

Habilidades no cognitivas en América Latina. Una medición desde pruebas estandarizadas.*

Cynthia G. Marchioni

Tesis de Maestría
Maestría en Economía
Universidad Nacional de La Plata

Directora: Mariana Marchionni

Co-directora: María Lucila Berniell

Junio 2017

Clasificación JEL: I21, I24, J24

* Este trabajo corresponde al avance de mi tesis de Maestría en Economía de la UNLP, dirigida por Mariana Marchionni (UNLP) y co-dirigida por Lucila Berniell (CAF), a quienes agradezco enormemente por su valiosa guía y apoyo. Mis agradecimientos también a Dolores De La Mata, Juan Dip, Gustavo Fajardo, Agustina Hatrick, Julieta Vera y Emmanuel Vázquez, por los productivos intercambios, como a los asistentes del seminario de avances de tesis y de la LI Reunión Anual de la Asociación Argentina de Economía Política por sus comentarios. Cualquier error del presente trabajo es de mi entera responsabilidad.

Resumen

La medición de habilidades es compleja, puesto que las mismas son inobservables. Este problema es particularmente importante en la medición de habilidades no cognitivas, donde usualmente se emplean métodos basados en el auto-reporte, que están sujetos a sesgos tales como la deseabilidad social de la respuesta y el sesgo de referencia. En este trabajo se calcula una medida numérica de habilidades no cognitivas que no sufre de estos problemas, aplicando la metodología propuesta por Borghans y Schils (2012). Esta estrategia extrae la medida directamente de la nota de los estudiantes en las pruebas PISA del año 2012, y estima habilidades no cognitivas como la diferencia entre las tasas de respuesta correcta al inicio del examen y la tasa al final del mismo. Esta tesis realiza un aporte en términos de evidencia empírica y de base de datos para el estudio más profundo acerca de las habilidades de América Latina. No solo se compara el stock de habilidades no cognitivas de la región con el resto del mundo, sino que también se exploran heterogeneidades (de género y tipo de escuela). Los resultados muestran un rezago de América Latina tanto en habilidades cognitivas como en las no cognitivas. Además, se encuentran brechas de género significativas en habilidades no cognitivas (favoreciendo a las mujeres) para la gran mayoría de los 65 países participantes de la prueba, mientras que los resultados de las brechas entre escuelas públicas y privadas no son concluyentes.

Campos temáticos (Clasificación JEL): I21, I24, J24.

Palabras clave: habilidades no cognitivas, decaimiento, PISA, América Latina.

Abstract

Measuring skills is difficult, because they are unobservable. This problem is even harder for the case of non-cognitive skills, as compared to cognitive, since for them the measurement methods typically rely heavily on self-reporting, which suffers from several biases such as social desirability of response and reference bias. This paper quantifies non-cognitive skills with a measure that overcomes these problems, applying the methodology proposed by Borghans and Schils (2012). This strategy extracts the measure directly from the students' scores in PISA 2012 tests, and estimates non-cognitive skills as the difference between the correct response rates at the beginning and at the end of the exam. The contribution of this paper is to provide empirical evidence and a database for a deeper understanding of skills in Latin America. This paper presents a comparison between non-cognitive skills in this region and the rest of the world, and also examines heterogeneities (gender, and public versus private schools). The results show that Latin America lags behind in cognitive as well as in non-cognitive skills. In addition, there are significant gender gaps in non-cognitive skills (favoring women) for most of the 65 countries participating in the test, and no conclusive results regarding a gap by school type.

JEL Classification: I21, I24, J24.

Índice

I. Introducción.....	1
II. Relevancia de las habilidades no cognitivas.....	3
III. Medición de las habilidades no cognitivas.....	8
IV. Datos.....	12
V. Estrategia empírica.....	15
VI. Resultados.....	18
VI.I. Resultados Generales: Comparación entre América Latina y otras Regiones.....	18
VI.II. Heterogeneidades.....	23
VI.II.i. Resultados según género.....	23
VI.II.ii. Resultados según tipo de escuela: privadas vs. públicas.....	25
VI.II.iii. Extensión: evolución del decaimiento.....	27
VII. Consideraciones finales.....	28
VIII. Referencias Bibliográficas.....	32
IX. Apéndice.....	36

I. Introducción

El conjunto de habilidades de una persona consiste en sus aptitudes y capacidades para desarrollar distintas actividades, incluyendo sus conocimientos, sus destrezas físicas y también su carácter o temperamento (CAF, 2016). Diversas disciplinas se han ocupado de estudiar las habilidades, usando diferentes definiciones y categorizaciones. Una de las clasificaciones utilizadas, especialmente en economía, reduce las habilidades a dos categorías: las cognitivas y las no cognitivas.¹ Por supuesto, ambas componen el capital humano, un concepto más general y sumamente importante en la investigación económica.

Las habilidades cognitivas son las asociadas al procesamiento de información, la capacidad de adquirir conocimientos, de dominar una nueva tarea y de tomar decisiones considerando sus posibles consecuencias (Borghans *et al.*, 2008a). Ejemplos de este tipo de habilidades son la memoria, el razonamiento, el cálculo, la resolución de problemas y el entendimiento del lenguaje. En cambio, las habilidades no cognitivas incluyen herramientas socioemocionales, tales como la perseverancia, autoeficacia, la motivación intrínseca y la resiliencia (Heckman *et al.*, 2006; Heckman y Kautz, 2013).²

La principal contribución de esta tesis consiste en aportar una medición actualizada de habilidades no cognitivas comparable para América Latina y otras regiones del mundo, aplicando una reciente metodología propuesta por Borghans y Schils (2012). Este es un avance en cuanto a la medición de este tipo de habilidades, sobre todo en la región, ya que a pesar de la importancia que la literatura le adjudica a las habilidades no cognitivas, no existen grandes mediciones internacionales (que incluyan a países latinoamericanos) que permitan extraer conclusiones similares a las que se obtienen sobre habilidades cognitivas utilizando pruebas estandarizadas como las de PISA (*Programme for International Student Assessment*, OCDE)³. Esta prueba evalúa las habilidades de los estudiantes de 15 años en

¹ Sin embargo, tanto a nivel conceptual como en la práctica, es difícil describir algún comportamiento humano que no incluya ningún proceso cognitivo. En todo aprendizaje o creación de nuevo conocimiento los factores cognitivos y no cognitivos interactúan continuamente (Bransford *et al.*, 2000).

² Cabe notar que el hecho de compartimentar en solo dos dimensiones las habilidades de una persona deja las habilidades físicas también en el conjunto de las no cognitivas. Sin embargo, en este trabajo se considerarán dentro de las habilidades no cognitivas solo a aquellas que tienen que ver con aspectos socioemocionales de la conducta humana.

³ Otros ejemplos de iniciativas a nivel internacional que incluyen países latinoamericanos y que evalúan sus habilidades son la Encuesta CAF 2015, que recogió medidas de habilidades cognitivas y no cognitivas en una encuesta en 10.000 hogares a personas de entre 15 y 55 años en 10 países en América Latina; y las pruebas del LLECE (Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, UNESCO) que evalúan el desempeño escolar en tercer y sexto grado de escuela primaria en las áreas de matemática, lenguaje y para sexto grado también en el área de ciencias naturales. Sin embargo, ninguna de las dos permiten realizar comparaciones con otras regiones.

las áreas de lectura, matemática y ciencias de manera comparable a nivel internacional, y ha incluido desde sus inicios a países de América Latina.

Además de proveer una medida de habilidades no cognitivas comparable entre los 65 países que participaron en la prueba PISA en su edición 2012, esta tesis investiga las posibles razones detrás de los resultados encontrados, analizando las diferencias en habilidades por género y por tipo de escuela (privadas vs. públicas).

Otro aporte de este trabajo es la base de datos construida para la aplicación de la metodología. Su armado requirió del emparejamiento de cada uno de los ítems que forman parte de la prueba PISA con el conjunto de preguntas al que corresponde, usando como base el manual y el reporte técnico de la prueba (OCDE, 2012). De esta manera, se pudo reconstruir cada uno de los cuadernillos, lo que permitió asociar el resultado obtenido en cada pregunta con el individuo que efectivamente la había respondido.

La estrategia empleada en este estudio consiste en obtener las tasas de respuestas correctas para la primera y la última pregunta de la prueba, y luego calcular la diferencia entre las mismas. Se considera esta diferencia como el decaimiento del desempeño a lo largo de la prueba, y un indicador *negativo* de habilidades no cognitivas (es decir, a mayor diferencia, menores habilidades no cognitivas). De esta manera, se obtiene una medida numérica de habilidades no cognitivas a partir de la cual los países pueden ser rankeados de acuerdo a esta dimensión, tal como habitualmente son ordenados de acuerdo a la nota final promedio de los alumnos en la prueba (lo que suele considerarse una medida de habilidades cognitivas).

La principal ventaja de esta estrategia es la de superar las limitaciones de las mediciones tradicionales. Al extraer las medidas de habilidades directamente de los resultados de los exámenes que toman los estudiantes, se evitan los sesgos de respuesta de los individuos, que merman las mediciones comúnmente utilizadas en la literatura. Estos sesgos, derivados del auto-reporte, incluyen la deseabilidad social de la respuesta, el sesgo de referencia desde el cual las personas contestan preguntas de autopercepción, y la dependencia de la capacidad de comprensión de los individuos para reportar correcta y rigurosamente sus habilidades no cognitivas.

Otras ventajas de este método son que permite realizar comparaciones entre países con distintos niveles de desarrollo y también que el método es robusto a incluir o no preguntas de opinión de los estudiantes, y a tomar como punto de corte diferentes momentos del

examen (tomar por ejemplo el promedio de las cinco primeras preguntas vs. el promedio las cinco últimas, en lugar de la primera y la última).

En cuanto a las desventajas de la estrategia, la principal es que no es posible diferenciar dentro del componente de habilidades no cognitivas los efectos de los sistemas de incentivos o la cultura, y por otro lado, no se mide una habilidad no cognitiva en particular, sino que se asume que aquellas que más influyen en el desarrollo de un examen -perseverancia, autoeficacia, motivación intrínseca y resiliencia- son las capturadas.

El resto de la tesis cuenta con siete secciones. La sección II recorre los aportes recientes de la literatura que demuestra la importancia de las habilidades no cognitivas en distintos aspectos de la vida de las personas, reflejando por qué es fundamental entender el nivel de desarrollo no cognitivo de la población de estudiantes de América Latina en relación a otras regiones. La sección III presenta cuáles son los principales desafíos en la medición de las habilidades no cognitivas. La sección IV describe los datos que serán utilizados en el análisis, mientras que en la sección V se explica la estrategia empírica empleada. La sección VI resume los principales resultados encontrados, a la vez que analiza ciertas heterogeneidades y presenta una extensión de la investigación. La sección VII culmina con las consideraciones finales.

II. Relevancia de las habilidades no cognitivas

La literatura presenta evidencia de la importancia del capital humano en distintas dimensiones del bienestar individual y agregado. A nivel agregado, trabajos como el de Barro (1991) y el de Sala-i Martin *et al.* (2004) han asociado los años de escolaridad con la evolución del PBI, concluyendo que esta particular medida de capital humano es importante para el desarrollo económico. Sin embargo, utilizar los años de educación para medir habilidades implica asumir que en todos los países se adquiere la misma cantidad y calidad de conocimientos durante un año educativo. Por esta razón, estudios recientes han buscado opciones más sofisticadas de aproximar el capital humano a fin de poder correlacionarlo luego con el crecimiento económico. Una de las alternativas ha sido utilizar las notas obtenidas en pruebas estandarizadas, como lo hace Hanushek y Woessmann (2012). Estos autores utilizan un índice basado en los resultados de nueve pruebas internacionales de desempeño⁴ tomadas entre 1964 y 2003, como medida de

⁴ Las pruebas utilizadas por Hanushek y Woessmann (2012) son: First International Mathematics Study (FIMS, 1964), First International Science Study (FISS, 1970-71), First International Reading Study (FIRS, 1970-72), Second International Mathematics Study (SIMS, 1980-82), Second International Science Study (SISS, 1983-84),

habilidades cognitivas, y encuentran una fuerte relación positiva y causal entre el puntaje medio alcanzado por los países en las pruebas y su crecimiento de largo plazo entre los años 1960 y 2000.

Por otro lado, las habilidades cognitivas influyen en diversos ámbitos como el laboral, el logro educativo, la salud, y la inclusión cívica de las personas. Por ejemplo, una gran cantidad de evidencia empírica asocia las habilidades cognitivas con mayores ingresos laborales, ya sea cuando dichas habilidades son aproximadas por años de escolaridad (Angrist y Krueger, 1991) o con mediciones específicas resultantes de encuestas (Hanushek y Zhang, 2009; Acosta *et al.*, 2015; CAF, 2016)⁵. A su vez, existe evidencia de que los individuos más educados reportan mejores indicadores de salud y satisfacción con la vida (Oreopoulos, 2007; Conti *et al.*, 2010; CAF, 2016) y mayor involucramiento social, en actividades como el voluntariado y la participación política (OECD, 2013).

Sin embargo, en los últimos años ha crecido el gran interés en el rol de las habilidades no cognitivas dentro del capital humano (ver por ejemplo Levin, 2012)⁶, y por lo tanto en su importancia para el bienestar de las personas y de los países. Por ejemplo, Balart *et al.* (2015) demuestra que las habilidades no cognitivas están asociadas al desempeño económico de las naciones, al encontrar una correlación significativa y positiva entre estas habilidades y el crecimiento económico. En este estudio los autores retoman el esquema de Hanushek y Woessmann (2012) pero incorporan medidas de habilidades no cognitivas resultantes de la metodología desarrollada por Borghans y Schils (2012), la misma que será empleada en el presente trabajo y será explicada en detalle en una sección V. El principal resultado de Balart *et al.* (2015) señala que cuando se incorpora el factor de habilidades no cognitivas al análisis, el efecto de las habilidades cognitivas sobre el crecimiento estimado por Hanushek y Woessmann (2012) se reduce un 40%.

A nivel individual, para países desarrollados las investigaciones de Mueller y Puig (2006) y Heineck y Anger (2010), para Estados Unidos y Alemania respectivamente, muestran la importancia de la amabilidad (*agreeableness*), la responsabilidad (*conscientiousness*), la estabilidad emocional, la extraversión y la apertura a nuevas experiencias en los salarios. En América Latina, Bassi *et al.* (2012) utilizan datos de la Encuesta sobre Trayectorias y

Second International Reading Study (SIRS, 1990-91), Third International Mathematics and Science Study (TIMSS, 1994-95, 1999 y 2003), Programme for International Student Assessment (PISA, 2000 y 2003) y Progress in International Reading Literacy Study (PIRLS, 2001).

⁵ Ver por ejemplo Mincer, 1974; Card, 1999; Murnane *et al.*, 2000; Psacharopoulos y Patrinos, 2004; entre otros.

⁶ Este interés no es nuevo en la economía (ver Bowles y Gintis, 1976; Heckman y Rubinstein, 2001), pero en los últimos años se ha incrementado de manera notable a partir de varios trabajos, tanto teóricos como empíricos, que han reposicionado al desarrollo socioemocional de las personas como un factor determinante del bienestar.

Habilidades para una muestra de jóvenes de entre 25 y 30 años en Argentina y Chile, y encuentran una fuerte asociación de las habilidades socioemocionales con la participación laboral y con la probabilidad de tener un empleo.

