

DOCUMENTO DE DISCUSIÓN

DD/11/05

“Diferencias étnicas y de sexo en el progreso educativo en Perú: ¿para quién y cuándo son los riesgos mayores?”

Juan F. Castro, Gustavo Yamada y Roberto Asmat



UNIVERSIDAD DEL PACÍFICO
CENTRO DE INVESTIGACIÓN

Documento de Discusión

“Diferencias étnicas y de sexo en el progreso educativo en Perú: ¿para quién y cuándo son los riesgos mayores?”

Juan F. Castro, Gustavo Yamada y Roberto Asmat ⁺

Septiembre 2011

Resumen Ejecutivo

En esta investigación analizamos las diferencias étnicas y de sexo en el progreso educativo en el Perú, en los últimos 50 años. Usamos un pool de datos de la Encuesta Nacional de Hogares del 2004-2009 e identificamos cuatro etnias: blanca/mestiza, amazónica, quechua/aymara, y afrodescendiente. Modelamos los años de educación acumulados con un enfoque de *duración* y calculamos la “brecha” de progreso educativo a través del riesgo de abandonar el sistema educativo del grupo medido respecto al de los blancos/mestizos (para las comparaciones por grupos étnicos) y respecto al del hombre (para las comparaciones de sexo). Empleamos tanto un enfoque de riesgos proporcionales como no proporcionales. Este último nos permite analizar los efectos de la etnia y el sexo sobre la probabilidad de deserción en cada etapa del proceso educativo. Encontramos que el riesgo de deserción de los grupos quechua, amazónico y afrodescendiente es mayor que el del grupo de blancos/mestizos. Las restricciones materiales (recursos monetarios y oferta educativa) explican buena parte de la brecha quechua. La brecha afrodescendiente depende más de factores estructurales relacionados con el contexto familiar y educativo de la persona y que no son directamente observables. Encontramos que la etnia afrodescendiente se encuentra en clara desventaja respecto al resto de grupos, con un elevado riesgo de deserción concentrado en los primeros años de la educación primaria. Asimismo, detrás de una brecha hombre-mujer promedio prácticamente cerrada, subsiste aún una diferencia importante entre los riesgos de deserción (a favor de los hombres) concentrada en los primeros años de la educación básica y al interior de los grupos quechuas y amazónicos.

Palabras clave: Progreso educativo, brechas étnicas, brechas de sexo, modelos de duración, Perú.

Correo de los autores: Castro_JF@up.edu.pe, Yamada_GA@up.edu.pe y Asmat_RC@up.edu.pe

⁺ Agradecemos a la Fundación Ford por el apoyo financiero para la realización de este estudio. Todas las opiniones aquí vertidas son de exclusiva responsabilidad de los autores y no comprometen ni a la Universidad del Pacífico ni a la Fundación Ford.

*Las opiniones expresadas en los Documentos de Discusión son de exclusiva responsabilidad de los autores y no expresan necesariamente aquellas del Centro de Investigación de la Universidad del Pacífico. Los Documentos de Discusión difunden los resultados preliminares de las investigaciones de los autores con el propósito de recoger comentarios y generar debate en la comunidad académica.

I. INTRODUCCION

Todavía uno de cada cinco jóvenes peruanos de las minorías étnicas¹ deserta para siempre del sistema educativo sin haber acabado siquiera la educación primaria. En efecto, al 2009, el 22% de las personas identificadas como amazónicas entre 18 y 25 años lograron menos de 6 años de educación total. Las figuras análogas para los grupos quechua/aymara y afrodescendiente son de 20% y 16%, respectivamente. En contraste, solo el 8% de blancos/mestizos tiene dicha desventaja².

Claramente, los grupos no-mestizos enfrentan desventajas en el acceso y progreso en el sistema educativo del Perú. Estas restricciones en la acumulación de capital humano los ponen en clara desventaja en el mercado laboral y, por lo tanto, afectan su capacidad de consumo y bienestar no sólo de por vida, sino también con consecuencias sobre sus hijos³. Aunque la exclusión de grupos sociales es más amplia y se evidencia en el acceso a otros servicios públicos como salud, agua y saneamiento (véase Benavides y Valdivia, 2004; Montero y Yamada, 2011), en el mercado laboral de manera directa (véase Anderson y Shapiro, 1996) y en la participación política y sistema judicial, esta investigación se centra en examinar las diferencias étnicas en el progreso educativo.

La cobertura de la educación básica en el Perú (los 11 primeros años que involucran los niveles de primaria y secundaria) ha mejorado en las últimas 5 décadas. De acuerdo a estimaciones de Castro y Yamada (2011) sobre la base de encuestas nacionales, al comparar el sistema educativo de finales de la décadas de los 90's y 50's del siglo pasado, la cobertura en el acceso a la educación primaria pasó de 87% a 98%, mientras que para la educación secundaria pasó de 46% a 80%. Sin embargo, los mismos autores encuentran que estos avances difieren entre grupos étnicos. Si bien es cierto las brechas étnicas disminuyeron con el tiempo (inclusive al punto de casi cerrarse en la matrícula en primaria), los grupos amazónico, quechua/aymara y afrodescendientes desertan a lo largo del proceso educativo mucho más que el grupo mestizo/blanco, sobre todo en el tránsito de la conclusión primaria a la matrícula en secundaria.

Nuestro interés por el estudio de estas diferencias se debe a la estrecha relación que guardan con las brechas étnicas, ampliamente reportadas, que se encuentran a nivel de ingresos laborales (véase Yamada, Lizarzaburu y Samanamud, 2011). Diferencias en el nivel

¹ En la medida que representan una proporción bastante reducida de la población total.

² Cálculos realizados sobre la base de la Encuesta Nacional de Hogares (2004-2009).

³ Innumerables estudios dan cuenta de la relación positiva entre los años de educación de los padres y la asistencia escolar y logro educativo de los hijos.

y calidad de la educación de los individuos tienen, sin duda, un efecto sobre su productividad y desempeño en el mercado laboral (véase Heckman, 2006). Desde un enfoque de igualdad de oportunidades, la literatura también encuentra que la educación es el factor más importante para eliminar las barreras a las oportunidades de tal manera que el resultado final en el bienestar de los individuos se deba a méritos personales (véase Roemer, 1998).

A partir de lo anterior, resulta fundamental examinar dichas diferencias étnicas en el logro educativo para entender por qué algunos grupos progresan y otros no salen del atraso y la pobreza persistente. Este estudio toma relevancia en una sociedad caracterizada por la diversidad cultural y busca promover el debate sobre exclusión y discriminación étnica, la propuesta de políticas que promueven la diversidad cultural y la construcción de una democracia multicultural en el marco de la igualdad de oportunidades.

