (2016), Prada (2013), Hanushek, Schwerdt, Wiederhold, & Woessmann (2013); entre otros papers relacionados.

3. Data y Habilidades

La base STEP comprende una muestra de observaciones ($n = 2617$), tanto hombres como mujeres de zonas urbanas de distintos países en desarrollo, con información relacionada a cuestiones domésticas e información relacionada al desarrollo educacional, desempeño en el ámbito laboral, test de rasgos de personalidad, habilidades relacionadas al trabajo, datos económicos de cada individuo y su familia. La encuesta posee una serie de instrumentos apuntados a obtener datos de las habilidades cognitivas y no cognitivas, la cual es realizada a un individuo al azar del hogar que cumpla con los requisitos mencionados. Así también encontramos datos relacionados al contexto de la familia y de la formación tanto de la madre como del padre, cuestiones étnicas y geográficas, nivel socio-económico en diferentes etapas, y también desempeño escolar, esfuerzo en el ámbito escolar y tiempo hasta llegar a la escuela para ser tratadas como variables instrumentales para escolaridad y habilidades medidas. Adicionalmente a dicha encuesta en habilidades, el individuo al azar realiza un test avanzado específico de competencias en lectura desarrollado por ETS.³ Dada nuestra intención de replicar el análisis realizado en el paper de cabecera, no utilizaremos los resultados del test ETS en el presente trabajo. Para un análisis en detalle de dicha variable como indicador agregado de las habilidades cognitivas en Colombia, se recomienda Acosta, Muller & Sarzosa (2015).

Desde el punto de vista cognitivo, utilizaremos las variables derivadas del cuestionario de doméstico STEP y de la información de la evaluación de conocimientos prácticos, generando dichas variables agregadas para los distintos tipos de habilidad. Así, obtenemos nuestras variables parciales cognitivas respecto las respuestas obtenidas en escritura, lectura y aritmética que luego utilizaremos para generar un proxy de la inteligencia del individuo. Estas variables parciales corresponden a habilidades auto reportadas en los test ante los entrevistadores, respecto a cada una de las dimensiones cognitivas como son los conocimientos de lectura, escritura y aritmética que el individuo posee fuera del ámbito laboral. El análisis de componentes principales es un método multivariado que busca maximizar la

³ Ver Educational Testing Services (2014).

varianza de una combinación lineal de variables, donde básicamente buscamos una dimensión de componentes sobre la cual las observaciones maximizan su separación, ortogonal a las demás dimensiones. Generamos entonces nuestro indicador agregado de habilidades cognitivas que utilizaremos para determinar la relación entre los retornos y la habilidad intelectual del individuo, mediante las variables parciales cognitivas descriptas anteriormente.

Por el otro lado, deseamos estudiar la relación entre los salarios y el nivel de habilidades socio-emocionales de los individuos encuestados. Para ello, el cuestionario STEP posee instrumentos para medir los rasgos de personalidad, como el modelo *Big Five Personality Factors*.⁴ A estos se los suele resumir con la sigla *OCEAN* que concentra diferentes dimensiones de inteligencia: *openness to experience* (apertura al cambio), *conscientiousness* (responsabilidad), *extraversion* (extroversión), *agreeableness* (afabilidad) y *neuroticism* (inestabilidad emocional).

Openness to experience hace referencia a la capacidad creativa de la persona y a la curiosidad intelectual, cultural y estética. *Conscientiousness* es asociado a rasgos de responsabilidad, busca objetivos a largo plazo, autodisciplina, como opuesto a actitudes de procrastinación y espontáneas. *Extraversion* refiere al interés del individuo por cuestiones externas y de interacción con otros individuos. *Agreeableness* se caracteriza por actitudes de participación, entendimiento y cooperación con otros individuos. Por último, *Neuroticism* refleja rasgos de ansiedad, desestabilidad emocional y se relaciona con sentimientos negativos.

Así también, STEP recopila información respecto a otros aspectos de la personalidad del individuo como son los rasgos de perseverancia *Grit*, aversión al riesgo, entre datos respecto a habilidades socio-emocionales.

El modelo de *Grit* propone determinar una serie de dimensiones referidas a la perseverancia y al afán por alcanzar objetivos a largo plazo, como medidas de los niveles de personalidad de un individuo.⁵ Siguiendo a *Duckworth et al.*, utilizamos la información recopilada en nuestra muestra mediante una batería de preguntas referidas a la dimensión no-cognitiva de perseverancia.

De forma similar, proponemos la utilización de la aversión al riesgo como un proxy de rasgos de la personalidad para asociarla a la propensión de un individuo a realizar transiciones empleo-empleo, siguiendo modelos de dispersión salarial como son *Burdett-Mortensen (1998)* y *Harris-Todaro (1970)*, pero focalizando el análisis en los retornos obtenidos por variaciones en dicha dimensión.

⁴ Para mayores referencias, remitirse a *Goldberg (1993)*.

⁵ Para mayores referencias vea *Duckworth, Peterson, Matthews & Kelly (2007)*.

Para el estudio en cuestión utilizamos una muestra de nuestra base original conformada por $n = 1245$ para Colombia y $n = 1193$ para Bolivia; de entre 15 y 55 años de edad, empleados asalariados en zonas urbanas que poseen ingresos positivos y que poseen información completa respecto a trayectoria educativa, rasgos de habilidades cognitivas y características de habilidades no cognitivas. Tomamos en este caso solo a aquellos individuos asalariados, dado que suponemos que la relación entre ingresos y habilidades cognitivas y no cognitivas difiere respecto a aquellos ocupados por cuenta propia.⁶ Dado que la misma base corresponde a una muestra de la población, se establecen una serie de pasos y criterios para la determinación de una muestra representativa siguiendo los lineamientos propuestos por el Banco Mundial. Siguiendo dichos lineamientos, obtenemos el peso que cada unidad, observación o *PSU* representa en nuestra muestra respecto a la población de cada país. Mediante estos *weights*, obtenemos la inversa de la probabilidad de que dicha observación haya sido muestreada y utilizamos esta probabilidad en nuestras regresiones para ponderar cada observación del análisis. Luego, estandarizamos mediante *z-score* las variables cognitivas y no-cognitivas y resumimos todas las variables a considerar en los siguientes cuadros estadísticos reportados en la **Tabla 1** para Colombia.

Tabla 1: Resumen Estadístico para Colombia

	N	Mean	SD	Min	Max
log of hourly usd earnings	1245	1.27	0.80	0.02	5.25
years of educ	1245	10.76	3.53	1.00	20.00
read	1245	0.00	1.00	-1.66	1.19
write	1245	0.00	1.00	-1.30	2.22
num	1245	0.00	1.00	-1.78	1.68
aggregate cognitive measure	1245	0.00	1.33	-2.71	2.95
work experience	1245	16.53	10.28	0.00	42.00
work experience squared	1245	378.74	390.14	0.00	1764.00
extraversion	1245	0.00	1.00	-3.16	1.51
agreeableness	1245	0.00	1.00	-3.20	1.50
conscientiousness	1245	0.00	1.00	-3.43	1.29
emotional stability	1245	0.00	1.00	-2.18	1.95
openness	1245	0.00	1.00	-3.71	1.52
grit	1245	0.00	1.00	-3.32	1.68
gender	1245	0.48	0.50	0.00	1.00
urban	1245	1.00	0.00	1.00	1.00
mother educ below primary	1245	0.07	0.26	0.00	1.00
mother educ lower secondary	1245	0.54	0.50	0.00	1.00
mother educ upper secondary	1245	0.25	0.44	0.00	1.00
mother educ tertiary	1245	0.06	0.24	0.00	1.00
father educ lower secondary	1245	0.47	0.50	0.00	1.00
father educ upper secondary	1245	0.23	0.42	0.00	1.00
father educ tertiary	1245	0.08	0.27	0.00	1.00
barranquilla	1245	0.05	0.22	0.00	1.00
bello	1245	0.03	0.16	0.00	1.00

⁶ Vea Hartog, van Praag & van der Sluis (2010).

bucaramanga	1245	0.03	0.17	0.00	1.00
cali	1245	0.12	0.32	0.00	1.00
copacabana	1245	0.00	0.04	0.00	1.00
cucuta	1245	0.07	0.25	0.00	1.00
envigado	1245	0.02	0.15	0.00	1.00
floridablanca	1245	0.02	0.16	0.00	1.00
giron	1245	0.01	0.10	0.00	1.00
ibague	1245	0.07	0.25	0.00	1.00
itagui	1245	0.01	0.09	0.00	1.00
la estrella	1245	0.00	0.06	0.00	1.00
manizalez	1245	0.04	0.19	0.00	1.00
medellin	1245	0.08	0.27	0.00	1.00
piedecuesta	1245	0.00	0.06	0.00	1.00
soledad	1245	0.03	0.16	0.00	1.00
villamaria	1245	0.00	0.05	0.00	1.00
villavicencio	1245	0.04	0.20	0.00	1.00

Nota: Estadísticos sin ponderar.

En la **Tabla 2** presentamos el resumen estadístico de las variables a utilizar en nuestra base de datos *STEP* para Bolivia.

Tabla 2: Resumen Estadístico para Bolivia

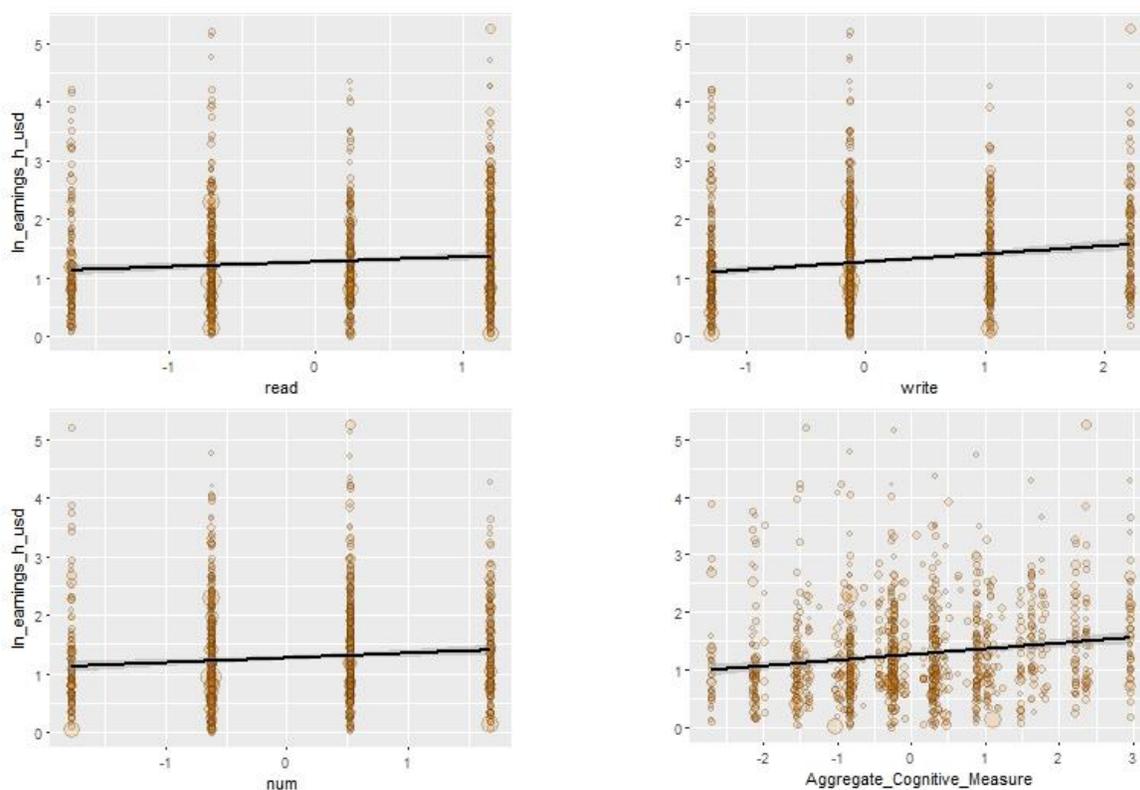
	N	Mean	SD	Min	Max
log of hourly usd earnings	1193	1.34	0.87	0.01	6.55
years of educ	1193	11.99	4.22	1.00	19.00
read	1193	0.00	1.00	-1.98	0.93
write	1193	0.00	1.00	-1.38	1.77
num	1193	0.00	1.00	-2.00	1.44
aggregate cognitive measure	1193	0.00	1.34	-3.07	2.38
work experience	1193	13.96	9.72	0.00	43.00
work experience squared	1193	289.45	338.85	0.00	1849.00
extraversion	1193	0.00	1.00	-2.99	1.43
agreeableness	1193	0.00	1.00	-3.21	1.50
conscientiousness	1193	0.00	1.00	-3.54	1.63
emotional stability	1193	0.00	1.00	-2.15	2.21
openness	1193	0.00	1.00	-3.58	1.47
grit	1193	0.00	1.00	-3.15	1.63
gender	1193	0.51	0.50	0.00	1.00
urban	1193	1.00	0.00	1.00	1.00
mother educ below primary	1193	0.44	0.50	0.00	1.00
mother educ lower secondary	1193	0.08	0.27	0.00	1.00
mother educ upper secondary	1193	0.21	0.41	0.00	1.00
mother educ tertiary	1193	0.15	0.36	0.00	1.00
father educ lower secondary	1193	0.08	0.27	0.00	1.00
father educ upper secondary	1193	0.25	0.43	0.00	1.00
father educ tertiary	1193	0.21	0.41	0.00	1.00
castellano	1193	0.98	0.13	0.00	1.00
aymara_quechua	1193	0.02	0.13	0.00	1.00
la paz	1193	0.27	0.44	0.00	1.00
el alto	1193	0.24	0.43	0.00	1.00
santa cruz	1193	0.34	0.47	0.00	1.00
cochabamba	1193	0.15	0.36	0.00	1.00

Nota: Estadísticos sin ponderar.

4. Análisis Econométrico

Para comenzar a analizar la relación entre los ingresos laborales y las habilidades cognitivas utilizamos métodos gráficos que nos permitan apreciar rápidamente la naturaleza de dicha relación. Mediante diagramas de dispersión entre dichas habilidades cognitivas respecto al logaritmo de los ingresos laborales encontramos diferentes resultados respecto a Colombia y Bolivia ponderando por sus respectivos *weights* por observación. A su vez, analizamos dicha relación respecto a la variable agregada cognitiva. Como ya mencionamos anteriormente, dicha variable se obtiene mediante el análisis de componentes principales. Este es un método multivariado que busca maximizar la varianza de una combinación lineal de variables, donde básicamente buscamos una dimensión de componentes sobre la cual las observaciones maximizan su separación, ortogonal a las demás dimensiones. Generamos entonces nuestro indicador agregado de habilidades cognitivas, mediante las variables parciales escritura, lectura y aritmética.

Figura 1: Correlograma del logaritmo de retornos vs habilidades cognitivas en Colombia

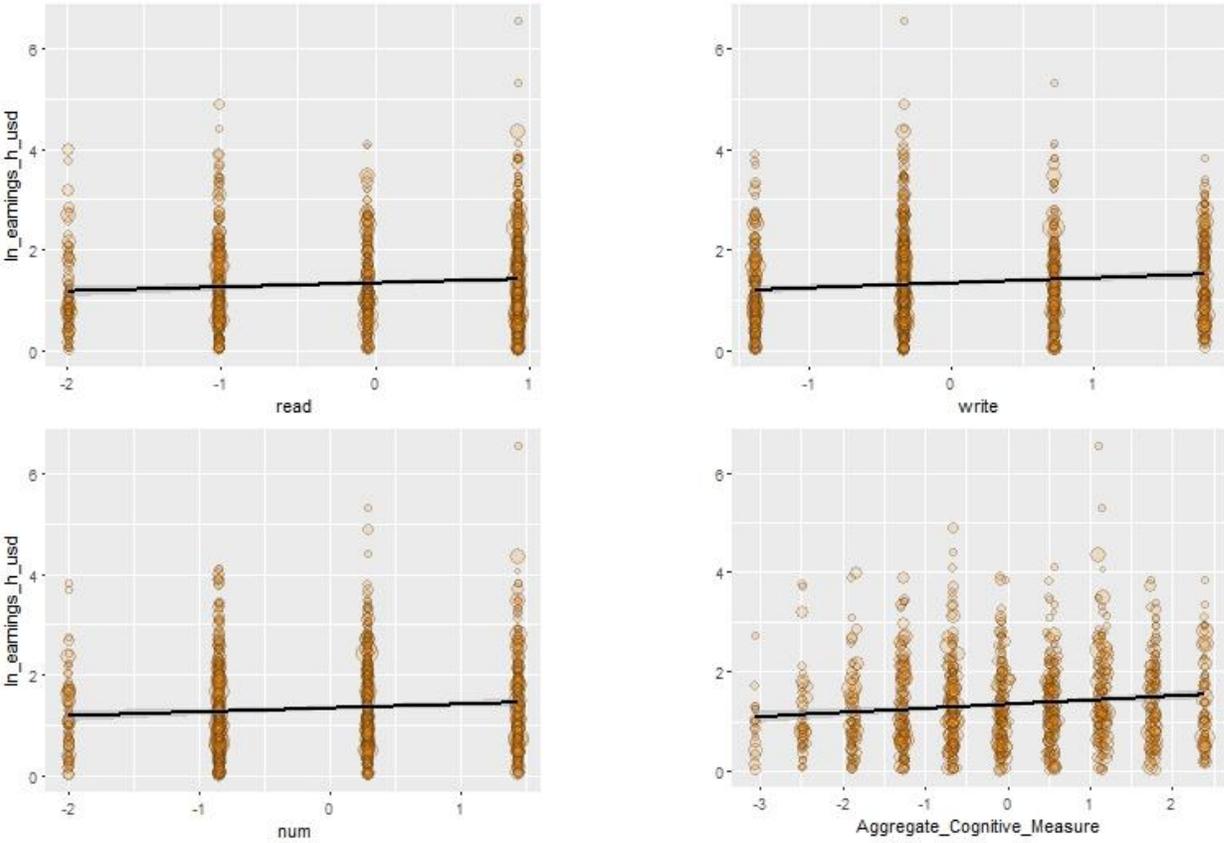


Nota: Ponderación por observación

Como se muestra en la **Figura 1**, Colombia posee una relación positiva en todos los casos respecto a las variables cognitivas y el log de los ingresos laborales; entonces, mayor nivel de habilidades cognitivas parecería estar acompañado por mayores retornos, sin tener en cuenta otras variables y su correlación.

Por otro lado, en la **Figura 2** se reporta el gráfico de dispersión respecto a los ingresos laborales y los niveles de habilidades cognitivas en Bolivia. También aquí encontramos en todos los casos una relación positiva, corroborando lo observado en Colombia.