Estas asociaciones también se reflejan en CAF (2016) para el caso de 10 países de América Latina, donde los resultados muestran que mientras una habilidad no cognitiva (*Grit*) es la que más se asocia con la participación laboral y con la probabilidad de encontrar un empleo, son las habilidades cognitivas las que se relacionan de manera más fuerte con los ingresos salariales y con la calidad del empleo (medida por el estatus de formalidad laboral). A su vez, Heckman *et al.* (2006) encuentran que ambos tipos de habilidades son importantes para explicar salarios más altos y mayor empleabilidad. Díaz *et al.* (2012) estiman con variables instrumentales el efecto de las habilidades cognitivas y no cognitivas en los salarios para el caso de Perú, utilizando la Encuesta de Habilidades y Mercado Laboral (ENHAB), y hallan que el mercado laboral peruano valora ambos tipos de habilidades: un incremento de una desviación estándar en los años de educación se asocia a un 15% más de salarios, mientras que cambios similares en el estabilidad emocional y perseverancia se encuentran asociados con incrementos de 5% y 8% en los salarios, respectivamente.

Para este trabajo, es relevante la evidencia que vincula las habilidades con el logro educativo. La literatura de diversas disciplinas que se interesan por esta materia (psicología, pedagogía, y también la economía) sostiene que el desempeño en los exámenes no depende solamente de las habilidades cognitivas sino también de las habilidades no cognitivas de los estudiantes, así como de los incentivos involucrados y del marco institucional en el que son implementados (Borghans *et al.*, 2008b; Heckman y Kautz, 2013). Un alumno que presta atención en clase, es responsable con sus tareas escolares y busca apoyo cuando encuentra dificultad para entender algún concepto, tiene altas probabilidades de lograr sus metas y lograr buenos resultados en sus exámenes, de manera que tanto sus capacidades cognitivas como no cognitivas son determinantes de su éxito académico (Koch *et al.*, 2014).

En particular, estudios recientes han mostrado que uno de los factores que impactan de manera crucial en el desempeño de los alumnos es la motivación⁷. Al respecto, Duckworth *et al.* (2011) y Borghans *et al.* (2011) proveen evidencia de la importancia de la motivación de los estudiantes en los *tests* de inteligencia, y además encuentran que al incluir este factor en el análisis se reduce el efecto predictivo de los índices de coeficiente intelectual

⁷ No solo es importante la motivación propia (intrínseca), sino que también lo es la brindada por la escuela y por la familia (ver por ejemplo, Gneezy *et al.*, 2011 o Levitt *et al.*, 2012).

en distintos resultados relacionados tradicionalmente al desarrollo cognitivo. Asimismo, la denominada Teoría de la Atribución identifica como dos de los principales factores explicativos del éxito y el fracaso académico al esfuerzo y la motivación. Weiner (1986) explica que la cantidad de esfuerzo que alguien dedique a un examen o a la realización de una actividad particular, depende de cuánto éxito creen las personas que pueden tener en el desempeño de dicha tarea. Es decir, que el esfuerzo es un elemento que el individuo puede controlar y que la auto-percepción de las propias habilidades juega un rol fundamental en los resultados que los alumnos pueden obtener.

A pesar de la importancia de las habilidades no cognitivas en el logro educativo, escasean medidas comparables entre países para evaluar el desarrollo relativo alcanzado en esa dimensión del capital humano, sobre todo en la región, con excepción de la Encuesta CAF 2015, que recogió medidas de habilidades cognitivas y no cognitivas en una encuesta en 10 países en América Latina. Sin embargo, con los resultados de dicha encuesta no pueden realizarse comparaciones con regiones más desarrolladas, ya que solo releva países latinoamericanos. Por esta razón, esta tesis realiza un aporte para obtener medidas comparables entre regiones con distinto nivel de desarrollo económico, y en especial para la región, ya que se encuentra muy rezagada de acuerdo a los resultados obtenidos en las pruebas internacionales de desempeño académico.

En particular, si bien América Latina ha mejorado varios indicadores de cobertura del sistema educativo en las últimas dos décadas⁸ y también ha aumentado el gasto público en educación, continúa mostrando señales de bajo rendimiento académico en comparación con otras regiones del mundo. En la edición 2012 de las pruebas PISA, los países latinoamericanos que participaron (Argentina, Brasil, Chile, Colombia, Costa Rica, México, Perú y Uruguay) se ubicaron entre los últimos puestos del ranking internacional.⁹ De hecho, a partir de los resultados obtenidos en 2012, puede concluirse que un 64% de los alumnos latinoamericanos no alcanzan el nivel 2 de competencia en matemáticas¹⁰. Esto significa que aproximadamente dos de cada tres estudiantes de la región no son capaces de interpretar preguntas o de aplicar fórmulas y procedimientos para resolver problemas matemáticos básicos. La posición relativa de los países latinoamericanos en cuanto a la

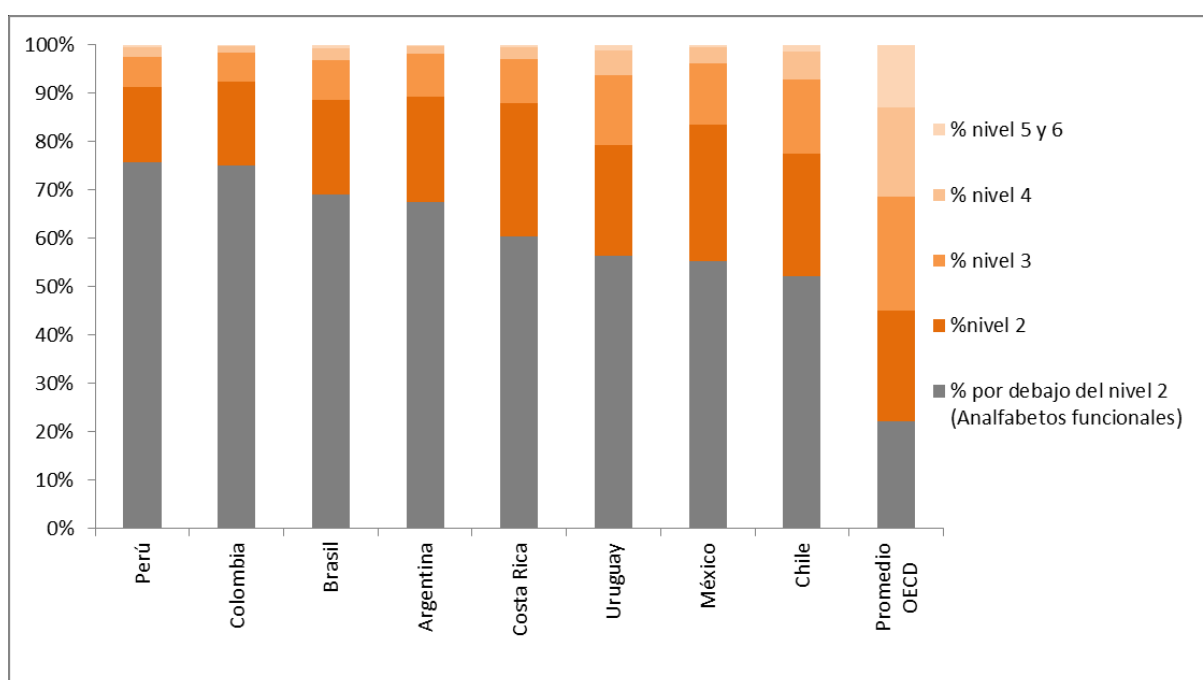
⁸ Por ejemplo, ampliando la cobertura en primera infancia y en secundaria, especialmente para niños y jóvenes de familias de escasos recursos (CAF, 2016).

⁹ En la edición 2012 del Programa los países latinoamericanos se ubicaron entre los 20 con peores resultados (de 65) para las tres materias, obteniendo la posición más alta Chile y la más baja Perú en las tres áreas evaluadas.

¹⁰ El porcentaje de alumnos que no llegan a ese nivel mínimo de conocimientos es del 23% para la OCDE y de 9% para Asia del Pacífico.

distribución de desempeño por nivel de competencias (niveles 1 al 6)¹¹ puede apreciarse en el Gráfico 1. De manera similar, los resultados de las pruebas de lectura y ciencias arrojan cifras del 46 y 50% de estudiantes por debajo de nivel 2 de competencias, respectivamente. Esto quiere decir que la mitad de los adolescentes de la región pueden ser considerados “analfabetos funcionales” por no tener los conocimientos de lectura, matemáticas y ciencias fundamentales para participar productivamente en la vida laboral.

Gráfico 1. Distribución de los niveles de competencias matemáticas obtenidos por los estudiantes de 15 años en PISA 2012 (América Latina y países de la OCDE).



Fuente: elaboración propia con base en PISA 2012.

Nota: los países que se incluyen en el promedio de la OCDE son: Alemania, Australia, Austria, Bélgica, Canadá, Dinamarca, Eslovaquia, Eslovenia, España, Estados Unidos, Estonia, Finlandia, Francia, Grecia, Hungría, Irlanda, Islandia, Israel, Japón, Luxemburgo, Noruega, Nueva Zelanda, Países Bajos, Polonia, Portugal, Reino Unido, República Checa, República de Corea, Suecia, Suiza y Turquía.

¹¹ Los niveles de competencia determinados en PISA reflejan diferentes niveles de capacidades que son crecientes en complejidad, siendo el nivel 1 el más bajo y el nivel 6 el más alto posible. Para más información respecto de los puntos de corte y el detalle de las características de cada nivel, ver el Reporte Técnico de PISA (OECD, 2012).

III. Medición de las habilidades no cognitivas

Como se ha mencionado, numerosos trabajos muestran que las habilidades cognitivas y las no cognitivas importan tanto para el bienestar de las personas y como para el de los países, por lo cual es importante contar con mediciones cada vez más precisas de las habilidades. Sin embargo, dado que las habilidades no son observables, medirlas implica resolver un complejo problema de identificación. Una de las principales dificultades para la identificación de una habilidad en particular reside en obtener medidas netas de los numerosos factores que están asociados a la habilidad que se intenta medir a través de un cierto indicador. Este es un problema común tanto para habilidades cognitivas como para las no cognitivas. Pero las habilidades no cognitivas son particularmente complejas de medir, ya que por su propia naturaleza, contienen múltiples aristas y no son fácilmente jerarquizables (es decir, ordenables de acuerdo a su importancia relativa) como sí lo son las habilidades cognitivas¹².

La relativa facilidad para medir habilidades cognitivas ha favorecido su primacía en la literatura y en las iniciativas de gran escala para medir habilidades. Las pruebas estandarizadas de desempeño académico son fruto de estas iniciativas, y permiten conocer en más detalle las capacidades cognitivas de la población de estudiantes en distintos países en el mundo.

Para el caso de las habilidades no cognitivas, los indicadores que generalmente se emplean para obtener medidas de habilidades son de tres tipos, y aunque ninguno de ellos es perfecto, han hecho considerables aportes al esclarecimiento de la formación de este tipo de habilidades, así como el papel que juegan en distintas dimensiones de la vida de las personas. Una primera estrategia utiliza magnitudes extraídas de reportes de los individuos, por ejemplo a través de *tests* psicométricos que tratan de recoger cuán responsable, o enfocada, o empática, es una persona que responde una encuesta¹³.

A pesar de ser ampliamente utilizadas, estas medidas presentan algunas limitaciones. La más importante es que derivan del auto-reporte, por lo que son altamente susceptibles a sesgos de referencia (Kautz *et al.*, 2014, West *et al.*, 2015). Este refiere a los distintos patrones desde los cuales cada persona responde a preguntas de autopercepción. Por

¹² En lo que respecta a las habilidades cognitivas, la literatura de la psicología propone que las mismas se ordenan jerárquicamente, y que pueden ser predichas según un factor general de inteligencia, llamado factor G, que es común a todas las habilidades intelectuales.

¹³ Ejemplos populares de auto-reporte que se diseñaron como *tests* psicométricos cortos administrados en encuestas son el TIPI (*Ten Item Personality Inventory*) para medir el modelo de los Big Five (Goldberg, 1990), el *test* de Grit (Duckworth y Quinn, 2009), la escala de autoeficacia (Schwarzer y Jerusalem, 1995), y la de autoestima (escala de Rosenberg), entre otros.

ejemplo, y particularmente importante para el desarrollo del presente trabajo, los individuos educados en diferentes culturas pueden interpretar de distinta manera una misma pregunta relacionada con conceptos como el “esfuerzo” o haber “trabajado duro” para alcanzar una meta¹⁴. Otra limitación de las medidas auto-reportadas son los sesgos debidos a la deseabilidad social de las respuestas (Paulhus, 1991). Existe una tendencia a dar una respuesta considerada socialmente aceptable, aunque ella no se condiga con el comportamiento real del encuestado. Adicionalmente, existe el obstáculo de la no respuesta en absoluto, y el de problemas comprensión de las preguntas, por lo que la capacidad de los encuestados de auto-reportar correcta y rigurosamente sus habilidades no cognitivas es vital. Por último, en cuanto la utilidad práctica de estas medidas, un problema adicional es que no suelen existir mediciones hechas con el mismo instrumento en diferentes países.

El trabajo de Butler y Adams (2007) utiliza los datos de las pruebas PISA 2003 para explotar una pregunta particular realizada ese año a los estudiantes: el Termómetro del Esfuerzo. En ese ejercicio los alumnos debían contestar cuánto esfuerzo habían dedicado a la prueba. Los autores no encuentran grandes variaciones entre países en cuanto al promedio de esfuerzo auto-reportado, independientemente de los resultados finalmente obtenidos en las pruebas. De hecho, los alumnos japoneses, que obtuvieron en promedio notas dentro del *top ten* del ranking, fueron los que reportaron menor esfuerzo, mientras que los estudiantes mexicanos reportaron mayores niveles de esfuerzo y se ubicaron en esa edición de PISA en los últimos escalafones del ranking de notas. El paper interpreta esto como una muestra de la deseabilidad social de la respuesta¹⁵.