El resto del documento está organizado como sigue. En la sección II se realiza una revisión de la literatura referida al estudio de las diferencias étnicas en el progreso y logro educativo, y se discuten las dificultades empíricas que enfrenta un análisis de esta naturaleza.. En la parte III se describe la información que se utilizó en este estudio para analizar las brechas étnicas, así como la construcción de la base de datos y variables usadas. En la sección IV se discute la metodología a emplear. Al respecto, esta investigación realiza dos aportes importantes. En primer lugar, aplica el análisis de duración para capturar el efecto de la etnicidad y el sexo sobre el progreso educativo. Para esto, se emplea la metodología de riesgo proporcional desarrollada en Cox (1972) y que es muy utilizada para el estudio de las tasas de sobrevivencia en las áreas de la salud pública y el desempleo. Asimismo, propone la adaptación de la metodología desarrollada por Royston y Parmar (2001) sobre riesgos no proporcionales para analizar la evolución de estas brechas a lo largo del proceso educativo. En la sección V se muestran los resultados aplicando ambos métodos y, finalmente, en la sección VI se resumen las conclusiones y la discusión de algunas implicancias de políticas públicas a partir de las evidencias encontradas.

II. ETNICIDAD Y EDUCACIÓN: QUÉ SE HA HECHO Y CÓMO MEDIR SU RELACIÓN

La literatura histórica indica que la discriminación racial y étnica en América Latina y el Caribe toma forma y se remonta a la conquista y colonización a la que fueron sometidos estos pueblos⁴. Sin embargo, como señalan Bello y Hopenhayn (2001), ya en Estados modernos, la raza y etnicidad, junto con el sexo y clase económica, generan sistemas a través de los cuales se impide el acceso equitativo a oportunidades socio-económicas. Esto quiere decir que la problemática actual de los grupos indígenas y minorías étnicas ya no se explica por la dominación externa sino más bien por procesos de exclusión interna. Por esto, la literatura social, incluyendo el presente estudio, se interesa en estudiar las brechas racial/étnicas en el sistema educativo.

Aunque las investigaciones de tipo social abordan la raza y etnicidad de forma variada, la raza se suele asociar a distinciones biológicas⁵, mientras que la etnicidad se vincula a características culturales, aunque ambas se encuentran altamente relacionadas. Así, podemos encontrar estudios que se centran en el aspecto racial o étnico dependiendo de la disponibilidad de información.

Diversas investigaciones han buscado explicar la relación entre raza/eticidad y la acumulación de capital humano sobre la base de un análisis de las brechas raciales/étnicas en la matrícula y progreso educativo. Por ejemplo, para el caso de Estados Unidos, Hauser (1993) y Buchmann, *et al.* (2009) analizan las brechas raciales/étnicas a lo largo de varias décadas. En el primer caso, se estudia la tasa de matrícula y culminación de la educación básica para los blancos, negros e hispanos luego de controlar por un conjunto más amplio de características como *proxys* de oferta (densidad poblacional) y demanda (recursos del hogar). En el segundo trabajo, se analiza la tasa de culminación de la educación superior luego de controlar por las características (edad, año del censo y categoría dicotómica “blanco/negro” y sus interacciones). Ambos utilizan regresiones logísticas para analizar la probabilidad de completar o acceder a un determinado nivel educativo, y encuentran brechas significativas entre grupos étnicas.

Para el caso de Canadá, Abada, *et al.* (2008) examinan las diferencias en el logro educativo entre hijos de inmigrantes según la nación de origen de los padres. Encuentran que los hijos de padres inmigrantes (en particular de China e India) logran mayores tasas de

⁴ Esto debido a que la pirámide social de estas sociedades se construyó sobre una jerarquización racial. Los esclavos africanos ocupaban el escalón más bajo, en el medio se encontraban los indígenas (los dominados) y en la cúspide los luso-hispanos (los dominantes). Véase Bello y Hopenhayn (2001).

⁵ Distinciones biológicas atribuidas a genotipos y fenotipos, especialmente con relación al color de la piel. Bello y Hopenhayn (2001, pág. 7)

culminación de la educación superior que los hijos de padres nacidos en Canadá, luego de controlar por las características individuales y familiares. Al igual que los estudios mencionados anteriormente, construyen un modelo probabilístico multivariado para modelar la tasa de culminación de la educación superior para cada grupo de origen y encuentran que el nivel educativo de los padres es el factor que explica el mayor porcentaje de la tasa de culminación, aunque dicho efecto es diferenciado según la nación de origen de los padres, lo cual indica diferencias culturales no observables que operan en el proceso de acumulación de capital humano.

Ñopo y Winder (2008), para el caso mexicano, encuentran que los individuos indígenas tienen en promedio 1.5 veces menos años de educación que los individuos no-indígenas. Mediante el uso de matrices de transición demuestran que la movilidad de la educación de los individuos indígenas es menor. Para explicar esto, se realizan regresiones separadas para el grupo indígena y el no-indígena, y encuentran que las características ocupacionales de los padres y la infraestructura en el hogar afectan por igual a la movilidad educativa, aunque las características individuales y familiares sí presentan efectos diferenciados.

Para el caso peruano, Maldonado y Rios (2006) reportan brechas educativas significativas según origen étnico. Así, la población caucásica o blanca tiene en promedio 5 años de educación más que la población quechua. Dichos autores, mediante el uso de matrices de movilidad educativa intergeneracional y del índice de movilidad de Andersen, encuentran que la variable más importante en la explicación en la brecha educativa es el número de años de estudio del jefe de hogar, la cual explica el 9.6% de la variación en el caso indígena y el 8.6% en el caso no indígena. Por otro lado, Benavides y Valdivia (2004) reportan mejoras en la matrícula primaria sin diferencias significativas por origen étnico⁶, así como una mejora en la permanencia en el sistema educativo durante las últimas décadas, aunque encuentran una brecha étnica que se manifiesta más claramente cuando se compara la proporción de la población que continua y llega a terminar la secundaria.

Una forma alternativa de entender la relación entre la etnicidad/raza y la acumulación de capital humano se logra observando las diferencias en cuanto a rendimiento educativo o logro de aprendizajes. Así, Patacchini y Zenou (2006) investigan la brecha étnica en las pruebas en estudiantes en Inglaterra y encuentran que existe una brecha étnica creciente en

⁶ Dichos autores plantean que, si bien en el pasado la problemática de la matrícula en primaria fue un factor importante a nivel global, para las generaciones más recientes, más del 90% de los niños terminan la primaria, aunque muchas veces lo hacen con un retraso significativo que afecta la probabilidad que continúen en secundaria.

el intervalo de 7 a 11 años, y decreciente entre los 11 y 16 años, a favor de la población blanca. Al analizar dicha brecha, encuentran que existe una parte que no es explicada por las características observables en la encuesta utilizada (estructura familiar, nivel socioeconómico y calidad de los colegios). Esta brecha es atribuida a las diferencias culturales en las actitudes hacia la educación entre las razas bajo análisis. Argumentan que una posible explicación es que los estudiantes negros, en escuelas donde predominan los estudiantes blancos, son objeto de burla, además de ser acusados por sus pares de “actuar como blancos”. Al analizar la evolución de la brecha, encuentran que el cambio en el interés de los padres a través del tiempo es el factor que modifica el progreso académico de sus hijos. Esto se debe, argumentan, a que los padres negros ofrecen insuficiente orientación, tiempo y atención al progreso educativo de sus hijos.