Figura 2: Correlograma del logaritmo de retornos vs habilidades cognitivas en Bolivia



Nota: Ponderación por observación

Dado que dichos gráficos únicamente muestran la relación entre dichas variables, dejando de lado los efectos de otras variables de relevancia en relación a las habilidades cognitivas y al log de las ganancias laborales, regresamos el *log* de ingresos laborales por hora respecto a las habilidades

cognitivas incluidas individualmente, controlando por experiencia laboral, genero, localización geográfica y educación de la madre y del padre.

En la **Tabla 3** para Colombia observamos que, al incluir de forma individual cada una de las habilidades cognitivas respecto a sus covariantes pero sin incluir años de educación, todas resultan significativas y positivas excepto por la variable lectura. Así, un desvío en estándar en escritura alcanza estima un 10.7% de mayores ingresos, mientras que un desvío estándar en aritmética se asocia a un aumento de los ingresos en 5.5% con un leve nivel de significancia.

Tabla 3: Estimaciones MCO de retornos a habilidades cognitivas sin incluir años de escolaridad, en Colombia

	(1)	(2)	(3)
read	0.058 [1.61]		
write		0.107 [3.61]***	
num			0.055 [1.85]*
R^2	0.07	0.08	0.07
N	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Realizamos el mismo set de regresiones para Bolivia, incluyendo de a una habilidad cognitiva en nuestra regresión a la vez que incluimos las diferentes covariantes pero sin tener en cuenta aún los años de educación. En la **Tabla 4** podemos observar que no todas las habilidades cognitivas resultan significativas, siendo que resolución de problemas matemáticos no resulta significativa estadísticamente. Sin embargo, los retornos a un desvío estándar respecto a lecto-compresión en Bolivia se estiman en un 17.2% de mayores ingresos y un desvío estándar respecto a escritura estima aumentar los ingresos laborales en un 12.8%.

Tabla 4 : Estimaciones MCO de retornos a habilidades cognitivas sin incluir años de escolaridad, en Bolivia

	(1)	(2)	(3)
read	0.172 [3.01]***		
write		0.128 [2.72]***	
num			0.081 [1.54]
R^2	0.21	0.20	0.19
N	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Como mencionamos anteriormente, omitimos en dichas regresiones incluir a la variable años de educación. Con lo cual procedemos a regresar cada una de las variables cognitivas por separado, pero en esta ocasión agregando al set de covariantes inicial la variable años de educación. En la **Tabla 5** reestimamos los resultados obtenidos en Colombia, y observamos que sólo escritura se mantiene estadísticamente significativo y su coeficiente ve reducido su magnitud a 8,6%, al igual que se ven reducidos los demás coeficientes.

Tabla 5: Estimaciones MCO de retornos a habilidades cognitivas incluyendo años de escolaridad, en Colombia

	(1)	(2)	(3)
read	0.027 [0.74]		
write		0.086 [2.83]***	
num			0.032 [1.08]
R^2	0.09	0.10	0.09
N	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Por otro lado, y siguiendo la misma lógica que en Colombia, incluimos a la variable años de escolaridad en nuestras regresiones individuales para Bolivia. Las habilidades cognitivas escritura y lectura en este caso pierden cierto grado de significancia estadística y ven reducidas sus magnitudes, quedando sólo escritura asociado a un 8.9% y con niveles aceptables estadísticamente. Dichos resultados pueden visualizarse en la **Tabla 6**.

Tabla 6: Estimaciones MCO de retornos a habilidades cognitivas incluyendo años de escolaridad, en Bolivia

	(1)	(2)	(3)
read	0.076 [1.62]		
write		0.089 [2.10]**	
num			0.023 [0.48]
R^2	0.26	0.26	0.26
N	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Siguiendo la literatura y trabajos anteriores, la variable años de escolaridad pareciera estar correlacionada y capturar cierta magnitud del efecto de las habilidades cognitivas respecto a los ingresos laborales. Para poder analizar dichas habilidades cognitivas de forma más robusta, generamos un indicador de inteligencia cognitiva donde cada una de las dimensiones cognitivas analizadas previamente puedan ser contenidas en una única variable. Mediante *principal components analysis*, obtenemos nuestro indicador agregado que utilizaremos como proxy de las habilidades cognitivas de cada individuo. El análisis de componentes principales es un método multivariado que busca maximizar la varianza de una matriz de variables, donde buscamos una dimensión de componentes sobre la cual las observaciones maximizan su separación, ortogonal a las demás dimensiones de la matriz construida. Dado que solo el primer componente presenta un *eigenvalue* mayor a 1, retenemos dicho componente y lo utilizamos como nuestro indicador cognitivo.⁷ En la **Tabla 7** y **8** presentamos la matriz de correlación de dicho componente agregado respecto a las variables cognitivas (que corresponden a medidas parciales de la inteligencia de cada individuo), para Colombia y Bolivia respectivamente.

Para el caso de Colombia, los valores de correlación varían entre 0.28 para lecto-compresión ~ aritmética y 0.49 para lecto-compresión ~ escritura. Como es de esperar, todas las habilidades cognitivas se encuentran altamente correlacionadas con nuestro indicador cognitivo agregado por encima de 0.69.

Tabla 7: Matriz de correlación de habilidades cognitivas para Colombia

	read	write	num	aggregate measure
read	1.0000			
write	0.4917	1.0000		
num	0.2846	0.3765	1.0000	
aggregate measure	0.7804	0.8305	0.6901	1.0000

Consecuentemente, en Bolivia los valores de correlación varían entre 0.328 para lecto-compresión ~ aritmética y 0.499 para lecto-compresión ~ escritura. Por otro lado, todas las habilidades cognitivas se encuentran altamente correlacionadas con nuestro indicador cognitivo agregado por encima de 0.688 y siendo significativas estadísticamente en todos los casos.

⁷ Para una explicación más detallada respecto a métodos multivariados, remitirse a *A. Rencher (2002)*.

Tabla 8: Matriz de correlación de habilidades cognitivas para Bolivia

	read	write	num	aggregate measure
read	1.0000			
write	0.4990	1.0000		
num	0.3281	0.3376	1.0000	
aggregate measure	0.8063	0.8114	0.6883	1.0000

Siguiendo con nuestros análisis, procedemos a estimar las regresiones anteriores pero en este caso incluimos a todas las habilidades cognitivas en dicha regresión, con y sin años de escolaridad, para luego comprar dichos resultados con las estimaciones de los retornos a nuestro componente agregado cognitivo, aquí también con y sin incluir a la variable años de educación controlando por experiencia laboral, género, localización geográfica y educación de los padres.

Presentamos los resultados para Colombia en la **Tabla 9**. Sin controlar por años de educación, en la primera columna incluimos a todas las variables cognitivas a la vez y estimamos que un desvío estándar se asocia a un aumento en el ingreso laboral en torno al 9.5% para escritura, mientras que resolución de problemas matemáticos y lecto-compresión no resultan estadísticamente significativas. Nuevamente sin controlar por años de escolaridad, en la segunda columna estimamos los retornos respecto al indicador agregado de habilidades cognitivas. Dicho coeficiente resulta altamente significativo estadísticamente y estima una magnitud del coeficiente en 8.5% de mayores retornos por cada desvío estándar adicional. Las columnas 3 y 4 replican lo analizado en las primeras columnas, pero incluyendo años de educación en ambas regresiones. La columna 3 entonces estima los retornos asociados a las dimensiones parciales cognitivas incluidas al mismo tiempo, controlando por años de educación. Sólo escritura resulta estadísticamente significativa con un coeficiente estimado en 8.6%, capturando años de escolaridad el efecto marginal en los ingresos laborales en 2.8% por año adicional de educación y siendo estadísticamente significativa. Finalmente en la última columna regresamos el log de los ingresos laborales por hora respecto a nuestro indicador agregado de medidas cognitivas, junto al resto de covariables y controlando por años de educación. Estimamos que, por cada desvío estándar adicional, años de educación mantiene casi su misma magnitud y significancia; mientras que nuestro indicador cognitivo resulta levemente menos significativo y positivamente correlacionado, incrementando los ingresos en 6.1% por *sd*.

Tabla 9: Estimaciones MCO de retornos a las habilidades cognitivas, incluyendo a todos los indicadores parciales al mismo tiempo y sin controlar por años de educación en Colombia

	(1)	(2)	(3)	(4)
read	0.014 [0.37]		-0.008 [0.20]	
write	0.095 [2.98]***		0.086 [2.63]***	
num	0.033 [1.13]		0.019 [0.64]	
aggregate cognitive measure		0.085 [3.22]***		0.061 [2.24]**
years of educ			0.028 [2.54]**	0.027 [2.34]**
R^2	0.09	0.08	0.10	0.09
N	1,245	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

En la **Tabla 10** replicamos el mismo análisis y estudiamos los resultados obtenidos para Bolivia. Lecto-comprensión resulta estadísticamente significativo y positivamente correlacionado, estimando un 13.1% por cada desvío estándar según podemos observar en la primer columna de la tabla. En la columna siguiente estimamos el retorno asociado a nuestro indicador cognitivo y obtenemos que, estadísticamente significativo al 1%, por cada desvío estándar el ingreso laboral aumenta con una magnitud de 14.5%, excluyendo de la regresión a la variable años de educación. A continuación en la columna 3 incluimos junto con nuestras habilidades cognitivas parciales a la variable años de educación a nuestras estimaciones, captando la misma parte de los efectos de las medidas cognitivas y generando un retorno de 6.0% por cada desvío estándar con un alto nivel de significancia. Por su parte escritura, aritmética y lecto-comprensión no resultan significativas siquiera al 10%. La columna 4 muestra los retornos a los años de educación y a nuestra variable agregada cognitiva, incluyendo el resto de nuestras variables de control. Por cada año adicional de educación se estima que los ingresos laborales aumentan en un 5.9% con un alto nivel de significancia, mientras que los retornos asociados a nuestro indicador cognitivo resulta moderadamente significativos al 5% y con un coeficiente de 7.7% por desvío estándar.

Tabla 10: Estimaciones MCO de retornos a las habilidades cognitivas, incluyendo a todos los indicadores parciales al mismo tiempo y sin controlar por años de educación en Bolivia

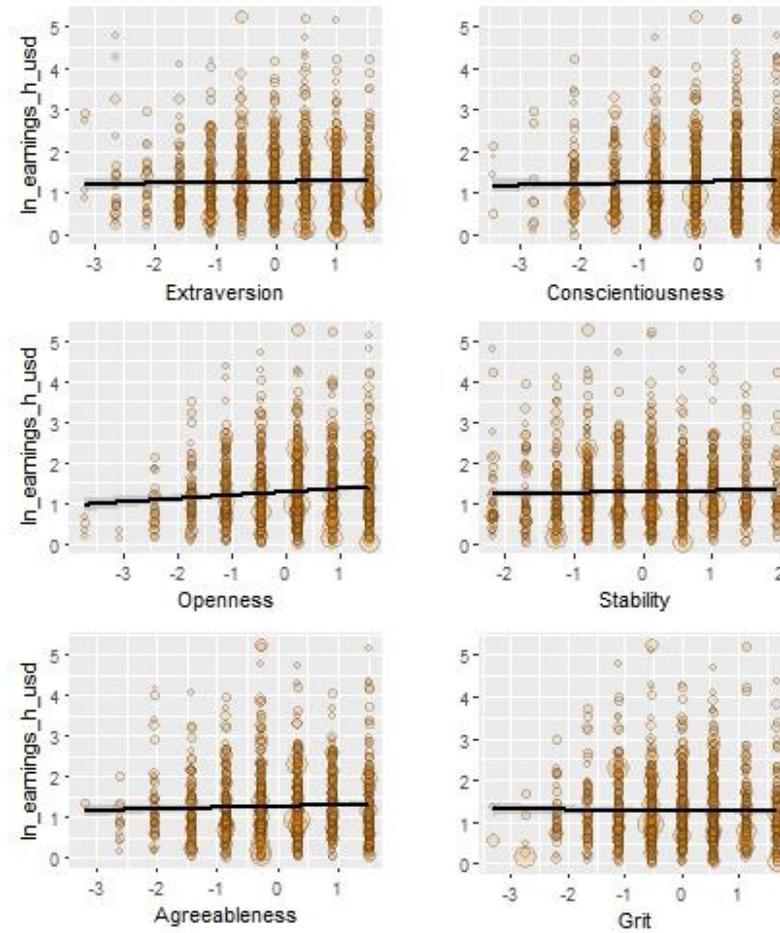
	(1)	(2)	(3)	(4)
read	0.131 [2.01]**		0.043 [0.83]	
write	0.074 [1.41]		0.074 [1.60]	
num	0.037 [0.73]		0.008 [0.17]	
aggregate cognitive measure		0.145 [3.58]***		0.077 [2.11]**
years of educ			0.060 [4.11]***	0.059 [3.93]***
R^2	0.22	0.22	0.26	0.26
N	1,193	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Podemos realizar el mismo set de regresiones precedentes para estudiar y entender la relación entre el ingreso laboral y otra de las dimensiones de la inteligencia de cada individuo, las habilidades socio-emocionales. Dentro de nuestros instrumentos de medición de dichas habilidades no cognitivas contamos con los cinco factores de personalidad o *Big Five Factors* de *Goldberg*, como también con los rasgos de *Grit* para determinar perseverancia de esfuerzo.

Nuevamente recurrimos a métodos gráficos para poder tener una primera aproximación a la relación entre las habilidades no-cognitivas normalizadas y sus retornos laborales. Utilizamos gráficos de dispersión para cada una de las dimensiones de los *Big Five* y para nuestro rasgo *Grit* de perseverancia. En la **Figura 3** reportamos los resultados para Colombia, observando que todos los *Big Five Factors* presentan una relación cuasi plana respecto al logaritmo del ingreso laboral por hora. Haciendo mayor foco en cada una, observamos en casi todas las variables no cognitivas una relación de tendencia levemente positiva, excepto en el caso de perseverancia de *Grit* el cual presenta una tendencia muy leve negativa.

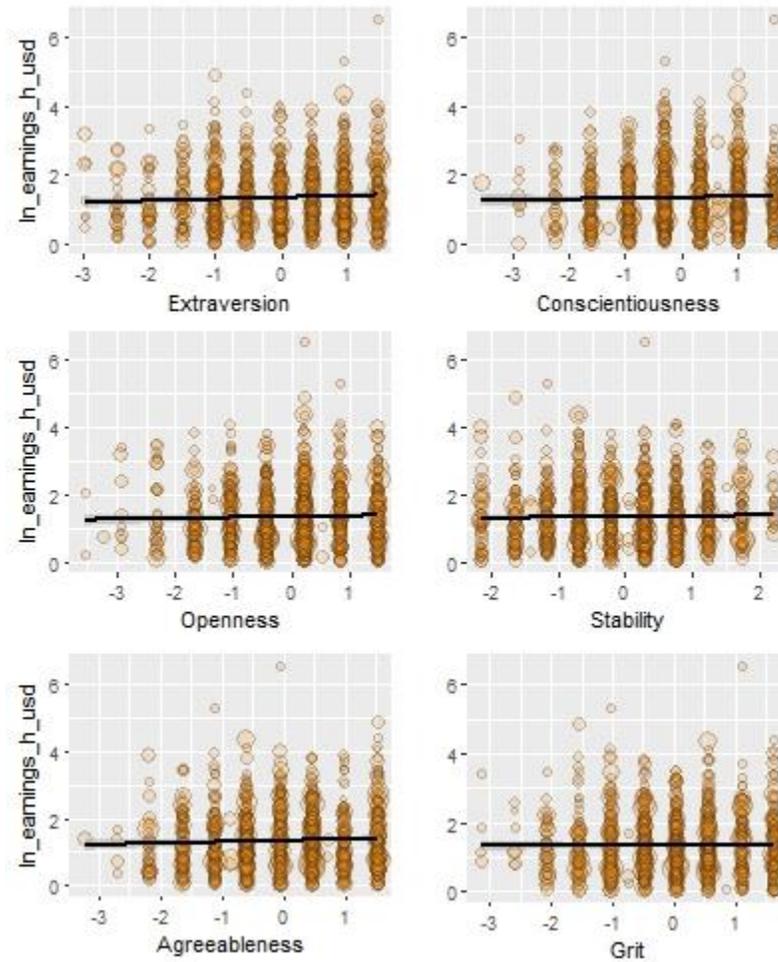
Figura 3: Correlogramas respecto al logaritmo de retornos vs habilidad no cognitivas en Colombia



Nota: Ponderación por observación

Por otro lado, en la **Figura 4** graficamos la relación entre retornos y cada una de las habilidades no-cognitivas en Bolivia. En este caso, encontramos relaciones dispares entre las diferentes dimensiones de las *Big Five*: extraversión, afabilidad, responsabilidad, estabilidad emocional y apertura al cambio; como así también para los rasgos *Grit*; pero manteniendo al igual que en el caso de Colombia una relación cuasi plana a primera vista.

Figura 4: Correlogramas respecto al logaritmo de retornos vs habilidad no cognitivas en Bolivia



Nota: Ponderación por observación

Entonces, y al igual que con las habilidades cognitivas, presentamos los resultados de regresar los retornos respecto a las habilidades no cognitivas incluidas individualmente, controlando por experiencia laboral, género, localización geográfica y educación de los padres. En la columna 1 de la **Tabla 11** presentamos los resultados de dicha regresión para Colombia, excluyendo a la variable años de escolaridad del análisis. Ninguna de las *Big Five* presenta niveles aceptables de significancia. Similarmente, en la columna 2 exponemos los resultados de estimar el retorno a las habilidades socio-emocionales incluidas de a una a la vez, pero en este caso controlando también por años de educación. Los resultados se mantienen en su mayoría consistentes a los encontrados en la columna 1, pero en este caso el rasgo de perseverancia de Grit resulta significativo levemente: por cada desvío estándar, se estima una reducción del *log* de ingresos laborales en un 5.7%.

Tabla 11: Estimaciones MCO de retornos a las habilidades no-cognitivas, incluyendo una a la vez las variables socio-emocionales en Colombia

	Sin controlar por años de educación	Incluyendo años de educación
extraversion	-0.015 [0.46]	-0.013 [0.41]
agreeableness	0.028 [0.77]	0.005 [0.13]
conscientiousness	0.010 [0.26]	0.005 [0.11]
emotional stability	-0.001 [0.01]	-0.011 [0.27]
openness	0.030 [0.80]	0.014 [0.37]
grit	-0.050 [1.62]	-0.057 [1.85]*

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Paralelamente, corremos la misma serie regresiones respecto al retorno de las habilidades no cognitivas en Bolivia. La **Tabla 12** presenta las estimaciones de incluir cada de una las habilidades no cognitivas de manera individual, sin controlar y luego controlando por años de educación. En la primera columna obtenemos los resultados sin controlar por años de escolaridad, apreciando que ninguna de las variables de forma individual presenta niveles significativos estadísticamente. Lo mismo sucede en la columna segunda donde controlamos por años de educación, siendo también que ninguno de nuestros indicadores socio-emocionales de las grandes cinco ni los rasgos *Grit* muestran niveles estadísticos significativos.