Con el propósito de observar si esto sucede también en la base de datos a utilizar en este trabajo, se grafica el esfuerzo promedio reportado para cada país con datos de la prueba PISA 2012. Como puede observarse en el gráfico, las respuestas no varían demasiado, moviéndose entre los valores 8,1 y 9,6 para los 64 países con datos para esta pregunta¹⁶. De hecho, el cálculo del desvío estándar es tan solo de 0,29. Del mismo modo que Butler y Adams (2007), se puede concluir que el esfuerzo reportado es prácticamente constante

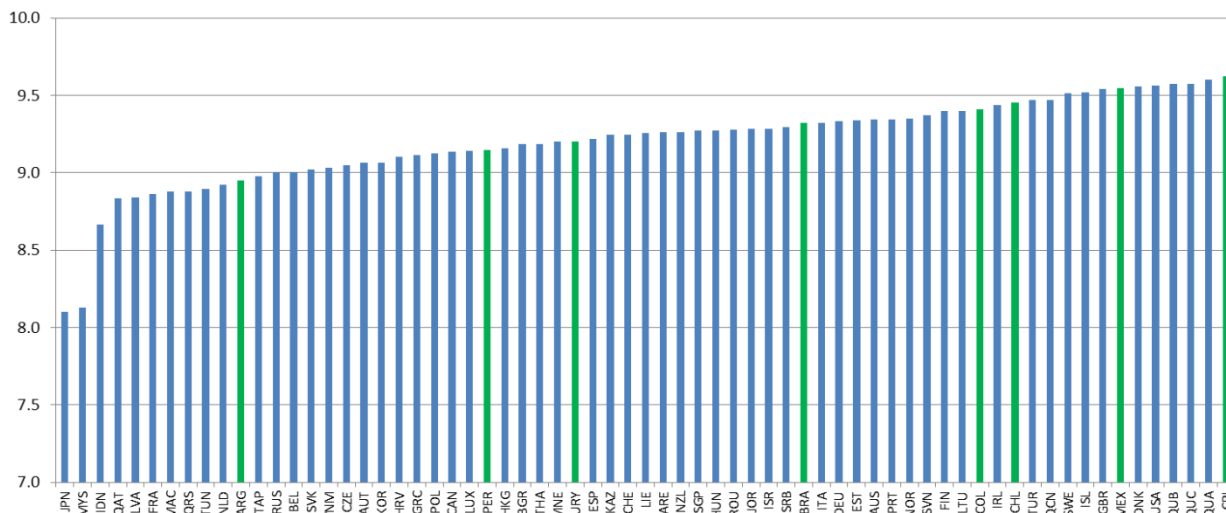
¹⁴ Al respecto, Heine *et al.* (2008) demuestran que el promedio de diligencia (*conscientiousness*, dentro de la escala de los *Big Five*) auto-reportado por los individuos de ciertos países se correlaciona negativamente con numerosas medidas objetivas de diligencia a nivel país, dentro de las que incluyen la eficiencia del servicio postal y la precisión de los relojes en los bancos públicos, reflejando la inconsistencia entre lo percibido y la realidad.

¹⁵ Podría argumentarse que en realidad la prueba simplemente fue demasiado fácil para los alumnos japoneses, pero esto no es válido si se considera que no todos los alumnos japoneses obtuvieron niveles 5 o 6 de competencia. Por esa razón, se sostiene que la deseabilidad social de la respuesta tiene que ver con lo que se espera de adolescente en Japón y lo que la sociedad espera de uno mexicano, además de la definición del concepto de “esfuerzo” que puede tener cada cultura.

¹⁶ Albania no presenta datos.

entre países. Este es un ejemplo de por qué algunas de las medidas auto-reportadas por alumnos en los cuestionarios de PISA no resultan muy informativas¹⁷.

Gráfico 2. Promedio del esfuerzo dedicado a la prueba PISA 2012 reportado por los alumnos de cada país.



Fuente: Elaboración propia en base a PISA 2012.

El segundo tipo de medidas de habilidades no cognitivas comúnmente utilizadas se obtiene por observación de conductas en la vida real. Así, la participación en actividades negativas -como actividades delictivas o de abuso de sustancias psicoactivas- o actividades positivas -como la participación en voluntariados-, puede utilizarse como *proxy* del nivel de determinada habilidad no cognitiva. Por ejemplo, Agan (2011) encuentra, utilizando la base National Longitudinal Study of Youth-Children and Young Adults Survey (CNLSY), que los individuos más propensos a cometer crímenes resultan ser los que son más impulsivos, tienen preferencia por el riesgo, buscan la satisfacción inmediata sin mediar las consecuencias de sus actos y no respetan las normas, y lo interpreta como la falta de ciertas habilidades no cognitivas.

El tercer tipo de indicadores construye medidas a partir del desempeño observado en la ejecución de una tarea en particular, generalmente en ambientes controlados. Este es el caso de algunas pruebas de laboratorio, donde se mide la reacción a distintos estímulos en la consecución de una tarea simple. Por ejemplo, el famoso *test* de la golosina (Mischel,

¹⁷ En el Anexo puede verse el Gráfico A.1. donde se realiza el mismo ejercicio para el ítem “Creo que puedo triunfar si me esfuerzo lo suficiente” (que se considera una pregunta relacionada con habilidades no cognitivas, ver Weiner, 1986). Se aprecian resultados similares a los del Gráfico 2.

2014), que intenta medir postergación en la gratificación (auto-regulación) en niños pequeños.

En este trabajo se utiliza una estrategia que puede incluirse en este tercer grupo, pues se trata de una medida que cuantifica habilidades no cognitivas extrayendo información del comportamiento de los individuos durante la resolución de un examen o *test* de desempeño académico¹⁸.

Además de tener cada una de estas estrategias de identificación sus falencias, la elección de qué medida utilizar para computar habilidades no cognitivas puede no ser trivial. Humphries y Kosse (2017) ponen a prueba diversas medidas de habilidades no cognitivas tradicionalmente usadas en la literatura y las comparan tanto entre sí (encontrando correlaciones incluso negativas entre distintas medidas) en cuanto a su poder predictivo del promedio escolar obtenido (*Grade Point Average*) para una muestra de estudiantes alemanes. Los autores encuentran que la elección de la medida utilizada como indicador de habilidades no cognitivas tiene una influencia significativa tanto en los resultados y las conclusiones de una estimación como en la interpretación que pueda hacer un *policy-maker* del papel de este tipo de habilidades.

Lo anterior da una idea de la importancia de medir con precisión algo tan abstracto como las habilidades no cognitivas. En esta tesis se utiliza la estrategia metodológica de Borghans y Schils (2012) quienes explotan la variabilidad exógena del orden en que se responden las preguntas en las pruebas PISA para diferenciar habilidades cognitivas y no cognitivas, reconociendo un patrón que indica que los estudiantes suelen obtener menores puntajes en las últimas preguntas. Así, se saca provecho de esta variación sostenida de los puntajes calculando la frecuencia de respuestas correctas para las preguntas al inicio de la prueba y contrastándola con la frecuencia de haber respondido correctamente las preguntas ubicadas al final del cuadernillo de pruebas (lo que se explica con más detalle en la siguiente sección). De esta manera, puede identificarse el decaimiento en la nota de las preguntas según su orden (la diferencia de tasas de respuesta correcta entre la primera y la última pregunta) como una medida numérica de habilidades no cognitivas, considerando que la motivación intrínseca, la perseverancia, la autoeficacia y la resiliencia, son habilidades necesarias para mantener el desempeño a lo largo del examen. Así, se obtiene una medida numérica de habilidades derivada directamente de las respuestas de

¹⁸ Existen otras estrategias que también utilizan el desempeño académico para extraer mediciones de la dimensión no cognitiva. Por ejemplo, se han diseñado modelos para identificar este tipo de habilidades teniendo en cuenta el tiempo que le lleva a un individuo responder el examen y el dedicado a cada pregunta (Borghans *et al.*, 2008a), las preguntas no contestadas (Hitt *et al.*, 2016) o las preguntas “salteadas” (Hernández y Hershaff, 2014).

los alumnos en las pruebas, lo que supera los sesgos de las propuestas tradicionales ya mencionados. Además la medida es comparable entre países, por basarse en pruebas internacionales.

IV. Datos

El presente trabajo utiliza una base de datos construida a partir de información proveniente de la edición del año 2012 del Programa para la Evaluación Internacional de Alumnos (PISA, por sus siglas en inglés) llevado a cabo por la OCDE cada tres años desde el año 2000. En el año 2012 participaron aproximadamente 510.000 estudiantes de entre 15 años y 3 meses y 16 años y 2 meses de edad de 18.139 escuelas pertenecientes a 65 países.

El objetivo de PISA es evaluar, a través de pruebas estandarizadas, los conocimientos y competencias de los estudiantes, en las áreas de lenguaje, matemática y ciencias, junto con algunos otros módulos en años particulares, como por ejemplo un módulo de educación financiera en la edición de 2012. La intención de dicha prueba no solo es medir cuánto los alumnos han aprendido y si son capaces de reproducirlo, sino también lo que pueden hacer con ello en diferentes circunstancias, por lo que los ejercicios incluyen tareas de razonamiento, interpretación de la información y resolución de problemas. Además, los estudiantes responden a un cuestionario en el cual se les solicita información sobre sus hogares, su escuela, su familia, sus bienes materiales y sus experiencias de aprendizaje.¹⁹

Una particularidad de las pruebas PISA, es que las preguntas se distribuyen en 13 conjuntos (o grupos) de preguntas según el área involucrada. Cada uno de estos conjuntos tiene diferente cantidad²⁰ y tipos de preguntas (a desarrollar o de opciones múltiples, por ejemplo) pero son diseñados de tal manera que sean equivalentes en nivel de dificultad (media del conjunto de preguntas) y de extensión, de modo que cada conjunto tiene una duración esperada de resolución de 30 minutos²¹.

Dado que responder todas las preguntas implicaría que la duración de la prueba fuera de 6 horas y media, cada alumno responde solamente 4 conjuntos de preguntas. Los conjuntos

¹⁹ En el Tabla A.1 del Apéndice puede verse el cronograma de una prueba PISA con los lapsos de duración de cada etapa.

²⁰ Ver Tabla A.2 del Apéndice.

²¹ Si bien cada una de las áreas que evalúa PISA tiene una determinada cantidad de conjuntos, cada año una de las áreas tiene mayor cantidad de preguntas y conjuntos que las demás, debido a que ha sido seleccionada para ser el área foco de PISA ese año. A saber, en el año 2012 las pruebas se focalizaron en matemática, por lo cual hubo siete conjuntos de preguntas de matemática, tres conjuntos de preguntas de lectura y tres conjuntos de preguntas de ciencias.

son distribuidos aleatoriamente entre 13 cuadernillos, y cada alumno recibe aleatoriamente uno al inicio de la prueba. De esta manera el contenido y el orden de las preguntas que cada alumno debe responder son aleatorios. Por ejemplo, como puede verse en la Tabla 1, el primer cuadernillo consta de un primer grupo de preguntas de matemática, seguido por preguntas de ciencias, otro grupo de preguntas de matemáticas y por último, otro de ciencias. Con este ejemplo se pueden hacer explícitas dos cosas. En primer lugar, que no necesariamente todos los alumnos responden preguntas de todas las áreas. Y en segundo lugar, que el alumno que recibe el, por ejemplo, cuadernillo #1, contestará en primer lugar una pregunta de matemáticas, y en último lugar una pregunta de ciencias, y esto cambiará para cada alumno de acuerdo al cuadernillo que le toque.

Tabla 1. Contenido de los cuadernillos de la edición 2012 de la prueba PISA.

Cuadernillo	Conjuntos				Set Estándar	Set de Menor dificultad
1	Matemática_5	Ciencias_3	Matemática_6A	Ciencias_2	*	
2	Ciencia3	Lenguaje_3	Matemática_7A	Lenguaje_2	*	
3	Lenguaje_3	Matemática_6A	Ciencias_1	Matemática_3	*	
4	Matemática_6A	Matemática_7A	Lenguaje_1	Matemática_4	*	
5	Matemática_7A	Ciencias_1	Matemática_1	Matemática_5	*	
6	Matemática_1	Matemática_2	Lenguaje_2	Matemática_6A	*	
7	Matemática_2	Ciencias_2	Matemática_3	Matemática_7A	*	
8	Ciencias_2	Lenguaje_2	Matemática_4	Ciencias_1	*	*
9	Lenguaje_2	Matemática_3	Matemática_5	Lenguaje_1	*	*
10	Matemática_3	Matemática_4	Ciencias_3	Matemática_1	*	*
11	Matemática_4	Matemática_5	Lenguaje_3	Matemática_2	*	*
12	Ciencias_1	Lenguaje_1	Matemática_2	Ciencias_3	*	*
13	Lenguaje_1	Matemática_1	Ciencias_2	Lenguaje_3	*	*
21	Matemática_5	Ciencias_3	Matemática_6B	Ciencias_2		*
22	Ciencias_3	Lenguaje_3	Matemática_7B	Lenguaje_2		*
23	Lenguaje_3	Matemática_6B	Ciencias_1	Matemática_3		*
24	Matemática_6B	Matemática_7B	Lenguaje_1	Matemática_4		*
25	Matemática_7B	Ciencias_1	Matemática_1	Matemática_5		*
26	Matemática_1	Matemática_2	Lenguaje_2	Matemática_6B		*
27	Matemática_2	Ciencias_2	Matemática_3	Matemática_7B		*

Fuente: Elaboración propia en base a OECD (2012)

a/ La columna refleja los cuadernillos que forman parte de la modalidad “estándar” y por lo tanto tienen conjuntos con el mismo nivel de dificultad que las pruebas en años anteriores.

b/ La columna muestra los cuadernillos que forman parte de la modalidad “Menor dificultad” y por lo tanto incluyen dos conjuntos de matemática con preguntas más fáciles (los conjuntos Matemática_6B y Matemática_7B en reemplazo de los conjuntos Matemática_6A y Matemática_7A).

A su vez, cada conjunto (y por lo tanto cada pregunta) aparece en 4 de los 13 cuadernillos entregados en cada país²² ²³, aunque nunca en el mismo orden. Es decir que un mismo conjunto, ocupa una única vez las 4 posiciones posibles en 4 de los 13. Por ejemplo, tomando el caso del conjunto Matemática_5: en el cuadernillo 1 Matemática_5 aparece como primer conjunto, en el cuadernillo 11 aparece como segundo conjunto, en el 9 es el tercero y en el cuadernillo 5 aparece en el cuarto lugar, y ya no vuelve a aparecer otro cuadernillo estándar. Esto quiere decir, que hay 13 posibles primeras preguntas y 13 posibles últimas preguntas.

Además, cada cuadernillo es asignado aleatoriamente a cada uno de los participantes, con una distribución balanceada de cuadernillos por país (ver Tabla A.3 en el Apéndice). El tipo de muestreo usado en PISA asegura que cada cuadernillo es entregado a un número de estudiantes lo suficientemente grande y variado, de manera que la información recogida sea suficiente para obtener una estimación adecuada del rendimiento de los estudiantes tanto a nivel país como dentro de los subgrupos de población más relevantes: varones y mujeres, alumnos procedentes de distintos entornos sociales y económicos, etc. (OECD, 2009a) (Ver Tabla A.4 en el Apéndice). Las características de aleatoriedad y balanceo de cuadernillos en PISA son de vital importancia para la estrategia empírica de este trabajo, que se detalla en la sección V.

Por otro lado, vale hacer un punto en lo que respecta a los puntajes de PISA publicados por la OCDE. PISA utiliza métodos de imputación a través de valores plausibles, para reportar el desempeño de los estudiantes, en lugar de calcular el desempeño particular de cada uno. Esto se debe en parte a que no todos los participantes de la prueba reciben preguntas de todas áreas, y en parte a que el objetivo de PISA es medir las capacidades a nivel poblacional, y no particular. Así, para cada área, PISA estima una distribución de puntajes para cada grupo de estudiantes y luego toma aleatoriamente cinco valores de esa distribución, calculando el resultado para cada alumno según la probabilidad de contestar correctamente la pregunta dadas sus características, tales como género y raza (OCDE, 2009b).

²² En la prueba PISA 2012 algunos países (dentro de los que se incluyen los 8 países latinoamericanos participantes en esta edición, ver Tabla A.3 del Apéndice) hicieron uso de la opción de administrar un set de cuadernillos de menor dificultad a sus estudiantes (columna “Menor dificultad”, en Tabla 1). Esta posibilidad se ofreció a aquellos países que se incorporaban por primera vez al Programa PISA y se esperaba que tuvieran un puntaje bajo, y también para aquellos que habían obtenido un puntaje promedio de 450 o menos en el área foco de la edición anterior, es decir, en el área de lectura del año 2009 (OECD, 2012).