Nuestro trabajo se centra en el progreso educativo y, por lo mismo, busca medir el efecto que tiene la etnicidad (y el sexo) sobre la probabilidad de que un individuo se matricule en el nivel de instrucción básico y progrese hasta la educación superior. Al respecto, es frecuente encontrar una fuerte correlación entre la etnicidad y otras variables que determinan el logro educativo como pueden ser: la disponibilidad de recursos monetarios en el hogar mientras el niño/joven transita por el sistema educativo, el ambiente provisto por el hogar para la formación de habilidades básicas durante periodos críticos para este proceso (1-3 años), o la oferta de servicios educativos.

Grupos étnicos históricamente excluidos y que concentran altos índices de pobreza normalmente no cuentan con los recursos monetarios para hacer frente a los costos directos y de oportunidad de la educación, no han sido capaces de ofrecer un ambiente favorable para el desarrollo de habilidades en sus hijos, y se encuentran marginados (geográficamente) de la oferta (incluso pública) del servicio educativo. Al respecto, Torero, *et al.* (2002) encuentran que los individuos con altos niveles de intensidad de etnicidad blanca tienen bajos niveles de pobreza, altos índices de escolaridad y mayor acceso a servicios públicos. Asimismo, encuentran que dichos individuos tienden a permanecer por más tiempo en el sistema educativo.

Atendiendo a lo anterior, un análisis de la relación entre etnicidad y logro educativo en Perú debe necesariamente aspirar no sólo a documentar las brechas existentes sino también a evaluar hasta qué punto estas brechas tienen que ver, por ejemplo, con restricciones monetarias, el contexto familiar, restricciones por el lado de la oferta (tanto en cantidad como calidad), o mecanismos de discriminación presentes en el sistema educativo que

terminan afectando el proceso de aprendizaje de estudiantes de determinados grupos étnicos.

Para lograr lo anterior, uno de los principales retos a nivel empírico lo constituye la falta de información: varios de los elementos arriba indicados no son directamente observables y/o no se cuenta con información sobre los mismos en las encuestas multipropósito (como la Encuesta Nacional de Hogares – ENAHO) en la medida en que deben hacer referencia a la realidad del individuo mientras se encontraba estudiando.

Para este trabajo, hacemos uso de dos variables que intentan recoger la disponibilidad de recursos monetarios en el hogar en el momento en que la persona estaba (o debía estar) estudiando, así como potenciales restricciones por el lado de la oferta. Para lo primero empleamos el nivel educativo del padre. Para lo segundo usamos el ámbito (urbano o rural) asociado al lugar de nacimiento del individuo. Con esto, la idea es controlar por la presencia de restricciones materiales inmediatas (capacidad del hogar para hacer frente a los costos directos y de oportunidad de la educación y/o oferta suficiente del servicio educativo) que impiden que el niño o adolescente continúe acumulando años de educación. Esto implica que la variable que identifica la etnicidad (o el sexo) está recogiendo otros factores más estructurales asociados al contexto familiar y educativo de la persona (la presencia de un entorno favorable para el desarrollo de habilidades básicas durante su infancia, preferencias y percepciones familiares respecto al retorno a la educación, mecanismos de discriminación en el proceso educativo que incrementan los costos psíquicos de la educación y/o no favorecen la acumulación de habilidades de determinados grupos, etc.).

III. LOS DATOS Y HECHOS ESTILIZADOS

En la presente investigación, el progreso educativo de un individuo se mide como el número de años que este acumuló en el sistema educativo formal peruano: desde la educación básica (primaria, secundaria) hasta la educación superior (técnica o universitaria). Esta información se extrae de la Encuesta Nacional de Hogares (ENAHO) a nivel nacional. Así, se construyó un pool de información de dichas encuestas del 2004 al 2009 con el fin de incrementar el número de observaciones muestrales y lograr una mejor representación de las etnias minoritarias, en especial, de los grupos afrodescendientes y amazónicos los que, como veremos, representan una proporción bastante reducida de la población total.

Construcción de la base de datos

La unión de estas cinco encuestas es analizada como un gran corte transversal. De esta manera, las observaciones reflejan las condiciones socio-demográficas del Perú agregadas durante el sexenio 2004-2009. Se incluyó a individuos desde los 25 a los 60 años. Se asume que a la edad de 25 años el tránsito por el sistema educativo debe haber culminado para el grueso de la población, si se tiene en cuenta un ingreso a la edad de 5 años y una educación máxima típica de 16 años efectivos (asumiendo que la educación superior toma 5 años⁷).

Para medir las diferencias en el tiempo, se agrupan a los individuos en cuatro cohortes de edades: de 25 a 30 años, de 31 a 40 años, de 41 a 50 años y de 51 a 60 años. De este modo, se puede analizar en retrospectiva el progreso educativo en distintas épocas de la historia peruana contemporánea: por ejemplo, el análisis de los niveles educativos alcanzados por la cohorte más antigua nos aproxima al desempeño del sistema educativo peruano de hace 50 años; mientras que los resultados para la cohorte más joven se refieren a las características del sistema educativo de la década pasada.

Construcción de variables

Con el fin de homogeneizar la variable principal de estudio, se corrige el sesgo que podría tener la repitencia escolar o atraso en la formación superior sobre la medición de los años de educación. Es así que se imputan los años normativos a todo individuo que reportó haber culminado determinado nivel educativo⁸.

La ENAHO, en su módulo de Gobernabilidad, Democracia y Transparencia, recoge información sobre la etnia de origen (autoreportada) y el nivel de educación del padre del jefe de hogar. La primera variable constituye un insumo para nuestra variable de etnicidad. La segunda, tal como fue explicado en la sección anterior, constituye en un determinante fundamental de la demanda por educación, ya que un mayor nivel educativo del padre se relaciona con mayores recursos disponibles en el hogar.

Siguiendo a Castro y Yamada (2010), la definición de etnicidad utilizada en esta investigación tiene a la lengua materna como rasgo más exógeno y primario, y a la respuesta del auto-reporte como rasgo secundario. Se clasifica como “quechua/aymara” a los individuos cuya lengua materna es quechua/aymara, independientemente de la respuesta en

⁷ Los años de estudios de posgrado no se incluyen dentro del análisis.

⁸ La primaria completa, la secundaria completa, la superior no universitaria y la superior universitaria involucran 6, 5, 3 y 5 años, respectivamente.

el auto-reporte.⁹ Del mismo modo ocurre para aquellos con lenguas nativas amazónicas, en este caso se les clasifica como “amazónico”¹⁰. Sin embargo, para aquellos peruanos que tienen como lengua materna el castellano u otras lenguas, la respuesta de auto-reporte es crucial para la identificación.