Tabla 12: Estimaciones MCO de retornos a las habilidades no-cognitivas, incluyendo una a la vez las variables socio-emocionales en Bolivia

	Sin controlar por años de educación	Incluyendo años de educación
extraversion	-0.029 [0.66]	-0.028 [0.67]
agreeableness	0.021 [0.46]	0.001 [0.02]
conscientiousness	0.032 [0.70]	0.000 [0.00]
emotional stability	-0.043 [0.76]	-0.052 [1.11]
openness	-0.026 [0.64]	-0.034 [0.87]
grit	0.003 [0.05]	-0.034 [0.78]

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Para tener en cuenta el efecto total de las habilidades no-cognitivas en su conjunto, realizamos la regresión de todas las variables no cognitivas al mismo tiempo, por grupos, respecto al log de ingresos por hora, controlando por experiencia laboral, género, localización geográfica, educación de los padres e incluyendo y excluyendo a la variable escolaridad de nuestras regresiones. En la **Tabla 13** se expone el análisis realizado para Colombia, donde la primer columna de la serie presenta los coeficientes y sus niveles de significancia al estimar el grupo de las *Big Five* incluidas simultáneamente en grupo pero sin controlar por años de educación. Al igual que en la tabla 10, ninguna de las habilidades no cognitivas muestran umbrales significativos estadísticamente. Por su lado, los rasgos de Grit fueron incluidos individualmente en la tabla 10, sin presentar niveles significativos. La siguiente columna manifiesta los efectos de regresar ambos grupos de variables socio-emocionales al mismo tiempo respecto al ingreso laboral y sin controlar por la variable años de escolaridad. Aquí la dimensión no-cognitiva de perseverancia resulta únicamente significativa al 5%, y donde su coeficiente se estima en -6.3% por desvío estándar. La columna 3 presenta la misma regresión que la primer columna pero agregando como control a la variable años de educación y donde, nuevamente, las *Big Five* no muestran niveles de significancia estadística. La última columna toma a todas las habilidades no cognitivas a la vez y estima que el rasgo de Grit es significativo con una relación negativa de 6.5% por desvío estándar.

Tabla 13: Estimaciones MCO de retornos a las habilidades no-cognitivas, incluyendo por grupos las variables socio-emocionales en Colombia

	(1)	(2)	(3)	(4)
extraversion	-0.021 [0.60]	-0.018 [0.51]	-0.015 [0.42]	-0.011 [0.32]
agreeableness	0.021 [0.56]	0.028 [0.75]	0.001 [0.03]	0.008 [0.21]
conscientiousness	0.003 [0.09]	0.010 [0.28]	0.003 [0.07]	0.010 [0.26]
emotional stability	-0.002 [0.06]	-0.009 [0.23]	-0.011 [0.29]	-0.018 [0.47]
openness	0.026 [0.73]	0.034 [0.93]	0.017 [0.45]	0.025 [0.66]
grit		-0.063 [2.02]**		-0.065 [2.13]**
years of educ			0.035 [3.30]***	0.035 [3.41]***
R^2	0.07	0.08	0.09	0.09
N	1,245	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

La **Tabla 14** muestra el resultado análogo en Bolivia de estimar los retornos a las habilidades socio-emocionales al incluir las mismas por grupos, controlando por experiencia laboral, género, localización geográfica, educación de los padres e incluyendo y excluyendo años de educación. La primer columna, que incluye al grupo de los *Big Five Factors* exclusivamente junto al resto de variables de control y excluyendo a años de escolaridad, no presenta coeficientes que resulten estadísticamente significativos ni siquiera al 10%. Así también, como habíamos estimado en la Tabla 11, perseverancia de *Grit* tampoco logra superar el umbral y por ende no resulta significativo. Al incluir a todas las dimensiones no-cognitivas al mismo tiempo en nuestra regresión, sin controlar por años de educación, observamos en la columna 2 que no se generan efectos de significancia en el retorno laboral del individuo. Controlando la regresión por años de educación, la columna 3 expresa los retornos a cada una de los indicadores socio-emocionales de los *Big Five Factors* estimando entonces que ninguno de ellos resulta estadísticamente significativa al 10%. Dicho resultado se mantiene aún al agregar ambos grupos de indicadores no cognitivos y controlando por años de educación, como se observa en la columna 4. Con lo cual, resulta de interés el hecho de que el mercado laboral en Bolivia no valore en mayor o menor medida a individuos que posean diversos niveles de habilidades no cognitivas.

Tabla 14: Estimaciones MCO de retornos a las habilidades no-cognitivas, incluyendo por grupos las variables socio-emocionales en Bolivia

	(1)	(2)	(3)	(4)
extraversion	-0.014 [0.34]	-0.014 [0.34]	-0.012 [0.27]	-0.011 [0.27]
agreeableness	0.018 [0.40]	0.018 [0.40]	0.003 [0.07]	0.007 [0.16]
conscientiousness	0.033 [0.76]	0.033 [0.74]	0.004 [0.10]	0.008 [0.17]
emotional stability	-0.044 [0.79]	-0.044 [0.80]	-0.052 [1.13]	-0.054 [1.16]
openness	-0.031 [0.75]	-0.030 [0.74]	-0.036 [0.89]	-0.030 [0.75]
grit		-0.001 [0.03]		-0.034 [0.75]
years of educ			0.068 [4.72]***	0.069 [4.70]***
R^2	0.19	0.19	0.26	0.26
N	1,193	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

El análisis preliminar llevado a cabo hasta el momento arroja resultados acordes a los encontrados en nuestro paper de cabecera, pero con magnitudes y relaciones diferentes en muchos casos. Al parecer,

cada mercado laboral en América Latina valora y repaga de forma diversa las distintas dimensiones cognitivas y no cognitivas. En particular, desde el punto de vista cognitivo, ambos países presentan relaciones positivas y significativas en relación a su indicador cognitivo agregado como ocurre en el caso de Perú. Desde el enfoque de las habilidades socio-emocionales, el rasgo de perseverancia de *Grit* resulta estadísticamente significativa y con retornos negativos en Colombia; mientras que las dimensiones socio-emocionales no resultan significativa en Bolivia.

Sin embargo, en nuestras estimaciones no estamos teniendo en cuenta el problema de endogeneidad que supone nuestro análisis. En la sección siguiente nos encargaremos de capturar los efectos parciales de cada indicador y de la variable años de educación una vez que se tiene en cuenta la relación entre dichas covariables. Para ello utilizamos nuestro vector de habilidades cognitivas y no cognitivas y, mediante un procedimiento econométrico, capturamos los efectos exógenos en el retorno al ingreso laboral por hora mediante la utilización de un set de variables instrumentales y los residuos de sus regresiones.

5. Corrección de sesgos y problemas potenciales

a. Problema de endogeneidad

Dado que las habilidades cognitivas y no cognitivas se van gestando y desarrollando a lo largo de la educación del individuo, nos encontramos con el problema potencial de que dichas habilidades se encuentren afectadas por los años de educación del individuo. Con lo cual, de existir una relación positiva entre dicha variable y nuestras dimensiones cognitivas y no cognitivas, no estaríamos capturando el efecto total de la variable años de escolaridad en detrimento de la misma en nuestros coeficientes estimados en las regresiones anteriores.

Por otro lado, de no existir dicha relación supondríamos entonces que tanto años de educación como el set de habilidades son exógenas, y podríamos realizar entonces una regresión de *Mincer* para estimar los retornos, extendiendo por habilidades cognitivas y no cognitivas dicho modelo. La típica regresión de *Mincer* estima los retornos al *log* del ingreso laboral por hora, respecto a años de educación y otras covariables.⁸

La **Tabla 15** nos muestra los resultados de una regresión por *MCO* del *log* del ingreso laboral horario respecto a años de educación y controlando en todos los casos por experiencia laboral, género y

⁸ Para mayores referencias, remitirse a *J. Mincer* (1958).

localización geográfica para Colombia. En primera instancia, por cada año adicional de educación, los ingresos aumentan en 3.7%, donde dicho coeficiente es estadísticamente significativo al 1%. Así también, si controlamos dicha regresión por el historial académico de sus padres para tener en cuenta efectos exógenos (ya sea mediante la transición de conocimiento entre padres e hijos u otros efectos posibles), encontramos que el retorno a los años de educación cae a 3.5% por cada año adicional y mantienen su alto nivel de significancia. En la última columna agregamos a nuestra regresión tanto nuestras habilidades cognitivas como no cognitivas, para evitar el sesgo que puede darse por excluir variables relevantes. Nuevamente vemos que nuestro coeficiente estimado de años de educación cae hasta alcanzar una magnitud de 2.7% por cada año adicional. Nuestra medida cognitiva agregada estima un retorno de 6.0% por cada desvío estándar adicional, con un nivel de significancia al 5%; mientras que desde el punto de vista socio-emocional, el rasgo de perseverancia de esfuerzo es la única que es estadísticamente significativa (al 5%) con un retorno esperado por cada desvío estándar de -6.3%.

Tabla 15: Estimaciones MCO de retornos a años de educación, habilidades cognitivas y no-cognitivas para Colombia

	Sin incluir educación de los padres	Controlando por educación de los padres	Controlando por educación de los padres y habilidades
years of educ	0.037 [3.40]***	0.035 [3.21]***	0.027 [2.53]**
aggregate cognitive measure			0.060 [2.26]**
extraversion			-0.010 [0.30]
agreeableness			0.008 [0.22]
conscientiousness			0.006 [0.16]
emotional stability			-0.026 [0.67]
openness			0.020 [0.54]
grit			-0.063 [2.04]**
R^2	0.06	0.09	0.10
N	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Nota: Controlamos en todos los casos por experiencia laboral, género y localización geográfica.

Replicamos dicha secuencia de regresiones de *Mincer* para el caso de Bolivia y volcamos los resultados en la **Tabla 16**. La columna 1 estudia los retornos vinculados a los por años de educación, controlando adicionalmente por experiencia laboral, género y localización geográfica. La relación que emerge resulta positiva y altamente significativa estadísticamente, estimando un aumento del 8% al log de los ingresos laborales por hora por cada año de educación adicional. Agregando como variable de control a nuestra regresión las variables nivel educativo de los padres, reestimamos nuestro coeficiente académico en la columna 2, y vemos que el mismo cae en magnitud pero mantiene su nivel estadístico: un año adicional de educación se estima en un aumento del retorno esperando en 6.8%. Luego en la columna 3, corremos nuevamente nuestra regresión pero controlando adicionalmente por nuestros grupos de variables cognitivas y no cognitivas. Según observamos, nuestro coeficiente de años de educación vuelve a caer y estima un retorno asociado por cada año de educación adicional en el orden del 5.9%, siendo altamente significativo estadísticamente. Por el lado de nuestro set de habilidades, nuestro componente principal agregado cognitivo resulta significativo siquiera al 5% estimando un retorno por cada año adicional de educación en 8.5%. Por el lado socio-emocional no encontramos evidencia de que los retornos asociados a dichas habilidades sean estadísticamente diferentes a 0.

Tabla 16: Estimaciones MCO de retornos a años de educación, habilidades cognitivas y no-cognitivas para Bolivia

	Sin incluir educación de los padres	Controlando por educación de los padres	Controlando por educación de los padres y habilidades
years of educ	0.080 [6.21]***	0.068 [4.70]***	0.059 [3.89]***
aggregate cognitive measure			0.085 [2.44]**
extraversion			-0.020 [0.49]
agreeableness			0.009 [0.21]
conscientiousness			0.014 [0.31]
stability			-0.063 [1.36]
openness			-0.027 [0.67]
grit			-0.033 [0.75]
R^2	0.18	0.26	0.27
N	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

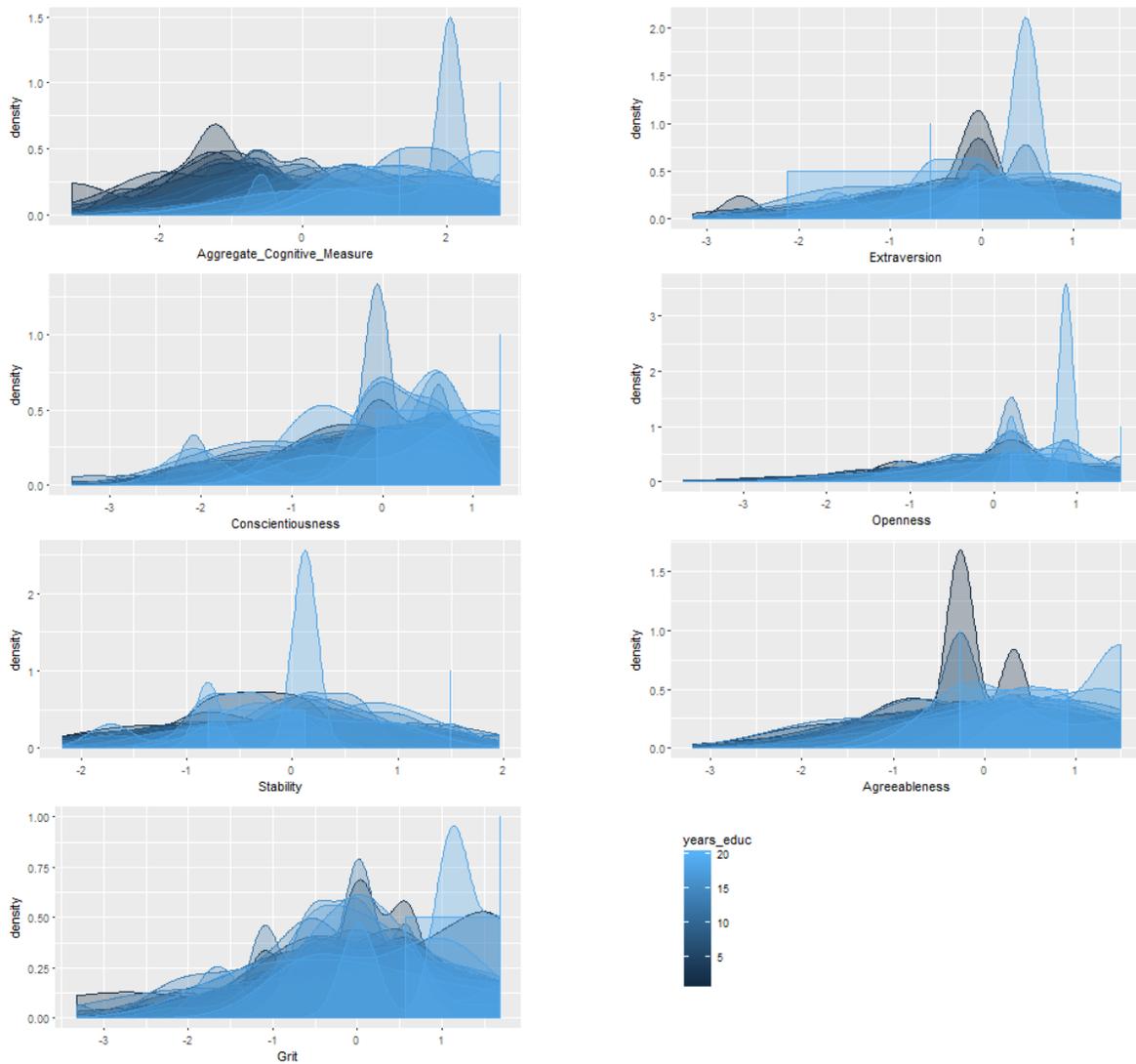
Nota: controlando en todos los casos por experiencia laboral, género y localización geográfica.

Con ello observamos que nuestros resultados anteriores se mantienen en línea con los valores estimados mediante regresiones de *Mincer* y suponiendo que las variables años de educación y habilidades resultan exógenas. Pero si estas fueran endógenas, entonces los años de educación estarían afectando a nuestras habilidades y siendo así, reduciendo la magnitud del efecto que genera cada año adicional de educación en el individuo. Por tanto en nuestras regresiones anteriores no incluyendo las habilidades cognitivas y no-cognitivas, no se puede aislar el efecto real de cada año de educación respecto a dichas habilidades. La significancia estadística de los parámetros del modelo no necesariamente implica que se utilizan las variables correctas y que los efectos medidos son válidos, por eso nos apoyamos en la teoría relevante existente y desarrollamos la metodología apropiada para nuestro modelo.

Entonces, para corroborar la relación entre años de educación y nuestro set de habilidades, la **Figura 5** muestra en forma esquemática mediante gráficos de densidad la relación existente entre nuestro indicador cognitivo agregado, los 5 factores socio-emocionales y los rasgos de *Grit* respecto a años de escolaridad para aquellos individuos en Colombia.⁹

⁹ Dichos gráficos fueron realizados mediante el paquete estadístico R, utilizando su librería ggplot2. Para mayores referencias respecto al tema, puede remitirse a *Matloff* (2011).

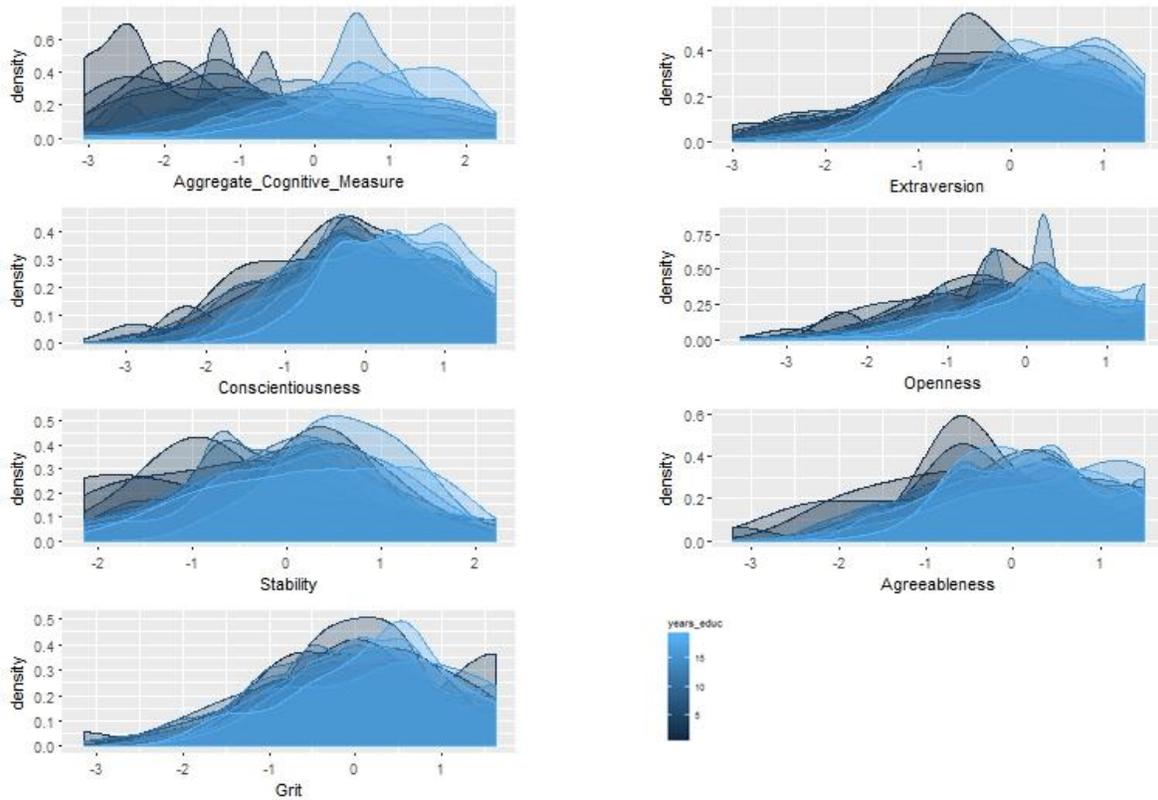
Figura 5: Gráficos de densidad del conjunto de habilidad por cada año de educación en Colombia



En la **Figura 6** graficamos los resultados de cómo los años de educación afectan a las habilidades medidas para Bolivia, donde nuevamente analizamos tanto las habilidades cognitivas como las no cognitivas.