²³ La cantidad total de cuadernillos en PISA 2012 asciende a 25 (ver Tabla 1), donde 20 corresponden a las pruebas estándar (cuadernillos 1 al 13) o menor dificultad (cuadernillos 8 al 13 y 21 al 27), uno de menor duración para alumnos con necesidades especiales (cuadernillo 20, excluido de la Tabla 1 y del análisis en este trabajo), y 4 cuadernillos adicionales de conceptos financieros (cuadernillo 70 a 74), que solo algunos países decidieron tomar, por lo que también son excluidos de este análisis.

La idea detrás de la utilización de los valores plausibles es considerar en el resultado final, errores de medición producto de las diferentes condiciones en que los alumnos toman el examen, y que pueden afectar su desempeño en la prueba, de modo de poder obtener estimadores insesgados de las variables asociadas al desempeño en la prueba (Marchionni *et al.*, 2013). Sin embargo, esto da lugar a que ocurra, por ejemplo, que un alumno asiático y uno latinoamericano respondan idénticamente las preguntas de cierta área, pero que se reporten menores capacidades para el estudiante latinoamericano dado que el desempeño promedio histórico de ambos grupos en la prueba PISA es diferente.

Para esta tesis se construyó una base de datos a nivel alumno-pregunta con la nota que cada uno obtuvo, basada en los datos publicados por OCDE, pero que innova en cuanto a la imputación de los resultados obtenidos por cada alumno. Esta base empareja cada uno de los ítems de la prueba, con el conjunto de preguntas al que corresponde. Luego se ordenan, pudiéndose reconstruir cada uno de los cuadernillos, y determinar entonces el orden de las preguntas en el mismo. Esto fue necesario para poder asociar el resultado de cada pregunta, con el individuo que efectivamente la respondió. Es decir, que en este trabajo, por la naturaleza misma de la metodología a aplicar, se usan los “*raw values*” (lo que efectivamente respondió cada alumno) y no los imputados. Luego, se estandarizaron las notas de tal manera que valieran 1 en caso de ser correcta y 0 en caso de ser incorrecta o parcialmente correcta²⁴.

V. Estrategia empírica

La metodología utilizada es la planteada en Balart *et al.* (2015) en base a Borghans y Schils (2012). La misma consiste en estimar habilidades no cognitivas con el decaimiento en desempeño a lo largo del examen. Se considera esta caída en el rendimiento como una medida (negativa) de perseverancia, concentración, auto-eficacia y motivación intrínseca. Se estima el decaimiento con un modelo probit que incluye efectos fijos por pregunta. De esta manera, para cada país se estima el siguiente modelo:

$$P[Y_{ij} = 1] = F(\alpha_0 + \alpha_1 Q_{ij} + \sum_{j=2}^J \mu_j) \quad (1)$$

donde Y_{ij} es el puntaje obtenido por el alumno i en la pregunta j (0 si la respuesta es incorrecta y 1 si es correcta), Q_{ij} es el número de orden (la posición dentro del examen) de

²⁴La decisión de tomar como incorrectas a las parcialmente correctas o incompletas no afecta las conclusiones del presente trabajo. Se realizaron ejercicios considerándolas como correctas y no hubo cambios en el ordenamiento de los países.

la pregunta j para el alumno i , y μ_j es el efecto fijo de las preguntas que controla por la dificultad o el tipo (a desarrollar o de opción múltiple) de la pregunta. La variable Q_{ij} es reescalada o normalizada, de manera que para la primera pregunta de cada cuadernillo toma valor 0, mientras que toma valor 1 para la última pregunta²⁵. De esta manera, y por la aleatoriedad en la asignación de cuadernillos, la variable Q_{ij} es exógena. La ecuación 1 se estimó separadamente para cada país aplicando los ponderadores de la base de PISA.

La medida de habilidades no cognitivas será entonces la diferencia en desempeño promedio del país al comienzo del *test* (la tasa de respuestas correctas promedio de responder correctamente la primera pregunta) y el desempeño hacia el final (tasas de respuestas correctas promedio de la última pregunta):

$$\text{Decaimiento} = F(\widehat{\alpha}_0) - F(\widehat{\alpha}_0 + \widehat{\alpha}_1) \quad (2)$$

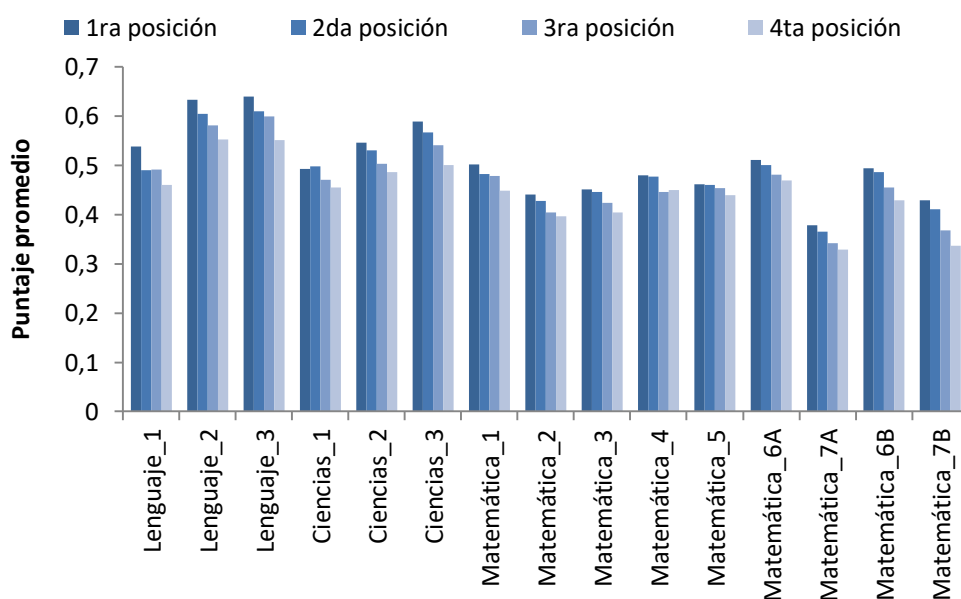
para cada país, considerando que la función de distribución es una función acumulada normal.

En las secciones anteriores se ha planteado que tanto las habilidades cognitivas como las no cognitivas juegan un rol fundamental y simultáneo en el desempeño en un examen. Vale notar que el supuesto fundamental bajo el cual esta metodología es válida, es que el nivel de habilidades no cognitivas que se requieren para avanzar en la prueba es creciente e influye cada vez más a medida que transcurre el tiempo durante la evaluación.

Para indagar respecto de cuán sostenible es ese supuesto, en primer lugar se presenta el Gráfico 3 que muestra la variación del puntaje promedio obtenido en cada conjunto de preguntas, dependiendo del orden en que aparece a lo largo del cuadernillo. Esta evidencia permite observar que un mismo conjunto, es decir, un mismo grupo de preguntas y por tanto un mismo nivel de dificultad, se responde de peor manera si está ubicado en las últimas posiciones del cuadernillo que si está al inicio. Este patrón es general para todos los conjuntos que son parte de la prueba, incluso para los dos conjuntos de menor dificultad (Matemática_6B y Matemática_7B). Esto permite ver que efectivamente existe un cierto decaimiento en el desempeño a lo largo de la prueba independientemente del área temática.

²⁵ Esta transformación se debe a que, tal como puede verse en la Tabla A.2 del Apéndice, no todos los cuadernillos tienen la misma cantidad de preguntas. De esta manera entonces, siempre la pregunta con valor $Q_{ij}=0$ es la que ocupa la primera posición y la pregunta con valor $Q_{ij}=1$ es la última.

Gráfico 3. Puntaje promedio obtenido en cada conjunto de preguntas, según su ubicación en el cuadernillo.



Fuente: elaboración propia con base en PISA 2012.

Establecido esto, es válido preguntarse si el orden en que aparece un determinado conjunto tiene que ver con su dificultad. Dicho de otro modo, esta inquietud refiere a si el decaimiento se debe a que las preguntas más difíciles aparecen sistemáticamente al final del examen. Pero esto no sucede en la práctica, debido a que todos los conjuntos varían entre las posiciones 1era. a 4ta. en los 13 cuadernillos que se distribuyen en igual porcentaje entre los alumnos, como se explicó en la sección anterior. Otra preocupación tiene que ver con que los cuadernillos puedan ser estratégicamente distribuidos entre los alumnos, de acuerdo a las capacidades que quien administra la prueba conozca ex-ante de cada estudiante. Este potencial problema también es evitado, ya que los cuadernillos son aleatoriamente distribuidos entre los alumnos que toman la prueba. Las Tablas A.3 y A.4 en el Apéndice muestran la distribución de cuadernillos es uniforme entre países, y también al interior de cada país de acuerdo al género y tipo de escuela (pública o privada).

Por último, también podría argumentarse que si en promedio, los individuos de un país no muestran variación en la tasa de respuesta correcta de las preguntas o presentan tasas con valores extremos muy bajos para la primera pregunta, el indicador de habilidades no cognitivas podría ser artificialmente alto (por ser la diferencia muy pequeña). Es decir, la diferencia entre las tasas de respuesta correcta para la primera y la última pregunta podría ser la misma, y la diferencia ser cero confundiendo un rendimiento constante (malo

o bueno a lo largo de toda la prueba) con un nivel alto habilidades no cognitivas. Este no es el caso ya que, como se verá en la siguiente sección todos los países muestran variación en el desempeño entre el inicio y el final del examen.

La principal ventaja de utilizar esta estrategia es que al extraer las medidas de habilidades directamente de los resultados de los exámenes que toman los estudiantes, se evita el efecto de los sesgos de deseabilidad social de la respuesta, de referencia y la capacidad de comprensión de los individuos, derivados del auto-reporte. Otras ventajas son que se pueden realizar comparaciones entre países con distintos niveles de desarrollo y también que el método ha demostrado ser robusto en ciertas dimensiones. Por ejemplo, se realizaron diversas pruebas y no se evidenciaron cambios sustanciales en los rankings ni al incluir las preguntas de opinión²⁶ ni al cambiar el punto de corte donde se realizan las mediciones. Es decir, tomar por ejemplo el promedio de las cinco primeras preguntas vs. el promedio las cinco últimas, en lugar de la primera y la última.

En cuanto a las desventajas de la estrategia, la principal es que no es posible diferenciar dentro del componente de habilidades no cognitivas los efectos de los sistemas de incentivos o la cultura, y por otro lado, no se mide una habilidad no cognitiva en particular, si no que se asume que aquellas que más influyen en el desarrollo de un examen - perseverancia, autoeficacia, motivación intrínseca y resiliencia- son las capturadas.

VI. Resultados²⁷

VI.I. Resultados Generales: Comparación entre América Latina y otras regiones

La Tabla 2 a continuación presenta las estimaciones de las habilidades no cognitivas a partir de la ecuación (1) para todos los países participantes de la prueba PISA en su edición del año 2012. Esta tabla muestra para cada país la diferencia entre las tasas de respuesta correcta de la primera y la última pregunta, medida que se interpreta como el decaimiento del rendimiento, que es el indicador de *falta* de habilidades no cognitivas, o sea que un mayor decaimiento está asociado a menores habilidades no cognitivas.

Como puede observarse, los países de América Latina se ubican en los últimos lugares en el ranking de habilidades no cognitivas. De hecho, los primeros cinco puestos están

²⁶ Estas preguntas nunca eran ni la primera ni la última en ningún cuadernillo.

²⁷ Los resultados obtenidos han demostrado ser robustos ante diferentes ejercicios tales como controlar por nivel de riqueza y cambiar la cantidad de preguntas que se toman como indicadores de habilidades (promedio de las primeras 5 y las últimas 5 preguntas en lugar de la primera y la última). Por cuestiones de espacio no están aquí presentados pero pueden solicitarse dichos ejercicios a la autora.

ocupados por cinco países latinoamericanos, y los restantes tres países de la región que tomaron la prueba están en el quintil de países con mayor decaimiento.

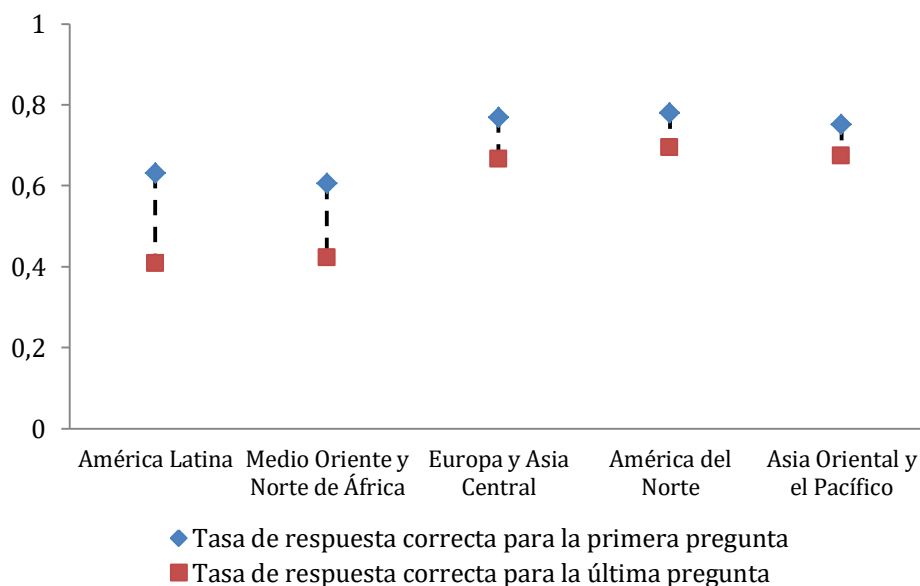
Tabla 2. Decaimiento en el desempeño de la prueba PISA 2012 por país, ordenados de mayor a menor.

País	Indicador de Habilidades no cognitivas (negativo)	País	Indicador de Habilidades no cognitivas (negativo)
Colombia	0,279	Luxemburgo	0,098
Uruguay	0,251	Eslovaquia	0,097
México	0,245	Reino Unido	0,097
Perú	0,244	Lituania	0,096
Brasil	0,239	Hungría	0,093
Túnez	0,226	Polonia	0,091
Grecia	0,219	Latvia	0,090
Catar	0,195	Canadá	0,087
Costa Rica	0,193	Nueva Zelanda	0,087
Israel	0,185	Estados Unidos	0,083
Argentina	0,179	Eslovenia	0,082
Jordania	0,173	Croacia	0,079
Chile	0,166	Países Bajos	0,077
Albania	0,163	Macao-China	0,077
Bulgaria	0,148	Dinamarca	0,072
Tailandia	0,147	Australia	0,071
Montenegro	0,142	Estonia	0,069
Rusia	0,141	República Checa	0,069
Serbia	0,140	Bélgica	0,068
Rumania	0,137	República de Corea	0,065
Indonesia	0,135	Alemania	0,065
Portugal	0,135	Japón	0,065
Emiratos Árabes Unidos	0,133	Liechtenstein	0,063
Suecia	0,132	Irlanda	0,063
Perm (Rusia)	0,123	Suiza	0,060
España	0,121	Vietnam	0,057
Italia	0,119	Austria	0,054
Kazajistán	0,119	Singapur	0,052
Noruega	0,118	Hong Kong-China	0,049
Malasia	0,114	Shanghai-China	0,046
Francia	0,107	Finlandia	0,045
Islandia	0,106	Taipéi	0,045
Turquía	0,101		

Fuente: elaboración propia con base en PISA 2012.