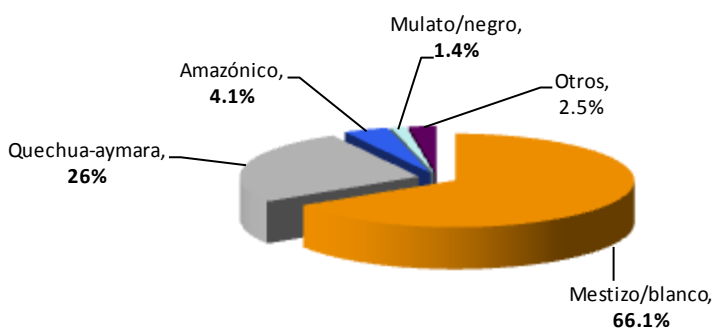
Cuadro 1
Identificación de etnicidad con lengua materna y auto-reporte

Segunda etapa: Auto-reporte	Primera etapa: Lengua materna			
	<i>Quechua/aymara</i>	<i>Castellano</i>	<i>Nativa</i>	<i>Otras</i>
<i>De la Amazonía</i>	Quechua	Amazónico	Amazónico	Amazónico
<i>Mestizo/blanco</i>	Quechua	Mestizo/blanco	Amazónico	Mestizo/blanco
<i>Negro/mulato/zambo</i>	Quechua	Negro	Amazónico	Negro
<i>Quechua/aymara</i>	Quechua	Quechua	Amazónico	Quechua

Fuente: Castro y Yamada (2010)

El Gráfico 1 presenta la distribución de la población considerada para este estudio utilizando esta definición operativa de etnicidad.

Gráfico 1
Distribución de la población peruana 2004-2009 por grupos étnicos (de acuerdo a definición combinada de lengua materna y auto-reporte)



Fuente: INEI. Encuesta Nacional de Hogares (ENAH)
Elaboración propia

Como se puede apreciar, nuestra clasificación ha permitido identificar, dentro del grupo de hispano hablantes, a mestizo/blancos (la gran mayoría, 66%), quechua/aymaras,

⁹ Si bien las lenguas quechua y aymara son distintas, la agrupación de dichas lenguas indígenas se debe a su origen e historia comunes asociada a la región andina.

¹⁰ La ENAHO no ofrece la posibilidad de distinguir entre lenguas amazónicas. Aunque que las lenguas amazónicas son heterogéneas, se agrupan por ser autóctonas de la región amazónica.

amazónicos y mulato/negros¹¹. Nótese, además, que en la medida en que la lengua materna ha primado para el caso de quechuas y amazónicos, la estrategia en dos etapas ha minimizado los potenciales sesgos causados por el hecho de que personas de estas etnias se identifiquen como mestizos.

Finalmente, el ámbito geográfico (urbano o rural) a nivel distrital se utiliza como un *proxy* de la cantidad y calidad de la oferta educativa, ya que en las zonas rurales existe mucho menos cobertura educativa relativa a las zonas urbanas. Asumimos que el distrito donde nació el individuo es aquel donde se educó, por lo tanto, interesa el ámbito del distrito de nacimiento. Debido a que la ENAHO no proporciona dicha información fue necesario imputar el ámbito geográfico al distrito de nacimiento de acuerdo con el nivel de ruralidad que actualmente presenta el distrito¹².

Algunos hechos estilizados

Sobre la base de los datos arriba descritos, este estudio busca examinar las diferencias étnicas y de sexo en el progreso educativo en el Perú a través de un estudio retrospectivo que abarque un periodo de tiempo amplio. En el siguiente cuadro se muestra el número de años de educación promedio alcanzado según el grupo étnico para las cuatro cohortes de edades consideradas. Asimismo, se muestran los promedios, según sexo, dentro de cada etnia.

En la antepenúltima fila del Cuadro 2 se observa el progreso a nivel nacional del logro educativo, entendido como el número de años de educación acumulados. Así, un peruano entre 51 y 60 años de edad tiene en promedio 6.6 años de educación (logró culminar la educación primaria), mientras que un peruano entre 25 y 30 años acumuló 10.1 años de educación (a punto de culminar la educación secundaria).

A pesar de este importante progreso educativo, los promedios nacionales esconden diferencias marcadas entre grupos étnicos. Aún para la cohorte más reciente (de 25 a 30 años), un individuo mestizo/blanco posee en promedio 2.5 años de educación adicionales comparado a un individuo perteneciente a la etnia amazónica. De hecho, todo el resto de

¹¹ Siguiendo a Yamada y Castro (2011), dentro del grupo hispanohablante que no se auto-reporta como quechua o amazónico, se puede identificar a dos grupos adicionales (blancos/mestizos y afrodescendientes), los cuales se distinguen claramente por su cultura y rasgos físicos. Debido a esto, surge la necesidad de separarlos y analizar si dichas diferencias tienen un efecto significativo en el progreso educativo.

¹² El nivel de ruralidad se calcula a nivel distrital como el porcentaje de hogares que pertenecen a las zonas rurales. Si el distrito tiene un nivel de ruralidad mayor al 50%, se considera como rural.

etnias no-mestizas presenta promedios menores de logro educativo que la etnia mestiza/blanca (que tiene 10.7 años de educación): las etnias quechua y afrodescendiente poseen 8.4 y 9.0 años de educación en promedio, respectivamente. Si comparamos estos promedios con la realidad del sistema educativo peruano de hace 50 años notaremos que esta brecha de alrededor de 2 años es una característica persistente.

Cuadro 2
Años de educación promedio según etnicidad, sexo y cohorte de edades

	25 a 30 años	31 a 40 años	41 a 50 años	51 a 60 años
Etnia Mestiza/Blanca	10.7	9.9	9.1	7.7
Hombre	10.8	10.2	9.6	8.6
Mujer	10.5	9.5	8.5	6.9
Etnia Quechua/Aymara	8.4	7.7	6.5	4.8
Hombre	9.4	8.8	8.0	6.6
Mujer	7.7	6.8	5.1	3.1
Etnia Amazónica	8.2	7.8	6.9	5.4
Hombre	9.0	8.6	7.8	6.5
Mujer	7.5	6.9	6.0	4.2
Etnia Afrodescendiente	9.0	8.2	7.2	5.5
Hombre	9.0	8.5	7.5	6.5
Mujer	9.0	7.8	7.0	4.6
Promedio nacional	10.1	9.1	8.1	6.6
Hombre	10.4	9.7	9.0	7.8
Mujer	9.7	8.6	7.3	5.4

Fuente: ENAHO 2004-2009.

Elaboración: Propia.

Con respecto a las diferencias según sexo, el promedio nacional indica que la brecha casi se ha cerrado a nivel nacional: 10.4 y 9.7 años de educación en la cohorte de 25 a 30 años para los hombres y mujeres, respectivamente. Este resultado contrasta con los 7.8 y 5.4 años en la cohorte de 51 a 60 años. Sin embargo, nuevamente, la brecha promedio según sexo y su evolución esconden grandes diferencias entre etnias. Si bien para la etnias mestiza/blanca y afrodescendiente la brecha de sexo prácticamente no existe (10.8 años en hombres frente a 10.5 años en mujeres), lo mismo no sucede para el resto de las etnias no-mestizas. En la etnia quechua/aymara y amazónica aún persisten diferencias según sexo en el logro educativo cercanas a los 2 años.