Dichas representaciones nos muestran que parecería existir una relación positiva para ambos países, donde a mayor nivel educativo mayor nivel de nuestro componente cognitivo agregado como también de nuestras habilidades no cognitivas, como supone así también la literatura relacionada. Es por esto que suponemos que existe un problema de endogeneidad entre la variable años de educación y nuestras habilidades y que, entonces, la estimación de las habilidades medidas captura parte del efecto de la correlación entre la elección del individuo respecto a educación y estos parámetros.

Figura 6: Gráficos de densidad del conjunto de habilidad por cada año de educación en Bolivia



Siguiendo entonces a *Neal and Johnson (1996)*, *Heckman and Mullen (2004)*, *Heckman, Stixrud and Urzua (2006)*, no es posible distinguir el efecto causal entre mayores habilidades y mayores retornos, respecto del efecto causal de mayor cantidad de años de educación en mayores habilidades y mayores retornos. Si analizamos nuestras regresiones anteriores, cada vez que se incluye a la variable años de educación, todos los coeficientes cognitivos y no cognitivos caen, dado que solo mantienen su efecto directo en los retornos. De modo similar sucede cuando incluimos en nuestra regresión nuestro set de habilidades, el coeficiente de años de educación cae excluyendo el efecto indirecto en los retornos a causa de dichas habilidades.

Para hacer frente a este problema de factor de confusión, se propone un procedimiento en dos partes tal como se desarrolla en nuestro paper de cabecera “*The Returns to Cognitive and Non-Cognitive Skills in Urban Peru*”. *D. Vera Tudela, JJ Díaz & O. Arias (2012)* mediante la utilización de los residuos de una serie de regresiones y la implementación de variables instrumentales. Por lo cual, en el primer paso estimamos una serie de regresiones instrumentales de nuestras habilidades cognitivas y no cognitivas respecto a años de educación y otras variables de control, para luego obtener los residuos de dicha regresión y utilizar los dichos residuos como el efecto exógeno de dicha habilidad

respecto a años de educación para ser utilizada en nuestra regresión de retornos. Entonces, realizamos una regresión para cada una de nuestras métricas de habilidades mediante variables instrumentales respecto a años de educación y otras variables de control, obteniendo así los residuos de cada una de dichas regresiones para conservarlas como nuestras nuevas variables de habilidades cognitivas y socio-emocionales que luego utilizaremos en nuestra regresión de retornos por hora. Mediante este primer proceso, logramos eliminar el efecto endógeno que años de educación genera en cada una de nuestro set de habilidades y por ende evitamos nuestro sesgo de endogeneidad manteniendo únicamente el efecto exógeno de cada una de las variables en cuestión. Sin embargo, para poder realizar este primer paso del procedimiento es necesario poseer un conjunto de variables que se encuentren correlacionados de manera positiva con años de educación, pero que no se encuentre correlacionado con nuestras habilidades y que además resulten robustas para ser consideradas como variables instrumentales.

Así de nuestra base *STEP* tomamos 4 indicadores como nuestras variables instrumentales: desempeño escolar, tiempo para llegar al colegio, esfuerzo en el trabajo y nivel socioeconómico familiar. Entonces, para poder trabajar con dichas variables transformamos éstas en *dummies* donde: desempeño escolar toma el valor de 1 si el individuo manifiesta que su desempeño en el colegio fue superior a la media o 0 en el caso contrario; tiempo para llegar al colegio toma el valor de 1 cuando el individuo tenía 30 o más minutos de viaje hasta el colegio o 0 en el caso contrario; y nivel socioeconómico familiar toma el valor de 1 si el individuo manifiesta un alto nivel socioeconómico mientras obtenía su educación o 0 en el caso contrario. Adicionalmente en nuestro paper de cabecera utiliza la variable esfuerzo realizado en el colegio como un cuarto instrumento, sin embargo este instrumento no se encuentra presente en nuestra base. Por lo cual, proponemos utilizar las respuestas obtenidas respecto de seguimiento de resultados escolares por parte de los padres como proxy de la variable instrumental esfuerzo realizado en el colegio. Pero, antes de tomar a nuestros 4 instrumentos debemos realizar una serie de test de identificación y de robustez instrumental a modo de obtener residuos apropiados para nuestras habilidades cognitivas y socio-emocionales. Para cada uno de nuestros instrumentos, realizamos el test estadístico *Kleibergen-Paap rk LM* para comprobar si nuestra ecuación se encuentra identificada (y por ende nuestro instrumento es relevante), correlacionado con nuestro regresor endógeno años de educación y entonces su matriz es rango columna completo. Luego testeamos si estos instrumentos identificados se encuentran sólo correlacionados de manera leve con la regresora endógena y por tal presenta problemas de instrumentos débiles. *Stock and Yogo* (2002, 2005) presentan un análisis profundo y detallado respecto al tema. Así, el estadístico *Kleibergen-Paap rk Wald F* testea el potencial problema de debilidad de nuestros instrumentos para diferentes valores críticos según el estadístico de *Cragg-*

Donald F. Finalmente testeamos las restricciones por sobre-especificación de nuestros instrumentos mediante el test de *Sargan-Hansen*, también conocido como estadístico *Hansen J.* Buscamos entonces que nuestros instrumentos no se encuentren correlacionados con el término de error de nuestra regresión y que los instrumentos excluidos sean correctamente excluidos de nuestra ecuación estimada. Implementando los sucesivos test estadísticos descritos, encontramos diferencias para cada país de estudio respecto a los instrumentos que resultan válidos para nuestras regresiones instrumentales de habilidades cognitivas y no cognitivas tomando como variable endógena años de educación.¹⁰

La **Tabla 17** compila entonces para Colombia la primer parte de nuestro procedimiento, estimando mediante *Limited Information Maximum Likelihood (LIML)* los coeficientes y obteniendo los residuos para cada una de nuestras habilidades, un set de covariables y al utilizar como variables instrumentales tiempo al colegio, desempeño escolar y nivel socioeconómico; instrumentos que no presentan problemas de identificación ni *weak instruments*. Por cada habilidad medida encontramos una columna en nuestra tabla, con sus correspondientes coeficientes estimados, desvío estándar y niveles de significancia. Así, un año adicional de educación resultan en un aumento de nuestro parámetro cognitivo agregado en 0.342 desvío estándar, donde también se estima un aumento de 0.345 desvíos estándar para nuestro parámetro socio-emocional de apertura al cambio por cada año adicional de educación; siendo ambos parámetros los de mayor magnitud y siendo altamente significativos estadísticamente. Tanto las habilidades no-cognitivas perseverancia de *Grit* y extraversión presentan valores altamente significativos pero de menor dimensión; mientras que afabilidad y concientización presentan niveles bajos de significancia, a la vez que estabilidad emocional no resulta de interés estadístico.

¹⁰ Para mayores especificaciones respecto a la validez de los instrumentos, vea Apéndice I y Apéndice II.

Tabla 17: Regresiones instrumentales de las habilidades cognitivas y no cognitivas respecto al log de ingresos laborales por hora para Colombia

	Aggregate Cognitive Measure	Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Openness	Grit
years of educ	0.342 [3.38]***	0.170 [2.66]***	0.163 [2.38]**	0.175 [2.27]**	0.049 [0.86]	0.345 [4.42]***	0.328 [3.92]***
work_exp	-0.084 [5.01]***	-0.013 [0.96]	-0.000 [0.01]	0.011 [0.82]	-0.013 [1.04]	-0.043 [2.72]***	-0.006 [0.37]
work_exp_sq	0.002 [3.49]***	0.001 [1.66]*	0.001 [1.30]	0.000 [0.82]	0.000 [0.68]	0.002 [3.64]***	0.001 [2.36]**
gender	-0.029 [0.37]	-0.119 [1.85]*	-0.074 [1.20]	0.000 [0.00]	-0.527 [9.42]***	-0.201 [2.67]***	-0.156 [2.05]**
mother_educ_below_primary	0.091 [0.48]	0.054 [0.31]	-0.048 [0.29]	-0.074 [0.45]	0.013 [0.09]	0.297 [1.28]	0.219 [1.03]
mother_educ_lower_secondary	-0.334 [1.36]	-0.216 [1.18]	-0.312 [1.66]*	-0.356 [1.81]*	0.019 [0.12]	-0.423 [1.80]*	-0.572 [2.41]**
mother_educ_upper_secondary	-0.325 [1.25]	-0.186 [0.97]	-0.481 [2.43]**	-0.315 [1.52]	0.072 [0.41]	-0.515 [2.08]**	-0.592 [2.38]**
mother_educ_tertiary	-0.431 [1.14]	-0.199 [0.76]	-0.629 [2.33]**	-0.419 [1.46]	0.027 [0.11]	-0.594 [1.85]*	-0.957 [2.89]***
father_educ_lower_secondary	0.088 [0.76]	-0.117 [1.20]	-0.010 [0.10]	-0.062 [0.62]	0.099 [1.21]	-0.244 [2.02]**	-0.295 [2.53]**
father_educ_upper_secondary	0.044 [0.25]	-0.259 [1.86]*	0.024 [0.18]	-0.184 [1.28]	0.145 [1.28]	-0.334 [2.10]**	-0.492 [3.03]***
father_educ_tertiary	-0.085 [0.25]	-0.491 [2.18]**	0.011 [0.05]	-0.493 [1.96]**	-0.082 [0.42]	-0.759 [2.94]***	-0.794 [2.93]***
barranquilla	-0.173 [0.96]	0.158 [1.22]	0.040 [0.27]	-0.014 [0.09]	-0.183 [1.33]	-0.417 [2.27]**	-0.333 [1.71]*
bello	0.002 [0.01]	-0.265 [1.31]	0.214 [1.10]	-0.169 [0.69]	-0.117 [0.69]	-0.053 [0.20]	-0.022 [0.09]
bucaramanga	-0.104 [0.44]	-0.361 [1.66]*	-0.269 [1.52]	-0.119 [0.65]	-0.282 [1.79]*	-0.588 [2.97]***	-0.767 [3.52]***
cali	0.055 [0.43]	-0.089 [0.84]	-0.026 [0.27]	0.286 [2.83]***	-0.064 [0.68]	-0.143 [1.14]	-0.263 [2.13]**
copacabana	-0.667 [1.47]	-0.746 [1.04]	-1.670 [7.57]***	-0.489 [1.87]*	-0.535 [2.39]**	-1.422 [4.61]***	-0.613 [2.58]***

cucuta	0.372	0.125	-0.117	-0.053	-0.126	0.115	0.050
	[2.34]**	[0.94]	[0.76]	[0.37]	[1.06]	[0.61]	[0.28]
envigado	-0.273	-0.456	-0.051	-0.121	-0.052	-0.632	0.224
	[1.03]	[2.55]**	[0.25]	[0.50]	[0.27]	[2.76]***	[0.93]
floridablanca	0.438	-0.040	-0.374	-0.367	-0.081	-0.351	-0.556
	[1.92]*	[0.24]	[2.00]**	[1.99]**	[0.46]	[1.31]	[2.03]**
giron	-0.115	-0.171	0.527	0.188	0.521	-0.313	-0.299
	[0.44]	[0.99]	[1.86]*	[0.73]	[2.43]**	[0.70]	[0.79]
ibague	0.125	0.309	0.149	0.347	0.024	-0.014	-0.048
	[0.79]	[2.40]**	[1.16]	[3.08]***	[0.19]	[0.09]	[0.30]
itagui	-0.191	-0.165	0.032	0.035	-0.195	0.256	0.503
	[0.50]	[0.42]	[0.11]	[0.13]	[0.91]	[0.68]	[1.19]
la_estrella	-0.747	-0.200	0.051	-1.045	-0.489	0.407	1.442
	[2.32]**	[0.55]	[0.16]	[3.02]***	[2.73]***	[1.63]	[5.19]***
manizales	-0.003	-0.035	0.029	0.002	-0.000	-0.436	-0.468
	[0.01]	[0.19]	[0.14]	[0.01]	[0.00]	[1.89]*	[1.88]*
medellin	-0.355	-0.300	-0.067	-0.029	0.075	-0.371	0.105
	[2.28]**	[2.18]**	[0.56]	[0.24]	[0.71]	[2.54]**	[0.74]
piedecuesta	0.164	0.741	0.044	0.013	-0.429	0.568	-0.409
	[0.24]	[2.50]**	[0.11]	[0.03]	[2.21]**	[1.38]	[0.50]
soledad	-0.204	0.163	0.016	0.202	0.140	-0.154	-0.056
	[0.95]	[0.85]	[0.09]	[1.09]	[0.78]	[0.74]	[0.24]
villamaria	-0.462	0.452	0.233	0.044	0.287	0.337	-0.789
	[1.24]	[1.69]*	[0.52]	[0.35]	[0.73]	[0.85]	[5.51]***
villavicencio	-0.120	-0.052	0.013	-0.328	-0.155	-0.208	0.086
	[0.58]	[0.40]	[0.10]	[1.69]*	[1.43]	[1.13]	[0.46]
constant	-2.930	-1.524	-1.619	-1.852	-0.248	-2.976	-3.018
	[3.36]***	[2.66]***	[2.74]***	[2.79]***	[0.51]	[4.29]***	[4.09]***
R^2	0.11	-0.13	-0.03	-0.11	0.12	-0.64	-0.65
N	1,245	1,245	1,245	1,245	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Nota: Se toma como IV tiempo al colegio, desempeño escolar y nivel socioeconómico.

Para el caso de Bolivia, la **Tabla 18** reporta los resultados de la primer parte de nuestro procedimiento, estimando nuevamente mediante *LIML*. Para el caso de Bolivia, utilizamos como variables instrumentales tiempo al colegio, desempeño y esfuerzo realizado en el colegio¹¹; dado que estos resultan instrumentos robustos y válidos dados nuestra serie de test. Años de educación se estima para todos los casos en forma positiva y altamente significativa tanto para las habilidades cognitivas como para los no-cognitivas. Las habilidades que mayor magnitud muestran son nuestro indicador agregado cognitivo, donde un año más de educación aumenta éste en 0.322 desvío estándar adicional; extraversión y neuroticismo, donde un año adicional de educación genera una variación de 0.218 y 0.148 desvíos estándar respectivamente. Mientras que apertura al cambio, concientización y afabilidad estiman magnitudes menores variando entre 0.087 y 0.143 y de menor significancia estadística, siendo que el rasgo de perseverancia de esfuerzo Grit resulta altamente significativo pero con un coeficiente estimado de 0.130.

Obtenidos entonces los residuos de cada una de nuestras habilidades medidas mediante la primer parte de nuestro procedimiento al utilizar los distintos sets de instrumentos para capturar los efectos directos de cada una de dichos parámetros cognitivos y no cognitivos, procedemos con la segunda parte en la cual se utilizan dichos residuos como regresoras en nuestra estimación de retornos. Por tanto, en el primer paso obtenemos lo que suponemos corresponde al efecto directo que cada habilidad cognitiva y no cognitiva posee respecto a los retornos estimados. Mediante mínimos cuadrados ordinarios estimamos el *log* de retornos por hora respecto de años de educación y nuestro conjunto de residuos de habilidades, corrigiendo entonces por problemas potenciales de endogeneidad en Colombia y Bolivia.

Cabe recordar que en el contexto de las regresiones instrumentales (tal es el caso al utilizar *ivregress2* en *Stata*), el R^2 no tiene realmente un significado estadístico. Para mayores especificaciones puede remitirse a *J.M. Wooldridge (2015)*. Pero vale mencionar que dichos parámetros no presentan problema alguno para nuestras estimaciones y cálculos posteriores.

Así también, es importante remarcar la utilización del estimador de máxima verosimilitud de información limitada (*LIML*) dado que el mismo evita el sesgo de sobre-especificación de efectos constantes en los modelos. El mismo provee la misma distribución asintótica de *2SLS*, pero presenta una reducción finita muestral de dicho sesgo. Para mayores referencias, remitirse a *Hausman, Newey, Chao, Woutersen & Swanson (2007)* o *Davidson & MacKinnon (2001)*.

¹¹ Como mencionamos anteriormente, tomamos a la variable de seguimiento y apoyo de los padres durante el periodo escolar como un proxy de la variable esfuerzo realizado en el colegio.

