



**“HABILIDADES COGNITIVAS Y SOCIOEMOCIONALES, Y
SALARIOS: EL ROL DE LAS HABILIDADES LATENTES EN LA
BRECHA SALARIAL DE GÉNERO EN EL PERÚ”**

**Trabajo de Investigación presentado
para optar al Grado Académico de
Magíster en Economía**

Presentado por

Srta. Luciana Velarde Arrisueño

Asesor: Profesor Pablo Lavado Padilla

2013

Para mi mamá, Mónica, mi motivación y mi razón de ser.

“En la vida no gana el más grande ni el más fuerte, sino aquel que cree poder hacerlo” (adaptación de la frase de Rudyard Kipling).

Agradezco la contribución de mi asesor Pablo Lavado, cuyo apoyo fue fundamental para el desarrollo de este documento. También agradezco los comentarios de Sergio Urzúa y de los participantes de la conferencia “Conference on Skills, Education and Labor Market Outcomes” llevada a cabo en la Universidad de Maryland; de la conferencia “Inequalities in Children’s Outcomes in Developing Countries” llevada a cabo en la Universidad de Oxford, y del “Seminario de Investigación” del Banco Central de Reserva.

Resumen ejecutivo

En los últimos años, la literatura ha analizado la relación entre las habilidades cognitivas y no cognitivas y los resultados del mercado laboral. Más aún, se ha estudiado la correlación entre diferencias en estas habilidades y diferencias en salarios entre hombres y mujeres, pero la mayoría de los estudios en torno al tema se han enfocado en países desarrollados.

El principal objetivo de este estudio es analizar el rol de las habilidades cognitivas y no cognitivas en la brecha salarial de género en el Perú. Para ello, proponemos utilizar información longitudinal de medidas de habilidad para estimar componentes inobservables (latentes) de ambas habilidades y analizar el rol de dichos componentes en el salario.

Los resultados indican que existe una brecha salarial de género significativa en el Perú y que la misma puede ser explicada por diferencias en el retorno de habilidades cognitivas latentes y por diferencias en la dotación de habilidades no cognitivas latentes. Luego de estimar un modelo conjunto de educación, empleo, ocupación y salarios, se observa que aun cuando las habilidades latentes explican la brecha de salarios, las mismas son más importantes para explicar la diferencia en la decisión tomada por hombres y mujeres en cuanto a su ocupación.

Índice

Índice de tablas.....	vii
Índice de anexos	viii
Resumen ejecutivo.....	iv
Capítulo I. Introducción	1
Capítulo II. Revisión de literatura.....	3
Capítulo III. Modelo	8
1. Modelo de educación	9
2. Modelo de empleo.....	9
3. Modelo de ocupación.....	10
4. Modelo de salarios	10
Capítulo IV. Fuentes de información y muestra	12
Capítulo V. Implantación econométrica	16
1. Estimación de habilidades latentes	17
2. Brecha salarial del género y el método de descomposición de Blinder-Oaxaca.....	19
Capítulo VI. Resultados.....	20
1. Salarios y medidas de habilidad.....	20
1.1 Método de Descomposición de Blinder-Oaxaca.....	21
1.2 Estimación conjunta: educación, empleo, ocupación y salarios.	23
2. Salarios y habilidades latentes	25
2.1 Aproximando habilidades latentes	25
2.2 Método de Descomposición de Blinder-Oaxaca.....	29
2.3 Estimación conjunta: educación, empleo, ocupación y salarios	31
3. Pruebas de robustez: soporte común.....	32
4. Limitaciones.....	33
Conclusiones y recomendaciones	35

Bibliografía	38
Anexos	40
Nota biográfica	46

Índice de tablas

Tabla 1.	Estadísticas descriptivas: Estudio Niños del Milenio.....	13
Tabla 2.	Estadísticas descriptivas: ENHAB	15
Tabla 3.	Ecuación de Mincer con medidas de habilidad	21
Tabla 4.	Descomposición de Blinder-Oaxaca con medidas de habilidad.....	23
Tabla 5.	Estimación conjunta con medidas de habilidad.....	24
Tabla 6.	Estimación de la primera etapa (modelo de efectos fijos – medidas de habilidad).....	26
Tabla 7.	Estimación de la segunda etapa (habilidades latentes sobre características permanentes, NM).....	27
Tabla 8.	Estadísticas de la tercera etapa (habilidades latentes predichas en ambas bases de datos).....	28
Tabla 9.	Ecuación de Mincer con habilidades latentes.....	29
Tabla 10.	Descomposición de Blinder-Oaxaca con habilidades latentes	31
Tabla 11.	Estimación conjunta con habilidades latentes	32

Índice de anexos

Anexo 1.	Componentes psicosociales – muestras de ENM	41
Anexo 2.	Habilidades latentes y soporte común	43

Capítulo I. Introducción

La existencia de brechas salariales no es un tema nuevo, en particular aquella que implica salarios distintos entre hombres y mujeres. A lo largo de los años, diversos expertos han analizado los factores que podrían estar explicando ésta y otras brechas. Un tema que ha llamado la atención de estudiosos en este campo es el rol que juegan las diferencias en habilidades cognitivas y no cognitivas en las brecha salariales. En particular, se ha documentado qué diferencias en habilidades cognitivas entre hombres y mujeres se encuentran fuertemente relacionadas con diferencias en los salarios de los mismos (Neal y Johnson 1996: 869-895; Ritter y Taylor 2011: 30-42). El mayor nivel de este tipo de habilidades entre los hombres parece contribuir a los mayores salarios que reciben en comparación con las mujeres. Más aún, las brechas en este tipo de habilidad podrían contribuir a brechas de salario no solo por las diferencias en el nivel de este tipo de habilidad, pero también por diferencias en el retorno de las mismas en términos de salario; un punto adicional en una prueba cognitiva podría significar una mayor ganancia en términos de salario para un hombre que para una mujer, *ceteris paribus*.

Recientemente la literatura se ha enfocado en la relación entre las habilidades no cognitivas (ámbitos de la personalidad) y la productividad de los individuos (Heckman *et al.* 2006: 411-482). En cuanto a la relación entre las pruebas que miden habilidades no cognitivas e indicadores de mercado laboral, se ha encontrado que existe una relación positiva entre salarios y algunas medidas de habilidad no cognitiva (Fortin, 2008; Grove *et al.* 2011: 827-874; Cobb-Clark y Tann 2009). En cuanto a cómo se forman este tipo de habilidades, se ha cuestionado la idoneidad de las medidas de habilidad comúnmente utilizadas ya que parecen ser malas aproximaciones de las habilidades latentes debido a los potenciales problemas de error de medición y causalidad reversa a los que son susceptibles (Heckman *et al.* 2006: 411-482). En efecto, uno de los principales factores que determinan los salarios son las habilidades latentes. No obstante, muy pocos estudios han analizado el rol de las habilidades no cognitivas latentes en la brecha salarial de género. Más aún, este tema en particular no ha sido abordado para países en vías de desarrollo y, en particular, para la región de América Latina.

El principal objetivo de este documento es analizar el rol de las habilidades cognitivas y no cognitivas latentes en la brecha salarial de género en un país en vías de desarrollo por medio de aproximaciones de las habilidades latentes.

Para ello, proponemos un procedimiento econométrico que permite estimar aproximaciones de las habilidades cognitivas y no cognitivas latentes sobre la base del modelo de habilidades latentes propuesto por Heckman *et al.* (2006). A diferencia de dichos autores, nosotros explotaremos la disponibilidad de información de datos de panel para utilizar la persistencia a lo largo del tiempo en las pruebas de habilidades como fuente de identificación. Luego, complementaremos los estimados obtenidos con información de salarios con el fin de estimar el rol de las habilidades latentes en la brecha salarial de género. En concreto, partimos de un modelo simple al cual aplicamos el método de descomposición de Blinder-Oaxaca, para luego estimar un modelo conjunto de educación, empleo, ocupación y salarios. Asimismo, compararemos los resultados obtenidos haciendo uso de medidas de habilidad y de las habilidades latentes estimadas como resultado del procedimiento econométrico propuesto.

Capítulo II. Revisión de literatura

Durante décadas los investigadores han centrado su atención en la relación entre el puntaje de pruebas de habilidad e indicadores del mercado laboral. Dichos estudios se concentran en el puntaje de pruebas que evalúan habilidades cognitivas (pruebas de razonamiento matemático, verbal, entre otros). Murnane *et al.* (1995: 251-266) evalúan el rol de las habilidades matemáticas de jóvenes próximos a graduarse de la secundaria en sus salarios a los 24 años y encuentran una relación positiva y creciente de dichas habilidades cognitivas en los salarios (especialmente cuando las habilidades fueron medidas próximas a la graduación). En un estudio más reciente, Cunha *et al.* (2006) proponen que la habilidad cognitiva afecta la probabilidad de alcanzar un mayor nivel educativo y el retorno económico del mismo.

Recientemente se ha incrementado el interés por analizar el rol de las habilidades no cognitivas y su impacto en variables asociadas al mercado laboral. Trabajos precursores como el de Bowles y Gintis (1976), Edwards (1976: 125-138) y Klein *et al.* (1991: 929-954) muestran que habilidades no cognitivas como la dependencia y persistencia son altamente valoradas por los empleadores. Estudios más recientes como el de Heckman *et al.* (2006) apoyan la existencia de dicha relación al encontrar una relación positiva entre el puntaje de pruebas que miden habilidades no cognitivas e indicadores asociados al mercado laboral.

En cuanto a la contribución de las diferencias en habilidades a las brechas en diversos indicadores del mercado laboral, gran parte de los estudios abordan el rol de las habilidades cognitivas en brechas salariales raciales o de género (Neal y Johnson, 1996; Ritter y Taylor, 2011). Pocos estudios abordan la contribución de habilidades no cognitivas a brechas salariales de género.

En este sentido, Fortin (2008) investiga el impacto de aspectos de la personalidad como autoestima, *locus* de control, la importancia del dinero/trabajo y la importancia de las personas/familia en la brecha salarial de género. Haciendo uso de dos encuestas longitudinales, el Estudio Longitudinal Nacional de 1972 (por sus siglas en inglés, NLS72) y el Estudio Educativo Longitudinal Nacional de 1988 (por sus siglas en inglés, NELS88), se encuentra cuáles son los factores no cognitivos que explican una parte pequeña pero importante -alrededor de dos puntos en el logaritmo del salario- de la brecha salarial de género entre trabajadores de 30 años. En particular, en cuanto a la importancia otorgada al "dinero/trabajo" y a las "personas/familia" (asociadas a autoestima y confianza) se encontró que mientras los hombres

suelen ser más ambiciosos y valorar más aquellos empleos que ofrecen salarios altos, las mujeres suelen optar por trabajos con un componente altruista.

Grove *et al.* (2011) exploran el poder explicativo de variables no cognitivas y variables de capital humano en la brecha salarial de profesionales con Maestrías en Administración de Negocios (por sus siglas en inglés, MBA). Haciendo uso de una encuesta longitudinal de individuos que se registraron para el Examen de Admisión para Estudios de Postgrado en Gerencia (por sus siglas en inglés, GMAT) entre 1990 y 1998, se encuentra que el 82% de la brecha salarial de género es explicada por habilidades no cognitivas y preferencias relacionadas con la familia, carrera y empleo. Al parecer, las mujeres en la muestra experimentan una penalidad en el salario por revelar preferencias hacia trabajos altruistas, los cuales -en promedio- otorgan menores salarios. Los hombres atribuyen una mayor importancia a la riqueza que las mujeres, pero ello no está asociado con la brecha salarial entre ambos. La brecha está explicada principalmente por la experiencia educativa y la tenencia de empleo por parte de los hombres.

Cobb-Clark y Tann (2011) evalúan si las habilidades no cognitivas de hombres y mujeres ejercen influencia sobre su ocupación y si, por tanto, contribuyen a la disparidad en salarios relativos. Utilizando información de empleados (no autoempleados) de 25 a 65 años contenida en la Encuesta de Hogares, Ingreso y Dinámica Laboral en Australia (HILDA) de 2001 y 2006 encuentran una relación entre el nivel de habilidades no cognitivas y la ocupación en la que dichos empleados laboran. No obstante, si bien la brecha salarial de género es de 0,143 (en términos del logaritmo del salario), el 96,6% es atribuible a diferencias en el salario de hombres y mujeres empleados en la misma ocupación. Por tanto, el componente más importante de la brecha salarial de género ocurre al interior de cada ocupación y permanece aún inexplicado.

Gran parte de los estudios en torno al rol de las pruebas que evalúan habilidades cognitivas y no cognitivas en indicadores de mercado laboral se han realizado en países desarrollados. A nuestro entender, muy pocos han abordado el tema para países en desarrollo y, en particular, para países de la región.