IV. METODOLOGIA: LA EDUCACIÓN COMO DURACIÓN

En lo que sigue, nuestro interés recae en evaluar la importancia que tienen la etnicidad y el sexo como factores determinantes de la permanencia en el sistema educativo. Como se discutió en la sección II, la idea es evaluar la capacidad explicativa de estas variables una vez que se ha controlado por la existencia de restricciones materiales inmediatas relacionadas con la disponibilidad de recursos en el hogar y una oferta suficiente del servicio educativo. De esta forma, la significancia y magnitud de los efectos relacionados con las variables de etnicidad y sexo serán informativas de la importancia que tienen otros factores relacionados al contexto familiar y educativo que no son directamente observables.

Más allá de la inclusión de controles que capturen características socioeconómicas (algo que es imprescindible hacer y han hecho todos los estudios reseñados en la sección II), nuestro principal aporte metodológico tiene que ver con el hecho de analizar los años de educación acumulados como una “duración”. Es decir, desde el punto de vista de una variable aleatoria cuya probabilidad de quedar debajo de determinado valor puede ser analizada en función de los años acumulados hasta el momento y otras características observables de la persona.

Nótese que si condicionamos la probabilidad de que los años de educación queden por debajo de determinado valor al tiempo transcurrido o años acumulados hasta el momento, estaremos modelando la probabilidad de abandonar el proceso educativo (o desertar) en ese instante. Como se verá más adelante, nuestro análisis se centrará en esto último y, en particular, en evaluar cómo cambia esta probabilidad en diferentes etapas del proceso educativo y para diferentes características del individuo (entre ellas, obviamente, la etnicidad y sexo).

La modelación del progreso educativo como una variable que exhibe dependencia temporal (la probabilidad de que se detenga depende del tiempo transcurrido hasta el momento) es la primera contribución del análisis. Los estudios reseñados en la sección II también emplean modelos probabilísticos pero no toman en cuenta esta dependencia temporal: por ejemplo, se recoge la culminación de determinado nivel a través de una variable dicotómica (1 si culminó el nivel, 0 de otro modo).

La segunda contribución tiene que ver con explotar esta dependencia temporal y permitir que la probabilidad de abandonar el proceso educativo sea diferente en diferentes etapas y, a su vez, que estas diferencias estén función de las características de la persona. Tal como

se discute a continuación, esto último no es posible en los modelos de duración más simples.

Modelos de duración

La metodología de duración consiste en modelar el espacio de tiempo en un determinado estado antes de la transición hacia otro. Para nuestro caso, la transición ocurre cuando el individuo pasa de pertenecer al sistema educativo a no pertenecer nunca más a dicho sistema¹³.

Siguiendo a Cameron y Trivedi (2005), la duración es una variable aleatoria positiva denotada por T . Si $F(t)$ es la distribución acumulada de T y $f(t) = \frac{\partial F(t)}{\partial t}$ es la función de densidad, entonces la probabilidad de que la duración sea menor o igual a t es:

$$F(t) = \Pr[T \leq t] = \int_0^t f(s) ds \dots [1]$$

El primer concepto fundamental en esta metodología es la *función de supervivencia*, que por complemento, es la probabilidad de exceder t :

$$S(t) = \Pr[T > t] = 1 - F(t) \dots [2]$$

Es decir, la *función de supervivencia* modela la probabilidad de que un individuo se mantenga un determinado número de años en el sistema educativo.

El segundo concepto importante se refiere a la *función de riesgo*, que es la probabilidad instantánea de que el individuo abandone el sistema educativo en un determinado momento del proceso, dado que sobrevivió hasta dicho momento. Formalmente:

¹³ Usualmente, esta metodología enfrenta el problema de censura ya que la duración es observada de manera parcial hasta, digamos, un valor “c”. Por ejemplo, en un estudio de desempleo, la duración observada no será equivalente a la realmente experimentada ya que para un individuo que está desempleado al momento de recolectar la información, se necesitará tiempo adicional para observar por completo la duración. Sin embargo, en la presente investigación, la censura no constituye un problema ya que los individuos concluyeron su tránsito por el sistema educativo, por lo tanto, el espacio de tiempo es igual al observado. Las preguntas correspondientes en la ENAHO indagan sobre el máximo nivel educativo alcanzado por el individuo y el número de años de educación obtenido en dicho nivel. Por tanto, estamos seguros de que se trata de salidas permanentes del sistema educativo.

$$\lambda_t = \lim_{h \rightarrow 0} \frac{\Pr(t \leq T < t+h | t \leq T, X; \beta)}{h} = \frac{f(t | X; \beta)}{1 - F(t | X; \beta)} = \frac{f(t | X; \beta)}{S(t | X; \beta)} \dots [3]$$

Finalmente, la *función de riesgo acumulada* se define como:

$$H(t) = \int_0^t \lambda(s) ds = -\ln S(t) \dots [4]$$

Esta última función se interpreta como el número esperado de deserciones que ocurren hasta un determinado momento del proceso. Aunque los resultados de esta función no serán reportados en la presente investigación, sí es importante para el desarrollo del modelo de riesgos no proporcionales que será discutido más adelante.

Ahora bien, el análisis de duración examina la relación entre la sobrevivencia, generalmente en la forma de la función de riesgo, y un conjunto de variables explicativas (regresores). Los métodos de estimación se pueden dividir en no paramétricos y paramétricos. En este estudio no utilizamos los primeros debido a que su naturaleza dificulta la incorporación de variables de control. En el segundo grupo se encuentran los métodos que consisten en asumir que la distribución del tiempo de sobrevivencia adopta una determinada función matemática. Entre ellas se encuentran la función de duración exponencial, de Weibull, de Gompertz, entre otras. El problema con estas distribuciones es que derivan en formas determinadas restrictivas para las funciones de sobrevivencia y riesgo. Por ejemplo, la función de duración exponencial deriva en una función de riesgo constante, lo cual puede originar problemas de especificación.

Buena parte de los estudios que emplean el análisis de duración utilizan modelos de riesgo proporcional debido a que pueden ser interpretados con facilidad. En estos, la función de riesgo condicional se puede descomponer de la siguiente forma:

$$\lambda(t | x) = \lambda_0(t) \phi(x, \beta) \dots [5]$$

Donde $\lambda_0(t)$ se denomina *riesgo base*, y es una función de t , mientras que $\phi(x, \beta)$ es una función de x (los regresores) y sus parámetros. Se denomina modelo de riesgo proporcional debido que la función de riesgo es proporcional al riesgo base a través de un factor de escala $\phi(x, \beta)$, que no es una función del tiempo.