El fenómeno por regiones puede apreciarse en el Gráfico 4. América Latina resulta ser la región que, en promedio, presenta el mayor decaimiento en comparación con las demás regiones participantes de la prueba PISA 2012 (el detalle de los países incluidos en cada una de las regiones se encuentra en el Tabla A.5. del Apéndice).

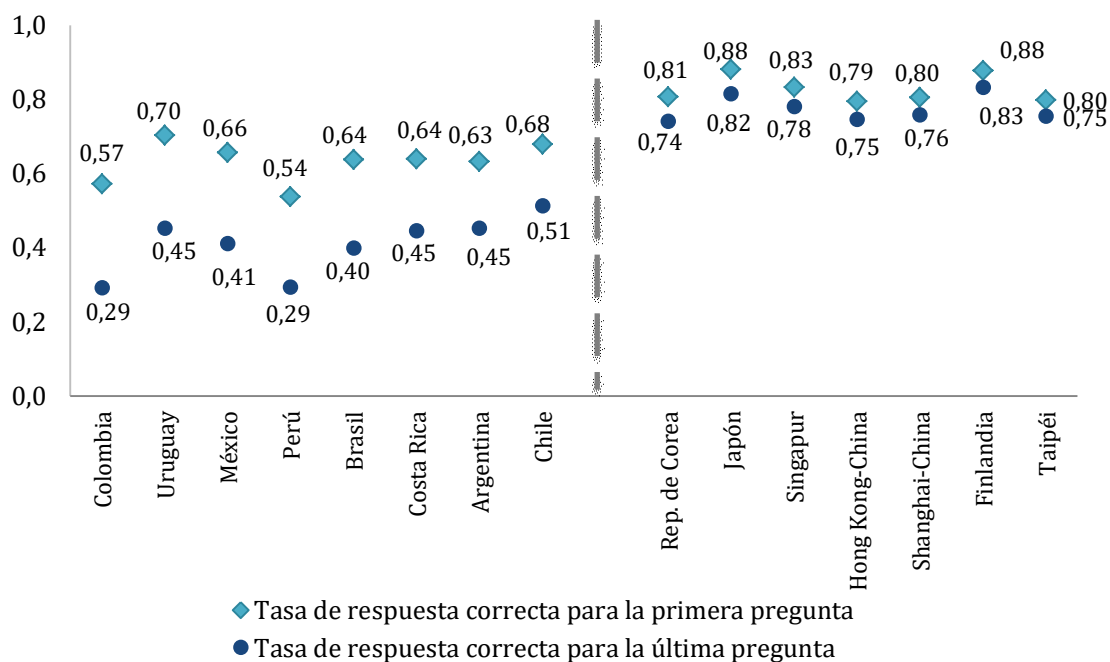
Gráfico 4. Tasas de respuesta correcta promedio para la primera y la última pregunta de la prueba PISA 2012, por regiones.



Fuente: elaboración propia con base en PISA 2012.

Para analizar la posición relativa de los países latinoamericanos respecto de los demás participantes, el Gráfico 5 presenta las tasas de respuesta correcta para la primera y la última pregunta para 8 los países de la región y para algunos de los países que mayor puntaje promedio obtuvieron en la prueba PISA. La diferencia de alturas entre el punto y el rombo es la representación gráfica del decaimiento estimado. Este gráfico permite notar que además de tener grandes caídas en el desempeño, el nivel inicial desde el que parten los países de América Latina es considerablemente menor que el de los países seleccionados. Por lo tanto, aquellos países que obtienen mejores resultados en las pruebas PISA son aquellos que presentan elevados indicadores tanto de habilidades cognitivas como no cognitivas, en línea con la literatura reciente que sostiene que los resultados de las pruebas estandarizadas no reflejan solo los conocimientos sino una batería de otras capacidades (Borghans *et al.*, 2008b).

Gráfico 5. Tasas de respuesta correcta para la primera y la última pregunta de la prueba PISA 2012, para países de América Latina y países seleccionados.



Fuente: elaboración propia con base en PISA 2012.

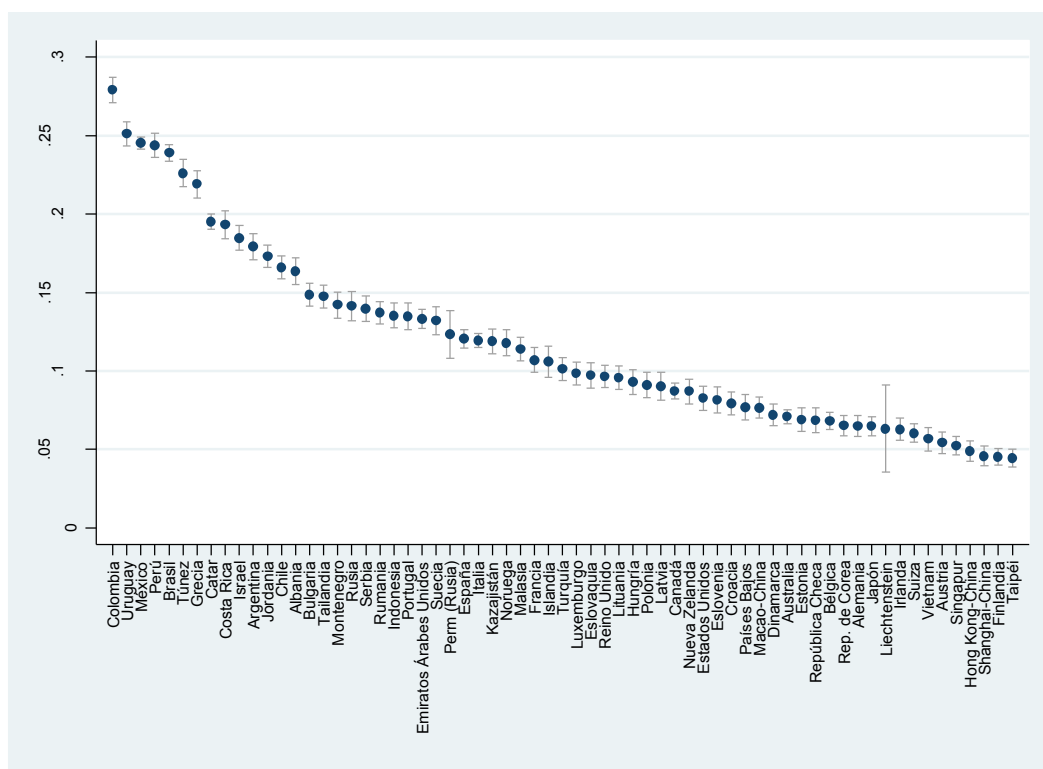
Respecto de estos resultados, ciertamente en los países menos desarrollados, las condiciones de aprendizaje son menos favorables que en los países desarrollados, no solo en cuanto a infraestructura o medios de aprendizaje, sino también en cuanto al rol educativo de las escuelas donde el personal docente y no docente debe lidiar con realidades sociales y económicas complejas. Particularmente en América Latina, se observa una gran distancia respecto de los resultados obtenidos en otras regiones, aún a pesar de, como ya se dijo, haber aumentado el gasto público en educación y haber ampliado la cobertura.

Vale recordar en este punto que en la edición de PISA 2012 se utilizaron para ciertos países, entre ellos los 8 participantes de América Latina, cuadernillos con conjuntos de preguntas más fáciles. Aun así América Latina figura entre los peores resultados, por lo que al repetir las estimaciones tomando los datos solamente para aquellos cuadernillos

comunes para todos los países (cuadernillos del 8 al 13, ver Tabla 1) los resultados son peores para la región, pero los rankings no cambian²⁸.

El Gráfico 6 presenta la diferencia entre las dos tasas, es decir, el decaimiento de cada uno de los países, acompañado de su intervalo de confianza al 95%. Puede observarse que ninguno de los valores atraviesa el cero y que los intervalos son considerablemente pequeños para todos los países (a excepción de Liechtenstein que cuenta con muy pocas observaciones en la muestra), por lo que las diferencias resultan ser estadísticamente significativas.

Gráfico 6. Decaimiento de prueba PISA 2012 para todos los países participantes con intervalos de confianza al 95%.



Fuente: elaboración propia con base en PISA 2012.

²⁸ En la Tabla A.6 del Apéndice, puede verse que el puntaje promedio obtenido por los países latinoamericanos no muestra variaciones entre cuadernillos del set estándar y los cuadernillos con preguntas más sencillas.

VI.II. Heterogeneidades

Para abordar las razones que pueden estar detrás de estas diferencias en decaimiento del rendimiento, se plantea una caracterización de esta medida de falta de habilidades no cognitivas de acuerdo a dos dimensiones: género y tipo de escuela (privado vs. pública).

VI.II.i. Resultados según género

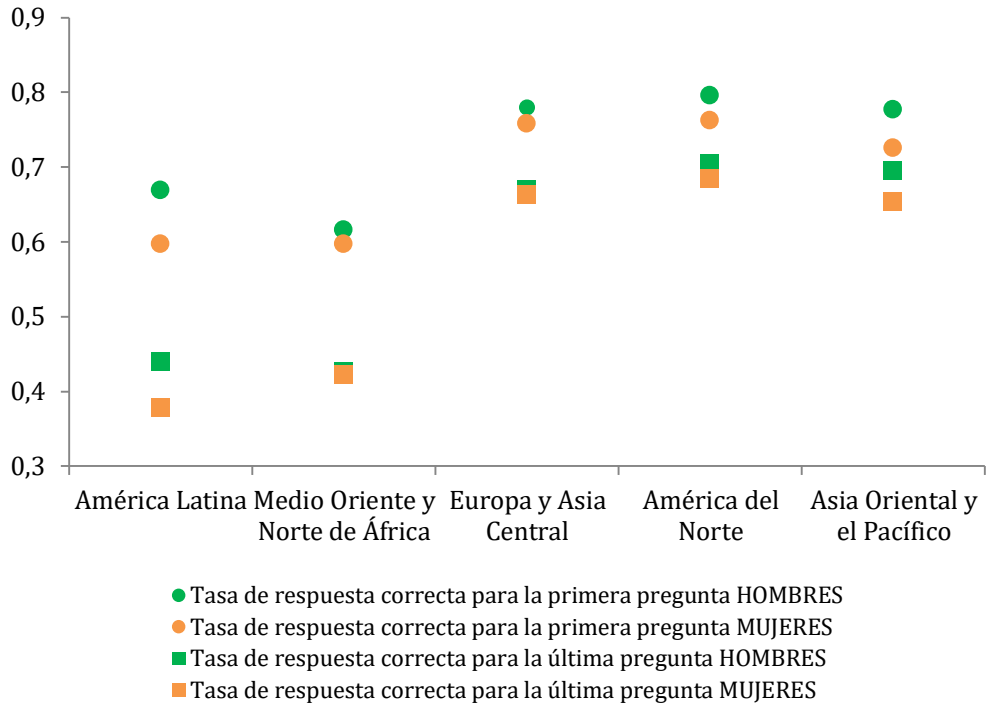
Respecto de la dimensión de género, se presentan los Gráficos 7 y 8. El Gráfico 7 muestra, de manera análoga al Gráfico 4, las tasa de respuesta correcta promedio para la primera y la última pregunta por regiones, pero diferenciadas según género, las cuales son derivadas de la estimación de la ecuación (1) en dos sub-muestras separadas (para hombres y mujeres, por país)²⁹. Mientras que las brechas de género en habilidades cognitivas pueden apreciarse en la distancia vertical entre los dos puntos, las brechas de género en habilidades no cognitivas surgen de la comparación de las diferencias verticales entre cuadrado y punto para cada caso, puesto que nuevamente esta distancia refleja el decaimiento en desempeño a lo largo de la prueba. El Gráfico 7 entonces permite observar que América Latina es la región que presenta mayores brechas en ambos tipos de habilidades.

Para investigar respecto de la significatividad estadística de estas diferencias, se calcularon los intervalos de confianza al 95% a partir de las estimaciones obtenidas de los parámetros de la ecuación (1)³⁰. La brecha de género en las habilidades no cognitivas se calculó como la diferencia del decaimiento de los hombres y el decaimiento de las mujeres para cada país, de tal manera que un resultado positivo indicaría un menor decaimiento de las mujeres a lo largo de la prueba, y por lo tanto las caracterizaría con una medida de habilidades no cognitivas más alta que la de los hombres, y lo contrario si el resultado es negativo. El Gráfico 8 permite observar que para la gran mayoría de los países la brecha es positiva, y los resultados de brechas negativas no son significativos, indicando que las mujeres presentan mayores habilidades no cognitivas.

²⁹ Estos resultados se presentan en las Tablas A.7 y A.8 del Apéndice.

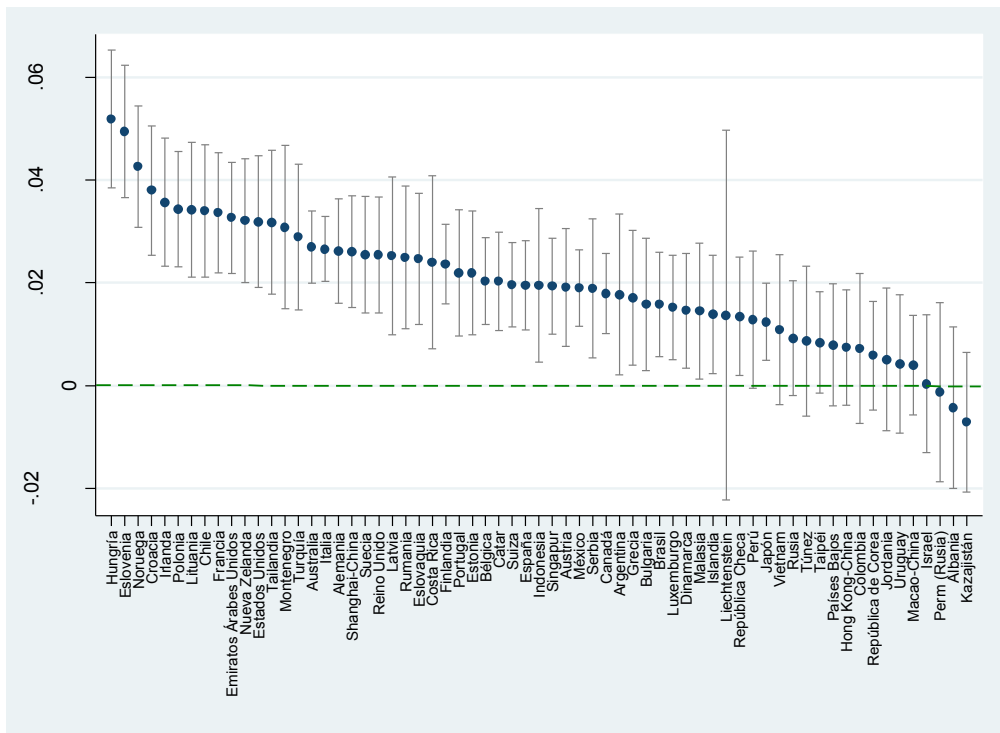
³⁰ Los intervalos de confianza fueron calculados bajo supuestos de comportamiento asintótico normal (método delta), estimando el modelo sin interactuar los efectos fijos de pregunta con la variable dicotómica de interés (género o tipo de escuela, para los resultados a presentarse más adelante).

Gráfico 7. Tasa de respuesta correcta para la primera y la última pregunta de PISA 2012, por género y región.



Fuente: elaboración propia con base en PISA 2012.