Bassi y Galiani (2009) utilizan una encuesta nacional con información sobre adultos de 25 a 30 años con el fin de explorar el rol que ejercen las pruebas que miden habilidades cognitivas y no cognitivas en el logaritmo de los salarios. Los autores encuentran coeficientes significativos para ambos tipos de habilidad así como que los mismos se reducen luego de controlar por el

nivel educativo de los individuos. Este resultado probablemente esté asociado al hecho que al utilizar medidas de habilidad (a diferencia de habilidades latentes) el efecto de los años de educación genera un sesgo en los coeficientes de dichas medidas en aquellas regresiones de salarios (por ejemplo) en las que no se controla por años de educación.

Díaz *et al.* (2012) estiman retornos a la educación, habilidades cognitivas y no cognitivas en el Perú utilizando la Encuesta de Habilidades y Mercado Laboral (ENHAB) que comprende una muestra de personas en edad de trabajar, por medio de una aproximación de variables instrumentales para abordar los problemas asociados a la endogeneidad de los años de educación presentes en la ecuación de salarios. Los autores hallaron que el mercado laboral peruano valora la educación y las habilidades cognitivas y no cognitivas. En particular, encuentran que un incremento de una desviación estándar en los años de educación se encuentra asociado con un incremento de 15% en los salarios, mientras que cambios similares en el nivel de habilidades cognitivas y no cognitivas se encuentran asociados con incrementos de 9% y 5% a 8% en los salarios, respectivamente.

Urzúa *et al.* (2009) profundizan en el análisis de discriminación de género en el mercado laboral para el caso de Chile, utilizando información sobre indicadores de mercado laboral, logro, desempeño educativo y otras variables relacionadas con el entorno familiar de cada individuo. Los autores trabajan a base de estudios previos que abordan la estimación de modelos de mercado laboral con múltiples fuentes de heterogeneidad no observada generada por habilidades cognitivas y no cognitivas. Sin embargo, por limitaciones en la disponibilidad de la información, solo consideran una fuente de heterogeneidad no observada, es decir, capturan el efecto de las habilidades cognitivas y no cognitivas en un solo componente (como una sola variable). Los resultados obtenidos sugieren la existencia de brechas de género en variables como experiencia, empleo, horas trabajadas y salarios por hora que no pueden ser explicadas por características observables o mecanismos de selección que generan endogeneidad. No obstante, hallan evidencia que las mujeres que adquieren bajos niveles educativos suelen ser víctimas de discriminación salarial. Hasta donde tenemos conocimiento, este estudio es el único que aborda el rol de habilidades cognitivas y no cognitivas (latentes) en la brecha salarial de género en un país en vías de desarrollo.

De acuerdo con muchos de los estudios presentados líneas arriba, un segundo tema de interés es el proceso de formación de las habilidades cognitivas y no cognitivas. Heckman *et al.* (2006) argumentan que las pruebas que miden este tipo de habilidades no reflejan el verdadero nivel de

habilidad (lo denominado "habilidad latente") ya que, al ser función del nivel educativo, se encuentran medidas con error. Por tanto, utilizar el puntaje de dichas pruebas en regresiones de salarios y años de educación es problemático. Al controlar por el nivel educativo, tanto las habilidades cognitivas como las no cognitivas pueden predecir los salarios. Sin embargo, el nivel educativo es una variable de elección y, por tanto, implica problemas de endogeneidad que deben ser considerados. Omitir el nivel educativo de una ecuación de salarios incrementa la correlación entre ambos tipos de habilidades y el salario percibido. Los efectos estimados comprenden tanto el efecto directo (sobre la productividad) como el indirecto (a través del nivel educativo) de las habilidades en los salarios. No obstante, existe una importante diferencia entre pruebas que miden las habilidades no cognitivas y cognitivas, como las de Coeficiente Intelectual (CI), y pruebas que miden logros de aprendizaje. Si bien el CI tiende a estar establecido alrededor de los ocho años de edad, los resultados de pruebas de logros de aprendizaje han evidenciado ser cambiantes y crecientes según el nivel educativo adquirido.

Hansen *et al.* (2004: 39-98) desarrollan dos métodos para estimar el efecto del nivel educativo en las pruebas de logros de aprendizaje que controlan por la endogeneidad asociada. Plantean que tanto el nivel educativo adquirido como los puntajes de pruebas de habilidades son generados por un componente no observado común: las habilidades latentes. De esta manera, encuentran que el efecto del nivel educativo en el puntaje de pruebas de habilidad es lineal a través de niveles educativos y mayor para bajos niveles de habilidad. Alcanzar un mayor nivel educativo incrementa el puntaje obtenido en la prueba Test de Calificación de las Fuerzas Armadas (por sus siglas en inglés, AFQT), considerada en el estudio en cuestión, entre dos y cuatro puntos porcentuales. Los autores contribuyen a la estimación del impacto del nivel educativo en las habilidades medidas en varios quintiles de la distribución de habilidades latentes. Asimismo, presentan evidencia que la medida de coeficiente intelectual (CI) utilizada por Herrnstein y Murray (1994) se encuentra fuertemente afectada por el nivel educativo de los individuos. Modelan los puntajes de pruebas de habilidades como función de las habilidades latentes (además de otros factores), y modelan el nivel educativo como función de las habilidades latentes (y otros determinantes). Consideran efectos de truncamiento (en pruebas sencillas, el puntaje máximo puede ser alcanzado por niños con distintos niveles de habilidad) y endogeneidad del nivel educativo (elecciones de fecha de entrada y salida del sistema educativo).

Helmers y Patnam (2011: 252-266) investigan los factores que determinan los puntajes en pruebas de habilidades cognitivas y no cognitivas de niños en Andhra Pradesh, India, utilizando

una base de datos que contiene información sobre dos cohortes de niños de hasta 12 años de edad. Aprovechando la disponibilidad de datos de panel estiman un modelo de Relaciones Estructurales Lineales (LISREL) que permite estimar habilidades latentes (cognitivas y no cognitivas) e inversión familiar, y permite relacionar estas variables a otras características observables asociadas al niño, sus padres y el hogar al que pertenece. Trabajan sobre la base de lo desarrollado por Cunha y Heckman (2007: 31-47) con el fin de examinar la dinámica que gobierna la formación de habilidades cognitivas y no cognitivas y la relación entre ambas a lo largo del tiempo. En particular, centran su atención en explorar los determinantes en el proceso de formación de ambos tipos de habilidad. Los autores encuentran evidencia a favor de la importancia de la inversión familiar (cuidado por parte de los padres durante el embarazo y los primeros meses del niño) y la salud de los niños durante el primer año de edad.

De esta manera, la principal contribución del presente estudio es analizar el rol de las habilidades latentes cognitivas y no cognitivas y del retorno de las mismas en términos de salarios en la brecha salarial de género en un país en vías de desarrollo. En este sentido, uno de nuestros principales objetivos es estimar las habilidades latentes que son no observables para el economista pero que son persistentes en el tiempo.

Capítulo III. Modelo

El modelo planteado está construido sobre la base de lo propuesto por Heckman *et al.* (2006), Cunha *et al.* (2010) y Cunha y Heckman (2007). El modelo plantea que existen dos factores subyacentes: la habilidad latente cognitiva y la no cognitiva. Así, controlando por factores observables, estos factores son capaces de explicar la dependencia entre decisiones y resultados. Ambos factores son conocidos por cada individuo pero no por el econométrista. Asimismo, el nivel de habilidades latentes se encuentra fijo para el momento en que el individuo toma decisiones en el mercado laboral.

La estrategia de identificación es similar a la planteada en Heckman *et al.* (2006). La habilidad cognitiva latente (f^C) afecta solo a la medida de habilidad cognitiva (C) y la habilidad no cognitiva latente (f^N) afecta solo a la medida de habilidad no cognitiva (N).

$$\begin{aligned}C &= \beta_C X_C + \alpha_C f^C + e_C \\N &= \beta_N X_N + \alpha_N f^N + e_N\end{aligned}$$

El modelo se sustenta en el hecho de que, condicional a la información contenida en el vector de variables X , la dependencia temporal entre las medidas de habilidad es atribuible a las habilidades latentes. Controlar por esta dependencia implica controlar por la endogeneidad del modelo. Heckman *et al.* (2006) estiman la distribución de habilidades latentes aprovechando la disponibilidad de dos o más medidas de habilidad. En contraste, la estrategia de especificación que utilizaremos aprovecha la disponibilidad de datos de panel para las medidas de habilidad disponibles, específicamente, el tener información sobre la misma medida de habilidad en dos momentos del tiempo. También asumimos un modelo lineal en los parámetros. No obstante, ello puede ser interpretado como una aproximación de un modelo de comportamiento más flexible tal como se sustenta en Heckman *et al.* (2006).

Al analizar el rol de las habilidades en la brecha salarial de género se debe tomar en cuenta que parte de lo que podría interpretarse como discriminación en realidad puede deberse a la dinámica comprendida en la elección de ocupaciones; hombres y mujeres podrían auto-seleccionarse hacia ocupaciones que premian distintos tipos de habilidad. Asimismo, si bien controlar por los años de educación de un individuo en una ecuación de salarios genera endogeneidad, se debe considerar el rol de las habilidades en la elección de un nivel educativo.

Por tanto, de manera similar a lo aplicado en Heckman *et al.* (2006), trabajaremos sobre la base de un modelo que considere las decisiones de educación, empleo y ocupación.

El modelo tiene cuatro etapas. El individuo elige un nivel educativo considerando el acervo de habilidades latentes; luego, dado el nivel educativo y sus habilidades latentes, el mismo elige participar o no del mercado laboral. Así, luego de tomar la decisión de participar, el individuo debe elegir el tipo de ocupación en el que desea trabajar, considerando su nivel de habilidades, y finalmente, dadas todas las elecciones tomadas hasta el momento y el acervo de habilidades, el individuo es asignado un determinado salario. A continuación se presenta el modelo de manera formal.

1. Modelo de educación

Cada individuo elige el nivel de educación que maximiza su beneficio tomando en cuenta el nivel de habilidades cognitivas y no cognitivas latentes que tiene. Así, considerando una especificación lineal en los parámetros, y sea I_S el beneficio neto asociado con haber logrado el nivel educativo S :

$$I_S = \beta_S X_S + \alpha_S^C f^C + \alpha_S^N f^N + e_S$$

donde S es el nivel educativo elegido por el individuo entre dos opciones: secundaria completa como máximo nivel educativo o al menos un año de educación superior; X_S es un vector de variables observables que afectan la decisión de educación; β_S es el vector de parámetros asociados; α_S^C y α_S^N son las ponderaciones de las habilidades cognitiva y no cognitiva latentes, respectivamente, y e_S representa un error idiosincrático que se asume como independiente de f^N , f^C y X_S . Los términos de error correspondientes a ambos niveles educativos son independientes.

2. Modelo de empleo

Luego de elegir el nivel educativo y considerando el resultado de dicha elección, el individuo debe tomar la decisión acerca de participar o no en el mercado laboral. Así, el individuo decidirá participar si el beneficio neto asociado es mayor que si no participara, considerando también su nivel de habilidades latentes. Sea I_E el beneficio asociado a trabajar y asumiendo una especificación lineal en los parámetros:

$$I_E = \beta_E X_E + \alpha_E^C f^C + \alpha_E^N f^N + e_E$$

donde E es la decisión tomada por el individuo: trabajar o no trabajar; X_E es un vector de variables observables que afectan la decisión de empleo; β_E es el vector de parámetros asociados, α_E^N y α_E^C son las ponderaciones de las habilidades cognitiva y no cognitiva latentes, respectivamente, y e_E representa un error idiosincrático que se asume como independiente de f^N , f^C y X_E . Los términos de error correspondientes a ambas opciones laborales (empleo o desempleo) son independientes.

3. Modelo de ocupación

Una vez tomada la decisión de participar del mercado laboral, el individuo debe elegir en qué ocupación desempeñarse. Así, considerando las decisiones tomadas hasta el momento y su nivel de habilidades latentes, decidirá entre una de dos opciones de ocupación: una ocupación que requiere de capacitación especializada (denominada "white collar") u otra que no lo requiera y esté más orientada al trabajo manual (denominada "blue collar"). Así, sea I_0 la utilidad asociada a elegir una ocupación determinada, el modelo lineal que representa a dicha decisión es:

$$I_0 = \beta_0 X_0 + \alpha_0^C f^C + \alpha_0^N f^N + e_0$$

donde 0 representa a la ocupación elegida por el individuo; X_0 es un vector de variables observables que afectan la decisión de ocupación; β_0 es el vector de parámetros asociados; α_0^C y α_0^N son las ponderaciones de las habilidades cognitiva y no cognitiva latentes, respectivamente, y e_0 representa un error idiosincrático que se asume como independiente de f^N , f^C y X_0 . Los términos de error correspondientes a ambas opciones ocupacionales son independientes.