A. El modelo de riesgos proporcionales de Cox

Cox (1972) propone una forma de evitar el problema de especificación que surge al utilizar modelos paramétricos. En particular, propone un modelo semi-paramétrico donde la descomposición de la función de riesgo le permite prescindir de la definición de una forma funcional para $\lambda_0(t)$, mientras que la forma de $\phi(x, \beta)$ puede quedar totalmente especificada. Formalmente:

$$\lambda_t = \lambda_0(t)\phi(x, \beta) \dots [6]$$

Es frecuente utilizar una función exponencial para $\phi(x, \beta)$ debido a que sus resultados pueden ser interpretados con facilidad, además de que asegura que $\phi(x, \beta) > 0$. Formalmente:

$$\lambda_t = \lambda_0(t)\exp(x, \beta) \dots [7]$$

Si trabajamos con esta función y suponemos que el regresor x_j se incrementa en una unidad manteniendo los demás constantes (para un nuevo vector de regresores x'), tenemos:

$$\begin{aligned}\lambda(t | x'; \beta) &= \lambda_0(t)\exp(x\beta + \beta_j) \\ &= \lambda_0(t)\exp(x\beta)\exp(\beta_j) \\ &= \lambda(t | x; \beta)\exp(\beta_j)\end{aligned}$$

De esta manera, el nuevo riesgo es $\exp(\beta_j)$ veces el riesgo original. Entonces, el aumento porcentual en el riesgo viene dado por: $(\exp(\beta_j) - 1) * 100\%$. Esto requiere la estimación de β , pero no de la función riesgo base $\lambda_0(t)$.

B. El modelo de riesgos no proporcionales

Asumir que los efectos de los regresores son proporcionales a un riesgo base puede ser un supuesto muy restrictivo y su invalidez significaría una especificación incorrecta de nuestro modelo. En términos de nuestra variable de interés y los controles propuestos, imponer esta restricción equivaldría a suponer, por ejemplo, que el efecto que tiene la etnicidad sobre la probabilidad de deserción es la misma en todas las etapas del proceso educativo.

Al respecto, nuestro trabajo apunta no sólo a medir la brecha étnica y de sexo, o el efecto promedio de cambios en los regresores sobre el riesgo de desertar. Nuestro análisis busca, además, observar la evolución de dicha brecha a lo largo del proceso educativo y explorar en qué etapas del tránsito por el sistema educativo el riesgo de deserción aumenta y resulta más crítico para el individuo, luego de controlar por ciertas características que aproximen las condiciones de demanda y oferta educativa.

Royston y Parmar (2001) proponen una versión paramétrica del modelo de Cox (1972) que supera los problemas mencionados líneas arriba. Esta consiste en modelar el logaritmo de la función de riesgo acumulada mediante el uso de las funciones *spline* cúbicas (véase Anexo 1). En su forma más sencilla y general, este modelo involucra una transformación de la función del riesgo acumulado. La función de transformación elegida es el logaritmo natural:

$$\ln H(t; x) = \ln H_0(t) + \beta' x = s(t) + \beta' x \dots [8]$$

La función $s(t)$ no está especificada. Los autores eligen representarla mediante una función *spline* cúbica:

$$s(t) = \gamma_0 + \gamma_1 \ln t + \gamma_2 v_1(\ln t) + \dots + \gamma_{m+1} v_m(\ln t) \dots [9]$$

La no proporcionalidad es inducida al permitir que los coeficientes $\gamma_1, \dots, \gamma_m$ dependan de los regresores. La desventaja de este modelo es que la interpretación de los coeficientes suele ser complicada. Por esto, se estiman las funciones de riesgo para los grupos de estudio y luego se construyen funciones de riesgo relativo para apreciar las diferencias resultantes en la probabilidad de abandonar el proceso educativo.

V. RESULTADOS Y SU INTERPRETACIÓN

En esta sección se muestran los resultados de la estimación de las brechas étnicas y de sexo y su evolución a lo largo del proceso educativo, utilizando los dos tipos de modelos descritos en la sección anterior.

Debido a que el modelo de riesgos proporcionales ofrece (a través del riesgo relativo) una aproximación sencilla a la magnitud de dichas brechas resulta ilustrativo comparar los resultados antes y después de incluir los controles relacionados con la disponibilidad de recursos en el hogar y oferta educativa. Esta comparación permitirá descomponer las brechas étnicas según un porcentaje atribuible a restricciones materiales inmediatas y otro atribuible al contexto familiar y educativo.

Posteriormente, verificaremos el supuesto de proporcionalidad de los regresores en el modelo de duración. Y finalmente, a la luz de lo anterior, desarrollaremos el modelo de riesgos no proporcionales para examinar la evolución de las brechas étnicas, resolviendo el problema de especificación impuesto por el supuesto de proporcionalidad.

Modelo de riesgos proporcionales

A continuación se muestra el resultado de las estimaciones con y sin controles que aproximan las características de oferta (con el ámbito geográfico donde nació el individuo) y restricciones monetarias del hogar (a través del nivel de educación del padre).

**Cuadro 3: Efectos proporcionales sobre la tasa de riesgo:
 $\exp(\beta_j)$ para cada variable j**

Variable Explicativa	(A) Modelo sin controles				(B) Modelo con controles			
	25-30	31-40	41-50	51-60	25-30	31-40	41-50	51-60
Quechua	1.59***	1.46***	1.45***	1.47***	1.14***	1.13***	1.08***	1.12***
Amazónico	1.69***	1.47***	1.46***	1.37***	1.42***	1.27***	1.19***	1.13*
Afrodescendiente	1.34***	1.34***	1.36***	1.46***	1.42***	1.29***	1.28***	1.35***
Sexo	1.05***	1.11***	1.15***	1.22***	1.01	1.07***	1.13***	1.24***
Rural	-	-	-	-	1.63***	1.68***	1.65***	1.53***
Primaria Padre	-	-	-	-	0.62***	0.58***	0.53***	0.50***
Secundaria Padre	-	-	-	-	0.42***	0.40***	0.35***	0.32***
Superior Padre	-	-	-	-	0.29***	0.29***	0.28***	0.25***

Nota: (***) significancia al 99%, (**) significancia al 95% y (*) significancia al 90%.

Los valores reportados en el Cuadro 3 indican cuántas veces mayor es el riesgo de deserción del individuo con la característica indicada por la variable explicativa, respecto a un individuo en la categoría base. Las categorías base son: mestizo/blanco, hombre, zona urbana y padre sin educación para las variables relacionadas con la etnicidad, el sexo, el lugar de nacimiento y el grado educativo del padre, respectivamente.

Si omitimos los controles propuestos, notaremos que la etnia amazónica es que la muestra el mayor riesgo relativo de deserción, seguida por la etnia quechua y la etnia afrodescendiente. Por ejemplo, para la cohorte de 25 a 30 años, se observa que la etnia amazónica presenta un riesgo de desertar a lo largo del proceso educativo 69% mayor al riesgo del grupo blanco/mestizo. Para los casos quechua y afrodescendiente los riesgos son 59% y 37% mayores, respectivamente.

El riesgo de deserción que enfrentan las etnias quechua y amazónica, relativo al grupo blanca/mestiza, aumenta a medida que analizamos las cohortes más jóvenes. Por ejemplo, en la cohorte de 51 a 60 años, la etnia quechua presentaba un 47% más de riesgo, mientras que la cohorte de 25 a 30 años presenta 59% más de riesgo de deserción. Por el contrario, para la etnia afrodescendiente, el riesgo relativo parece haber disminuido con el tiempo.