Tabla 18: Regresiones instrumentales de las habilidades cognitivas y no cognitivas respecto al log de ingresos laborales por hora para Bolivia

	Aggregate Cognitive Measure	Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Openness	Grit
years of educ	0.322 [5.03]***	0.218 [3.91]***	0.087 [1.93]*	0.143 [3.04]***	0.148 [2.64]***	0.115 [2.51]**	0.130 [2.81]***
work_exp	-0.123 [5.80]***	-0.069 [3.63]***	-0.003 [0.20]	0.008 [0.52]	-0.027 [1.48]	-0.013 [0.80]	-0.009 [0.58]
work_exp_sq	0.003 [4.38]***	0.003 [3.73]***	0.001 [1.06]	0.000 [0.79]	0.001 [1.77]*	0.001 [1.18]	0.001 [1.74]*
gender	-0.066 [0.81]	0.091 [1.29]	-0.125 [2.03]**	0.053 [0.85]	-0.440 [6.87]***	-0.038 [0.63]	0.030 [0.48]
mother_educ_below_primary	0.106 [0.84]	-0.011 [0.09]	-0.190 [1.83]*	-0.010 [0.09]	-0.042 [0.42]	-0.001 [0.01]	-0.084 [0.80]
mother_educ_lower_secondary	0.086 [0.43]	-0.067 [0.39]	-0.276 [1.91]*	-0.079 [0.50]	-0.076 [0.50]	-0.176 [1.19]	-0.406 [2.80]***
mother_educ_upper_secondary	0.363 [2.25]**	0.157 [1.10]	-0.316 [2.59]***	0.061 [0.48]	-0.152 [1.23]	-0.075 [0.63]	-0.243 [1.92]*
mother_educ_tertiary	0.157 [0.71]	0.012 [0.06]	-0.361 [2.15]**	0.041 [0.24]	-0.223 [1.22]	-0.222 [1.33]	-0.368 [2.16]**
father_educ_lower_secondary	-0.222 [1.27]	0.019 [0.13]	-0.032 [0.25]	-0.210 [1.56]	-0.181 [1.33]	-0.010 [0.08]	-0.300 [2.24]**
father_educ_upper_secondary	-0.160 [1.01]	-0.299 [2.28]**	-0.044 [0.41]	-0.148 [1.30]	-0.067 [0.56]	-0.006 [0.06]	-0.156 [1.34]
father_educ_tertiary	-0.599 [2.35]**	-0.564 [2.65]***	-0.163 [0.90]	-0.428 [2.35]**	-0.284 [1.31]	-0.030 [0.17]	-0.333 [1.82]*
aymara_quechua	0.264 [0.83]	-0.122 [0.46]	0.016 [0.07]	0.089 [0.38]	0.311 [1.28]	-0.105 [0.53]	0.244 [1.27]
la_paz	-0.333 [2.46]**	-0.291 [2.55]**	-0.067 [0.70]	-0.087 [0.88]	-0.213 [2.22]**	-0.215 [2.21]**	-0.153 [1.54]
el_alto	-0.047 [0.36]	-0.084 [0.75]	-0.210 [2.16]**	-0.161 [1.66]*	-0.020 [0.22]	-0.263 [2.69]***	-0.226 [2.39]**
santa_cruz	-0.192 [1.47]	-0.055 [0.51]	-0.048 [0.52]	-0.105 [1.11]	0.033 [0.35]	-0.407 [4.28]***	-0.236 [2.52]**
constant	-2.916 [4.40]***	-2.184 [3.85]***	-0.750 [1.57]	-1.758 [3.63]***	-1.295 [2.28]**	-1.046 [2.16]**	-1.258 [2.62]***
R ²	0.00	-0.31	0.00	-0.05	-0.08	0.01	-0.06
N	1,193	1,193	1,193	1,193	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Nota: Se toma como IV tiempo al colegio, desempeño y esfuerzo realizado en el colegio.

La **Tabla 19** reporta nuestras regresiones finales para Colombia, donde la columna 1 corresponde a nuestra regresión por *MCO* original; las columnas 2 a 4 utilizan el procedimiento en dos partes pero sólo tomando un instrumento (tiempo al colegio, desempeño escolar y nivel socioeconómico, respectivamente) para obtener nuestros residuos de habilidades; y finalmente la columna 5 utiliza nuestro procedimiento en dos partes tomando todos los instrumentos válidos dados nuestro set de test estadísticos. En todos los casos, se mantienen las relaciones encontradas entre las distintas regresiones, variando las magnitudes de nuestros coeficientes según los instrumentos que se utilicen y en algunos casos variando su nivel de significatividad. Dentro de los coeficientes que resultan significativos, la variable años de educación encuentre su piso en 0.017 al utilizar como instrumento nivel socioeconómico, y su techo en 0.031 al utilizar como instrumento tiempo de viaje hasta el colegio. Por su parte, nuestra medida cognitiva agregada varía entre 0.055 y 0.064. Respecto a las habilidades no cognitivas, sigue resultando interesante resaltar las estimaciones reportadas para el rasgo de perseverancia de *Grit* en todos los casos dado que la misma se estima significativa al 5% y al 10%, y reporta una relación negativa de entre 0.049 y 0.069 por cada desvío estándar.

Tabla 19: Estimaciones MCO del segundo paso del retorno a años de educación, habilidades cognitivas y no cognitivas para Colombia

	(1)	(2)	(3)	(4)	(5)
Regresión:	MCO	En dos pasos	En dos pasos	En dos pasos	En dos pasos
IV:		Tiempo al Colegio	Desempeño escolar	Nivel Socioeconómico	Set Completo
years of educ	0.027 [2.53]**	0.031 [2.43]**	0.030 [2.03]**	0.017 [1.16]	0.029 [2.00]**
aggregate cognitive measure	0.060 [2.26]**	0.064 [2.31]**	0.058 [2.09]**	0.055 [2.15]**	0.061 [2.20]**
extraversion	-0.010 [0.30]	-0.020 [0.52]	-0.020 [0.53]	-0.005 [0.14]	-0.019 [0.49]
agreeableness	0.008 [0.22]	0.002 [0.06]	-0.006 [0.16]	-0.014 [0.37]	-0.004 [0.11]
conscientiousness	0.006 [0.16]	0.013 [0.29]	0.004 [0.09]	0.002 [0.05]	0.006 [0.14]
emotional stability	-0.026 [0.67]	-0.035 [0.92]	-0.032 [0.79]	-0.034 [0.85]	-0.033 [0.81]
openness	0.020 [0.54]	0.020 [0.59]	0.025 [0.69]	0.044 [1.21]	0.026 [0.73]
grit	-0.063 [2.04]**	-0.061 [1.91]*	-0.069 [2.17]**	-0.049 [1.72]*	-0.067 [2.15]**
R^2	0.10	0.04	0.04	0.04	0.04
N	1,245	1,245	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

En efecto, para el caso de Bolivia, la **Tabla 20** reporta los resultados de las regresiones finales al estimar los *log* retornos laborales por hora respecto a años de educación y los residuos obtenidos de las habilidades cognitivas y socio-emocionales luego del primer paso del procedimiento econométrico implementado. La primera columna replica los resultados obtenidos en nuestra regresión inicial de *Mincer* que nos servirá de parámetro para las demás regresiones. Las columnas subsiguientes muestran las estimaciones utilizando sólo uno de los instrumentos por cada regresión (tiempo al colegio, desempeño escolar y esfuerzo realizado en el colegio; respectivamente), para terminar en la columna 5 utilizando todas las variables instrumentales en las regresiones del primer paso. Estadísticamente, años de educación mantiene su nivel de significancia y su magnitud aumenta al utilizar nuestro procedimiento al incluir un instrumento pero no cuando se utiliza el set completo, capturando los efectos directos de dicha variable y estimando desde 0.055 hasta 0.062 por año de educación adicional.

Por su parte nuestro indicador cognitivo agregado se mantiene significativo, aunque con variaciones en sus niveles críticos, estimando retornos desde 0.042 hasta 0.10 por desvío estándar. Dicha magnitud de nuestro agregado cognitivo presenta variaciones significativas en las magnitudes de los coeficientes estimados mediante nuestro procedimiento y los efectos directos de nuestro agregado cognitivo solo aumentan su magnitud al utilizar los residuos de nuestro indicador mediante regresiones instrumentales.

Desde el punto de vista no cognitivo, el factor de estabilidad emocional o neuroticismo resulta significativo al 10% al utilizar el set completo de instrumentos, estimando un retorno negativo de 0.044 por desvío estándar. Resulta interesante observar que el mercado laboral en Bolivia parece pricear de forma negativa las aptitudes más relajadas y tranquilas de los empleados. Seguramente estas habilidades se encuentren correlacionadas respecto a otros aspectos laborales que no resultan implícitamente desarrolladas en nuestro instrumento de recolección de datos utilizado. El resto de nuestras dimensiones no cognitivas no presentan resultados de interés desde un punto de vista estadístico, estimando que ninguno de los coeficientes resulta diferente de cero.

Tabla 20: Estimaciones MCO del segundo paso del retorno a años de educación, habilidades cognitivas y no cognitivas para Bolivia

	(1)	(2)	(3)	(4)	(5)
Regresión:	MCO	En dos pasos	En dos pasos	En dos pasos	En dos pasos
IV:		Tiempo al Colegio	Desempeño Escolar	Esfuerzo Escolar	Set Completo
year of educ	0.059 [3.89]***	0.060 [5.10]***	0.059 [3.73]***	0.062 [3.95]***	0.055 [7.99]***
aggregate cognitive measure	0.085 [2.44]**	0.097 [1.76]*	0.100 [2.18]**	0.096 [2.00]**	0.042 [2.01]**
extraversion	-0.020 [0.49]	-0.057 [1.21]	-0.057 [1.22]	-0.060 [1.33]	0.018 [0.74]
agreeableness	0.009 [0.21]	-0.009 [0.17]	-0.009 [0.18]	-0.009 [0.18]	0.015 [0.59]
conscientiousness	0.014 [0.31]	0.050 [0.69]	0.050 [0.72]	0.048 [0.70]	-0.027 [1.05]
emotional stability	-0.063 [1.36]	-0.041 [0.64]	-0.038 [0.58]	-0.040 [0.61]	-0.044 [1.76]*
openness	-0.027 [0.67]	-0.061 [1.07]	-0.059 [1.10]	-0.060 [1.03]	-0.013 [0.49]
grit	-0.033 [0.75]	-0.079 [1.56]	-0.078 [1.55]	-0.080 [1.56]	-0.009 [0.34]
R^2	0.27	0.14	0.14	0.14	0.07
N	1,193	1,193	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Como análisis adicional podríamos estimar nuevamente los residuos del primer paso del procedimiento al utilizar, para cada país, las variables instrumentales que no superaron los test estadísticos propuestos de instrumentos débiles e identificación. Una vez estimados los residuos de las habilidades cognitivas y no cognitivas, eliminando así el problema de endogeneidad potencial de nuestra ecuación, reestimamos las regresiones sobre el *log* retorno en el segundo paso de nuestro procedimiento. Para Colombia, la **Tabla 21** reporta las estimaciones del segundo paso al utilizar como variable instrumental esfuerzo en el colegio en el primer paso. Nuevamente la primera columna muestra las estimaciones de nuestra primera regresión para comparación, mientras que la segunda columna utiliza como *IV* para escolaridad únicamente esfuerzo escolar y la tercera columna utiliza a todo el conjunto de *IV*. Resulta interesante observar que al utilizar dicho instrumento, años de educación no resulta significativa. Así también, por el lado de las habilidades cognitivas y no cognitivas encontramos resultados similares a los hallados en las regresiones con las *IV* originales, pero con magnitudes menores en algunos casos.

Tabla 21: Segundo paso alternativo del retorno a años de educación, habilidades cognitivas y no cognitivas usando esfuerzo como IV para escolaridad en el primer paso en Colombia

	(1)	(2)	(3)
Regresión:	MCO	En dos pasos (Alternativo)	En dos pasos (Alternativo)
IV:		Esfuerzo Escolar	Set Completo
years of educ	0.027 [2.53]**	0.024 [1.35]	0.028 [1.95]*
aggregate cognitive measure	0.060 [2.26]**	0.065 [2.32]**	0.061 [2.21]**
extraversion	-0.010 [0.30]	-0.014 [0.37]	-0.019 [0.50]
agreeableness	0.008 [0.22]	-0.004 [0.11]	-0.004 [0.11]
conscientiousness	0.006 [0.16]	0.007 [0.21]	0.005 [0.13]
emotional stability	-0.026 [0.67]	-0.033 [0.88]	-0.032 [0.81]
openness	0.020 [0.54]	0.030 [0.74]	0.027 [0.73]
grit	-0.063 [2.04]**	-0.063 [1.99]**	-0.067 [2.14]**
R^2	0.10	0.04	0.04
N	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Consecuentemente, la **Tabla 22** reporta las estimaciones alternativas para Bolivia. En este caso se utiliza como variable instrumental al nivel socioeconómico de la familia del individuo al momento de terminar el colegio para el primer paso del procedimiento obteniendo así los residuos de las habilidades cognitivas y no cognitivas, y eliminando el problema de endogeneidad potencial de años de educación. Se procede entonces con el segundo paso alternativo, empleando únicamente nivel socioeconómico y luego empleando las 4 variables instrumentales descritas, agregando en la columna 1 nuestra regresión inicial para control. En la columna 2, utilizando únicamente a nivel socioeconómico como IV, no advertimos variaciones en lo que respecta a la variable años de educación. El indicador cognitivo agregado resulta significativamente mayor hasta 0.105, mientras que el conjunto de las habilidades no cognitivas no estiman retornos distintos a cero estadísticamente. Esto se mantiene al incluir el set completo de instrumentos y regresar con sus residuos, donde los coeficientes y sus niveles de confianza resultan similares a los estimados en el segundo paso original y pudiendo apreciar estos en la columna 3. Podría remarcarse que nuestro indicador agregado cognitivo estima un retorno por cada desvío estándar menor respecto a la columna 2, pero mayor a lo estimado mediante la regresión de *MCO* original. Así también, del punto de vista socioemocional, estabilidad emocional pierde magnitud y no alcanza niveles de confianza significativos.

Tabla 22: Segundo paso alternativo del retorno a años de educación, habilidades cognitivas y no cognitivas usando nivel socioeconómico como IV para escolaridad en el primer paso para Bolivia

	(1)	(2)	(3)
Regresión:	MCO	En dos pasos (Alternativo)	En dos pasos (Alternativo)
IV:		Nivel Socioeconómico	Set Completo
years of educ	0.059 [3.89]***	0.060 [3.51]***	0.060 [4.09]***
aggregate cognitive measure	0.085 [2.44]**	0.105 [2.12]**	0.099 [2.03]**
extraversion	-0.020 [0.49]	-0.039 [1.02]	-0.057 [1.26]
agreeableness	0.009 [0.21]	-0.003 [0.06]	-0.009 [0.18]
conscientiousness	0.014 [0.31]	0.028 [0.49]	0.050 [0.71]
emotional stability	-0.063 [1.36]	-0.028 [0.41]	-0.039 [0.58]
openness	-0.027 [0.67]	-0.077 [1.11]	-0.060 [1.07]
grit	-0.033 [0.75]	-0.064 [1.44]	-0.079 [1.56]
R^2	0.27	0.14	0.14
N	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

b. Sesgo de Selección

Una vez tratado el problema potencial de endogeneidad mediante el procedimiento propuesto por *Vera Tudela, Díaz & Arias* (2012), debemos tratar el posible problema de sesgo de selección. Para ello corregiremos nuestro análisis mediante el modelo de selección de *Heckman* (1976) en el cual la variable dependiente de la regresión sólo es observada dada cierta matriz de variables z adicional a un término de error, y donde el término error entre el modelo de regresión lineal de la variable y el término de error de la ecuación de selección debe ser 0. Cuando dicho coeficiente es distinto a cero, nos encontramos con un error sistémico y por tanto los resultados de la regresión se encuentran sesgados. Heckman explica que al intentar regresar el salario de las mujeres, se está suponiendo que ellas han tenido la opción aleatoria de trabajar o no y por ende se puede tomar sólo una muestra de la población para estimar un modelo de retornos salariales. Sin embargo, este supuesto de participación aleatorio puede no darse dada ciertas condiciones coyunturales, culturales, económicas y/o sociales estimando coeficientes por encima de los poblaciones. Esto es, las mujeres prefieren no trabajar dadas ciertas preferencias racionales y restricciones que no generan un nivel de utilidad en salarios suficiente como para ingresar al ámbito laboral. En consecuencia dichas mujeres que eligen no trabajar pueden tener salarios ofrecidos más altos que aquellas mujeres que eligen trabajar, pero sus salarios de reserva son aún mayores que estos salarios ofrecidos. Siguiendo el modelo de Heckman,

inicialmente propuesto por *Gronau* (1974) y *Lewis* (1974), se pueden obtener estimaciones robustas y asintóticamente eficientes sólo si existen una serie de variables que afecten las probabilidades del salario de reserva pero no del salario ofrecido por el empleador (*id est* las probabilidades de estar empleada).

Entonces se estima el modelo de Heckman mediante un procedimiento en dos pasos para Colombia y Bolivia respectivamente, con el objetivo de determinar si nuestra muestra presenta problemas de sesgo de selección. En el primer paso estimamos la probabilidad de que el individuo se encuentre empleado, como función de las variables iniciales de control y una variable adicional de identificación mediante un modelo *Probit*. Asumimos que esta variable adicional afecta negativamente la probabilidad estar empleado, pero no influye en el salario ofrecido estando empleado. Luego, calculamos el ratio inverso de Mills siguiendo a *Heckman* (1979) para luego incluir *lambda* en nuestra regresión original y observamos su significancia para entender si factores no observables que aumentan la probabilidad de participación en el empleo se asociación con mayores, menores o invariantes retornos. En nuestro caso utilizaremos las variables dicotómicas hijos en la familia y estado civil como variables de identificación, y aplicaremos el procedimiento descrito para conocer si es necesario corregir nuestro análisis dada la presencia de error sistémico testeado en Colombia. Los resultados se reportan en la **Tablas 23**, donde se implementa el modelo de Heckman en dos pasos sobre nuestra regresión de residuos de habilidades obtenidas en nuestro procedimiento econométrico anterior. Inicialmente podemos observar que, de las variables de selección, *married* resulta estadísticamente significativa al 5% y posee una relación negativa respecto a la probabilidad de estar empleado. Aun así el ratio inverso de Mills (*lambda*) no resulta significativo, donde dicho coeficiente captura el efecto esperado del error en la ecuación de ingresos laborales condicionales a estar empleado.

Tabla 23: Estimaciones en dos pasos del modelo de selección de Heckman para modelos de regresión con selección muestral en Colombia

Heckman residuos en dos pasos		
Log of hourly usd earnings	years of educ	0.083 [6.34]***
	resid aggregate cog measure	0.103 [2.81]***
	resid extraversion	0.049 [1.40]
	resid agreeableness	-0.024 [0.63]
	resid conscientiousness	0.034 [0.71]
	resid emotional stability	-0.024

		[0.63]
	resid openness	0.014
		[0.39]
	resid grit	-0.077
		[2.13]**
	constant	0.118
		[0.33]
emp	has_children	0.311
		[4.36]***
	married	-0.178
		[2.09]**
	years of educ	0.037
		[3.73]***
	resid aggregate cog measure	-0.084
		[2.77]***
	resid extraversion	-0.028
		[0.82]
	resid agreeableness	-0.052
		[1.47]
	resid conscientiousness	0.129
		[3.61]***
	resid emotional stability	0.043
		[1.19]
	resid openness	-0.003
		[0.09]
	resid grit	-0.001
		[0.02]
	constant	-0.354
		[3.03]***
mills	lambda	0.083
		[0.22]
N		1,470

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Nota: has_children y married resulta como variables de identificación y emp corresponde a una variable dicotómica que toma el valor de 1 si el individuo se encuentra empleado y 0 de otra manera.

Para cerciorarnos, realizamos el procedimiento en dos pasos de forma manual para poder así ponderar cada observación por su representación de la población en estudio y adicionalmente aplicamos el mismo procedimiento pero utilizando la regresión de Mincer como ecuación de salarios.