4. Modelo de salarios

Finalmente, el salario que recibe el individuo es producto de las decisiones previas de educación, empleo y ocupación, así como del nivel de habilidades cognitivas y no cognitivas latentes del individuo hasta ese momento. Por tanto, la ecuación de salarios sería:

$$Y = \beta_Y X_Y + \alpha_Y^C f^C + \alpha_Y^N f^N + e_Y$$

donde X_Y es un vector de características observables, β_Y es el vector de retornos asociados, α_Y^C y α_Y^N son las ponderaciones de las habilidades cognitivas y no cognitivas latentes, respectivamente, y e_Y representa un error idiosincrático que se asume como independiente de f^N , f^C y X_Y .

Capítulo IV. Fuentes de información y muestra

Con el objetivo de estimar el efecto de las habilidades en la brecha salarial de género es necesario contar con una base de datos que contenga información sobre ambas variables para un mismo individuo. Así, no es suficiente contar con información sobre medidas de habilidad, sino que nuestra estrategia de identificación requiere información sobre medidas de habilidad para distintos periodos de tiempo. Debido a la ausencia de una fuente de información que cumpla con estas características, haremos uso de dos bases de datos para llevar a cabo nuestro análisis.

La primera base de datos corresponde a la recogida en el “Estudio Niños del Milenio”¹ (ENM) para Perú (Boyden 2011a, Boyden 2011b, Huttly, 2011) Esta comprende información longitudinal para dos cohortes de niños (cohorte joven y cohorte adulta) para cada uno de los cuatro países: Etiopía, India (Andhra Pradesh y Telengana), Perú y Vietnam. En Perú la muestra representa al 95% de los niños peruanos (excluye al 5% que pertenece a familias con mayores ingresos). Los niños y sus cuidadores principales fueron encuestados tres veces, en los años 2002, 2006 y 2009. La encuesta recoge información sobre aspectos relacionados con el desarrollo de los niños, medidas de habilidad cognitiva y aspectos de la personalidad (actitudes y aspiraciones), medidas antropométricas y un conjunto de otras características a nivel individual y de hogares. En particular, incluye características como el nivel socio-económico de las familias (percibido por el niño), índices de riqueza, gasto en consumo y medidas de las habilidades de los cuidadores principales, entre otros.

Con el fin de analizar la distribución de habilidades entre niños peruanos, centramos nuestra atención en los 700 niños de la cohorte adulta, los mismos que tenían ocho años en la primera ronda del estudio (año 2002). Asimismo, mantenemos la submuestra de niños con información disponible para los ítems relacionados a medidas de habilidad (cognitiva y no cognitiva) e información individual relevante para las rondas 2 y 3, cuando los niños de la cohorte adulta tenían 12 y 15 años de edad, respectivamente. Finalmente, con fines de consistencia, trabajamos solamente con los niños que viven en zonas urbanas. La muestra final consta de 349 niños.

¹ Los datos usados en esta publicación provienen del estudio Niños del Milenio, conocido internacionalmente como Young Lives, una investigación longitudinal de 15 años que analiza la naturaleza cambiante de la pobreza infantil en Etiopía, India (estado de Andhra Pradesh y Telengana), Perú y Vietnam (www.ninosdelmilenio.org/ www.younglives.org.uk). Niños del Milenio es cofinanciado por UK AID del Departamento de Desarrollo Internacional (DFID, por sus siglas en inglés) y por el Ministerio de Relaciones Exteriores de los Países Bajos (2010-2014). Las opiniones aquí expresadas son de los autores y no necesariamente compartidas por el estudio Niños del Milenio / Young Lives, la Universidad de Oxford, DFID ni otros donantes.

La muestra con la que trabajaremos en adelante se encuentra distribuida entre niños y niñas (165 y 184, respectivamente), tienen una edad promedio de 149 meses y un promedio de seis años de educación durante la ronda 2. La tabla 1 presenta estadísticos descriptivos de las principales variables de interés para ambas rondas así como información sobre la lengua materna y el nivel educativo de los padres en la ronda 1 (lo que denominaremos "características permanentes").

Tabla 1. Estadísticas descriptivas: Estudio Niños del Milenio

	Ronda 2				Ronda 3			
	Muestra completa				Muestra completa			
	Media	DE	Mujeres	Hombres	Media	DE	Mujeres	Hombres
Habilidad cognitiva (puntaje TVIP)	76,92	13,84	75,964	77,777	101,083	14,6	99,788	102,245
Índice de autoeficacia	0,101	0,939	0,216	-0,003**	0,101	0,969	0,295	-0,073**
Índice de autoestima	0,139	0,897	0,088	0,185	0,147	0,94	0,208	0,093
Medida de autoeficacia del cuidador principal (estandarizado)	0,144	0,997	-0,012	0,284***	0,145	0,925	0,136	0,154
Medida de autoestima del cuidador principal (estandarizado)	0,046	1,003	0,058	0,036	0,044	1,028	0,13	-1,033
Talla para la edad (estandarizado)	-1,302	1,049	-1,337	-1,271	-1,306	0,863	-1,502	-
Índice de masa corporal	0,34	0,966	0,297	0,378	0,272	0,972	0,432	0,128***
Edad en meses	148,867	5,416	148,488	149,206	179,117	4,505	178,783	179,417
Años de educación	6,143	0,895	6,17	6,12	9,006	1,101	9,067	8,951
Ausencia a la escuela por enfermedad (>1 semana)	0,054	0,227	0,055	0,054	0,069	0,235	0,067	0,071
Índice de riqueza	0,602	0,19	0,596	0,608	0,661	0,15	0,665	0,658
log consumo del hogar per cápita	5,191	0,642	5,179	5,201	5,333	0,635	5,338	5,328
Lengua materna (español)	0,788	0,409	0,794	0,783				
Nivel educativo del padre	10,481	3,092	10,176	10,755				
Nivel educativo de la madre	9,602	3,499	9,648	9,56				
N			349	165 184			349	165 184

Notas: * indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** indica un nivel de significancia del 1% de la prueba de diferencia de medias entre hombres y mujeres.

Fuente: Boyden, 2011a, Boyden, 2011b.

Elaboración: Propia, 2013.

Cabe resaltar que en ambas rondas si bien los mujeres tienen un puntaje menor al promedio en medidas de habilidad, tal es el caso para los hombres en el caso del índice de auto eficacia (los resultados son mixtos para el caso de autoestima). No parece haber diferencias importantes entre géneros en cuanto a características del hogar y la familia. Las medidas de habilidad no cognitivas de los cuidadores principales difieren entre niños y niñas; los primeros evidencian altos niveles de auto eficacia durante la ronda 2 y bajos niveles de autoestima durante la ronda 3. Asimismo, diferencias importantes surgen entre rondas. Cabe resaltar que las medidas utilizadas para representar habilidades no cognitivas fueron construidas sobre la base del grado de acuerdo o desacuerdo de los niños respecto a enunciados relacionados con aspectos de la personalidad como autoestima y auto eficacia.

La segunda base de datos corresponde a una novedosa encuesta de hogares, recogida por el Banco Mundial, que contiene información de salarios y características individuales, así como medidas de habilidad cognitiva y no cognitiva para una muestra de individuos en edad de trabajar (14-50 años). La Encuesta Nacional de Habilidades y Mercado Laboral (ENHAB) (Banco Mundial 2010) es una encuesta representativa a nivel del Perú urbano. Dicha base de datos contiene información acerca de las condiciones de vida, información demográfica, logro académico, empleo/salarios e información nueva sobre medidas de habilidad cognitiva y no cognitiva, trayectoria educativa, participación temprana en el mercado laboral y características familiares. Las habilidades cognitivas fueron medidas por medio de pruebas cognitivas que evalúan capacidades numéricas y de resolución de problemas, memoria de trabajo, fluidez verbal y de lenguaje receptivo. Las habilidades no cognitivas fueron medidas de acuerdo a escalas GRIT (Duckworth *et al.*, 2007: 1087-1101) y los cinco grandes factores de la personalidad (Goldberg, 1990:1216-1229). Para el análisis que realizaremos, concentramos nuestra atención en siete medidas, los valores estandarizados de cada uno de los cinco grandes factores de la personalidad (estabilidad emocional, extroversión, amabilidad, cooperación, actitud concienzuda –fuerte- y apertura), una medida compuesta de los elementos de GRIT, y una medida compuesta de habilidades cognitivas. Algunas de las características individuales comprendidas son el pasado educativo del individuo, características familiares y estatus socio-económico (educación y ocupación de los padres, tamaño de la familia, información acerca del acceso y características de la escuela a la que asistieron los padres, entre otros).

Con el fin de aprovechar al máximo la información contenida en la base de datos de la ENHAB (Banco Mundial 2010), a lo largo de nuestro análisis trabajaremos con tres submuestras similares: i) muestra de individuos con información disponible sobre medidas de habilidad (puntaje de pruebas), N = 2421; ii) muestra de individuos que reportan información sobre salarios, N = 4063; iii) muestra de individuos con información disponible acerca de las principales características individuales, N = 7499. En términos generales, los individuos en la base de datos se encuentran distribuidos de manera similar entre hombres y mujeres, tienen, en promedio, 33 años de edad, ganan un salario de S/. 1.000 constantes del 2010 (aproximadamente, US\$ 350), trabajan 51 horas a la semana y han concluido estudios de secundaria. La tabla 2 ilustra algunos estadísticos descriptivos para las tres submuestras consideradas, así como la diferencia entre hombres y mujeres en cada una. Entre ellos, se encuentra el hecho que el hombre representativo de la muestra tiene salarios mayores (mensuales y por hora), trabaja más horas y tiene niveles más altos de habilidades cognitivas que la mujer representativa de la muestra. Sin embargo, los resultados son mixtos para el caso

de las habilidades no cognitivas. Mientras las mujeres muestran actitudes más consistentes, amables, cooperativas y concienzudas, los hombres muestran actitudes más persistentes, extrovertidas, emocionalmente estables y abiertas.

Tabla 2. Estadísticas descriptivas: ENHAB

	N	Media	DS	Mujeres	Hombres
Cognitive ability (PPV raw score)	2421	42,001	14,932	40,103	45,039***
Consistencia/interest	2421	-0,004	0,998	0,036	-0,069***
Persistencia/effort	2421	0,001	1,001	-0,046	0,076***
Grift (estandarizado)	2421	-0,002	0,997	-0,009	0,008
Extroversión	2421	0,002	1	-0,047	0,082***
Amabilidad	2421	0,002	1,006	0,039	-0,058***
Cooperación	2421	0,006	0,992	0,051	-0,065***
Actitud concienzuda (fuerte)	2420	0,000	1,001	0,056	-0,092***
Estabilidad emocional	2415	0,013	0,997	-0,072	0,149***
Apertura	2415	-0,004	0,999	-0,059	0,085***
Log salario por hora	4063	1,316	0,815	1,166	1,408***
Salario mensual	4063	972,155	1.284,643	738,7	1.116,299***
Salario por hora	4063	5,264	6,55	4,638	5,651***
Horas semanales de trabajo	4063	51,127	18,482	46,339	54,084***
Experiencia	4063	25,43	13,582	25,314	25,502
Edad	7499	33,514	15,282	33,305	33,736
Años de educación	7457	10,701	3,373	10,524	10,890***
Lengua materna (español)	7499	1,008	0,142	1,007	1,009
Nivel educativo del padre	7499	4,955	2,373	4,947	4,963
Nivel educativo de la madre	7499	4,284	2,333	4,253	4,324

Notas: * indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** indica un nivel de significancia del 1% de la prueba de diferencia de medias entre hombres y mujeres.

Fuente: Banco Mundial, 2010.

Elaboración: Propia, 2013.

Capítulo V. Implementación econométrica

El principal objetivo de este documento es identificar la contribución de las habilidades a la brecha salarial de género. Por tanto, a fin de identificar el rol de las habilidades en los salarios, la ecuación que se desearía estimar es:

$$\ln W_i = \alpha + \gamma S_i + \beta_A A_i + \mu_{it}$$

$$\ln W_i = \alpha + \gamma S_i + \beta_A A_i + \mu_{it}$$

donde $\ln W_i$ es el logaritmo de los salarios; S_i representa los años de educación y A_i la habilidad innata del individuo. Sin embargo, la poca disponibilidad de datos acerca de las habilidades innatas lleva a omitir A_i , generando endogeneidad en la ecuación de salarios. Si se asume que las habilidades generan incrementos en el salario y que las habilidades están positivamente relacionadas con los años de educación, γ estaría sobreestimando el verdadero valor del parámetro. La literatura empírica ha abordado este tema mediante la inclusión de medidas de habilidad (puntaje de pruebas que miden habilidades) y, por tanto, estimando la siguiente ecuación:

$$\ln W_i = \alpha + \gamma S_i + \beta_T T_i + v_{it}$$

donde T_i representa el puntaje de pruebas estandarizadas que miden habilidades cognitivas y/o no cognitivas. Esta especificación es problemática ya que es probable que los puntajes de pruebas de habilidades no solo estén determinados por las habilidades innatas (cuyo efecto desearíamos aislar) sino también por el nivel educativo del individuo, lo cual genera un problema puesto que ambos factores estarían asociados con incrementos en los salarios. Por tanto, β_T estaría capturando parcialmente el efecto indirecto de los años de educación en los salarios, a través de su efecto en las medidas de habilidad, e impidiendo obtener el verdadero efecto de los años de educación en los salarios.