Por otro lado, llama la atención que la brecha de sexo prácticamente se ha cerrado: si bien las mujeres presentaban un riesgo de deserción 22% mayor que el de los hombres hace 50 años, en la década pasada este riesgo fue sólo 4% mayor (aunque todavía estadísticamente significativo).

Varios de estos resultados cambian si controlamos por la presencia de restricciones materiales inmediatas que impiden que determinados grupos progresen en el sistema educativo. Tal como se aprecia al comparar los paneles (A) y (B) del Cuadro 3:

- (i) La brecha atribuible a cada etnia en relación con el grupo blanco/mestizo disminuye en los casos quechua y amazónico. Esto significa que en el primer modelo sin controles, las variables referidas a las etnias de origen estaban también capturando los efectos de estas restricciones. Al disminuir las brechas de los grupos étnicos, se está verificando que la pertenencia a estos grupos poblacionales se relaciona negativamente con el nivel socioeconómico y la disponibilidad de oferta educativa.
- (ii) La etnia afrodescendiente destaca por su elevado riesgo relativo, comparable ahora al de la etnia amazónica. Luego de incluir los controles propuestos, la etnia quechua experimenta una caída importante en su riesgo relativo lo que revela que buena parte del riesgo de abandonar el proceso educativo tiene que ver con las restricciones materiales recogidas en los controles.

- (iii) Sigue llamando la atención el fuerte incremento en el riesgo relativo que experimenta la etnia amazónica si comparamos los resultados para las distintas cohortes.
- (iv) La educación del padre exhibe un efecto mitigador creciente sobre el riesgo de deserción.. Cuando el padre de un individuo logra alcanzar la educación primaria como máximo, el individuo presenta un riesgo 39% menor con respecto a un individuo cuyo padre no posee nivel alguno de educación. Los cálculos análogos para un padre que logró la educación secundaria y superior son de una disminución de 58% y 71% en los riesgos, respectivamente.
- (v) Aquellos individuos que nacieron en un ámbito geográfico rural presentan 63% más riesgo de deserción para la cohorte más reciente. Este riesgo relativo se mantiene a lo largo de las cohortes de individuos, lo cual es consistente con la problemática histórica de la cobertura de educación en las zonas rurales.

Es posible descomponer las brechas totales halladas en el primer modelo (panel A) en una parte atribuible al contexto familiar y educativo (panel B) y otra atribuible a la presencia de restricciones materiales inmediatas (diferencia entre los resultados de ambos modelos)¹⁴. En el Gráfico 2 se muestran los resultados de esta descomposición para la cohorte más reciente (panel A) y la más antigua (panel B).

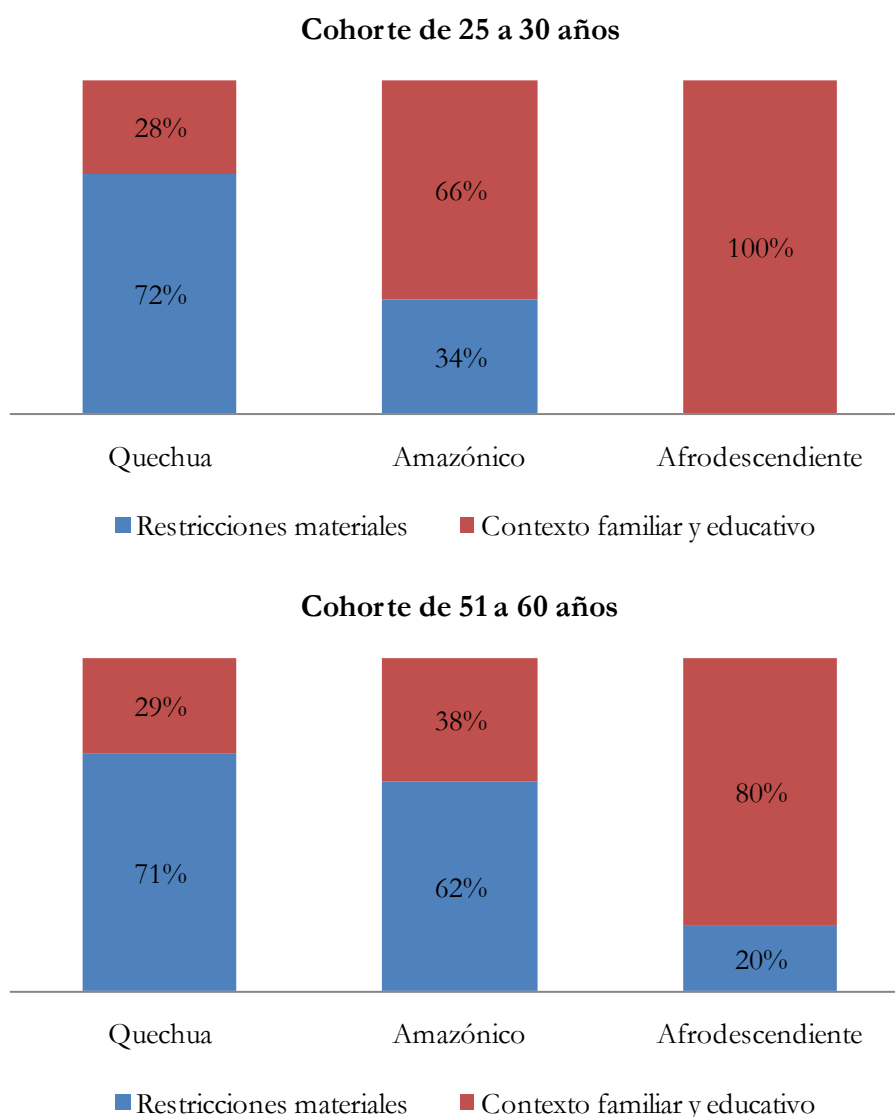
El Gráfico 2 revela lo siguiente:

- (i) La brecha quechua/aymara – mestizo/blanco en cuanto a progreso educativo se debe, principalmente, a que los primeros enfrentan restricciones de índole monetario y/o una oferta insuficiente del servicio educativo. Tanto hace 50 años como en la década pasada, estas restricciones materiales representan más del 70% de la brecha.
- (ii) Los factores relacionados con el contexto familiar y educativo que no son directamente observables han cobrado importancia en el tiempo para explicar las

¹⁴ Esta descomposición parte de la siguiente relación: $BT = (RM)(CFE)$. Donde BT hace referencia a la brecha total (reportada en el Panel A del Cuadro 3), RM se refiere a la brecha atribuible a restricciones materiales, y CFE a aquella atribuible al contexto familiar y educativo (reportada en el Panel B del Cuadro 3). Con esto, es posible hallar RM como $RM = (BT) / (CFE)$ y calcular la contribución porcentual de cada elemento como: $\%RM = \ln(RM) / \ln(BT)$ y $\%CFE = \ln(CFE) / \ln(BT)$.

brechas amazónica y afrodescendiente. Esto es particularmente significativo para la etnia afrodescendiente. Podemos decir que la totalidad de la brecha de este grupo es ahora atribuible a estos factores.