Gráfico 8. Brechas de género en habilidades no cognitivas para todos los países participantes de la prueba PISA 2012.



Fuente: elaboración propia con base en PISA 2012.

Si bien es escasa la literatura respecto de las diferencias en habilidades no cognitivas según género, y más aún para América Latina, sí existen trabajos como el de Lavado *et al.*, (2014) que estiman la distribución de habilidades socioemocionales por género para Perú, y van un paso más allá, calculando el efecto de estas diferencias en las brechas de diferentes outcomes entre hombres y mujeres. Los autores concluyen que parte de las brechas salariales, de empleo y de educación, se deben a que las mujeres presentan mayor acervo de habilidades no cognitivas que son valuadas negativamente por el mercado laboral.

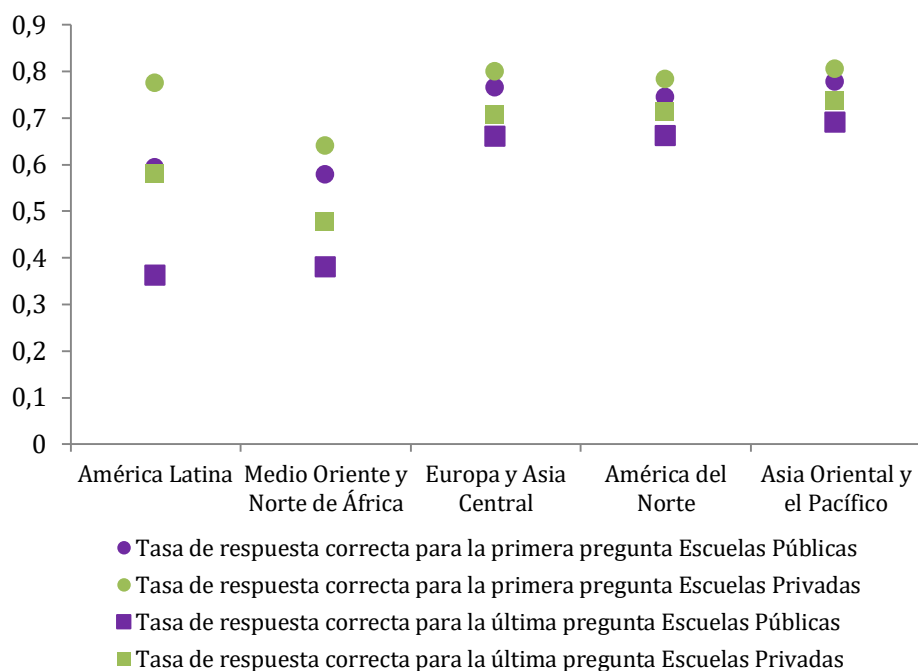
VI.II.ii. Resultados según tipo de escuela: privadas vs. públicas

En las Tablas A.9 y A.10 del Apéndice se presenta el indicador de habilidades no cognitivas para las escuelas públicas y privadas, respectivamente. En ambos rankings los países latinoamericanos continúan ubicándose en las peores posiciones.

En relación a la diferencia de ambos tipos de habilidades entre escuelas públicas y escuelas privadas, el Gráfico 9 presenta los promedios de las tasas pertinentes para distintas regiones, por tipo de escuela. En primer lugar, puede observarse que la tasa de respuesta correcta para la primera pregunta es siempre mayor en las escuelas privadas que en las públicas, para todas las regiones. Esto refleja una brecha en cuanto a esta medida de habilidades cognitivas, a favor de los jóvenes que asisten a escuelas privadas. Lo mismo sucede con la tasa de respuesta correcta para la última pregunta y con la distancia entre ambas tasas. Por ello puede deducirse que el decaimiento es, en promedio, mayor para los estudiantes de las escuelas públicas para cada una de las regiones, lo que implica concluir que la brecha también es a favor de los jóvenes que asisten a escuelas privadas en cuanto a las habilidades no cognitivas.

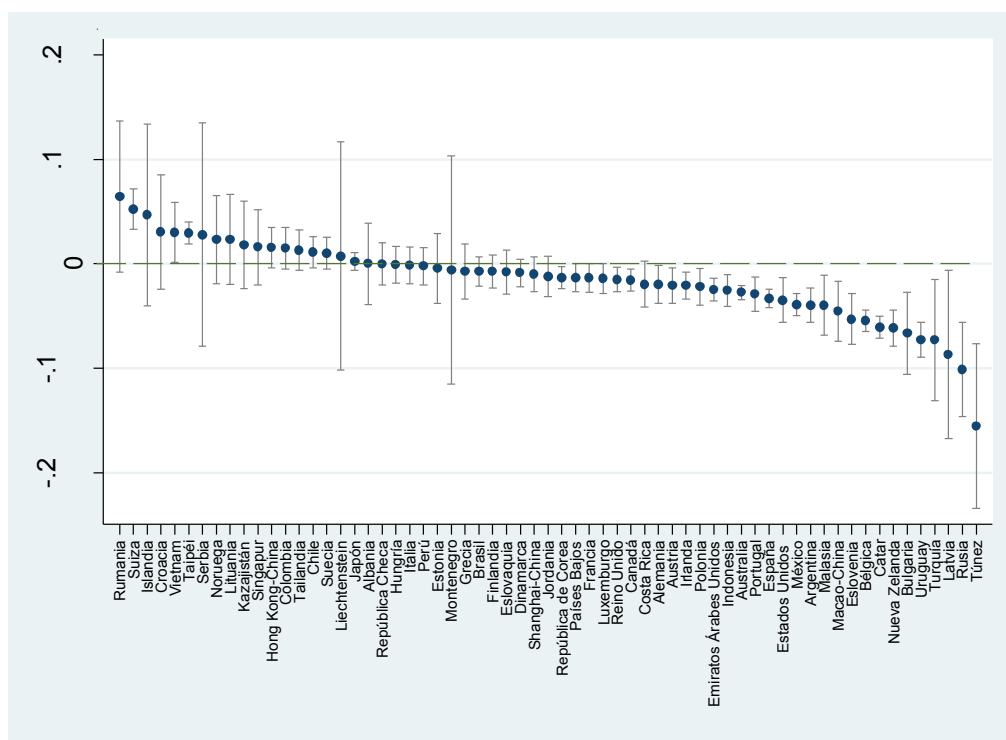
Sin embargo, al evaluar la significatividad estadística de estas brechas por tipo de escuela (es decir, la diferencia entre el decaimiento estimado por país para los alumnos de escuelas privadas menos el decaimiento de los alumnos de escuelas públicas), no se encuentran resultados concluyentes. Esto se debe a que para la mitad de los países las brechas no son estadísticamente significativas, mientras que para los países en que las brechas sí lo son, se encuentran resultados mixtos. En el Gráfico 10 puede verse que si bien la mayoría de los países en que la brecha es significativa, la misma refleja mayores habilidades no cognitivas para los estudiantes de escuelas privadas, también hay casos en que la brecha se invierte a favor de las escuelas públicas.

Gráfico 9. Tasas de respuesta correcta para la primera y la última pregunta de PISA 2012, por tipo de escuela y región.



Fuente: elaboración propia con base en PISA 2012.

Gráfico 10. Brechas por tipo de escuela en habilidades no cognitivas para todos los países participante de la prueba PISA 2012*.



Fuente: elaboración propia con base en PISA 2012.

*La base de datos original no presenta valores para escuelas privadas en Israel ni en Perm (Rusia).

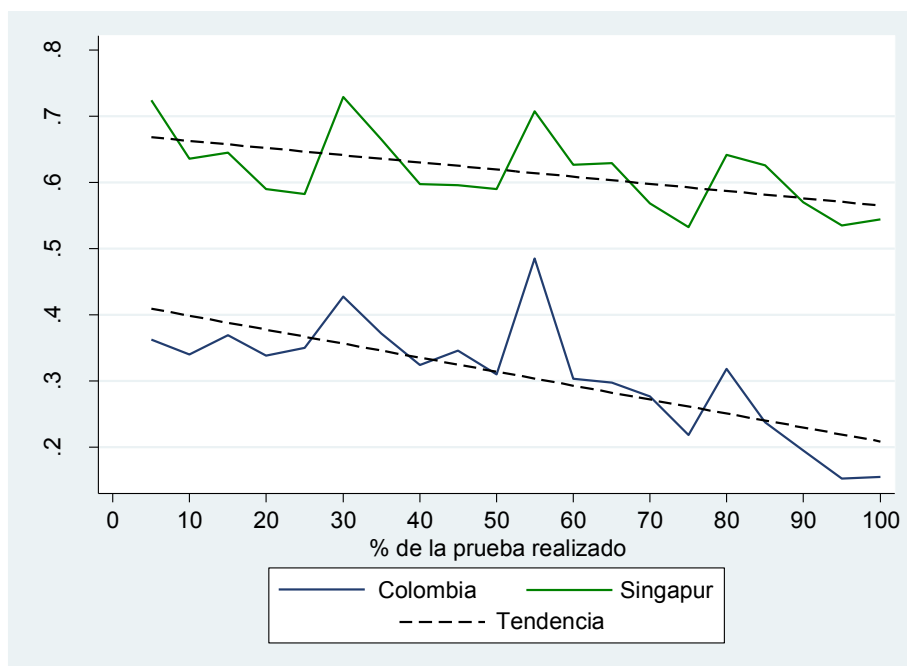
VI.II.iii. Extensión: evolución del decaimiento

Es interesante observar una de extensión a lo presentado hasta ahora: qué es lo que sucede a lo largo de la prueba, más allá de la comparación entre la tasa de respuesta correcta para el primer y el último ítem. La idea detrás de este ejercicio es advertir si hay cambios repentinos en la perseverancia, atención o grado de concentración de los estudiantes, tales que provocan que contesten peor las últimas preguntas, o si el decaimiento es una disminución progresiva en la tasa de respuestas correctas. Esto podría ser relevante para el diseño de estrategias que alivien este problema. Saber si son lapsos de desconcentración, si los mismos aparecen solo al cambiar de bloque de preguntas, si luego del descanso se nota un cierto aumento de la tasa de respuestas correctas, siendo que el recreo sirve para renovar las fuerzas, la concentración y la motivación del examen, entre otras posibilidades, puede permitir encarar el problema de la mejor manera posible. Incluso para adaptar las políticas o estrategias para cada país, puesto que podría darse el caso que dos países tuvieran el mismo decaimiento (o la misma falta de habilidades no cognitivas), pero que la situación sea más grave en uno de los dos, puesto que la concentración es mucho más variable, o en general los estudiantes no puedan responder un examen de más de x horas porque a partir de ese punto su desempeño cae drásticamente.

Con el fin de aportar luz sobre este interrogante, el Gráfico 11 presenta la evolución del desempeño en la prueba PISA para Colombia y Singapur, para contrastar un país con bajo puntaje y bajos niveles de habilidades no cognitivas versus un país con altos niveles de ambas habilidades. El gráfico representa para cada país el puntaje promedio obtenido por los estudiantes en diferentes puntos de avance en la realización de la prueba.

No solamente se puede ver que la línea sólida se ubica siempre en diferentes niveles de puntaje, mostrando que los alumnos de Singapur superan los resultados promedio obtenidos por los alumnos colombianos a lo largo de toda la prueba (habilidades cognitivas), sino que además la línea punteada, que representa la tendencia de estos puntajes, tiene más pendiente en el caso de Colombia que en el caso de Singapur, por lo cual la caída en las notas a lo largo de la prueba es mayor (habilidades no cognitivas).

Gráfico 11. Evolución del puntaje promedio (tasas de respuesta correcta) en la prueba PISA 2012, en diferentes momentos del examen.



Fuente: elaboración propia con base en PISA 2012.

Este gráfico, y los de algunos países más que pueden verse en el Gráfico A.2 del Apéndice, permiten analizar la forma en que se mueven los puntajes promedio a lo largo de la prueba³¹. En ambos gráficos se pueden distinguir claramente tres picos. Uno de estos, se mantiene para todos los países y es el ubicado luego de que el 50% de la prueba haya tenido lugar. Esto es cercano al momento en que los estudiantes tienen un recreo (ver Tabla A.1. en el Apéndice). Pero las variaciones son visiblemente más grandes para Colombia que para Singapur, incluyendo un importante decaimiento en el desempeño al final.

VII. Consideraciones finales

Las habilidades cognitivas y no cognitivas son importantes tanto a nivel individual como a nivel agregado, como lo muestra la abundante literatura que vincula el stock de este tipo de habilidades con resultados en múltiples dimensiones de la vida (como la salud, la

³¹ Puede notarse en ambos gráficos que los niveles son menores que los que se presentan en el Gráfico 5. Esto se debe a que los resultados de la regresión controlan por la dificultad de la pregunta. Sin embargo, la magnitud de las pendientes se mantiene de acuerdo a los resultados presentados. Esto se debe a que como se explicó, la dificultad es aleatoria a lo largo de la prueba.

felicidad, el mercado laboral y los logros académicos), y también con el crecimiento de los países.

Si bien se han desarrollado diversas alternativas cada vez más sofisticadas para poder obtener buenas mediciones de las habilidades cognitivas, la medición de las habilidades no cognitivas es incipiente y todavía no se conoce demasiado respecto de cuáles son las mejores herramientas para lograr mediciones apropiadas. Esto se debe a que, además del problema de identificación que debe sortearse para medir cualquier habilidad, las habilidades no cognitivas son particularmente difíciles de medir, porque contienen múltiples aristas y no son fácilmente jerarquizables como sí lo son las habilidades cognitivas.

Las alternativas más populares para obtener mediciones de las habilidades no cognitivas sufren de problemas relacionados con el auto-reporte, la deseabilidad social de la respuesta, el sesgo de referencia desde el cual las personas contestan preguntas de autopercepción y la dependencia de la capacidad de comprensión de los individuos al auto-reportar correcta y rigurosamente sus habilidades no cognitivas. Este trabajo utiliza una medida de habilidades no cognitivas exenta de estos problemas, para América Latina y todos los países participantes de la prueba PISA 2012. A través de la estrategia planteada, se extraen dos componentes de la nota de la prueba: el nivel inicial de desempeño y el decaimiento del desempeño a lo largo del examen (medido como la diferencia entre las tasas de respuesta correcta para la primera y la última pregunta), tomados como medidas de habilidades cognitivas y no cognitivas, respectivamente.

Los resultados obtenidos indican que América Latina se encuentra rezagada en su stock ambos tipos de habilidades, en comparación con lo observado para otras regiones. Además, este trabajo es pionero en intentar caracterizar el perfil de los estudiantes con habilidades no cognitivas más bajas. Para ello se analizan los resultados por género y por tipo de escuela (públicas vs. privadas). Al respecto, los resultados encontrados revelan una brecha de género positiva y significativa para la gran mayoría de los países, indicando mayores habilidades no cognitivas para las mujeres. Por su parte, los resultados relacionados brechas en las habilidades no cognitivas por tipo de escuela no son estadísticamente significativos.

Estos hallazgos proponen nuevas preguntas para continuar avanzando en esta línea de investigación como por ejemplo, ¿qué más hay detrás del bajo stock de habilidades en los estudiantes latinoamericanos? o ¿por qué se generan las brechas encontradas entre países? Existe evidencia de que un mayor nivel de desarrollo se asocia con la generación

de mejores habilidades de ambos tipos, pero es necesario profundizar el estudio de los mecanismos detrás de ello.