En este sentido, proponemos aplicar un procedimiento econométrico que permita estimar esta habilidad (latente) A_i e identificar cada uno de los efectos deseados en los salarios. Para ello, explotaremos la información longitudinal de medidas de habilidad contenida en el ENM para estimar proxies de las habilidades cognitivas y no cognitivas latentes. Luego, proponemos un método de imputación para predecir los efectos fijos estimados con la muestra del ENM en la muestra de la ENHAB. Al complementar estas habilidades latentes predichas con la

información de salarios recogida para la ENHAB, será posible analizar el rol de las habilidades cognitivas y no cognitivas latentes en los salarios. Específicamente, la implementación econométrica que aplicaremos se divide básicamente en dos etapas: i) la estimación de las habilidades latentes para la muestra del ENM, y ii) la estimación del efecto de las habilidades latentes (predichas) en los salarios para la muestra ENHAB.

1. Estimación de habilidades latentes

En la primera etapa trabajaremos con la base de datos del ENM. En primer lugar, hacemos uso de la variación (de la ronda 2 a la ronda 3) en las medidas de habilidad y los años de educación entre los niños en la muestra del ENM para recuperar los efectos fijos (no observables). En particular, intentaremos explicar la variación de las tres medidas de habilidad, dos habilidades no cognitivas (autoestima y autoeficacia) y una habilidad cognitiva (Test de Vocabulario e Imágenes Peabody, TVIP). El procedimiento de identificación requiere que se controle por características que puedan haber variado entre los 12 y 15 años y puedan explicar gran parte de la variación en las medidas de habilidad de los niños en ese mismo período. De esta manera, será posible explicar los cambios en las medidas de habilidad y aislar los efectos fijos no observables.

$$\Delta MA_{it} = \gamma_X \Delta X_{it} + \Delta \mu_{it}$$

Luego, estimamos la correlación entre las características que se mantienen constantes a lo largo de la vida del niño y los efectos fijos previamente hallados. Para ello, capturamos el efecto fijo o componente no observable de cada tipo de habilidad haciendo uso de los estimados obtenidos en la primera etapa para predecirlos, del valor promedio de los regresores en las rondas 2 y 3, y, luego, el valor predicho de la medida de habilidad respecto al valor observado correspondiente.

$$\widehat{MA}_{it} = \hat{\gamma}_0 + \hat{\gamma}_X X_{it}$$

$$\hat{LA}_i = \frac{1}{2} [(MA_{i1} - \widehat{MA}_{i1}) + (MA_{i2} - \widehat{MA}_{i2})]$$

$$\widehat{LA}_i = \gamma_0^{LA} + \gamma_1^{LA} Z_i + \mu_i^{LA}$$

Los valores hallados representan proxies de las habilidades latentes y, al ser efectos fijos, pueden ser modelados como función de variables que podrían determinar las habilidades latentes, permanecen constantes a lo largo de la vida del niño y sean comunes a ambas bases de datos. De esta manera, podremos predecir el valor de estas habilidades latentes en la muestra ENHAB. Algunos controles que podrían ser adecuados para modelar dichos efectos fijos son el género del niño y el nivel educativo de los padres.

La segunda etapa propone partir de estos efectos estimados y aprovechar la información sobre las “características permanentes” para la muestra ENHAB con el fin de obtener valores predichos de las habilidades latentes para dicha muestra. Así, el valor predicho de las habilidades latentes para la muestra ENHAB ($\widehat{L\bar{A}}_i$) es el resultado de las “características permanentes” observadas (Z_i) y los efectos previamente estimados ($\widehat{\gamma}_1^{LA}$).

$$\widehat{L\bar{A}}_i = \widehat{\gamma}_0^{LA} + \widehat{\gamma}_1^{LA} Z_i$$

Un supuesto importante en este procedimiento de “emparejamiento” es que las muestras del ENM y de la ENHAB son representativas a nivel nacional (es decir, evidencian características similares)². Así, con los valores predichos, es posible estimar una ecuación de salarios y, más aún, analizar las fuerzas detrás de la existencia de una brecha salarial de género. Para ello, modelamos los salarios como función de las habilidades cognitivas y no cognitivas latentes.

Con la muestra de ENHAB que contiene información de salarios, medidas de habilidad y habilidades latentes aplicaremos dos aproximaciones para analizar el rol de las habilidades en los salarios. La primera, consiste en aplicar el método de descomposición de Blinder-Oaxaca, el cual permite evaluar el rol de brechas en regresores sobre brechas en una variable de resultado. La segunda, en desarrollar el modelo completo de decisión del individuo incorporando interacciones entre las habilidades y una variable dicotómica de género para analizar el rol de brechas de habilidad en brechas de salario. Así, será posible enriquecer el análisis efectuando una comparación entre los resultados obtenidos mediante ambas aproximaciones y usando tanto las medidas de habilidad, como las habilidades latentes, para evaluar el rol de las habilidades en la brecha salarial de género.

² Debemos tomar en cuenta que el ENM no contiene información de aquellos niños que pertenecen a los hogares que perciben los ingresos más altos.

2. Brecha salarial de género y el método de descomposición de Blinder-Oaxaca

El método de descomposición de Blinder-Oaxaca es una metodología que permite descomponer diferencias en el salario promedio de dos grupos, en este caso, hombres y mujeres. Se parte de un modelo lineal que distingue entre la contribución de características observables y no observables:

$$Y_g = X\beta_g + \eta_g \quad \text{para } g = \text{hombre, mujer}$$

Luego, sea d una variable dicotómica que indica pertenencia a un grupo determinado, y^d el resultado de interés para un miembro del grupo d , X^d el vector de características observables (incluyendo una constante), $\hat{\beta}^d$ el vector columna de coeficientes de una regresión lineal de y^d sobre X^d , y las barras representan promedios. Se pueden re-expresar las diferencias salariales como diferencias en características observables o diferencias en los coeficientes de dichas características:

$$(\bar{Y}^1 - \bar{Y}^0) = (\bar{X}^1 - \bar{X}^0)\hat{\beta}^1 + \bar{X}^0(\hat{\beta}^1 - \hat{\beta}^0)$$

donde el primer y segundo términos del lado derecho de la ecuación representan los componentes explicados y no explicados de la brecha en el resultado promedio, respectivamente. A esto se le denomina “descomposición doble” ya que descompone brechas entre diferencias en nivel o diferencias en retornos. Una extensión de esta metodología es la denominada “descomposición triple” la cual incluye un tercer componente que interactúa (considera de manera simultánea) diferencias en niveles y retornos de los regresores considerados en el cual el último término del lado derecho de la ecuación representa el término de interacción:

$$\bar{Y}^1 - \bar{Y}^0 = (\bar{X}^1 - \bar{X}^0)\hat{\beta}^1 + \bar{X}^0(\hat{\beta}^1 - \hat{\beta}^0) + (\bar{X}^1 - \bar{X}^0)(\hat{\beta}^1 - \hat{\beta}^0)$$

Capítulo VI. Resultados

En esta sección comparamos los resultados de la estimación del efecto de las habilidades cognitivas y no cognitivas sobre los salarios obtenidos mediante el uso de puntajes de pruebas que miden dichas habilidades y aquellos obtenidos luego de considerar las habilidades latentes estimadas. En cada caso, empezamos presentando los resultados obtenidos tras estimar una ecuación de Mincer del logaritmo de los salarios controlando por el nivel educativo y las habilidades. Posteriormente, aplicamos el método de descomposición de Blinder-Oaxaca con el fin de estimar el impacto de las habilidades en la brecha salarial de género. Finalmente, en el afán de separar el efecto de las habilidades en la brecha salarial de género a través de cada una de las decisiones del individuo previas a la determinación del salario, estimamos un modelo conjunto de educación, empleo, elección de ocupación y salarios. Para ello, nos valemos del procedimiento explicado previamente para obtener estimados de las habilidades latentes cognitivas y no cognitivas, y presentamos los resultados correspondientes a cada una de las cuatro etapas que lo conforman.

1. Salarios y medidas de habilidad

A propósito de la discusión previa acerca de los problemas asociados a la estimación del efecto de las habilidades en los salarios, la tabla 3 muestra los resultados de una ecuación de Mincer bajo el supuesto (ingenuo) de que no existe correlación entre las medidas de habilidad y los años de educación. La columna 1 evidencia que luego de controlar por experiencia laboral, lugar de residencia y lengua materna, un año de educación adicional lleva a un incremento de 10,9% en los salarios. La columna 2 incorpora controles como el nivel educativo de los padres ya que ello podría explicar parte de la correlación entre salarios y educación. Con esto, el coeficiente estimado cae de 10,9% a 9,00%. El incorporar medidas de habilidad cognitiva y no cognitiva pone en evidencia que las medidas de habilidad cognitiva y estabilidad emocional llevan a salarios más altos, mientras que la amabilidad y la consistencia en términos de esfuerzo parecen tener el efecto contrario. Asimismo, incluir dichas medidas de habilidad genera una reducción en el retorno a cada año de educación, lo cual sugiere que el coeficiente de la columna 2 estaría sobre-estimando el efecto de la educación en los salarios.

Tabla 3. Ecuación de Mincer con medidas de habilidad

	[1]	[2]	[3]
Años de educación	0,1091***	0,0934***	0,0804***
Experiencia	0,0255	0,0318*	0,0329*
Experiencia ²	-0,0003	-0,0004	-0,0004
Reside en Lima	0,0397	0,0297	0,0091
Lengua materna (español)	0,2675	0,4296	0,5023**
Nivel educativo del padre		0,0148	0,0097
Nivel educativo de la madre		0,0386**	0,0366**
Goldberg, extroversión			0,0257
Goldberg, amabilidad			-0,0371
Goldberg, cooperación			-0,0616*
Goldberg, actitud concienzuda (fuerte)			-0,0005
Goldberg, estabilidad emocional			0,0810***
Goldberg, apertura			-0,0069
Grit 2, consistencia de intereses			-0,0453*
Grit 2, persistencia del esfuerzo			0,0031
Pruebas de habilidad cognitiva			0,0751***
Constante	-0,6028**	-0,9125**	-0,8131***
Número de observaciones	1.079	1.079	1.073
R-cuadrado	0,157	0,17	0,187

Nota: Errores estándar entre paréntesis se encuentran agrupados a nivel regional.* indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** un indica un nivel de significancia del 1%. La muestra corresponde a personas empleadas en el momento de la encuesta y con información disponible acerca de pruebas de habilidades y controles individuales.

Fuente: Banco Mundial, 2010.

Elaboración: Propia, 2013.

1.1 Método de descomposición de Blinder-Oaxaca

Con el fin de estimar el efecto de las medidas de habilidad, optamos por aplicar el método de descomposición de Blinder-Oaxaca. En la muestra utilizada, el logaritmo del salario promedio es 1,417 para hombres y 1,141 para mujeres, llevando a una brecha salarial estadísticamente significativa de 0,276. Sin embargo, nuestra atención se centra en la importancia relativa de los elementos que explican dicha brecha. Haciendo uso de los métodos de descomposición doble y triple (descritos previamente), encontramos que los datos favorecen el hecho de que la brecha salarial es atribuible a diferencias entre grupos de los retornos y niveles de las habilidades. No obstante, si bien habría un incremento significativo en el salario por hora de las mujeres si las mismas compartieran las características (valor promedio de los regresores) de los hombres, más del 80% de la brecha salarial de género se reduciría si las mujeres compartieran el valor de los coeficientes de los hombres, dadas sus propias características.

La brecha salarial de género, así como el efecto de la dotación y las diferencias en los coeficientes, son significativos aún luego de controlar por características individuales comunes

en este tipo de análisis. Las primeras dos columnas de la tabla 4 ilustran los resultados obtenidos mediante la descomposición doble y el uso de una especificación sencilla, y luego de agregar diversos controles, respectivamente. Las columnas 3 y 4 presentan los resultados obtenidos mediante la descomposición triple.

Un segundo ámbito de análisis corresponde a la contribución individual de los predictores al componente explicado de la brecha salarial de género. A partir de este momento, trabajaremos con la medida compuesta de GRIT como la prueba representativa que mide habilidades no cognitivas³. La tabla 4 presenta la proporción de la brecha salarial atribuidos a diferencias en medidas de habilidad cognitiva y no cognitiva, en términos de regresores (niveles) y sus respectivos coeficientes (retornos). Las cuatro aproximaciones presentadas (columnas 1 a 4) evidencian que el la diferencia en los niveles se debe, básicamente, a niveles heterogéneos de habilidades cognitivas (medidas) entre hombres y mujeres. En el caso de las habilidades no cognitivas, no se hallaron efectos significativos. Las brechas en términos de coeficientes (retornos) están relacionadas con los correspondientes a los controles individuales y heterogeneidad no observada entre géneros.