Tabla 24: Estimación alternativas en dos pasos del modelo de selección de Heckman para modelos de regresión con selección muestral en Colombia

	Heckman MCO	Heckman residuos
IMR	0.341	0.713
	[0.65]	[1.20]

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

En todos los casos obtenemos resultados similares a los encontrados anteriormente, pudiendo apreciar los mismos en las tablas 24 y 25. Con lo cual, nuestra muestra obtenida de la base de datos *STEP* pareciera ser consistente y no presenta sesgo de selección en Colombia. Por lo tanto puede suponerse que se mantiene la opción de aleatoriedad de trabajar en dicha población y por ende es correcto el análisis muestral para la estimación de retornos salariales.

Tabla 25: Estimación alternativas en dos pasos del modelo de selección de Heckman para modelos de regresión con selección muestral en Colombia

		Heckman con modelo de Mincer	
Log of hourly usd earnings	years of educ	0.080 [3.83]***	
	aggregate cog measure	0.072 [1.39]	
	extraversion	0.044 [1.14]	
	agreeableness	-0.054 [1.10]	
	conscientiousness	0.087 [1.25]	
	emotional stability	-0.011 [0.25]	
	openness	0.010 [0.26]	
	grit	-0.081 [2.12]**	
	work_exp	0.045 [1.61]	
	work_exp_sq	-0.001 [1.30]	
	mother_educ_below_primary	-0.468 [2.67]***	
	mother_educ_lower_secondary	-0.368 [2.51]**	
	mother_educ_upper_secondary	-0.252 [1.46]	
	mother_educ_tertiary	-0.292 [1.20]	
	father_educ_lower_secondary	0.019 [0.19]	
	father_educ_upper_secondary	0.100 [0.76]	
	father_educ_tertiary	0.300 [1.64]	
	constant	-0.478 [0.58]	
	emp	has_children	-0.024 [0.30]
		married	-0.224 [2.49]**
		years of educ	0.045 [3.66]***
		aggregate cog measure	-0.102 [3.21]***
		extraversion	-0.027

		[0.77]
	agreeableness	-0.075
		[2.07]**
	conscientiousness	0.151
		[4.16]***
	emotional stability	0.046
		[1.27]
	openness	-0.004
		[0.12]
	grit	-0.007
		[0.20]
	work_exp	0.069
		[7.38]***
	work_exp_sq	-0.001
		[7.31]***
	mother_educ_below_primary	0.166
		[1.10]
	mother_educ_lower_secondary	0.005
		[0.04]
	mother_educ_upper_secondary	0.039
		[0.25]
	mother_educ_tertiary	0.032
		[0.15]
	father_educ_lower_secondary	-0.039
		[0.42]
	father_educ_upper_secondary	-0.079
		[0.67]
	father_educ_tertiary	0.004
		[0.03]
	constant	-0.735
		[4.03]***
<hr/>		
mills	lambda	0.733
		[1.11]
<hr/>		
N		1,470

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Nota: has_children y married resulta como variables de identificación y emp corresponde a una variable dicotómica que toma el valor de 1 si el individuo se encuentra empleado y 0 de otra manera.

Por otro lado, testeamos en Bolivia mediante el mismo modelo de Heckman para chequear la necesidad de corregir nuestro análisis dado la presencia potencial de sesgo de selección. La **Tabla 26** reporta los resultados, donde se implementa el modelo de Heckman en dos pasos sobre nuestra regresión de residuos de habilidades obtenidas en nuestro procedimiento econométrico mediante instrumentos capturando los efectos directos de dichas habilidades.¹² Sin embargo, ninguna de nuestras variables de identificación cumplen con los supuestos necesarios del modelo de Heckman y por lo cual no podemos utilizar el *IMR* para testear la covarianza entre los errores de la ecuación de salarios y de la ecuación de selección dados los supuestos del modelo.

¹² Para mayores referencias respecto al método utilizado y sus aplicaciones en los distintos paquetes estadísticos se recomienda *Bushway* (2007).

Tabla 26: Estimaciones en dos pasos del modelo de selección de Heckman para modelos de regresión con selección muestral en Bolivia

Heckman residuos en dos pasos			
Log of hourly usd earnings	years of educ	0.078 [8.21]***	
	resid aggregate cog measure	0.124 [3.21]***	
	resid extraversion	-0.026 [0.67]	
	resid agreeableness	-0.018 [0.51]	
	resid conscientiousness	-0.003 [0.07]	
	resid emotional stability	0.006 [0.17]	
	resid openness	0.046 [1.17]	
	resid grit	-0.009 [0.25]	
	constant	0.319 [1.78]*	
	emp	has_children	0.659 [7.50]***
		married	-0.088 [1.00]
years of educ		0.025 [2.41]**	
resid aggregate cog measure		-0.150 [4.28]***	
resid extraversion		0.053 [1.33]	
resid agreeableness		-0.015 [0.38]	
resid conscientiousness		-0.053 [1.35]	
resid emotional stability		0.040 [1.02]	
resid openness		-0.000 [0.00]	
resid grit		0.021 [0.54]	
constant		-0.223 [1.68]*	
mills	lambda	-0.407 [1.75]*	
<i>N</i>		1,263	

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Nota: has_children y married resulta como variables de identificación y emp corresponde a una variable dicotómica que toma el valor de 1 si el individuo se encuentra empleado y 0 de otra manera.

Nuevamente aplicamos el modelo de Heckman pero utilizando la regresión de Mincer como ecuación de salarios y, en este caso, obtenemos resultados más interesantes que los reportados en la tabla

anterior. La **Tabla 27** presenta estimaciones del procedimiento en dos pasos de Heckman, donde ahora se cumplen los supuestos del modelo de Heckman. La variable de identificación *estado civil* resulta significativa y de relación negativa respecto a la probabilidad de ser mujer y encontrarse empleada, además de no presentar una relación estadísticamente significativa respecto al *log* de ingresos laboral por hora. Aun así, y al igual que en Colombia, el *IMR* no resulta significativo y por tanto no se puede rechazar la hipótesis nula de que nuestra muestra en Bolivia no presenta sesgo de selección.

Tabla 27: Estimación alternativa en dos pasos del modelo de selección de Heckman para modelos de regresión con selección muestral en Bolivia

Heckman con modelo de Mincer			
Log of hourly usd earnings	years of educ	0.063 [3.16]***	
	resid aggregate cog measure	0.052 [0.92]	
	resid extraversion	0.006 [0.15]	
	resid agreeableness	-0.030 [0.84]	
	resid conscientiousness	-0.025 [0.65]	
	resid emotional stability	0.029 [0.80]	
	resid openness	0.025 [0.66]	
	resid grit	0.005 [0.15]	
	work_exp	0.028 [1.69]*	
	work_exp_sq	-0.000 [1.43]	
	mother_educ_below_primary	-0.095 [0.83]	
	mother_educ_lower_secondary	0.202 [1.24]	
	mother_educ_upper_secondary	0.180 [1.12]	
	mother_educ_tertiary	0.476 [2.70]***	
	father_educ_lower_secondary	-0.204 [1.51]	
	father_educ_upper_secondary	-0.269 [2.57]**	
	father_educ_tertiary	-0.230 [1.70]*	
	constant	-0.044 [0.08]	
	emp	has_children	0.228 [2.13]**
		married	-0.269 [2.83]***
		years of educ	0.082 [6.91]***
		resid aggregate cog measure	-0.175 [4.53]***

	resid extraversion	0.076 [1.84]*
	resid agreeableness	-0.039 [0.98]
	resid conscientiousness	-0.038 [0.93]
	resid emotional stability	0.030 [0.74]
	resid openness	-0.008 [0.19]
	resid grit	0.042 [1.04]
	work_exp	0.052 [4.75]***
	work_exp_sq	-0.001 [4.13]***
	mother_educ_below_primary	0.081 [0.62]
	mother_educ_lower_secondary	-0.016 [0.08]
	mother_educ_upper_secondary	-0.238 [1.48]
	mother_educ_tertiary	-0.115 [0.60]
	father_educ_lower_secondary	0.194 [1.16]
	father_educ_upper_secondary	-0.056 [0.49]
	father_educ_tertiary	-0.195 [1.40]
	constant	-0.876 [4.82]***
mills	lambda	0.230 [0.48]
<i>N</i>		1,263

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Para poder tener en cuenta las ponderaciones por observación respecto a la población, realizamos el mismo análisis de Heckman en dos pasos y reportamos los resultados en la **Tabla 28**. La misma corrobora nuestros hallazgos, y por ende no se podría rechazar la hipótesis nula de que nuestra muestra en Bolivia no presenta sesgo de selección.

Tabla 28: Estimación alternativas en dos pasos del modelo de selección de Heckman para modelos de regresión con selección muestral en Bolivia

	Heckman MCO	Heckman residuos
IMR	-1.384 [1.85]*	-2.794 [2.94]***

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

6. Extensiones

a. Valor del diploma

En pos de utilizar características individuales de nuestra base de datos STEP, aplicamos una serie de extensiones que creemos pueden resultar de interés para replicar en otros países y profundizar en sus conceptos teóricos. En primer lugar planteamos un modelo de retornos respecto a la diferencia de poseer el título del estudio más alto alcanzado respecto a no poseerlo, donde individuos que posean su título en mano deberían obtener mayores ingresos dados ciertos incentivos de las empresas y ciertos rasgos en la personalidad de dichos individuos, *ceteris paribus* al respecto de las habilidades y variables de relevancia. Para ello generamos una serie de dummies por cada nivel educativo en cada país, excepto para el nivel más bajo para observar los valores marginales de las variables indicadoras.

La **Tabla 29** reporta los resultados en Colombia mediante mínimos cuadrados ordinarios respecto a diferentes set de covariables de control. La columna 1 estima el retorno en log del ingreso laboral por hora controlando por diploma del nivel más alto de estudio alcanzado, sumado a todas las variables de control de la *OLS* original en nuestro análisis inicial. La segunda columna reestima los resultados de la primer columna, pero tomando como set de covariables de control las habilidades residuales obtenidas en nuestro procedimiento en dos pasos donde sus instrumentos no poseen problemas de instrumentos débiles; mientras que la columna 3 reestima tomando las habilidades residuales incluyendo todas las variables instrumentales en nuestro procedimiento para obtener los residuos de las habilidades cognitivas y no cognitivas. En cuanto a las estimaciones en cada columna respecto a las demás covariables, las magnitudes y niveles de significancias de las habilidades cognitivas y no cognitivas varían por fuera de los rangos descriptos en las secciones anteriores, mientras que la variable años de educación aumenta considerablemente hasta 0.093 por cada año adicional de estudio, con un nivel de confianza del 99%. Por el lado cognitivo, nuestro indicador agregado reduce su magnitud y pierde su significancia estadística, mientras que por el lado de las habilidades no cognitivas, el rasgo de perseverancia de esfuerzo Grit aumenta su relación negativa y así también aumenta su nivel de confianza. Así también resulta significativa levemente la variable no cognitiva de apertura al cambio, estimando un aumento del retorno salarial en torno al 6.0% por cada desvío estándar adicional.

Tabla 29: Estimación MCO del valor del diploma respecto al nivel educativo más alto alcanzado en Colombia

Covariables de regresiones:	(1) MCO	(2) Residuos válidos	(3) Todos los residuos
general_hs	-0.155 [1.03]	-0.159 [0.91]	-0.159 [0.91]
vocational_hs	-0.229 [1.35]	-0.211 [1.10]	-0.211 [1.11]
teaching_hs	-0.394 [1.54]	-0.330 [1.16]	-0.330 [1.16]
technical	-0.135 [0.79]	-0.124 [0.64]	-0.124 [0.64]
technological	-0.202 [0.96]	-0.222 [0.99]	-0.221 [0.99]
bachelor	-0.032 [0.13]	-0.070 [0.27]	-0.069 [0.27]
specialization	0.571 [1.91]*	0.509 [1.65]*	0.510 [1.65]*
master	0.452 [1.60]	0.419 [1.47]	0.419 [1.47]
PhD	1.572 [4.30]***	1.736 [5.12]***	1.738 [5.11]***
years of educ	0.084 [2.54]**	0.092 [3.07]***	0.093 [3.08]***
aggregate cognitive measure	0.038 [1.28]	0.035 [1.23]	0.035 [1.22]
extraversion	-0.003 [0.10]	-0.008 [0.26]	-0.009 [0.27]
agreeableness	0.021 [0.70]	0.030 [0.90]	0.030 [0.90]
conscientiousness	0.029 [0.89]	0.047 [1.38]	0.046 [1.38]
emotional stability	0.033 [1.01]	0.031 [0.99]	0.031 [0.99]
openness	0.044 [1.28]	0.060 [1.76]*	0.060 [1.75]*
grit	-0.114 [3.62]***	-0.117 [3.74]***	-0.117 [3.74]***
R^2	0.18	0.16	0.16
N	906	906	906

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

En todos los casos, tanto poseer el diploma de *PhD* como de especialización resulta estadísticamente significativo y con una relación positiva respecto a los ingresos laborales. Tener el título de *PhD* estima un *log* retorno altamente significativo de entre 157.2% y 173,8%; mientras que poseer el diploma de especialización estima un *log* retorno entre el intervalo 50,9% y 57,1% por desvío estándar aunque con un nivel de confianza del 10%.

Por el lado de Bolivia, la **Tabla 30** reporta los resultados al incluir en nuestras distintas regresiones las variables dummies por cada nivel educativo siendo las mismas distintas a las presentadas en Colombia, especialmente en niveles superiores donde no se encuentran tantos estudios de diferentes

niveles de posgrado para los individuos presentes en nuestra muestra *STEP*. La primera columna presenta los resultados de nuestra regresión de Mincer, estimando que aquellos individuos que poseen diploma de especialidad, maestría o *Phd* generan en promedio mayores retornos en torno 68.3% con un nivel de confianza del 99%. Las columnas subsiguientes utilizan las regresiones de nuestro procedimiento econométrico descrito en la sección 5, aplicando como variables instrumentales sólo aquellas que resultan relevantes y luego aplicando el set completo de instrumentos. En ambos casos, el poseer diploma de estudios de posgrado se asocia a un 80,7% de mayores retornos, siendo nuevamente altamente significativo. Al parecer, nuestras dummies están captando parte del efecto que genera la variable años de educación en nuestras estimaciones y por tanto pierde sus niveles estadísticos. Por el lado de las habilidades cognitivas y no cognitivas, el indicador cognitivo agregado resulta significativo a niveles aceptables y su magnitud aumenta considerablemente en todos los casos analizados; mientras que nuevamente el rasgo de perseverancia de esfuerzo de *Grit* resulta significativo pero a niveles mucho menores y manteniendo su relación negativa. El resto de las variables no cognitivas se mantienen sin alcanzar niveles de significancia relevantes, como sucede en nuestras regresiones anteriores.

Tabla 30: Estimación MCO del valor del diploma respecto al nivel educativo más alto alcanzado en Bolivia

Covariables de regresiones:	(1) MCO	(2) Residuos válidos	(3) Todos los residuos
pre_univ_courses	-0.166 [1.21]	-0.004 [0.02]	-0.004 [0.02]
technical_courses	-0.160 [1.28]	-0.036 [0.15]	-0.036 [0.15]
technical_univ	-0.541 [2.84]***	-0.335 [1.18]	-0.335 [1.18]
bachelor	0.108 [0.60]	0.353 [1.11]	0.353 [1.11]
special_master_phd	0.683 [3.01]***	0.807 [2.45]**	0.807 [2.45]**
years of educ	0.031 [1.84]*	0.024 [0.75]	0.024 [0.75]
aggregate cognitive measure	0.083 [2.62]***	0.104 [2.33]**	0.104 [2.33]**
extraversion	-0.024 [0.66]	-0.053 [1.18]	-0.053 [1.17]
agreeableness	0.010 [0.25]	-0.016 [0.35]	-0.016 [0.35]
conscientiousness	0.016 [0.41]	0.045 [0.64]	0.045 [0.64]
emotional stability	-0.037 [0.94]	-0.000 [0.01]	-0.000 [0.01]
openness	-0.026 [0.64]	-0.051 [0.96]	-0.051 [0.96]
grit	-0.042 [1.11]	-0.087 [1.92]*	-0.087 [1.92]*
R^2	0.33	0.20	0.20
N	1,193	1,193	1,193

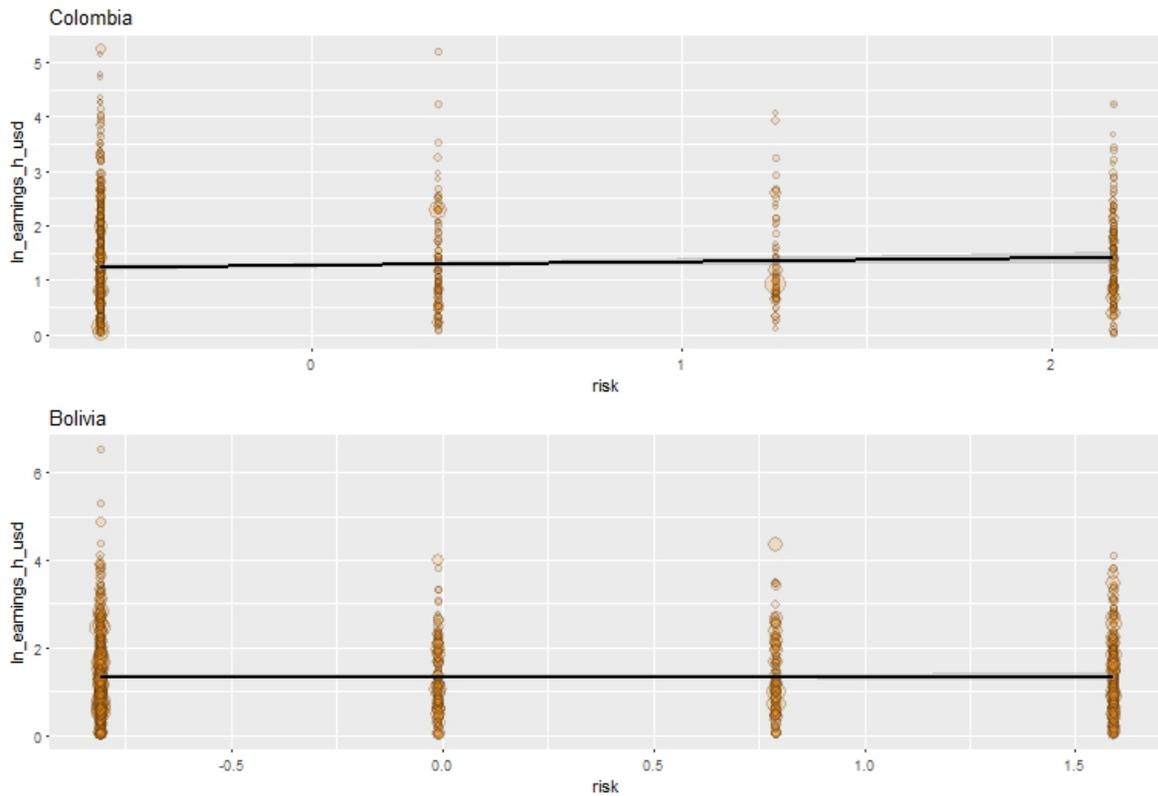
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

b. Retornos Laborales y Aversión al Riesgo

Como último análisis en el presente *paper* proponemos la inclusión de un rasgo adicional a nuestros modelos de estimaciones de retornos laborales, con la posible intersección de dicho rasgo con variables de modelos de dispersión salarial. Planteamos entonces la utilización de la aversión al riesgo como un proxy de rasgos de la personalidad asociados a la propensión de un individuo a realizar transiciones empleo-empleo, siguiendo modelos de dispersión salarial como son *Burdett Mortensen (1998)* y *Harris Todaro (1970)* focalizando en los retornos obtenidos por variaciones en dicha dimensión. En nuestro modelo propuesto suponemos que aquellos individuos que poseen mayor aversión al riesgo poseen ciertos rasgos y características de su personalidad que los llevan a reducir o evitar las transiciones empleo-empleo y por tanto sus ingresos deberían ser menores respecto a aquellos individuos con menor aversión al riesgo; como así también tomar una gama de decisiones que pueden o no generar mayores ingresos laborales. El rasgo de riesgo se construye de forma tal que a mayores niveles de dicho indicador, mayor es la propensión del individuo hacia el riesgo, y por ende menor es la aversión al riesgo. Siendo así, comenzamos analizando de forma gráfica la relación existente entre nuestro indicador de aversión al riesgo y el retorno logarítmico de los ingresos laborales por hora para Colombia y para Bolivia.