Asimismo, cabe preguntarse por las brechas de género. En la actualidad, la explicación que prima respecto de esta brecha en rendimiento académico a favor de los hombres, es la que argumenta que las diferencias se deben a distintos niveles de habilidades cognitivas y también a distintas presiones y aspiraciones culturales para hombres y mujeres. Los resultados de este trabajo, sugerirían que estos dos factores podrían tener efectos mayores de los que la literatura estima. Al presentar las mujeres mayores habilidades no cognitivas, pero igual obtener peores resultados en las pruebas, podría darse el caso de que las mujeres estén compensando con su esfuerzo y motivación las diferencias cognitivas y culturales, pero que los efectos de las mismas sean demasiado grandes para generar un cambio en los resultados académicos.

En lo referente al tipo de escuela, el hecho de no encontrar diferencias concluyentes en las medidas de habilidades no cognitivas entre ambos grupos, abre la puerta a varias interpretaciones. Por un lado, si uno considera al tipo de escuela como *proxy* de la situación socioeconómica de un estudiante, el resultado es intuitivo en los casos donde la brecha se da a favor de los estudiantes de escuelas privadas, pues el estatus socioeconómico familiar es un insumo previo importante en la formación de habilidades no cognitivas, tal como gran parte de la literatura sostiene. Pero por otro lado, si se piensa en el tipo de escuela como una *proxy* de dos tecnologías educativas distintas, la conclusión es que ambas tecnologías parecen tener diferentes productividades en la generación de habilidades socioemocionales en distintos países. Sin embargo, es posible también que los patrones de segregación socioeconómica por tipo de escuela sean lo que esté guiando la diversidad de resultados encontrados. Adicionalmente, si consideramos el hecho de que la prueba utilizada para realizar el análisis (PISA) evalúa solo a jóvenes matriculados de 15 años de edad, es probable que exista un efecto de selección. Es decir, si los alumnos menos motivados (con menos habilidades no cognitivas) de escuelas públicas (privadas) abandonan la educación antes de los 15 años, entonces la brecha encontrada en favor de las escuelas privadas estará subestimada (sobreestimada).

Estos resultados, son particularmente importantes para el diseño de políticas públicas que intenten mitigar las diferencias encontradas. Un estudio en mayor profundidad de las raíces de estas diferencias puede encaminar políticas más efectivas para aumentar el capital humano de los países, considerando simultáneamente la importancia de las habilidades cognitivas y la de las no cognitivas.

En general, la atención de los *policy-makers* de nuestra región se ha centrado en destinar recursos a la educación, principalmente a través de programas de inclusión, construcción de escuelas y más horas de clase o cambios en el currículo escolar, pero no se ha hecho mucho en cuanto la enseñanza y promoción de habilidades no cognitivas entre los alumnos. Se ha demostrado que las mismas son de vital importancia tanto para un mejor desempeño escolar como para muchos otros aspectos esenciales en el desarrollo de las personas y de las naciones (empleabilidad, calidad del empleo, PIB, salud, etc.). Estas habilidades pueden ser promovidas desde las aulas, incluyendo el desarrollo de habilidades no cognitivas en los currículos, y también desde los hogares, fomentando mejores ambientes de desarrollo a edad temprana. Dado que la inteligencia (cristalizada) es maleable hasta el final de la vida de las personas, las inversiones que se hagan en toda la población tendrán potenciales retornos positivos, que luego se volcarán en las siguientes generaciones.

Más investigación es necesaria para profundizar en las posibles causas y soluciones para la problemática de la región. Mediciones cada vez más precisas y de más amplia cobertura permitirán responder las preguntas que quedan pendientes y contribuirán al desarrollo de herramientas útiles para la formulación de políticas de educación y de trabajo integrales y de calidad, para una mejor formación del capital humano en América Latina.

VIII. Referencias Bibliográficas

Acosta, P., Muller, N., y Sarzosa, M. A. (2015). Beyond qualifications: returns to cognitive and socioemotional skills in Colombia. (Documento de trabajo N° 7430). Washington, D.C.: World Bank Policy Research.

Agan, Amanda Y. (2011). Non-cognitive skills and crime. University of Chicago, Mimeo.

Angrist, J. D., y Krueger, A. B. (1991). Does Compulsory School Attendance Affect Schooling and Earnings?. *The Quarterly Journal of Economics*, 106(4), 979-1014.

Balart, P., Oosterveen, M., y Webbink, D. (2015). Test Scores, Noncognitive Skills and Economic Growth (No. 9559). Institute for the Study of Labor (IZA).

Barro, R. J. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, 106(2), 407-443.

Bassi, M.; Busso, M.; Urzúa, S.; y Vargas, J. (2012). *Disconnected Skills, Education and Employment in Latin America*. Washington: IDB Publications

Borghans, L. y Schils, T. (2012). The leaning tower of Pisa. Decomposing achievement test scores into cognitive and noncognitive components. Disponible en: <http://www.sole-jb>

Borghans, L., B. H. Golsteyn, J. Heckman, y J. E. Humphries (2011). Identification problems in personality psychology. *Personality and Individual Differences* 51 (3), 315-320.

Borghans, L., Duckworth, A. L., Heckman, J. J., y Ter Weel, B. (2008a). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4), 72-1059.

Borghans, L., Meijers, H., y Ter Weel, B. (2008b). The role of noncognitive skills in explaining cognitive test scores. *Economic Inquiry*, 46(1), 2-12.

Bowles, S. y H. Gintis (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.

Bransford, J.D., Brown, A.L., y Cocking, R.R. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Butler, J., y Adams, R. J. (2007). The impact of differential investment of student effort on the outcomes of international studies. *Journal of applied measurement*, 8(3), 279.

CAF (2016). *Reporte de Economía y Desarrollo 2016: Más habilidades para el trabajo y la vida: los aportes de la familia, la escuela, el entorno y el mundo laboral*. Caracas: CAF.

- Card, D. (1999). The Causal Effect of Education on Earnings. *Handbook of Labor Economics*, vol. 3, 1801–1863.
- Conti, G., Heckman, J., y Urzúa, S. (2010). The education-health gradient. *The American economic review*, 100(2), 234.
- Díaz, J. J., O. Arias, and D. V. Tudela (2012). Does Perseverance Pay as Much as Being Smart?: The Returnsto Cognitive and Non-Cognitive Skillsin Urban Peru. Mimeo. The World Bank, Washington, DC.
- Duckworth, A. L., P. D. Quinn, D. R. Lynam, R. Loeber, y M. Stouthamer-Loeber (2011). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences* 108 (19), 7716-7720.
- Duckworth, A.L, y Quinn, P.D. (2009). Development and validation of the Short Grit Scale (GritS). *Journal of Personality Assessment*, 91, 166-174.
- Gneezy, U., Meier, S., y Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *The Journal of Economic Perspectives*, 25(4), 191-209.
- Goldberg, L. R. (1990). An alternative "description of personality": the big-five factor structure. *Journal of personality and social psychology*, 59(6), 1216.
- Hanushek, E. A., y Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4), 267-321.
- Hanushek, E. A., y Zhang, L. (2009). Quality-consistent estimates of international schooling and skill gradients. *Journal of Human Capital*, 3(2), 107-143.
- Heckman, J. J. y Y. Rubinstein (2001). The Importance of Noncognitive Skills: Lessons from the GED Testing Program. *American Economic Review* 91 (2), 145-149.
- Heckman, J. J., Stixrud J., y Urzúa S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24 (3), 411-482.
- Heckman, J. J., y Kautz, T. (2013). Fostering and measuring skills: Interventions that improve character and cognition (WP N° 19656). National Bureau of Economic Research.

Heine, S. J., Buchtel, E. E., y Norenzayan, A. (2008). What do cross-national comparisons of personality traits tell us? The case of conscientiousness. *Psychological Science* 19(4), 309–313.

Heineck, G., y Anger, S. (2010). The returns to cognitive abilities and personality traits in Germany. *Labour Economics*, 17(3), 535-546.

Hernández, M. y J. Hershaff (2014). Skipping questions in school exams: The role of socio-emotional skills on educational outcomes. Draft version: March 18, 2014

Hitt, C., Trivitt, J., y Cheng, A. (2016). When you say nothing at all: The predictive power of student effort on surveys. *Economics of Education Review*, 52, 105-119.

Humphries, J. E., y Kosse, F. (2017). On the interpretation of non-cognitive skills—what is being measured and why it matters. *Journal of Economic Behavior & Organization*, 136, 174-185.

Kautz T., Heckman JJ, Diris R., Ter Weel B., y Borghans L. (2014). *Fostering and Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success*. Paris: OECD.

Koch, A.; J. Nafziger y H. Nielsen (2014). *Behavioral Economics of Education*, Institute for the Study of Labor (IZA), IZA Discussion Papers 8470.

Lavado, P., Velarde, L., y Yamada, G. (2014). *Cognitive and Socioemotional Skills and Wages: The role of latent abilities on the gender wage gap in Peru (No, 2014-16)*. Peruvian Economic Association.

Levin, H. M. (2012). More than just test scores. *Prospects*, 42(3), 269-284.

Levitt, S. D., List, J. A., Neckermann, S., y Sadoff, S. (2012). *The Behavioralist Goes to School: Leveraging Behavioral Economics to Improve Educational Performance*. NBER Working Paper No. 18165. National Bureau of Economic Research.

Marchionni, M., F. Pinto y E. Vazquez (2013). “Determinantes de la Desigualdad en el Desempeño Educativo en la Argentina”. *Anales de la Asociación Argentina de Economía Política*, XLVIII Reunión Anual, Rosario, Argentina, pp. 1-34.

Mischel, W. (2014). *The Marshmallow Test: Mastering Self-control*. New York, NY: Hachette Book Group.

Mincer, J. A. (1974). *Schooling, Experience, and Earnings*. NBER Books.

- Mueller, G., y Plug, E. J. S. (2006). Estimating the effect of personality on male-female earnings. *Industrial & Labor Relations Review*, 60(1), 3-22.
- Murnane, R., Willett J., Duhaldeborde Y., y Tyler J. H. (2000). How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings? *Journal of Policy Analysis and Management* 19 (4): 547-68.
- OECD (2013). *Skills Outlook 2013: First Results from the Survey of Adult Skills*. OECD Publishing, París.
- OECD (2012). *PISA 2012 Technical Report*, Paris OECD.
- OECD (2009a). *PISA 2006 Data Analysis User Manual*, Paris: OECD.
- OECD (2009b). *PISA Data Analysis Manual SPSS. Second Edition*, Paris: OECD.
- Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling, *Journal of Public Economics* (97), 2213-2229.
- Paulhus, D. L. (1991). Measurement and control of response bias. En J. P. Robinson, P. R. Shaver, y L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego: Academic Press.
- Psacharopoulos, G., y Patrinos, H. A. (2004). Returns to investment in education: a further update. *Education economics*, 12(2), 111-134.
- Sala-i Martin, X., G. Doppelhofer, y R. Miller (2004). Determinants of long-term growth: A bayesian averaging of classical estimates (bace) approach. *The American Economic Review* 94(4), 813-835.
- Schwarzer, R. y Jerusalem, M. (1995). Generalized Self-Efficacy scale. En J. Weinman, S. Wright y M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Casual and control beliefs* (pp. 35-37). Windsor, UK: NFER-NELSON.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-Verlag.
- West, M. R., Kraft, M. A., Finn, A. S., Martin, R. E., Duckworth, A. L., Gabrieli, C. F., y Gabrieli, J. D. (2015). Promise and Paradox Measuring Students' Non-Cognitive Skills and the Impact of Schooling. *Educational Evaluation and Policy Analysis*.

IX. Apéndice

TABLAS

Tabla A.1. Cronograma de la prueba PISA.

Cronograma de la prueba PISA	
Tarea	Tiempo
Distribución del material y explicación de las consignas	10-15 minutos*
Cuadernillo de preguntas	1 hora**
Recreo corto	Hasta 5 minutos
Cuadernillo de preguntas	1 hora**
Recreo largo	15 minutos
Cuestionario para el estudiante	35 minutos*
Recolección del material y finalización del examen	3-5 minutos*
<i>Total</i>	3 horas y 15 minutos*

* Duración aproximada

** Duración exacta

Fuente: Elaboración propia en base a OECD (2012).

Tabla A.2. Cantidad de preguntas promedio de los dos tipos de sets de cuadernillos.

Cuadernillo	Cantidad de preguntas	
	Set Estándar	Set de menor dificultad
1	60	
2	58	
3	57	
4	51	
5	54	
6	52	
7	53	
8	63	63
9	53	53
10	53	53
11	49	49
12	60	60
13	58	58
21		60
22		58
23		57
24		51
25		54
26		51
27		53
<i>Total</i>	721	720
<i>Promedio</i>	55,46	55,38

Fuente: elaboración propia en base a datos PISA 2012.

Tabla A.4. Distribución de los cuadernillos por género y tipo de escuela en cada país.

Cuadernillo	Escuelas Públicas			Escuelas Privadas		
	Hombres	Mujeres	Total	Hombres	Mujeres	Total
1	10.507	10.607	21.114	2.528	2.588	5.116
2	10.386	10.572	20.958	2.694	2.467	5.161
3	10.699	10.312	21.011	2.569	2.602	5.171
4	10.627	10.381	21.008	2.649	2.538	5.187
5	10.586	10.464	21.050	2.591	2.588	5.179
6	10.637	10.573	21.210	2.544	2.646	5.190
7	10.469	10.604	21.073	2.578	2.632	5.210
8	14.846	15.104	29.950	3.561	3.656	7.217
9	14.777	14.937	29.714	3.592	3.635	7.227
10	13.588	14.067	27.655	3.404	3.484	6.888
11	13.656	13.838	27.494	3.442	3.462	6.904
12	13.618	13.923	27.541	3.400	3.444	6.844
13	13.407	14.045	27.452	3.383	3.538	6.921
20	454	315	769	220	100	320
21	4.122	4.565	8.687	1.026	1.067	2.093
22	4.093	4.658	8.751	977	1.058	2.035
23	4.186	4.511	8.697	1.017	1.064	2.081
24	4.191	4.542	8.733	1.013	1.070	2.083
25	4.178	4.549	8.727	1.018	1.041	2.059
26	4.142	4.520	8.662	1.006	1.064	2.070
27	4.082	4.551	8.633	1.012	1.073	2.085
Total	187.251	191.638	378.889	46.224	46.817	93.041

Fuente: elaboración propia en base a datos PISA 2012.

Tabla A.5. Clasificación de los países por región.

Región	Países
Asia Oriental y el Pacífico	Australia, Hong Kong-China, Indonesia, Japón, Macao-China, Malasia, Nueva Zelanda, República de Corea, Shanghai-China, Singapur, Tailandia, Taipéi, Vietnam,
Europa y Asia Central	Albania, Alemania, Austria, Bélgica, Bulgaria, Croacia, Dinamarca, Eslovaquia, Eslovenia, España, Estonia, Finlandia, Francia, Grecia, Hungría, Irlanda, Islandia, Italia, Kazajistán, Latvia, Liechtenstein, Lituania, Luxemburgo, Montenegro, Noruega, Países Bajos, Perm (Rusia), Polonia, Portugal, Reino Unido, República Checa, Rumania, Rusia, Serbia, Suecia, Suiza, Turquía,
Medio Oriente y Norte de África	Catar, Emiratos Árabes Unidos, Israel, Jordania, Túnez,
América del Norte	Canadá, Estados Unidos,
Latinoamérica	Argentina, Brasil, Chile, Colombia, Costa Rica, México, Perú, Uruguay,

Fuente: Banco Mundial.