La descomposición de Blinder-Oaxaca puede ser interpretada como un modelo simple del ciclo de vida desde una perspectiva que considera factores latentes. Las habilidades cognitivas y no cognitivas determinan el nivel educativo. No obstante, las habilidades latentes no son observables, lo que se observa son medidas de estas habilidades. Por tanto, los retornos estimados mediante dicho método de descomposición son función de los retornos a la educación y los parámetros que gobiernan la elección del nivel educativo alcanzado por el individuo. Asimismo, es plausible suponer que los parámetros que influyen en la decisión de ocupación también estén afectando el retorno promedio (general) estimado.

³ Se optó por trabajar con GRIT puesto que la literatura resalta la relevancia del rol de las habilidades no cognitivas en el salario. Asimismo, Díaz *et al.* (2012), autores que utilizaron también la ENHAB, hallan evidencia del rol fundamental que juega en ecuaciones de salario. No obstante, todas las estimaciones llevadas a cabo a lo largo del estudio se han realizado, también, con las demás medidas de habilidad y se han obtenido resultados similares.

Tabla 4. Descomposición de Blinder-Oaxaca con medidas de habilidad

	[1]	[2]	[3]	[4]
Brecha Salarial	0,276*** (0,065)	0,276*** (0,051)	0,276*** (0,065)	0,276*** (0,067)
Niveles	0,059*** (0,019)	0,045*** (0,019)	0,067*** (0,024)	0,058** (0,028)
Retornos	0,217*** (0,063)	0,231*** (0,048)	0,225 (0,067)	0,242*** (0,062)
Niveles				
Cognitiva	0,059*** (0,019)	0,058*** (0,015)	0,067*** (0,023)	0,067*** (0,023)
No Cognitiva	0,000 (0,002)	0,000 (0,001)	0,000 (0,002)	0,000 (0,001)
Retornos				
Cognitiva	-0,171 (0,167)	-0,192 (0,130)	-0,163 (0,160)	-0,184 (0,144)
Non cognitiva	0,007 (0,143)	-0,017 (0,165)	0,007 (0,144)	-0,017 (0,154)
Número observaciones	1.081	1.081	1.081	1.081
Controles	No	Sí	No	Sí
Interacciones	No	No	Sí	Sí

Nota: Errores estándar entre paréntesis se encuentran agrupados a nivel regional. * indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** un indica un nivel de significancia del 1%. La muestra corresponde a personas empleadas en el momento de la encuesta y con información disponible acerca de pruebas de habilidades y controles individuales. Los controles considerando son edad, edad al cuadrado y si el individuo reside en Lima.

Fuente: Banco Mundial, 2010.

Elaboración: Propia, 2013.

1.2 Estimación conjunta: educación, empleo, ocupación y salarios

Los resultados obtenidos previamente brindan información acerca del rol de las medidas de habilidad cognitiva y no cognitiva sobre los salarios, pero no consideran las decisiones intermedias tomadas por los individuos antes de recibir un nivel de salario determinado. Con el fin de separar los efectos de las medidas de habilidad sobre cada una de estas decisiones, procedemos a estimar un modelo conjunto que considera elecciones de educación, empleo y ocupación. El mismo es consistente con la serie de decisiones que debe tomar un individuo a lo largo de su vida profesional: consciente de su nivel de habilidades, un individuo elige cuál es el máximo nivel educativo que alcanza, luego, dado el nivel educativo que posee y sus habilidades latentes, decide participar o no del mercado laboral. Una vez que decide ser empleado, debe elegir su ocupación y, finalmente, recibir un salario acorde a sus decisiones previas y su nivel de habilidades. Para llevar a cabo un análisis de este tipo, consideramos dos niveles educativos

(secundaria completa es el punto de corte), la condición de ser empleado o no serlo, y dos clases de ocupación (aquellas que requieren capacitación especializada y aquellas que no).

Los resultados de la estimación por máxima verosimilitud del modelo conjunto se presentan en la tabla 5. El procedimiento de estimación requiere maximizar la verosimilitud conjunta de adquirir un determinado nivel educativo, estar empleado, haber elegido una ocupación que requiera capacitación especializada y ganar un cierto nivel de salarios. Por tanto, la contribución individual a la verosimilitud sería:

$$l_i = \underbrace{L_{si}(\theta_S|LA_i)}_{\text{Educación}} \overbrace{L_{hi}(\theta_h|LA_i, s_i)}^{\text{Empleo}} \underbrace{L_{oi}(\theta_o|LA_i, s_i, h_i = 1)}_{\text{Ocupación}} \overbrace{L_{wi}(\theta_w|LA_i, s_i, h_i = 1, o_i)}^{\text{Salarios}}$$

Cada columna de la tabla 5 corresponde a cada una de las decisiones involucradas en el modelo. Los resultados indican que mientras las medidas de habilidad cognitiva parecen importar más en la determinación de los años de educación y la elección de ocupación, las medidas de habilidad no cognitiva son más importantes para la determinación de los salarios y la empleabilidad. En términos de diferencias de género, los hombres parecen tener un mayor retorno que las mujeres en general, pero en particular para el caso de habilidades no cognitivas en términos de empleabilidad y mayores salarios. Las mujeres obtienen mayores retornos a las habilidades cognitivas solo para la elección del nivel educativo. No obstante, debemos considerar que los impactos estimados consideran medidas de habilidad (no habilidades latentes) lo cual podría estar capturando el efecto de otros factores correlacionados con las variables dependientes y las medidas de habilidad.

Tabla 5. Estimación conjunta con medidas de habilidad

Modelos	Educación	Empleo	Ocupación	Salarios por hora
Cognitiva (mujeres)	0,046***	-0,004*	0,043***	-0,03
	-0,003	-0,002	-0,004	-0,018
Interacción c/cognitiva	-0,006***	0,004	-0,007	0,040*
	-0,003	-0,003	-0,005	-0,022
No cognitiva (mujeres)	0,190***	0,122**	0,129	0,996**
	-0,067	-0,057	-0,109	-0,467
Interacción c/no cognitiva	0,134	0,195**	0,105	1,335**
	-0,093	-0,076	-0,134	-0,605
Número de observaciones	2.421			

Nota: * indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** indica un nivel de significancia del 1% de la prueba de diferencia de medias entre hombres y mujeres.

Fuente: Banco Mundial, 2010.

Elaboración: Propia, 2013.

2. Salarios y habilidades latentes

Con el objetivo de estimar el efecto neto de las habilidades sobre la brecha salarial de género, debemos considerar las habilidades latentes. A continuación presentamos los resultados del procedimiento propuesto para estimar dichas habilidades latentes (cognitivas y no cognitivas) y, luego, reestimamos los modelos introducidos previamente, considerando las habilidades latentes estimadas.

2.1 Aproximando habilidades latentes

La tabla 6 presenta los resultados de la primera etapa. Las columnas 1 a 3 corresponden a las estimaciones de la habilidad cognitiva (puntaje de la prueba TVIP), autoestima y autoeficacia, respectivamente. Cada regresión controla por la medida de habilidad correspondiente del cuidador principal del niño⁴, la medida estandarizada de la talla para la edad del niño, el índice de masa corporal del niño, su edad, la percepción del estatus socio-económico del hogar, un índice de riqueza del hogar, una variable que indica si el niño se ausentó a la escuela por problemas de salud durante más de una semana (variación en la educación) y el logaritmo del consumo real per cápita. Los errores estándar se encuentran agrupados por comunidad.

Cambios en la medida de autoestima del cuidador principal, la talla para la edad, el índice de masa corporal, haber faltado a la escuela por problemas de salud y el estatus socio-económico percibido evidenciaron una relación estadísticamente significativa con la medida de autoestima

⁴ La base de datos no contenía medida de habilidad cognitiva para el cuidador principal de los niños.

del niño, siendo el último factor el más importante. Los resultados son similares para el caso de la medida de auto eficacia, pero cambios en el estatus socio-económico percibido no son estadísticamente significativos. En cuanto a la medida de habilidad cognitiva, la variación en todos los controles incluidos fue estadísticamente significativa con excepción del índice de riqueza del hogar. La talla para la edad, el índice de masa corporal y el haber faltado a la escuela por problemas de salud evidenciaron un efecto negativo sobre el puntaje de la prueba.

Tabla 6. Estimación de la primera etapa (modelo de efectos fijos - medidas de habilidad)

	Autoestima	Auto eficacia	Habilidad cognitiva
Habilidad del cuidador principal	0,039*** (0,013)	0,099*** (0,023)	- -
Índice de riqueza	0,027 (0,180)	0,248 (0,261)	-2,793 (1,854)
Talla para la edad (estandarizado)	-0,136*** (0,047)	0,187*** (0,038)	-1,698*** (0,587)
Índice de masa corporal (estandarizado)	0,068* (0,023)	0,097 (0,075)	-1,377*** (0,302)
Edad (en meses)	-0,002 (0,001)	0,001 (0,001)	0,793*** (0,013)
Estatus socio-económico (percibido)	0,343*** (0,035)	0,007 (0,039)	1,144*** (0,376)
Ausencia a la escuela por problemas de salud	-0,040*** (0,048)	-0,271*** (0,049)	-1,416*** (0,462)
Log consumo del hogar	0,087 (0,060)	-0,043 (0,062)	-0,363 (0,553)
Número de observaciones	349	349	349
R-cuadrado	0,035	0,025	0,82

Nota: Se considera la muestra de niños que viven en zonas urbanas y tienen información sobre todas las variables incluidas en la regresión. * indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** un indica un nivel de significancia del 1% de la prueba de diferencia de medias entre hombres y mujeres. Errores estándares agrupados a nivel de comunidad.

Fuente: Elaboración propia, 2013.

La tabla 7 presenta los resultados de la segunda etapa del procedimiento propuesto. Los controles considerados son el género y el nivel educativo de los padres. La literatura en torno a la formación de habilidades sugiere que la habilidad “latente” corresponde a la habilidad innata y, por tanto, debería ser afectada por variables determinadas para el niño hasta los tres años de vida. En el presente estudio lo que estimamos como habilidades latentes en realidad es la habilidad formada hasta los 12 años del niño, por lo que es de suponer que las variables fijas hasta dicho momento serían importantes para determinar el nivel de habilidades latentes del niño.

Controles como el nivel educativo de los padres, el género del niño y la lengua materna son algunas de estas características importantes, pero no aquellas como características asociadas a la educación secundaria recibida (por problemas de endogeneidad), este hecho es el que motiva la forma reducida que desarrollamos. En el caso de las tres proxies de habilidad latente, todos los controles incluidos⁵ se mostraron estadísticamente significativos. Mientras las mujeres evidencian un mayor nivel de auto eficacia, lo opuesto ocurre para las habilidades cognitivas. Asimismo, el nivel educativo de los padres tiene un impacto positivo en todas las habilidades latentes predichas.

Tabla 7. Estimación de la segunda etapa (habilidades latentes sobre características permanentes, ENM)

	Autoestima	Autoeficacia	Habilidad cognitiva
Hombre	0,012 (,0016)	-0,361*** (0,013)	1,026*** (0,361)
Nivel educativo del padre	0,035*** (0,006)	0,018*** (0,004)	1,442*** (0,075)
Nivel educativo de la madre	0,016*** (,0002)	-0,008** (0,003)	0,461***+ (0,090)
Constante	-0,581*** (0,071)	0,43*** (0,044)	61,802*** (1,326)
Número de observaciones	349	349	349
R-cuadrado	0,048	0,067	0,169

Nota: * indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** un indica un nivel de significancia del 1% de la prueba de diferencia de medias entre hombres y mujeres. Errores estándares agrupados a nivel de comunidad.

Fuente: Elaboración propia, 2013.

La tabla 8 muestra estadísticas descriptivas de las predicciones realizadas en ambas bases de datos (Niños del Milenio y ENHAB) para la muestra completa y la brecha de género. Como se puede observar, ambas predicciones evidencian patrones similares en términos de la dirección (signo) de las brechas.

⁵ La variable de lengua materna no fue considerada para el análisis ya que alrededor del 90% de la muestra tiene el español como lengua materna lo que implica poca variabilidad en dicha característica.

Tabla 8. Estadísticos de la tercera etapa (habilidades latentes predichas en ambas bases de datos)

Niños del milenio	Mujeres	Hombres	Brecha de género	Muestra completa
Habilidad cognitiva latente	-42,676	-40,854	1,821	-41,716 (1,462)
Auto eficacia latente	0,532	0,182	-0,350***	0,347 (0,073)
Autoestima latente	-0,067	-0,036	0,031	-0,050 (0,074)
Número de observaciones	165	184		349

ENHAB	Mujeres	Hombres	Brecha de género	Muestra completa
Habilidad cognitiva latente	-47,409	-46,209	1,201***	-46,829 (0,201)
Auto eficacia latente	0,516	0,156	-0,361***	0,342(0,001)
Autoestima latente	-0,198	-0,193	0,005	-0,196 (0,005)
Número de observaciones	3.872	3.627		7.499

Nota: Los efectos fijos predichos para la muestra de ENHAB fueron construidos sobre la base de los estimados de la muestra urbana de ENM. Errores estándares a justados ante la presencia de correlación intra-grupos se reportan entre paréntesis. * indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** un indica un nivel de significancia del 1%.