La **Figura 7** muestra el correlograma para Colombia y para Bolivia, respectivamente y ponderando por sus *weights* correspondientes a cada observación. Gráficamente observamos que pareciera existir cierta relación positiva entre el nivel del factor riesgo y el log de los ingresos laborales por hora en Colombia; observación que no resulta evidente para el caso de Bolivia.

Figure 7: Correlograma del logaritmo de retornos vs propensión al riesgo, en Colombia y Bolivia



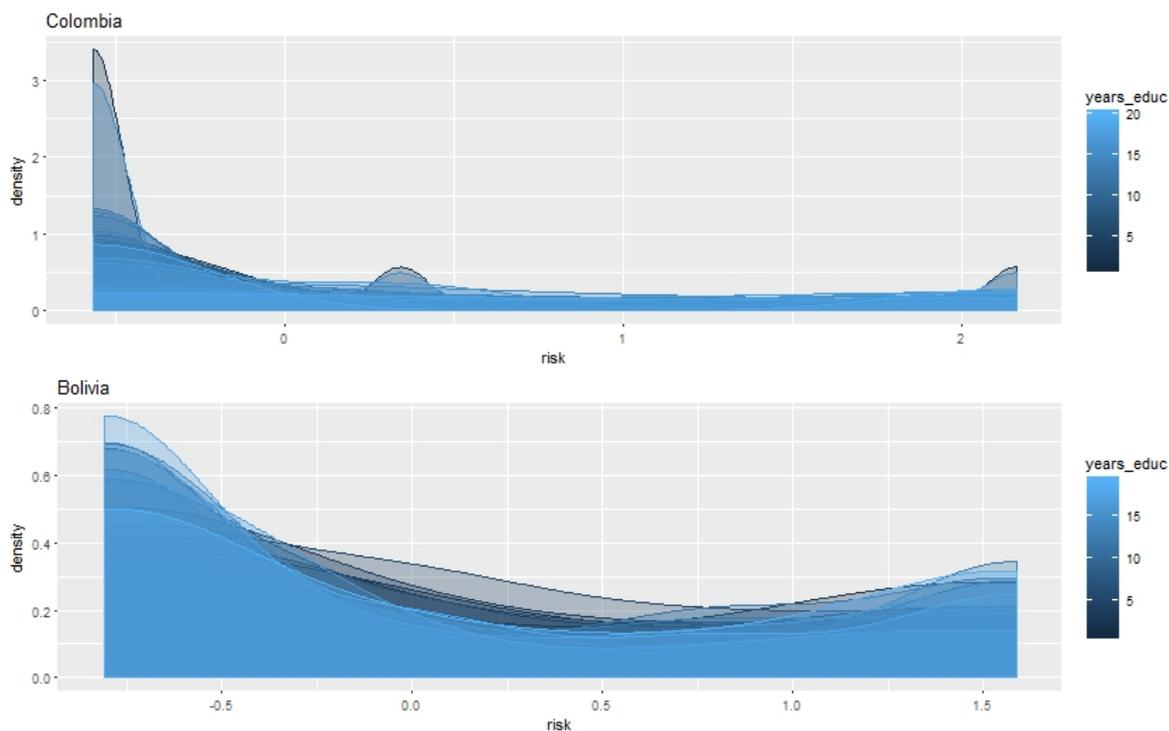
Por otro lado analizamos la matriz de correlaciones para ambos países, no encontrando coeficientes de correlación que supongan relación alguna entre la aversión al riesgo y nuestras habilidades cognitivas y no cognitivas. Así también los coeficientes en el caso de Bolivia respecto a la variable aversión al riesgo y dichas habilidades resultan negativos en la mayoría de los casos, pero con magnitudes que no resultan estadísticamente relevantes.¹³ Para un análisis más profundo en dicho sentido, se recomienda *Dohmen* (2010) donde se trata dicha relación junto con la variable adicional impaciencia.

Siendo así, cruzamos la información respecto de los niveles de aversión al riesgo en relación a la cantidad de años de educación en Colombia y Bolivia. Similar a lo reportado en los correlogramas anteriores, en Colombia se observa menores niveles de aversión al riesgo a mayor cantidad de años de educación del individuo. Lo opuesto sucede en Bolivia donde a mayor cantidad de años de

¹³ Tablas de correlaciones con la variable propensión al riesgo en el apéndice III.

educación, menor nivel de propensión al riesgo por parte del individuo. Dichos resultados pueden verse gráficamente para cada país en la **Figura 8**.

Figure 8: Gráficos de densidad de propensión al riesgo por cada año de educación, en Colombia y Bolivia



Continuando con nuestro análisis econométrico inicial para la variable aversión al riesgo, agregamos dicho rasgo a las regresiones originales controlando por experiencia laboral, genero, localización geográfica y educación de la madre y del padre. La **Tabla 31** reporta los resultados para Colombia, donde: la primer columna controla por el indicador cognitivo agregado junto al resto de covariables, la columna 2 controla adicionalmente por años de educación, la columna 3 reporta los resultados de controlar por el set completo de habilidades no cognitivas sin controlar por años de educación, mientras que la última columna sí lo hace. En todos los casos propuestos la variable de aversión al riesgo es significativa estadística y económicamente estimando un retorno asociado en torno al 6.8%-7.1%, y robusta a los diferentes controles de habilidades y covariables.

Tabla 31: Regresiones originales mediante MCO agregando como covariable propensión al riesgo a nuestros modelos en Colombia

	(1)	(2)	(3)	(4)
	Habilidades cognitivas sin años de educación	Habilidades cognitivas y años de educación	Habilidades no cognitivas sin años de educación	Habilidades no cognitivas y años de educación
risk	0.069 [1.73]*	0.068 [1.71]*	0.075 [2.00]**	0.071 [1.89]*
aggregate cognitive measure	0.080 [3.09]***	0.057 [2.12]**		
years of educ		0.027 [2.32]**		0.034 [3.31]***
extraversion			-0.015 [0.44]	-0.008 [0.25]
agreeableness			0.030 [0.86]	0.011 [0.31]
conscientiousness			0.010 [0.27]	0.010 [0.26]
emotional stability			-0.015 [0.40]	-0.023 [0.64]
openness			0.034 [0.93]	0.024 [0.66]
grit			-0.059 [1.91]*	-0.062 [2.02]**
R^2	0.09	0.10	0.09	0.10
N	1,245	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Así también, analizamos y presentamos los resultados de incluir la variable de aversión al riesgo en nuestras regresiones iniciales en Bolivia. A diferencia de Colombia, dicha variable parece estar relacionada de forma negativa respecto al *log* de los retornos salariales por hora. El set de regresiones es similar al realizado en la tabla anterior, con lo cual apreciamos que la variable de estudio no resulta significativa al controlar por las habilidades no cognitivas incluyendo o no en la regresión años de educación. No obstante, esta sí resulta significativa (aunque levemente) al controlar por el indicador cognitivo agregado y el resto de las covariables tanto en el caso de incluir años de educación como al no incluirla. En la **Tabla 32** se reportan los resultados obtenidos en cada caso según corresponde el análisis descripto.

Tabla 32: Regresiones originales mediante MCO agregando como covariable propensión al riesgo a nuestros modelos en Bolivia

	(1)	(2)	(3)	(4)
	Habilidades cognitivas sin años de educación	Habilidades cognitivas y años de educación	Habilidades no cognitivas sin años de educación	Habilidades no cognitivas y años de educación
risk	-0.091 [1.77]*	-0.082 [1.90]*	-0.063 [1.26]	-0.063 [1.54]
aggregate cognitive measure	0.156 [3.93]***	0.088 [2.56]**		
years of educ		0.058 [4.22]***		0.069 [4.98]***
extraversion			-0.016 [0.38]	-0.013 [0.31]
agreeableness			0.018 [0.40]	0.007 [0.16]
conscientiousness			0.031 [0.68]	0.005 [0.12]
emotional stability			-0.036 [0.68]	-0.046 [1.01]
openness			-0.027 [0.66]	-0.027 [0.67]
grit			0.000 [0.01]	-0.032 [0.73]
R^2	0.23	0.27	0.19	0.27
N	1,193	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Entonces, formalizando entonces nuestro modelo de retornos, habilidades cognitivas y no-cognitivas extendido, estandarizamos el rasgo de propensión al riesgo para incluirlo dentro de nuestra corrida de regresiones estándar del presente *paper*. La **Tabla 33** reporta los resultados de dicha extensión para el caso de Colombia.¹⁴ La primera columna estima el retorno en *log* de los ingresos laborales por hora incluyendo *risk* y utilizando todas las variables de control de la *OLS* original. La segunda columna estima dicho retorno pero utilizando nuestro procedimiento en dos pasos para obtener los residuos de las habilidades mediante variables instrumentales que no presentan problemas de debilidad. La tercera columna presenta los retornos asociados mediante la implementación del procedimiento en dos pasos para residuos pero mediante instrumentos que pueden no estar identificados de forma correcta en el

¹⁴ Para analizar los test de dichas variables instrumentales, remitirse al apéndice I.

modelo. En todos los casos se rechaza la hipótesis nula con un nivel de confianza del 10%, estimando un retorno positivo en torno al 7.0% por desvío estándar adicional de propensión al riesgo del individuo; mientras que el resto de las habilidades cognitivas y no cognitivas no sufren cambios ni en sus magnitudes ni en sus niveles de significancia. Con lo cual, parecería ser que aquellas personas que poseen mayores niveles de aceptación y propensión al riesgo, obtienen mayores ingresos laborales por hora. Dicha variable suponemos podría representarse como un proxy de la propensión de los individuos a realizar transiciones de empleo-empleo, con lo cual parecería interesante profundizar en dicho análisis y relación en *papers* futuros.

Tabla 33: Estimaciones MCO de nuestra ecuación de Mincer inicial, del segundo paso original y alternativo de nuestro procedimiento; incluyendo la variable riesgo para Colombia

	(1)	(2)	(3)
Covariables de regresiones:	MCO	Residuos válidos	Todos los residuos
years of educ	0.023 [2.25]**	0.029 [2.13]**	0.029 [2.09]**
aggregate cognitive measure	0.047 [1.77]*	0.057 [2.03]**	0.057 [2.04]**
extraversion	-0.006 [0.19]	-0.016 [0.43]	-0.017 [0.44]
agreeableness	0.016 [0.45]	-0.001 [0.02]	-0.001 [0.02]
conscientiousness	0.013 [0.33]	0.005 [0.13]	0.005 [0.12]
emotional stability	-0.031 [0.85]	-0.038 [0.99]	-0.038 [0.99]
openness	0.017 [0.45]	0.026 [0.73]	0.026 [0.74]
grit	-0.061 [1.98]**	-0.065 [2.08]**	-0.064 [2.07]**
risk	0.066 [1.76]*	0.070 [1.75]*	0.070 [1.75]*
R^2	0.10	0.05	0.05
N	1,245	1,245	1,245

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Replicamos dicha extensión en Bolivia, reportando los resultados en la **Tabla 34**.¹⁵ En todas las regresiones se estima que el rasgo de propensión al riesgo resulta significativa levemente, pero a diferencia de Colombia encontramos evidencia que a mayor propensión al riesgo en Bolivia menores son los retornos laborales esperados. Dicha relación negativa que emerge por parte del rasgo de propensión al riesgo y el retorno laboral, parece también captar parte del efecto de la habilidad no cognitiva de perseverancia al esfuerzo de *Grit*. Tomando dichos resultados, a mayor propensión al riesgo se estima un retorno de -12.4% por cada desvío estándar adicional. Por el lado de las habilidades cognitivas, apreciamos que el indicador agregado de las dimensiones cognitivas aumenta considerablemente su magnitud manteniendo sus niveles estadísticos; mientras que las habilidades socio-emocionales no presentan resultados de interés.

Tabla 34: Estimaciones MCO de nuestra ecuación de Mincer inicial, del segundo paso original y alternativo de nuestro procedimiento; incluyendo la variable riesgo para Bolivia

	(1)	(2)	(3)
Covariables de regresiones:	MCO	Residuos válidos	Todos los residuos
years of educ	0.058 [4.14]***	0.067 [5.22]***	0.067 [5.18]***
aggregate cognitive measure	0.095 [2.80]***	0.113 [2.57]**	0.113 [2.58]**
extraversion	-0.022 [0.58]	-0.069 [1.50]	-0.069 [1.50]
agreeableness	0.009 [0.22]	-0.009 [0.17]	-0.008 [0.17]
conscientiousness	0.012 [0.27]	0.037 [0.59]	0.037 [0.59]
emotional stability	-0.054 [1.20]	-0.026 [0.40]	-0.026 [0.41]
openness	-0.023 [0.57]	-0.048 [0.98]	-0.048 [0.98]
grit	-0.031 [0.71]	-0.075 [1.53]	-0.075 [1.53]
risk	-0.076 [1.84]*	-0.124 [1.97]**	-0.124 [1.97]**
R^2	0.28	0.16	0.16
N	1,193	1,193	1,193

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

¹⁵ Para analizar los test de dichas variables instrumentales, remitirse al apéndice II.

7. Conclusiones

A lo largo del presente trabajo nos hemos remitido a realizar distintas series de análisis con el objetivo de recolectar mayor evidencia respecto a resultados obtenidos en países en desarrollo respecto a los retornos salariales y su relación a las habilidades cognitivas y no cognitivas de cada individuo. De tal manera, hemos replicado los procedimientos econométricos desarrollados en Perú y en países desarrollados encontrando resultados de interés para cada país en concreto. A su vez, hemos extendido nuestro paper para tomar en cuenta otras dimensiones de la personalidad de una persona y hallamos que dichos rasgos generan impactos en los retornos del individuo. Por otro lado, chequeamos nuestras estimaciones ante potenciales sesgos y estimamos el valor marginal que genera en el mercado laboral el poseer el título de cada nivel educativo.

Desde el punto de vista muestral, los datos parecen ser consistentes y no presentar problemas potenciales respecto a sesgos de selección, habiendo testeado y corrigiendo mediante Heckman para poder sustentar nuestras hipótesis. En cada caso, testeamos y no rechazamos la hipótesis nula respecto a la existencia de factores no visibles que generan problemas de selección muestral en nuestros datos. Dicho resultado no parece sorprendente dado que el instrumento de toma de datos de STEP posee un marco establecido y detallado del proceso para la recolección de los datos y su posterior tratamiento para obtener la representación de dicha observación respecto a la población objetivo.

Al igual que en el caso de Perú, encontramos en Bolivia y Colombia evidencia respecto a la relación entre los retornos laborales y las habilidades cognitivas como también para las habilidades no cognitivas, aunque con diferencias significativas en cada caso.

Desde el punto de vista cognitivo, estimamos que un desvío estándar del indicador cognitivo agregado en Colombia se asocia a un aumento de los retornos laborales en un 6.1% neto de los efectos de la variable años de educación sobre dicha dimensión cognitiva, siendo levemente significativo y estimándose leves diferencias en su magnitud en las regresiones estándar y al utilizar instrumentos débiles para tratar los efectos de endogeneidad. En Bolivia encontramos resultados similares en niveles estadísticos, pero con magnitudes más dispersas en algunos casos. Un desvío estándar del indicador cognitivo agregado en dicho país se asocia a un aumento de los retornos laborales de entre 4.2% y 10.0%, descontando los efectos que años de educación produce en el indicador de referencia y siendo dichas estimaciones moderadamente significativas estadísticamente. Por tanto encontramos evidencia que sustenta los hallazgos en Perú respecto a la relación entre los salarios y las habilidades cognitivas, en consonancia con las estimaciones de trabajos similares en países desarrollados.

Nuestros resultados son estadísticamente y económicamente significativos, y robustos al controlar por distintas covariables.

Aquí también es interesante observar los efectos directos de la variable años de educación sobre los retornos laborales, observando en cada país las estimaciones obtenidas. Un año adicional de educación se estima aumentan los retornos salariales por hora en un 2.9% para el caso de Colombia, en comparación con las estimaciones halladas en Bolivia que estiman un aumento del retorno salarial en un 5.5%, siendo dichos resultados en ambos países altamente significativos estadísticamente.

Desde un criterio no cognitivo, nuestros resultados suponen estimaciones en contrapuesto con las halladas en Perú. Así, en Colombia se estima que la única habilidad socio-emocional que resulta significativa estadísticamente respecto a los retornos laborales es el rasgo de perseverancia al esfuerzo de Grit. Sin embargo, reportamos que dicha relación es negativa en torno a un 6.7% siendo moderadamente significativa y manteniéndose dichos resultados en cada set de regresiones realizadas. Por el lado de Bolivia, encontramos leves indicios respecto a relación de la habilidad no-cognitiva estabilidad emocional y los retornos salariales por hora. Se estima que por cada desvío estándar, los retornos salariales disminuyen en 4.4%, una vez que se elimina el efecto endógeno de la variable años de educación mediante el procedimiento econométrico en dos pasos reportado. Resulta entonces singular dichos hallazgos, dado que se estaría asumiendo que el mercado laboral en Colombia pricea de forma negativa aquellas aptitudes de esfuerzo y perseverancia en los individuos, mientras que el mercado laboral en Bolivia estima que aquellos rasgos asociados al nerviosismo generan mayores retornos por sobre aquellos individuos más relajados dentro de la fuerza laboral.