Tabla A.6. Puntaje promedio de las preguntas por cuadernillo. Comparación cuadernillos estándar versus cuadernillos de menor dificultad.

Cuadernillo	Puntaje promedio de las preguntas por cuadernillo						
	Argentina	Brasil	Chile	Colombia	México	Perú	Uruguay
8	0,352	0,314	0,434	0,327	0,357	0,274	0,351
9	0,330	0,307	0,427	0,316	0,370	0,278	0,352
10	0,307	0,271	0,400	0,277	0,339	0,252	0,326
11	0,329	0,316	0,429	0,312	0,361	0,285	0,360
12	0,314	0,283	0,410	0,298	0,336	0,259	0,333
13	0,324	0,316	0,448	0,316	0,360	0,273	0,342
21	0,301	0,280	0,426	0,295	0,330	0,245	0,344
22	0,383	0,350	0,498	0,379	0,420	0,321	0,394
23	0,356	0,322	0,437	0,326	0,372	0,293	0,362
24	0,301	0,287	0,415	0,314	0,355	0,272	0,331
25	0,290	0,260	0,378	0,275	0,318	0,245	0,320
26	0,308	0,285	0,405	0,286	0,340	0,250	0,322
27	0,302	0,275	0,405	0,296	0,328	0,257	0,333

Fuente: elaboración propia en base a datos PISA 2012.

Tabla A.7. Niveles iniciales y decaimiento en el desempeño de la prueba PISA 2012 por país. Muestra de hombres.

País	Indicador de Habilidades no cognitivas (negativo)	País	Indicador de Habilidades no cognitivas (negativo)
Colombia	0,281	Polonia	0,108
Perú	0,254	Kazajistán	0,108
Uruguay	0,250	Reino Unido	0,107
México	0,247	Eslovaquia	0,106
Brasil	0,242	Hungría	0,105
Túnez	0,237	Latvia	0,102
Grecia	0,216	Luxemburgo	0,102
Catar	0,204	Eslovenia	0,100
Costa Rica	0,198	Croacia	0,094
Argentina	0,186	Nueva Zelanda	0,093
Israel	0,182	Estados Unidos	0,093
Jordania	0,179	Canadá	0,090
Chile	0,177	Estonia	0,081
Albania	0,165	Países Bajos	0,080
Tailandia	0,158	Australia	0,077
Montenegro	0,155	Bélgica	0,076
Bulgaria	0,152	Irlanda	0,075
Rumania	0,148	Macao-China	0,074
Emiratos Árabes Unidos	0,147	Dinamarca	0,074
Rusia	0,147	República Checa	0,072
Indonesia	0,146	Alemania	0,072
Serbia	0,146	Japón	0,072
Suecia	0,143	Suiza	0,068
Portugal	0,139	República de Corea	0,066
Noruega	0,132	Singapur	0,062
Italia	0,128	Austria	0,062
España	0,128	Liechtenstein	0,059
Perm (Rusia)	0,122	Vietnam	0,059
Francia	0,120	Finlandia	0,056
Islandia	0,120	Shanghai-China	0,055
Malasia	0,114	Hong Kong-China	0,048
Turquía	0,110	Taipéi	0,048
Lituania	0,109		

Fuente: elaboración propia en base a datos PISA 2012.

Tabla A.8. Niveles iniciales y decaimiento en el desempeño de la prueba PISA 2012 por país. Muestra de mujeres.

País	Indicador de Habilidades no cognitivas (negativo)	País	Indicador de Habilidades no cognitivas (negativo)
<i>Colombia</i>	0,276	Islandia	0,091
<i>Uruguay</i>	0,255	Eslovaquia	0,089
<i>México</i>	0,242	Reino Unido	0,086
<i>Brasil</i>	0,236	Canadá	0,084
<i>Perú</i>	0,230	Lituania	0,083
Grecia	0,224	Macao-China	0,080
Túnez	0,214	Latvia	0,079
Catar	0,189	Nueva Zelanda	0,077
Israel	0,187	Hungría	0,075
<i>Costa Rica</i>	0,186	Polonia	0,075
<i>Argentina</i>	0,173	Países Bajos	0,074
Jordania	0,171	Estados Unidos	0,073
Albania	0,162	Dinamarca	0,070
<i>Chile</i>	0,155	República Checa	0,066
Bulgaria	0,146	República de Corea	0,065
Tailandia	0,138	Liechtenstein	0,064
Rusia	0,136	Croacia	0,064
Serbia	0,135	Australia	0,063
Portugal	0,131	Eslovenia	0,062
Montenegro	0,131	Bélgica	0,059
Kazajistán	0,130	Estonia	0,059
Rumania	0,127	Alemania	0,057
Perm (Rusia)	0,125	Japón	0,057
Suecia	0,124	Vietnam	0,054
Indonesia	0,123	Suiza	0,053
Emiratos Árabes Unidos	0,121	Irlanda	0,050
España	0,114	Hong Kong-China	0,048
Malasia	0,113	Austria	0,047
Italia	0,110	Singapur	0,042
Noruega	0,103	Taipéi	0,042
Luxemburgo	0,095	Shanghai-China	0,037
Francia	0,095	Finlandia	0,036
Turquía	0,092		

Fuente: elaboración propia en base a datos PISA 2012.

Tabla A.9. Niveles iniciales y decaimiento en el desempeño de la prueba PISA 2012 por país. Escuelas públicas.

País	Indicador de Habilidades no cognitivas (negativo)	País	Indicador de Habilidades no cognitivas (negativo)
Colombia	0,278	Francia	0,106
Uruguay	0,268	Reino Unido	0,104
México	0,252	Turquía	0,102
Perú	0,244	Luxemburgo	0,101
Brasil	0,244	Eslovaquia	0,096
Catar	0,239	Lituania	0,095
Túnez	0,227	Latvia	0,093
Grecia	0,221	Polonia	0,091
Costa Rica	0,198	Hungría	0,091
Argentina	0,196	Nueva Zelanda	0,090
Israel	0,184	Canadá	0,088
Jordania	0,178	Estados Unidos	0,086
Chile	0,167	Países Bajos	0,085
Albania	0,163	Australia	0,082
Emiratos Árabes Unidos	0,160	Eslovenia	0,081
Bulgaria	0,150	Croacia	0,078
Indonesia	0,146	Dinamarca	0,077
Tailandia	0,145	República de Corea	0,072
Montenegro	0,142	República Checa	0,072
Rusia	0,142	Irlanda	0,071
Serbia	0,142	Estonia	0,068
Portugal	0,138	Alemania	0,066
Rumania	0,137	Japón	0,066
España	0,131	Liechtenstein	0,064
Suecia	0,128	Suiza	0,057
Macao-China	0,127	Austria	0,057
Perm (Rusia)	0,123	Vietnam	0,054
Kazajistán	0,119	Singapur	0,051
Italia	0,118	Shanghai-China	0,047
Noruega	0,117	Finlandia	0,045
Malasia	0,117	Hong Kong-China	0,040
Bélgica	0,108	Taipéi	0,037
Islandia	0,106		

Fuente: elaboración propia en base a datos PISA 2012.

Tabla A.10. Niveles iniciales y decaimiento en el desempeño de la prueba PISA 2012 por país. Escuelas privadas*.

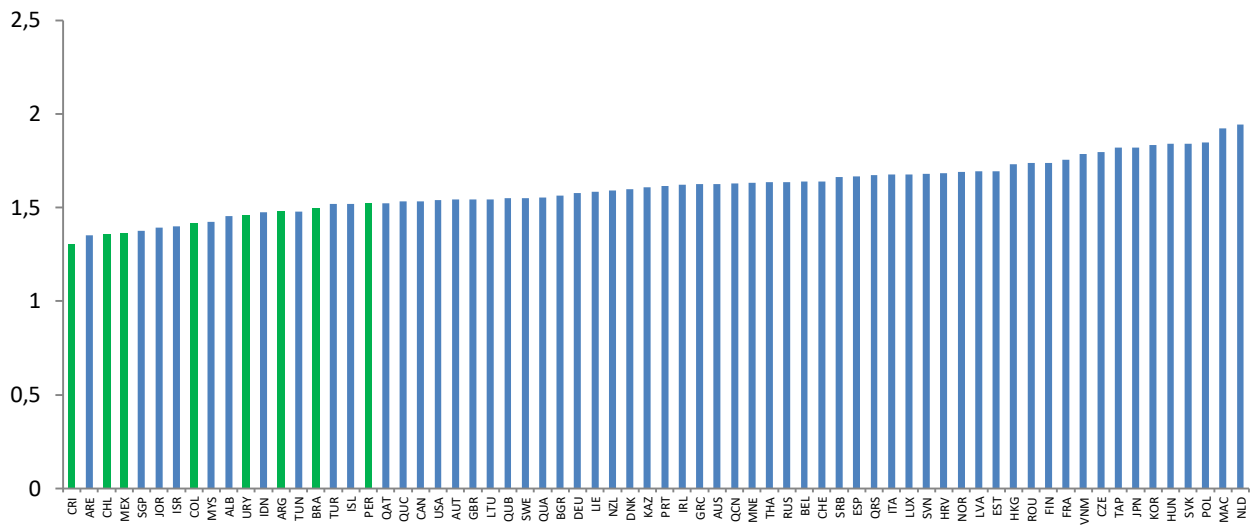
País	Indicador de Habilidades no cognitivas (negativo)	País	Indicador de Habilidades no cognitivas (negativo)
Colombia	0,277	Luxemburgo	0,087
Rumania	0,237	Vietnam	0,086
Perú	0,226	Polonia	0,086
Túnez	0,214	Canadá	0,084
Brasil	0,214	Liechtenstein	0,084
México	0,199	Francia	0,083
Uruguay	0,169	Montenegro	0,082
Chile	0,168	Singapur	0,077
Catar	0,166	Macao-China	0,075
Albania	0,162	República Checa	0,072
Tailandia	0,161	Países Bajos	0,071
Suecia	0,159	Dinamarca	0,062
Argentina	0,155	Japón	0,062
Jordania	0,154	Noruega	0,059
Costa Rica	0,152	Taipéi	0,059
Serbia	0,148	República de Corea	0,058
Croacia	0,146	Irlanda	0,058
Grecia	0,145	Australia	0,057
Kazajistán	0,139	Eslovenia	0,052
Indonesia	0,120	Malasia	0,052
Islandia	0,120	Estados Unidos	0,052
Rusia	0,118	Bélgica	0,052
Italia	0,117	Lituania	0,051
Emiratos Árabes Unidos	0,115	Hong Kong-China	0,050
Portugal	0,112	Turquía	0,048
Eslovaquia	0,109	Finlandia	0,042
Suiza	0,108	Alemania	0,038
Bulgaria	0,102	Austria	0,036
España	0,102	Shanghai-China	0,036
Hungría	0,102	Nueva Zelanda	0,020
Estonia	0,090	Latvia	-0,014
Reino Unido	0,089		

Fuente: elaboración propia en base a datos PISA 2012.

*La base de datos original no presenta valores para escuelas privadas en Israel ni en Perm (Rusia).

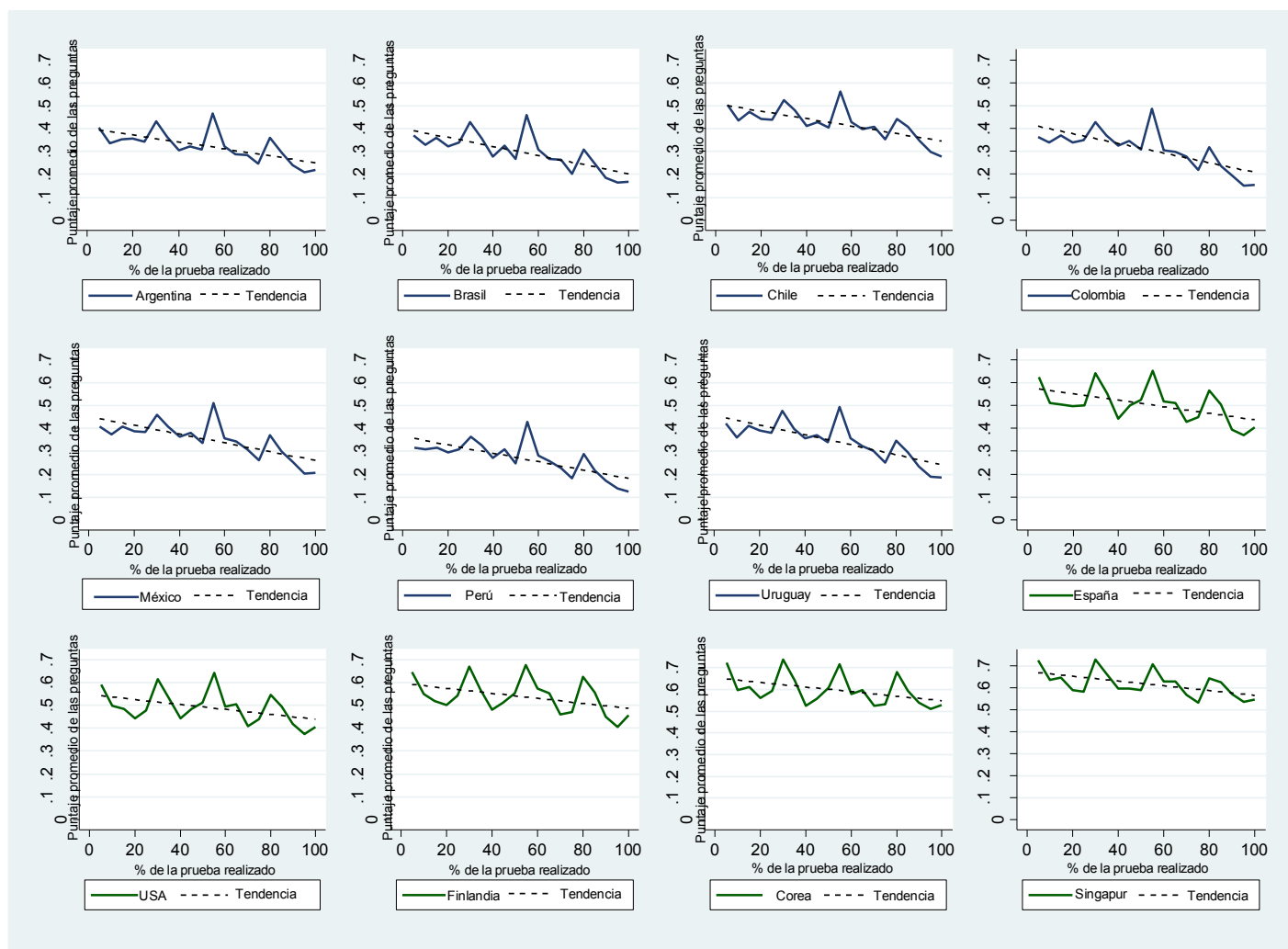
GRÁFICOS

Gráfico A.1. Respuestas al ítem “Creo que puedo triunfar si me esfuerzo lo suficiente” con datos de PISA 2012, Respuestas ponderadas promedio en base a lo reportado por los alumnos, para cada país.



Fuente: Elaboración propia en base a OECD (2012).

Gráfico A.2. Evolución del puntaje promedio en la prueba PISA 2012, a diferentes porcentajes de completitud de la prueba para países de América Latina y países seleccionados de la OCDE.



Fuente: Elaboración propia en base a OECD (2012).