Fuente: Elaboración propia, 2013.

Finalmente, se muestran los resultados de la estimación del modelo de salarios en función de habilidades latentes. La tabla 9 compara los resultados obtenidos tras estimar una ecuación de Mincer con medidas de habilidad (columna 3) y con habilidades latentes⁶ (columna 4). Vale la pena resaltar dos resultados. En primer lugar, el retorno a la educación considerando la especificación de la columna 4 es mayor que el correspondiente a la columna 3 y más similar al de la columna 2. Este hecho es consistente con la hipótesis de que las medidas de habilidad capturan parte del efecto de la educación sobre el salario (lo cual explicaría la caída en el retorno entre las columnas 2 y 3). En segundo lugar, cabe resaltar el cambio en la significancia estadística de las habilidades no cognitivas. Ello pone en evidencia que existe una relación entre habilidades no cognitivas y salarios, pero también implica que podemos considerar el efecto indirecto de las habilidades a través del nivel educativo al poder controlar por el nivel educativo y las habilidades latentes (netas).

⁶ A partir de este momento utilizaremos únicamente la autoeficacia como la habilidad no cognitiva representativa.

Tabla 9. Ecuación de Mincer con habilidades latentes

	[1]	[2]	[3]	[4]
Años de educación	0,1091***	0,0934***	0,0804***	0,0914***
Experiencia	0,0255	0,0316*	0,0329*	0,0333*
Experiencia ²	-0,0003	-0,0004	-0,0004	-0,0004
Reside en Lima	0,0397	0,0297	0,0091	0,0606*
Lengua materna (español)	0,2675	0,4296	0,5023**	0,2975
Nivel educativo del padre		0,0148	0,0097	0,0507
Nivel educativo de la madre		0,0386**	0,0365**	0,0312
Goldberg, extroversión			0,0257	
Goldberg, amabilidad			-0,0371	
Goldberg, cooperación			-0,0616*	
Goldberg, actitud concienzuda (fuerte)			-0,005	
Goldberg, estabilidad emocional			0,0810***	
Goldberg, apertura			-0,0069	
Grit 2, consistencia de intereses			-0,0453*	
Grit 2, persistencia del esfuerzo			0,0031	
Pruebas de habilidad cognitiva			0,0751***	
Habilidad no cognitiva latente (predicha)				-0,7259***
Habilidad cognitiva latente (predicha)				-0,0044
Constante	-0,6028**	-0,9125**	-0,8131***	-0,8876
Número de observaciones	1.079	1.079	1.073	1.079
R-cuadrado	0,157	0,17	0,187	0,193

Nota: Errores estándar robustos entre paréntesis se encuentran agrupados a nivel regional. * indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** un indica un nivel de significancia del 1%. La muestra corresponde a personas empleadas en el momento de la encuesta y con información disponible acerca de pruebas de habilidades y controles individuales.
Fuente: Elaboración propia, 2013.

2.2 Método de descomposición de Blinder-Oaxaca

La presente sección describe los resultados obtenidos tras aplicar el método de descomposición de Blinder-Oaxaca a la muestra completa en edad de trabajar⁷, pero considerando las diferencias en habilidades latentes (previamente predichas). Como se planteó previamente para el caso en que utilizamos medidas de habilidad, existe una brecha salarial de género significativa. A diferencia de los resultados obtenidos en la sección anterior, la brecha hallada por medio de la nueva especificación no es atribuible solo a diferencias en retornos o niveles de habilidades cognitivas, sino que las habilidades no cognitivas también parecen importantes para explicar la brecha salarial de género.

La tabla 10 muestra los resultados de la descomposición de Blinder-Oaxaca para la muestra de la ENHAB considerando brechas en habilidades cognitivas latentes así como en autoeficacia

⁷ Cabe resaltar que el tamaño de muestra en este caso es mayor al caso en que se aplicó Oaxaca-Blinder con medidas de habilidad ya que se considera también a aquellos individuos sin información para medidas de habilidad. Procedemos de esta manera con el fin de aprovechar al máximo la variabilidad de los datos disponibles.

(latente). Luego de aplicar las descomposiciones “doble” y “triple”, los datos muestran evidencia a favor del hecho de que la brecha salarial de género es atribuible a diferencias en los coeficientes y predictores.

En cuanto a diferencias en los retornos, es posible identificar que los hombres gozan de un mayor retorno a las habilidades cognitivas lo cual genera un incremento en la brecha salarial de género, pero no existen diferencias estadísticas en los retornos a las habilidades no cognitivas que afecten la brecha. En cuanto a las diferencias en los niveles de las habilidades, la tabla 10 refleja que diferencias en las habilidades cognitivas y no cognitivas favorecen a los hombres en términos de salario.

Si bien el mayor acervo de habilidades cognitivas entre los hombres contribuye a incrementar la brecha salarial de género, la evidencia muestra que si las mujeres tuvieran el mismo nivel de habilidades no cognitivas que los hombres, ganarían salarios significativamente mayores. A primera vista, ello parece contradecir lo observado mediante los estadísticos descriptivos calculados inicialmente en el que las mujeres mostraban mayores niveles de autoeficacia. No obstante, este fenómeno se explica por el hecho de que el retorno a las habilidades no cognitivas de las mujeres es negativo (como se hubiera podido observar si hubiéramos estimado la ecuación de Mincer con habilidades latentes por género). Por tanto, un mayor acervo de una habilidad que goza de un retorno negativo en el mercado genera, en efecto, un impacto negativo perjudicando a las mujeres e incrementando la brecha salarial de género.

Tabla 10. Descomposición de Blinder-Oaxaca con habilidades latentes

	[1]	[2]	[3]	[4]
Brecha salarial	0,241*** (0,032)	0,241*** (0,027)	0,241*** (0,032)	0,241*** (0,035)
Niveles	0,419*** (0,117)	0,607*** (0,134)	0,458*** (0,148)	0,618*** (0,152)
Retornos	-0,178 (0,124)	-0,366*** (0,136)	-0,160 (0,157)	-0,367*** (0,141)
Niveles				
Cognitiva	0,026*** (0,008)	0,037*** (0,010)	0,034*** (0,011)	0,043*** (0,013)
No cognitiva	0,393*** (0,115)	0,580*** (0,132)	0,424*** (0,144)	0,585*** (0,148)
Retornos				
Cognitiva	0,540*** (0,178)	0,477*** (0,254)	0,544*** (0,180)	0,481*** (0,187)
No cognitiva	0,049 (0,186)	0,004 (0,272)	0,063 (0,265)	-0,001 (0,270)
Número observaciones	4.079	4.079	4.079	4.079
Controles	No	Sí	No	Sí
Interacciones	No	No	Sí	Sí

Nota: Errores estándar entre paréntesis se encuentran agrupados a nivel regional.* indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** un indica un nivel de significancia del 1%. La muestra corresponde a personas empleadas en el momento de la encuesta y con información disponible acerca de pruebas de habilidades y controles individuales. Los controles considerados son edad, edad al cuadrado y si el individuo reside en Lima.

Fuente: Elaboración propia, 2013.

2.3 Estimación conjunta: educación, empleo, ocupación y salarios

La tabla 11 muestra los resultados de la estimación conjunta considerando las habilidades latentes cognitivas y no cognitivas. El considerar el rol de ambos tipos de habilidad en los procesos de elección considerados conlleva a resultados interesantes. En primer lugar, la habilidad cognitiva (latente) se muestra fundamental para lograr mayores niveles de educación, en similar magnitud para hombres y mujeres. Las habilidades no cognitivas parecen irrelevantes para esta primera decisión. Los siguientes tres procesos de elección deben ser interpretados de manera conjunta. En términos generales es posible identificar que hay diferencias de género importantes en lo retornos a las habilidades cognitivas y no cognitivas en cada una de las tres variables de resultado analizadas. Ello se encuentra explicado por la importancia de considerar la dinámica detrás de la elección de la ocupación y por el hecho de que las mujeres tienen un mayor nivel de habilidades no cognitivas, las cuales a su vez cuentan con un retorno negativo. A propósito de ello, es posible interpretar que los hombres ganan mayores salarios por haber sido empleados en ocupaciones que premian (más) las habilidades cognitivas. Ello, de la mano del hecho de que los hombres gozan de un mayor nivel de este tipo de habilidades, ayuda a explicar

la existencia de una brecha salarial de género que favorece a los hombres. Asimismo, el análisis conjunto pone en evidencia que el retorno a las habilidades no cognitivas en términos de empleo y salarios es mayor para los hombres aun cuando estos presentan un menor acervo de este tipo de habilidades.

Tabla 11. Estimación conjunta con habilidades latentes

Modelos	Educación	Empleo	Ocupación	Salarios por hora
Cognitiva (mujeres)	0,055*** (0,004)	-0,014*** (0,002)	0,065*** (0,004)	-0,076*** (0,018)
Interacción c/cognitiva	-0,005 (0,003)	-0,009*** (0,002)	-0,014*** (0,004)	-0,055*** (0,017)
No cognitiva (mujeres)	0,109 (0,443)	-0,117 (0,298)	-1,539*** (0,483)	-3,013 (2,314)
Interacción c/no cognitiva	-0,515 (0,438)	0,817** (0,319)	-1,347*** (0,512)	5,794*** (2,463)
Número observaciones	7.499			

Nota: * indica un nivel de significancia del 10%; ** indica un nivel de significancia del 5%; y *** un indica un nivel de significancia del 1%. El término de interacción equivale al producto: (Hombre)*(Habilidad Latente).
Fuente: Elaboración propia, 2013.

En contraste con los resultados de la estimación conjunta utilizando medidas de habilidad, en este caso es posible identificar que las diferencias entre géneros en habilidades cognitivas y no cognitivas a favor de los hombres se reducen cuando se consideran las habilidades latentes. En el caso de las habilidades cognitivas, las brechas se tornan significativas a favor de las mujeres. Más aún, el retorno a ambos tipos de habilidad latente gana significancia estadística en la ecuación de la decisión de la ocupación. Esta evidencia favorece la hipótesis de que gran parte de las diferencias atribuibles a brechas de habilidad ocurren en la dinámica que gobierna la elección de ocupaciones.

3. Pruebas de robustez: soporte común

A fin de obtener proxies de las habilidades cognitivas y no cognitivas latentes, el procedimiento propuesto requiere de un método de imputación (para la muestra de ENHAB) sobre la base de una estimación realizada en otra base de datos (ENM). Si bien se han presentado estadísticos descriptivos acerca de los estimados y predicciones de las habilidades latentes para las muestras de ENM y ENHAB, respectivamente, la literatura empírica sugiere que se trabaje con el soporte común. El soporte común se encuentra determinado por el área que comparten las distribuciones de las habilidades latentes estimadas y predichas.

Las figuras en los anexos 1 y 2 muestran las distribuciones de las habilidades cognitivas y no cognitivas latentes, respectivamente para las muestras de Niños del Milenio y ENHAB tanto para hombres como para mujeres. En todos los casos, la distribución correspondiente a la muestra de Niños del Milenio evidencia una mayor varianza que la muestra ENHAB, por tanto, el soporte común acota únicamente a la muestra de Niños del Milenio. Por tanto, luego de calcular los límites superior e inferior de cada una de las distribuciones correspondientes a la muestra de ENHAB se reestimaron los efectos fijos con la submuestra de Niños del Milenio cuyo efecto fijo inicial se encontraba comprendido en el soporte común inicial. Así, se predijo nuevamente hacia la muestra de ENHAB para luego comparar la distribución de los efectos fijos con el soporte común. No obstante, las figuras de los anexos 3 y 4 ponen en evidencia la poca ganancia que dicho procedimiento genera. Más aun, el acotar la muestra de Niños del Milenio genera una disminución significativa del número de observaciones lo cual impacta negativamente en la precisión de los estimados obtenidos con el soporte común.

Tras observar que acotar la muestra sobre la base del soporte común no contribuye a la obtención de mejores estimados sino que los perjudica, se optó por mantener el análisis desarrollado hasta la sección previa haciendo uso de la muestra completa de Niños del Milenio y la información correspondiente de las variables relevantes.

4. Limitaciones

El procedimiento de estimación propuesto y los resultados derivados del mismo no están libres de limitaciones. La principal limitación es la poca disponibilidad de información de los niños a temprana edad. La literatura en torno a la formación de habilidades y la relación entre habilidades, educación y salarios sugiere utilizar habilidades latentes como una aproximación a las habilidades innatas en tanto que estén limpias de cualquier evento experimentado más adelante en la vida. En este documento de investigación, las habilidades latentes que puedo estimar son aquellas que se encuentran fijas a partir de los 12 años de edad, las cuales probablemente difieran significativamente de las habilidades innatas (especialmente en el caso de habilidades no cognitivas). Por tanto, las diferencias estimadas en habilidades no observables podrían ser el resultado de estos factores no observables.