Nuestros resultados respecto a las habilidades no cognitivas parecerían no estar en línea con los hallados en Perú, donde *Vera Tudela, Díaz & Arias* encuentran relaciones significativas y positivas respecto al rasgo *Grit* de perseverancia de esfuerzo. Así también se trató el presente trabajo respecto a posibles *bias* que pudiera presentar nuestro análisis, adicionándose testeos posteriores respecto a omisión de variables mediante el test de error de especificación de *Ramsey* (1969) sin poder rechazar la hipótesis nula de no omisión de variables en nuestra regresión original. En dicho sentido se propone para trabajos futuros la comparación de los instrumentos de toma de datos de dichas habilidades no cognitivas para entender si dicha diferencia se debe a inconsistencias en los aspectos socioemocionales relevados en la base de datos *STEP*. Intuitivamente apreciamos que los instrumentos de evaluación de las distintas habilidades resultan de mayor profundidad en lo desarrollado por los autores en el *paper* de cabecera, en comparación con los instrumentos de *STEP* que resultan de evaluaciones auto reportadas. Es por ello que, con el objetivo de poder generar un

instrumento de toma de datos más precisos, resulta interesante la implementación de dichos test en los relevamientos de *STEP* para el ámbito externo al laboral.

En nuestra sección de extensiones propusimos la inclusión de un rasgo adicional de la personalidad de los individuos, tomando entonces la aversión al riesgo y buscando estimar el retorno salarial de dicha “habilidad”. En nuestro análisis encontramos evidencia entonces que aquellos individuos con aptitudes de mayor propensos al riesgo se asocian a retornos mayores en torno al 7.0% en Colombia, mientras que el mismo análisis en Bolivia estima que aquellos individuos con rasgos de mayor propensión al riesgo se asocian negativamente a retornos en torno al 12.4% según lo reportado. En ambos casos suponemos que dichos rasgos no solo se relacionan a retornos en torno a inversiones, sino también a rasgos del individuo que lo llevan a realizar transiciones empleo-empleo alcanzando (en el caso de Colombia) mayores retornos salariales. Se propone entonces para trabajos futuros analizar la relación entre ambas variables para obtener mayor evidencia en torno a lo propuesto, a su vez con la relación de dicho rasgo y el resto de las habilidades no cognitivas y cognitivas.

Desde otra perspectiva y continuando con nuestras extensiones, intentamos estimar el valor intrínseco de poseer el título del nivel educativo más alto alcanzado para poder cuantificar el valor que el mercado le asigna por sobre aquellos individuos que no lo poseen. En ambos países estudiados encontramos fuerte evidencia que respalda la obtención de títulos para los niveles más altos educativos, donde por ejemplo poseer el título de *Phd* en Colombia se asocia a retornos salariales mayores en torno al 173.8% mientras que en Bolivia se asocia a retornos salariales mayores en torno al 80.7%.

Concluimos el presente trabajo remarcando, por sobre las diferencias encontradas en los diferentes trabajos y estudios relacionados al nuestro, la importancia de un análisis económico respecto a las diferentes habilidades y su valoración por parte del mercado de trabajo con el objetivo de desarrollar políticas que busquen capacitar a la fuerza laboral en aquellas dimensiones que les genere mayor retorno y que sean las más solicitadas por las firmas. Mediante estas políticas se lograría aumentar la inserción de la fuerza laboral en el mercado y así también aumentar la productividad de dichos individuos y de las firmas. Por ende, y dado el contexto actual educativo en Argentina, se recomienda realizar el presente trabajo para su extensión en dicho país y reforzar las medidas que se pretenden llevar a cabo en pos de lograr focalizar los esfuerzos en políticas que resulten exitosas a largo plazo.

Referencias

- Acosta, Pablo Ariel; Muller, Noel; Sarzosa, Miguel Alonso. (2015). "Beyond qualifications: returns to cognitive and socio-emotional skills in Colombia" Policy Research working paper; no. WPS 7430. Washington, D.C.: World Bank Group.
- Brunello, G., and M. Schlotter. (2011). "Non Cognitive Skills and Personality Traits: Labour Market Relevance and their Development in Education & Training Systems". IZA Discussion Paper No 5743. *Institute for the Study of Labor*.
- Burdett, K., and Mortensen, D. (1998). "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, vol. 39, no. 2, pp. 257-273.
- Bushway, S., Johnson, B. D. and Slocum, L. A. (2007). "Is the Magic Still There? The Use of the Heckman Two-Step Correction for Selection Bias in Criminology". *Journal of Quantitative Criminology*, vol. 23, no. 2, pp. 151–178.
- Cueto, S., I. Muñoz, and A. Baertl. (2010). "Scholastic Achievement, Cognitive Skills and Personality Traits of Youths and Adults in Peru: A cross-sectional and intergenerational analysis." Unpublished manuscript. GRADE, Lima.
- Davidson, R., and J. G. MacKinnon. (2001). "Econometric Theory and Methods". Oxford University Press.
- Diaz, J.J., Arias, O., and Tudela, D.V. (2012). "Does perseverance pay as much as being smart? The returns to cognitive and non-cognitive skills in urban Peru." Working paper.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. (2010). "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review*, vol. 100, no. 3, pp. 1238-60.
- Duckworth, AL, C Peterson, MD Matthews and DR Kelly. (2007). "Grit: perseverance and passion for long-term goals." *Journal of personality and social psychology*, vol. 92, no. 6, pp. 1087.
- Educational Testing Services. (2014). "A Guide to Understanding the Literacy Assessment of the STEP Skills Measurement Survey". Princeton, NJ: IEA-ETS Research Institute
- Goldberg, L. (1990). "An alternative "description of personality": The Big-Five factor structure." *Journal of Personality and Social Psychology*, vol. 59, no. 6, pp. 1216-1229.

- Goldberg, L. (1993). "The structure of phenotypic personality traits." *American Psychologist*, vol. 48, no. 1, pp. 26-34.
- Gronau, R. (1974). 'Wage comparisons—a selectivity bias', *Journal of Political Economy*, vol. 82, no. 6, pp. 1119–1143.
- Guerra, N., K. Modecki, and W. Cunningham. (2014). "Social-emotional Skills Development across the Life Span: PRACTICE." Policy Research Working Paper 7123, World Bank, Washington, DC.
- Hansen, K. T., J. J. Heckman, and K. J. Mullen. 2004. "The Effect of Schooling and Ability on Achievement Test Scores." *Journal of Econometrics*, vol. 121, no. 1–2, pp. 39–98.
- Hanushek, E. A., G. Schwerdt, S. Wiederhold, and L. Woessmann. (2013). "Returns to Skills around the World: Evidence from PIAAC." NBER Working Paper 19762, National Bureau of Economic Research, Cambridge, MA.
- Harris, John R. & Todaro, Michael P. (1970), "Migration, Unemployment and Development: A Two-Sector Analysis", *American Economic Review*, vol. 60, no. 1, pp. 126–142.
- Hartog, J., M. van Praag, and J. van der Sluis. (2010). "If You Are So Smart, Why Aren't You an Entrepreneur? Returns to Cognitive and Social Ability: Entrepreneurs versus Employees." *Journal of Economics and Management Strategy*, vol. 19, no. 4, pp. 947–989.
- Hausman, Newey, Whitney, Woutersen, Chao and Swanson, Norman. (2007). "Instrumental Variable Estimation with Heteroskedasticity and Many Instruments". *Economics Working Paper Archive*. The Johns Hopkins University, Department of Economics.
- Heckman, J. (1976). "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models". *Annals of Economic and Social Measurement*, vol. 5, pp. 475-492.
- Heckman, J. (1979). "Sample Selection Bias as a Specification Error". *The Econometric Society. Econometrica*, Vol. 47, No. 1, pp. 153-161.
- Heckman, J., J. Stixrud and S. Urzua (2006). "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior". *Journal of Labor Economics*, vol. 24, no. 3, pp. 411-482.

- Heineck, G., and S. Anger. (2010). "The Returns to Cognitive Abilities and Personality Traits in Germany." *Labour Economics*, vol. 17, no. 3, pp. 535–46.
- Lavado, Pablo; Yamada, Gustavo; and Velarde, Luciana. 2014. "Habilidades no cognitivas y brecha de género salarial en el Perú", *Estudios Públicos*, Centro de Estudios Públicos, vol. 0, pp. 135.
- Lewis, H.G. (1974). 'Comments on selectivity biases in wage comparisons', *Journal of Political Economy*, vol. 82(6), pp. 1145–1155.
- Lindqvist, Erik, and Roine Vestman. (2011). "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment." *American Economic Journal: Applied Economics*, vol. 3, no. 1, pp. 101-28.
- Matloff, N.S. (2011). "The art of R programming: a tour of statistical software design". *No Starch Press*, pp 261-285.
- Miller, R. (2015) "The Noncognitive Returns to Private Schooling in India: Evidence on Self-Esteem and Self-Efficacy". Working paper.
- Mincer, J. (1958). "Schooling, Experience, and Earnings". Columbia University Press, New York.
- Neal, D. and W. Johnson. (1996). "The Role of Premarket Factors in Black-White Wage Differences". *Journal of Political Economy*, vol. 104, no. 5, pp. 869-895.
- Nordman, C.J., L.R. Sarr and S. Sharma (2015). "Cognitive, Non-Cognitive Skills and Gender Wage Gaps: Evidence from Linked Employer-Employee Data in Bangladesh." Working paper.
- Nyhus, E. K., and E. Pons. (2005). "The Effects of Personality on Earnings." *Journal of Economic Psychology* vol. 26, pp. 363–384.
- Osborne-Groves, M. (2005). "How Important Is Your Personality? Labor Market Returns to Personality for Women in the US and UK." *Journal of Economic Psychology*, vol. 26, no. 6, pp. 827-841.
- Prada, M. (2013). "Beyond Smart and Sociable: Rethinking the Role of Abilities on Occupational Choices and Wages." University of Maryland, College Park.
- Ramsey, J.B. (1969). "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis". *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 31, no. 2, pp. 350-371.

- Rencher, A. C. (2002). "Methods of Multivariate Analysis, Second Edition". *Wiley*, pp. 380-407.
- Stock, J. and M. Yogo (2005). "Testing for Weak Instruments in Linear IV Regression". *Andrews DWK Identification and Inference for Econometric Models*. New York: Cambridge University, pp. 80-108.
- Wooldridge, J.M. (2015). "Control Function Methods in Applied Econometrics". The *Journal of Human Resources*, vol. 50, pp. 420-445.
- World Bank. (2014). "STEP Skills Measurement Surveys: Innovative Tools for Assessing Skills." Social Protection and Labor Discussion Paper 1421, World Bank, Washington, DC.
- World Bank. (2016). "Skills for Work in Bulgaria: The Relationship between Cognitive and Socioemotional Skills and Labor Market Outcomes". World Bank, Washington, DC.

Apéndice I – Validez instrumental en Colombia

En el presente anexo presentamos los principales estadísticos respecto al análisis de la validez e instrumentos débiles de cada uno de los instrumentos utilizados en el presente paper en la tabla 17 y subsiguientes estimaciones, los cuales son resumidos en la **Tabla I**.

Tabla I: Test Estadísticos de variables instrumentales en Colombia

	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: tiempo al colegio	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	4.068	3.999	0.000	0.020
extraversion	4.068	3.999	0.000	0.230
agreeableness	4.068	3.999	0.000	0.658
conscientiousness	4.068	3.999	0.000	1.269
emotional stability	4.068	3.999	0.000	0.271
openness	4.068	3.999	0.000	1.579
grit	4.068	3.999	0.000	0.056
	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: desempeño escolar	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	25.351	25.285	0.000	11.704
extraversion	25.351	25.285	0.000	4.384
agreeableness	25.351	25.285	0.000	7.179
conscientiousness	25.351	25.285	0.000	10.844
emotional stability	25.351	25.285	0.000	0.020
openness	25.351	25.285	0.000	18.898
grit	25.351	25.285	0.000	19.425

	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: nivel socioeconómico	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	7.655	7.699	0.000	3.187
extraversion	7.655	7.699	0.000	2.361
agreeableness	7.655	7.699	0.000	0.409
conscientiousness	7.655	7.699	0.000	0.142
emotional stability	7.655	7.699	0.000	0.140
openness	7.655	7.699	0.000	7.287
grit	7.655	7.699	0.000	10.007

	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: esfuerzo escolar	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	0.151	0.147	0.000	0.314
extraversion	0.151	0.147	0.000	8.658
agreeableness	0.151	0.147	0.000	0.725
conscientiousness	0.151	0.147	0.000	4.782
emotional stability	0.151	0.147	0.000	2.930
openness	0.151	0.147	0.000	0.216
grit	0.151	0.147	0.000	0.087

	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: Tiempo al colegio, desempeño escolar y nivel socioeconómico	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	36.653	12.620	8.293	4.961
extraversion	36.653	12.620	0.330	6.697
agreeableness	36.653	12.620	4.793	2.654
conscientiousness	36.653	12.620	5.886	5.046
emotional stability	36.653	12.620	0.421	0.024

openness	36.653	12.620	0.225	27.517
grit	36.653	12.620	2.991	22.258
risk	36.653	12.620	3.087	0.019

IV: Tiempo al colegio, desempeño escolar, nivel socioeconómico y esfuerzo escolar	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	36.654	9.459	8.494	4.903
extraversion	36.654	9.459	7.067	6.366
agreeableness	36.654	9.459	5.636	2.792
conscientiousness	36.654	9.459	9.744	4.910
emotional stability	36.654	9.459	3.178	0.033
openness	36.654	9.459	0.241	27.512
grit	36.654	9.459	3.109	22.893
risk	36.654	9.459	3.117	0.021

Nota: Para la evaluación de los test de instrumentos débiles y sus niveles críticos, se recomienda Stock-Yogo (2005).

Apéndice II – Validez instrumental en Bolivia

En el presente anexo presentamos los principales estadísticos respecto al análisis de la validez e instrumentos débiles de cada uno de los instrumentos utilizados en el presente paper en la tabla 18 y subsiguientes estimaciones, los cuales son resumidos en la **Tabla II**.

Tabla II: Test Estadísticos de variables instrumentales en Bolivia

	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: tiempo al colegio	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	10.022	10.143	0.000	0.003
extraversion	10.022	10.143	0.000	0.411
agreeableness	10.022	10.143	0.000	0.777
conscientiousness	10.022	10.143	0.000	0.579
emotional stability	10.022	10.143	0.000	2.158
openness	10.022	10.143	0.000	0.010
grit	10.022	10.143	0.000	0.433
	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: desempeño escolar	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	25.923	26.464	0.000	19.193
extraversion	25.923	26.464	0.000	8.002
agreeableness	25.923	26.464	0.001	1.432
conscientiousness	25.923	26.464	0.002	2.808
emotional stability	25.923	26.464	0.003	7.916
openness	25.923	26.464	0.004	4.838
grit	25.923	26.464	0.005	6.458

	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: nivel socioeconómico	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	0.004	0.004	0.000	1.010
extraversion	0.004	0.004	0.000	2.756
agreeableness	0.004	0.004	0.000	0.028
conscientiousness	0.004	0.004	0.000	2.120
emotional stability	0.004	0.004	0.000	1.457
openness	0.004	0.004	0.000	0.544
grit	0.004	0.004	0.000	0.600

	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: esfuerzo escolar	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	10.987	10.999	0.000	7.777
extraversion	10.987	10.999	0.000	14.430
agreeableness	10.987	10.999	0.000	0.003
conscientiousness	10.987	10.999	0.000	2.660
emotional stability	10.987	10.999	0.000	5.999
openness	10.987	10.999	0.000	0.078
grit	10.987	10.999	0.000	0.033

	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
IV: Tiempo al colegio, desempeño escolar y esfuerzo escolar	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	41.245	14.742	4.451	18.573
extraversion	41.245	14.742	4.003	15.213
agreeableness	41.245	14.742	0.568	1.673
conscientiousness	41.245	14.742	0.379	5.198
emotional stability	41.245	14.742	8.049	6.422

openness	41.245	14.742	1.446	3.400
grit	41.245	14.742	1.517	5.374
risk	41.245	14.742	2.539	1.251

IV: Tiempo al colegio, desempeño escolar, esfuerzo escolar y nivel socioeconómico	Underidentification test	Weak identification test	Overident. test of all instruments	Endogeneity test
	<i>Kleibergen-Paap rk LM</i>	<i>Kleibergen-Paap rk Wald F</i>	<i>Hansen J statistic</i>	<i>Endog. regressors</i>
aggregate cognitive measure	41.454	11.108	5.213	18.307
extraversion	41.454	11.108	5.631	15.180
agreeableness	41.454	11.108	0.592	1.686
conscientiousness	41.454	11.108	2.281	4.874
emotional stability	41.454	11.108	9.475	6.551
openness	41.454	11.108	1.934	3.442
grit	41.454	11.108	2.062	5.341
risk	41.454	11.108	3.222	1.200

Nota: Para la evaluación de los test de instrumentos débiles y sus niveles críticos, se recomienda Stock-Yogo (2005).

Apéndice III – Tabla de correlaciones de propensión al riesgo

En el presente anexo adjuntamos las matrices de correlaciones de las habilidades cognitivas y no cognitivas respecto al rasgo de propensión al riesgo, para cada país en particular en las tablas III. A y III. B.

Tabla III. A: Matriz de correlación de propensión al riesgo en Colombia

	risk	aggregate measure	extraversion	agreeableness	conscientiousness	emotional stability	openness	grit
risk	1.0000							
aggregate measure	0.0878	1.0000						
extraversion	-0.0001	0.0865	1.0000					
agreeableness	-0.0136	0.0806	0.1173	1.0000				
conscientiousness	-0.0102	0.0481	0.0802	0.1522	1.0000			
emotional stability	0.0393	0.1490	0.1354	0.0487	0.0031	1.0000		
openness	-0.0295	0.1715	0.1652	0.2549	0.1346	0.0847	1.0000	
grit	0.0011	-0.0395	0.0626	0.1670	0.1789	-0.0601	0.1576	1.0000

Tabla III. B: Matriz de correlación de propensión al riesgo en Bolivia

	risk	aggregate measure	extraversion	agreeableness	conscientiousness	emotional stability	openness	grit
risk	1.0000							
aggregate measure	-0.0062	1.0000						
extraversion	-0.0258	0.1737	1.0000					
agreeableness	-0.0460	0.0473	0.0715	1.0000				
conscientiousness	-0.0224	0.0438	0.1126	0.1443	1.0000			
emotional stability	0.0217	0.1377	0.1248	0.0674	0.0376	1.0000		
openness	-0.0470	0.1678	0.2102	0.1607	0.1594	0.0250	1.0000	
grit	-0.0006	0.0182	0.0787	0.1580	0.1743	0.0315	0.2090	1.0000