Una segunda limitación es el hecho de que las medidas de habilidad no cognitiva recogidas en la base de datos del Estudio de Niños del Milenio no corresponden a los mismos aspectos de la personalidad que las recogidas en la base de datos de la ENHAB. La literatura relacionada

sugiere que esto no es un problema para las habilidades cognitivas, pero si podría serlo para las no cognitivas. En las regresiones por mínimos cuadrados ordinarios mostradas en la sección previa se podía observar cómo mientras algunas habilidades no cognitivas evidenciaban una relación positiva con los salarios, otras evidenciaban una negativa. Sin embargo, este tema es consecuencia de la poca disponibilidad de información y lo que se presenta en este documento es una primera aproximación.

Finalmente, una última limitación es que aquello que interpretamos como “brecha no explicada” puede ser explicado por heterogeneidad en preferencias u otros aspectos de la personalidad no observables.

Conclusiones y recomendaciones

1. Este documento muestra evidencia preliminar acerca del rol de las habilidades cognitivas y no cognitivas en la brecha salarial de género no explicada. En un primer intento por estimar el efecto de dichas habilidades en los salarios, se llevó a cabo una aproximación empírica que modela los salarios en función de medidas de habilidad. Se complementó el trabajo con un procedimiento que permite estimar proxies de habilidades latentes para estimar un modelo en función de habilidades latentes. El modelo está basado en lo propuesto por Heckman *et al.* (2006). Mientras dichos autores identifican las habilidades latentes en la dependencia de diferentes medidas de habilidad para un mismo periodo de tiempo, en este documento se utiliza la variación en el tiempo para una misma medida de habilidad. Esto es posible gracias a la disponibilidad de información de datos de panel acerca de medidas de habilidad cognitiva y no cognitiva. Asimismo, se estimó un modelo conjunto de educación, empleo, ocupación y salarios a fin de identificar los efectos de las habilidades latentes en la brecha salarial de género en las decisiones del individuo previas a la obtención de un determinado nivel de salario.
2. La principal contribución es analizar el rol de las habilidades latentes cognitivas y no cognitivas en la brecha salarial de género en un país en vías de desarrollo por medio de la estimación y consideración de proxies de habilidades latentes distinguiendo los efectos de dichas habilidades haciendo uso de un modelo conjunto de educación, empleo, ocupación y salarios.
3. El procedimiento econométrico y el modelo desarrollados en este documento evidencian algunos hechos importantes. En primer lugar, se confirma la existencia de una brecha salarial de género estadísticamente significativa para el Perú. Estimaciones con medidas de habilidad confirman lo propuesto en la literatura empírica respecto a los problemas de endogeneidad que surgen al utilizar puntajes de pruebas como medidas de habilidad. La descomposición de Blinder-Oaxaca en un modelo de medidas de habilidad evidencia la existencia de diferencias de género significativas en la dotación de habilidades cognitivas y la ausencia de diferencias relevantes en términos de habilidades no cognitivas (en dotación y retornos). La estimación de un modelo conjunto, por otro lado, muestra que las diferencias en habilidades cognitivas y no cognitivas entre hombres y mujeres son importantes, pero para las decisiones previas a la determinación del salario. Las habilidades cognitivas

parecen ser relevantes para determinar los años de educación y la elección de ocupación mientras que las no cognitivas lo son para el empleo y los salarios.

4. La aplicación de un modelo de efectos fijos para la identificación de proxies de habilidades latentes cognitivas y no cognitivas conllevan a una relación más clara entre habilidades y salario. Como consecuencia del procedimiento de estimación de las habilidades latentes fue posible llevar a cabo una descomposición más precisa de la brecha salarial de género. Al estimar la ecuación de salarios en función de medidas de habilidad se obtienen estimados sesgados como consecuencia de la dependencia de estas y el nivel educativo y la naturaleza endógena de esta última variable. Luego de identificar proxies para las habilidades latentes, es posible obtener estimados sin sesgos y, por tanto, la contribución real de las habilidades en los salarios (hasta el momento condicional a la elección de ocupación). Las habilidades latentes resultaron ser estadísticamente significativas para el salario promedio de la muestra así como para las diferencias de género. En particular, diferencias en la dotación de habilidades no cognitivas parecen contribuir negativamente a la brecha salarial de género. Las mujeres parecen recibir menores salarios que los hombres ya que cuentan con una mayor dotación de habilidades no cognitivas, las cuales son valoradas negativamente (castigadas) en el mercado laboral. No se encontraron diferencias de género significativas en el retorno de las habilidades cognitivas. Más aun, la estimación del modelo conjunto pone en evidencia que la brecha salarial de género observada es atribuible principalmente a diferencias en la elección de ocupación entre hombres y mujeres. Las habilidades cognitivas y no cognitivas son valoradas de forma distinta para hombres y mujeres en cuanto a las decisiones de educación, empleo y salario, pero principalmente los hombres ganan mayores salarios porque se autoseleccionan en ocupaciones que premian habilidades cognitivas, cuya dotación es mayor en hombres que en mujeres.
5. En síntesis, es posible concluir que el procedimiento propuesto para la estimación de las habilidades latentes, aun cuando no está libre de limitaciones, conlleva a resultados razonables a la luz de la literatura empírica relacionada en cuanto a la relación entre salario, educación y habilidades.
6. Algunas extensiones a raíz de este documento serían la estimación de los efectos marginales de las habilidades latentes en las decisiones de educación, empleo, ocupación y salarios, el análisis de la brecha salarial de género por cuartiles de ingreso, y la aplicación de la

estrategia de identificación de Heckman *et al.* (2006) que utiliza el puntaje de distintas pruebas para una misma habilidad en el mismo periodo de tiempo.

Bibliografía

Banco Mundial (2010) “Encuesta Nacional de Habilidades y Mercado Laboral” [Base de datos digital]. Lima: Banco Mundial.

Bassi, Marisa y Galiani, Sebastián. (2009). "Labor Market Insertion of Young Adults in Chile". Banco Interamericano de Desarrollo (BID). Mimeo.

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

Boyden, J. (2011a). “Young Lives: an International Study of Childhood Poverty: Round 2, 2006”. [Computer file]. Colchester, Essex: UK Data Archive [distributor]. SN: 6852. Fecha de consulta: 30/01/2013. Disponible en: <<http://dx.doi.org/10.5255/UKDA-SN-6852-1>>.

Boyden, J. (2011b). “Young Lives: an International Study of Childhood Poverty: Round 3, 2009”. [Computer file]. Colchester, Essex: UK Data Archive [distributor]. SN: 6853. Fecha de consulta: 30/01/2013. Disponible en: <<http://dx.doi.org/10.5255/UKDA-SN-6853-1>>.

Cobb-Clark, Deborah A. y Tan, Michelle. (2011). "Noncognitive Skills, Occupational Attainment, and Relative Wages". En; *Labour Economics*. Vol. 18, N°1

Cunha, Flavio y Heckman, James. (2007). "The Technology of Skill Formation". En: *American Economic Review*. Vol. 97, N° 2.

Cunha, Flavio; Heckman, James J.; Lochner, Lance; y Masterov, Dimitriy V. (2006). "Interpreting the evidence on life cycle skill formation". En: Hanushek, E. A. y Welch, F. (Eds.). (2006). *Handbook of the Economics of Education, Ch. 12*. Amsterdam: North-Holland.

Cunha, Flavio; Heckman, James J.; y Schennach, Susanne M. (2010) "Estimating the technology of cognitive and noncognitive skill formation". En: *Econométrica*. Vol. 78, N°3.

Díaz, Juan J.; Arias, Omar y Vera Tudela, David. (2012). “Does Perseverance Pay as Much as Being Smart?: The Returns to Cognitive and Non-cognitive Skills in urban Perú”. Mimeo.

Duckworth, A.L.; Peterson, C.; Mathews, M.D. y Kelly, D.R. (2007). “Grit: Perseverance and passion for long term goals”. En: *Journal of Personality and Social Psychology*, Vol. 92, N°6.

Edwards, Richard C. (1976). "Personal Traits and Success in Schooling and Work". En: *Journal of Human Resources*. Vol. 11, N°1.

Fortin, Nicole M. (2008). "The Gender Wage Gap among Young Adults in the United States: The Importance of Money versus People". En: *Journal of Human Resources*. Vol. 43, N°4.

- Goldberg, L.R. (1990). "An alternative 'description of personality': The Big-Five factor structure". En: *Journal of Personality and Social Psychology*. Vol. 59.
- Grove, Wayne A.; Hussey, Andrew; y Jetter, Michael. (2011). "The Gender Pay Gap Beyond Human Capital: Heterogeneity in Noncognitive Skills and in Labor Market Tastes". En: *Journal of Human Resources*. Vol. 46, N°4.
- Hansen, Karsten T.; Heckman, James J. y Mullen, Kathleen J. (2004). "The effect of schooling and ability on achievement test scores". En: *Journal of Econometrics*. Vol. 121, N° 1-2.
- Heckman, James J.; Stixrud, Jora y Urzúa, Sergio. (2006). "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior". En: *Journal of Labor Economics*. Vol. 24, N°3.
- Helmers, Christian y Patnam, Manasa. (2011). "The formation and evolution of childhood skill acquisition: Evidence from India". En: *Journal of Development Economics*. Vol. 95, N° 2.
- Herrnstein, Richard J. y Murray, Charles. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.
- Huttly, S. y Jones, N. (2011). "Young Lives: an International Study of Childhood Poverty: Round 1, 2002". [Computer file]. Cuarta edición. Colchester, Essex: UK Data Archive [distribuidor]. SN: 5307. Fecha de consulta: 30/01/2013. Disponible en: <<http://dx.doi.org/10.5255/UKDA-SN-5307-1>>.
- Klein, Roger; Spady, Richard; y Weiss, Andrew. (1991). "Factors Affecting the Output and Quit Propensities of Production Workers". En: *Review of Economic Studies*. Vol.58, N°2.
- Murnane, Richard J.; Willett, John B.; y Levy, Frank. (1995). "The growing importance of cognitive skills in wage determination". En: *The Review of Economics and Statistics*. Vol. 77, N°2. Boston: MIT Press.
- Neal, Derek A. y Johnson, William R. (1996). "The role of premarket factors in black-white wage differences". En: *Journal of Political Economy*. Vol. 104, N°5.
- Ritter, Joseph A. y Taylor, Lowell J. (2011). "Racial Disparity in Unemployment". En: *The Review of Economics and Statistics*. Vol. 93, N°1. Boston: MIT Press.
- Urzúa, Sergio; Bravo, David y Sanhueza, Claudia. (2009). "Ability, Schooling Choices and Gender Labor Market Discrimination: Evidence for Chile". En: Ñopo, Hugo; Chong, Alberto; y Moro, Andrea. (Eds.). (2009). *Discrimination in Latin America: An Economic Perspective*. Washington DC: Inter-American Development Bank.

Anexos

Anexo 1. Componentes psicosociales – muestras de ENM

Ronda 2

Variables	Muestra Completa		Hombres		Mujeres	
	Mean	SD	Mean	SD	Mean	SD
Auto-eficacia del niño						
Si me esfuerzo lo suficiente puedo mejorar mi situación de vida	1,987	0,157	1,981	0,188	1,994	0,113
Otras personas de mi familia tomando las decisiones respect a como gasto mi tiempo	1,267	0,941	1,320	0,924	1,206	0,959
Me gusta hacer planes sobre mi future la boral y acadrico	1,943	0,327	1,934	0,349	1,953	0,300
Si estudio lo suficiente sere premiado con un mayor empleo en el futuro	1,957	0,277	1,956	0,275	1,959	0,280
No tengo eleccion en cuanto al trabajo que ejerizo	0,859	0,976	0,841	0,978	0,887	0,979
Autoestima del niño						
Me siento orgulloso de mostrarle a mis amigos donde vivo	1,800	0,590	1,800	0,585	1,801	0,597
Me siento avergonzado de la ropa que visto	0,257	0,664	0,232	0,635	0,286	0,696
Me siento avergonzado de los zapatos que tengo	0,188	0,581	0,157	0,537	0,223	0,628
Me siento orgulloso del trabajo que ejercela cabeza del hogar	1,878	0,468	1,896	0,439	1,857	0,500
Nunca me he sentido avergonzado por no tener los libros o utiles correctos	0,453	0,829	0,439	0,817	0,468	0,843
Me siento preocupado porque no tengo el uniforme correcto	0,698	0,951	0,652	0,936	0,752	0,967
Me siento orgulloso de mis logros en el colegio	1,928	0,356	1,909	0,401	1,949	0,295
Me siento avergonzado del trabajo que ejerizo	0,196	0,587	0,150	0,521	0,268	0,675
Me siento orgulloso del trabajo que ejerizo	1,826	0,555	1,885	0,458	1,732	0,675
Auto-eficacia del cuidador principal						
Si me esfuerzo lo suficiente puedo mejorar mi situación de vida	1,942	0,315	1,962	0,242	1,918	0,382
Me gusta elaborar planes acerca de mi futuro	1,902	0,416	1,902	0,406	1,902	0,429
No tengo eleccion alguna respecto del colegio al que enviare a mi hijo	0,836	0,984	0,777	0,973	0,905	0,994
Si mi hijo enferma, no hay mucho que pueda hacer para que mejore/para ayudarlo	0,456	0,836	0,405	0,803	0,516	0,871
No puedo ayudar mucho a mi hijo sino le va bien en el colegio aun cuando me esfuerce	0,549	0,887	0,472	0,842	0,637	0,930
Autoestima del cuidador principal						
Me siento orgulloso de mostrar a mis amigos y familiares donde vivo	1,782	0,601	1,783	0,592	1,782	0,612
Me siento avergonzado de la ropa que visto	0,344	0,746	0,302	0,707	0,392	0,787
Me siento orgulloso del trabajo que ejercela cabeza del hogar	1,896	0,429	1,879	0,456	1,917	0,395
El trabajo que ejerizo me hace sentir orgulloso	1,939	0,328	1,943	0,313	1,934	0,346
Me siento orgulloso de mis hijos	1,972	0,231	1,970	0,237	1,975	0,224

Fuente: Boyden, 2011a.

Elaboración: Propia, 2013.

Anexo 1. Componentes psicosociales – muestras de ENM (continúa de la página anterior)

Ronda 3

Variables	Muestra Completa		Hombres		Mujeres	
	Media	DE	Media	DE	Media	DE
Auto-eficacia del niño						
Si me esfuerzo lo suficiente puedo mejorar mi situación de vida	1,191	0,544	1,155	0,554	1,232	0,531
Otras personas de mi familia toman todas las decisiones respect a como gasto mi tiempo	3,340	0,904	3,383	0,914	3,290	0,892
Me gusta hacer planes sobre mi future laboral y acadmico	1,137	0,660	1,039	0,684	1,249	0,613
Si estudio lo suficiente sere premiado con un mejor empleo en el futuro	1,359	0,588	1,330	0,608	1,392	0,564
No tengo eleccion en cuanto al trabajo que ejerzo	0,224	0,992	0,111	1,012	0,344	0,957
Autoestima del niño						
Me siento orgulloso de la ropa que visto	0,944	0,655	0,966	0,632	0,919	0,680
Me siento orgulloso de (tener) los zapatos que tengo	0,917	0,759	0,921	0,755	0,913	0,765
Nunca me he sentido avergonzado por no tener los libros o útiles correctos	0,618	0,940	0,602	0,943	0,635	0,938
Me siento orgulloso de tener el uniforme correcto	1,099	0,605	1,060	0,584	1,144	0,626
Me siento orgulloso del trabajo que ejerzo	1,042	0,572	1,052	0,571	1,031	0,574
Auto-eficacia del cuidador principal						
Si me esfuerzo lo suficiente puedo mejorar mi situación de vida	1,042	0,461	1,048	0,483	1,036	0,436
Me gusta elaborar planes acerca de mi futuro	0,578	0,837	0,545	0,803	0,616	0,873
No tengo eleccion alguna respecto del colegio al que enviare a mi hijo	0,578	0,837	0,545	0,803	0,616	0,873
Si mi hijo enferma, no hay mucho que pueda hacer para que mejore para ayudarlo	0,869	0,689	0,915	0,676	0,817	0,702
No puedo ayudar mucho a mi hijo sino le va bien en el colegio aun cuando me esfuerce mucho	0,767	0,811	0,847	0,800	0,675	0,816
Autoestima del cuidador principal						
Me siento orgulloso de mostrar a mis amigos y familiares donde vivo	0,880	0,643	0,904	0,645	0,852	0,641
Me siento avergonzado de la ropa que visto	0,736	0,660	0,734	0,705	0,740	0,606
Me siento orgulloso del trabajo que ejerco la cabeza del hogar	0,965	0,688	0,943	0,726	0,990	0,644
El trabajo que ejerco me hace sentir orgulloso	1,030	0,556	0,997	0,618	1,068	0,474
Me siento orgulloso de mis hijos	1,439	0,581	1,398	0,575	1,486	0,584

Fuente: Boyden, 2011b.
Elaboración: Propia, 2013.

Anexo 2. Habilidades latentes y soporte común

Figura 1. Distribución de habilidades cognitivas

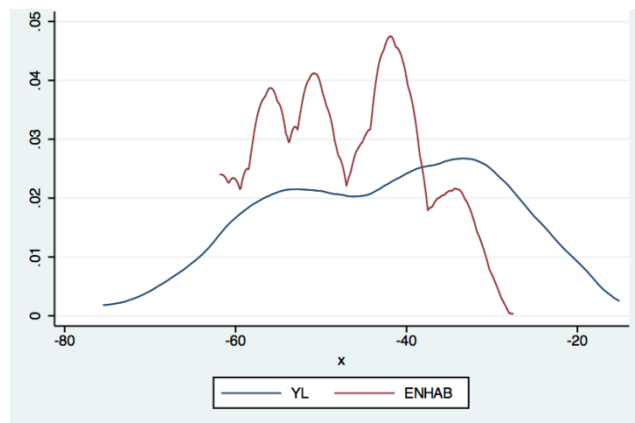
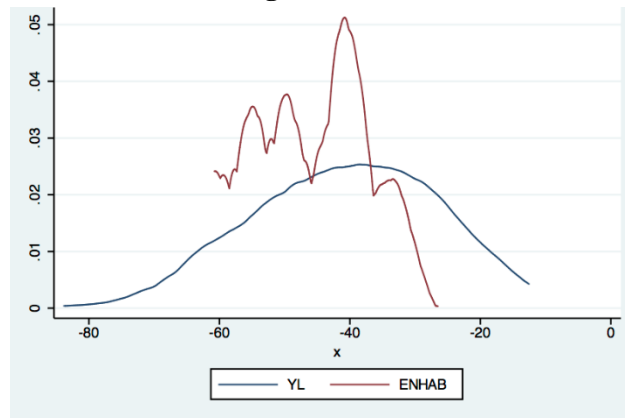
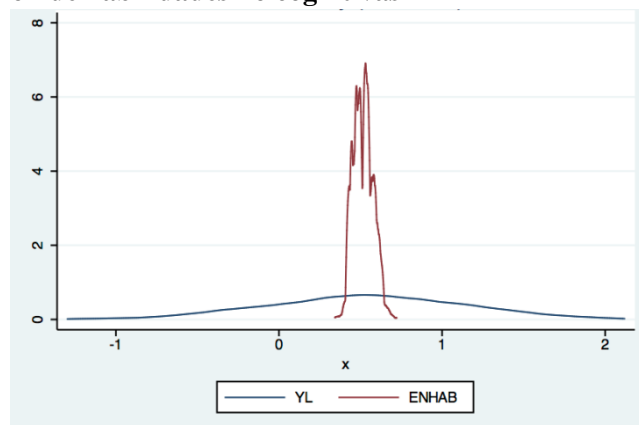
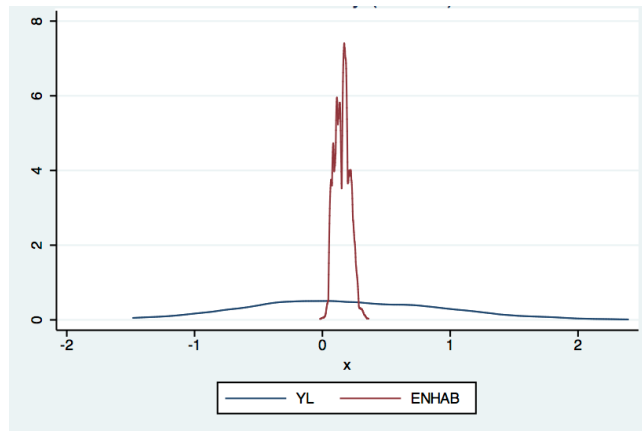


Figura 2. Distribución de habilidades no cognitivas



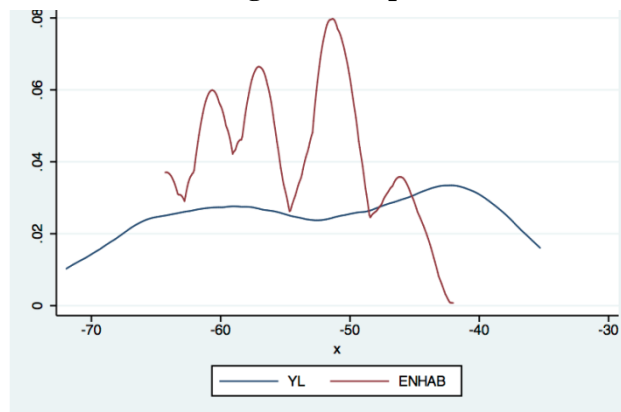
Anexo 2. Habilidades latentes y soporte común (continúa de la página anterior)

Figura 2. Distribución de habilidades no cognitivas (continúa de la página anterior)

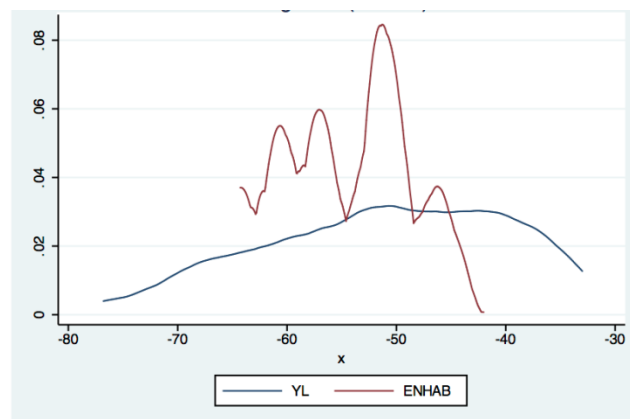


Panel B: hombres

Figura 3. Distribución de habilidades cognitivas (soporte común)



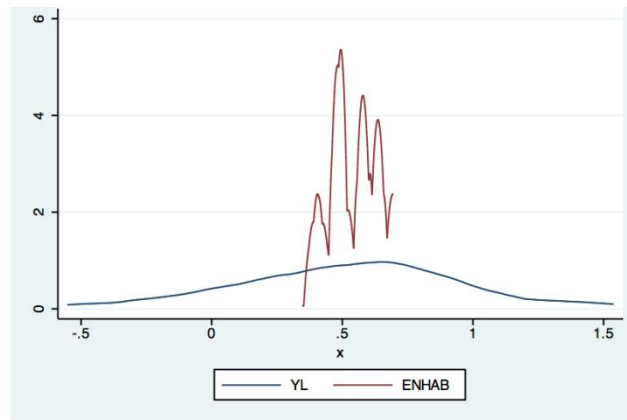
Panel A: mujeres



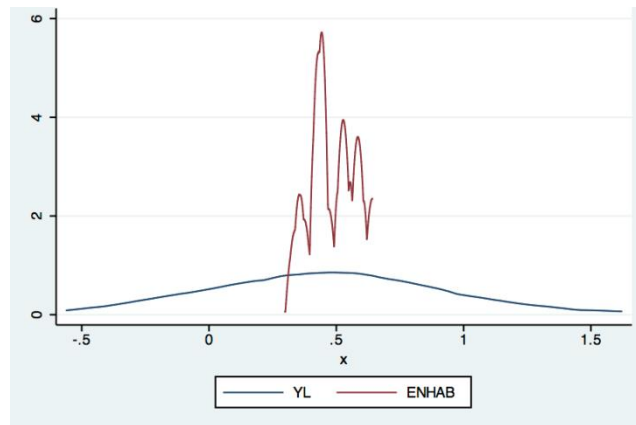
Panel B: hombres

Anexo 2. Habilidades latentes y soporte común (continúa de la página anterior)

Figura 4. Distribución de habilidades no cognitivas (soporte común)



Panel A: mujeres



Panel B: hombres

Nota biográfica

Luciana Velarde Arrisueño

Bachiller en Economía por la Universidad del Pacífico (2011) con concentración en el sector público y teoría económica; graduada del Curso de Extensión Universitaria de Economía del Banco Central de Reserva del Perú (2011) y egresada de la Maestría en Economía de la Universidad del Pacífico (2013). Se ha desempeñado como asistente de investigación en el Centro de Investigación de la Universidad del Pacífico (CIUP) en el área de Políticas Sociales, y como analista de Calidad de la Información en el Ministerio de Desarrollo e Inclusión Social. Desde el 2014 se desempeña como Profesional Asociada Junior en Educación del Banco Mundial en Perú.