Gráfico 2: Riesgo: Descomposición de la brecha según etnia de origen



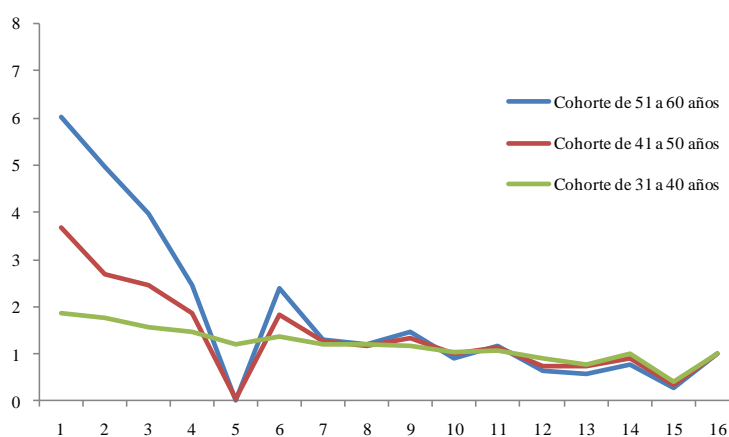
Si analizamos las diferencias entre cohortes en el panel (B) del Cuadro 2 notaremos que: (i) el riesgo relativo se incrementado para todas las etnias; y (ii) el riesgo relativo asociado al sexo femenino se ha reducido al punto de dejar de ser significativo para la cohorte más reciente. Al momento de interpretar el primer resultado es necesario tener en cuenta que se trata de un riesgo relativo y que, por lo mismo, su incremento no significa necesariamente

que los grupos minoritarios se encuentren peor que antes en términos absolutos con respecto a su progreso en el sistema educativo.

De hecho, si aproximamos el riesgo de deserción a través de la proporción de individuos que abandona el sistema educativo, notaremos que éste ha caído para todas las etnias. En el Gráfico 3 se reporta, para cada cohorte, la proporción de individuos que abandona el sistema educativo en cada año medida respecto a la proporción correspondiente de la cohorte más reciente¹⁵. Tal como se puede observar, la proporción de individuos que abandona el sistema educativo durante los primeros 5 años ha disminuido a través del tiempo para todas las etnias: el estadístico es mayor para las cohortes más antiguas en los cuatro casos. Sin embargo, el hecho de que el riesgo de los grupos étnicos relativo al grupo blanco/mestizo haya aumentado para las cohortes más recientes, indica que el ritmo de este progreso ha sido bastante desigual y ha favorecido al grupo blanco/mestizo. Tal como se observa en el Gráfico 3, las diferencias entre cohortes son mucho más marcadas para este grupo indicando que ha mejorado bastante más que el resto.

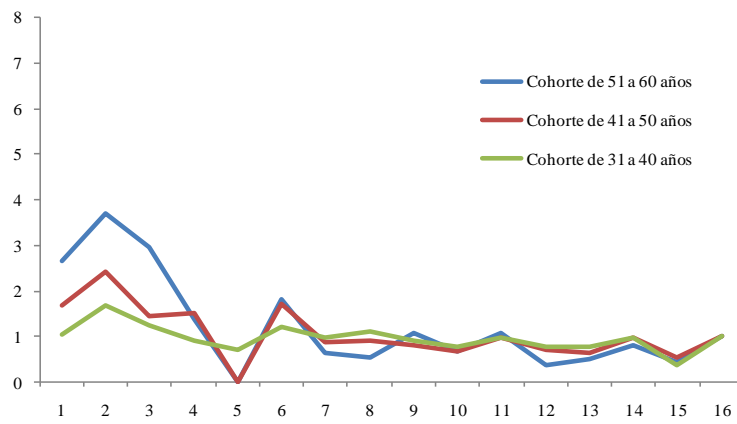
Gráfico 3: Proporción de individuos que abandona el sistema educativo según etnia de origen (relativo a la cohorte de 25 a 30 años)

(A) Etnia Blanca/Mestiza

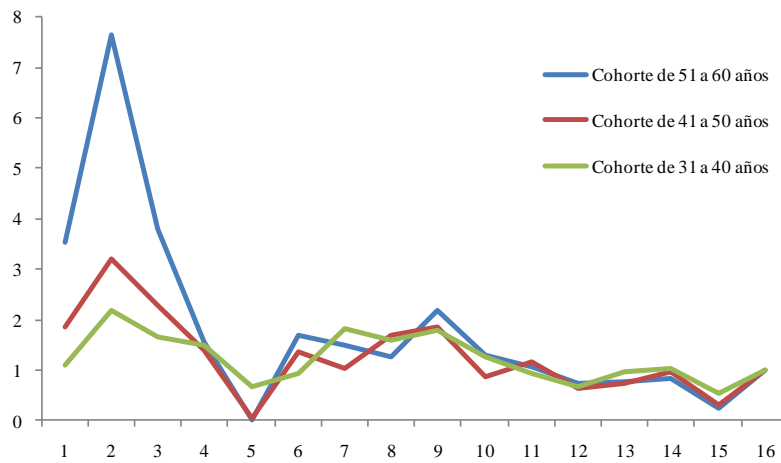


¹⁵ Este análisis tiene sólo un carácter descriptivo y no cuenta con el rigor estadístico ni los controles empleados para los riesgos relativos reportados en el Cuadro 3.

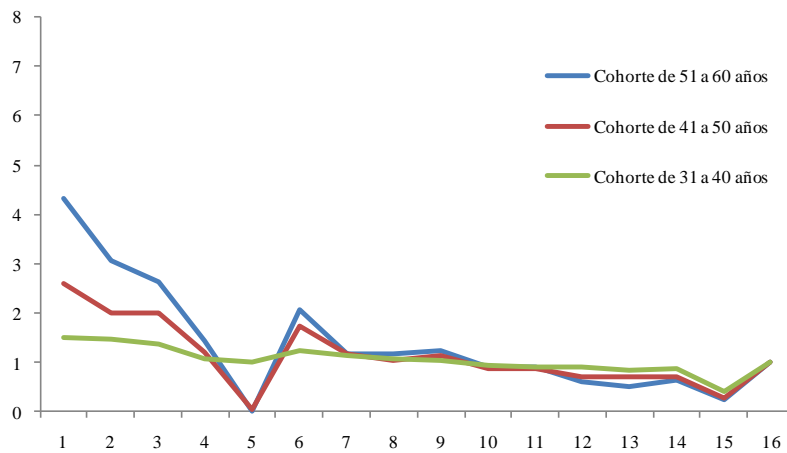
(B) *Etnia Amazónica*



(C) *Etnia Afrodescendiente*



(D) *Etnia Quechua*



En lo que respecta a la disminución y desaparición del riesgo relativo asociado al sexo femenino, cabe preguntarse qué tan cierto es este promedio nacional dentro de cada grupo étnico. Es decir, cabe preguntarse si es que aún subsisten brechas de sexo atribuibles al contexto familiar y educativo si analizamos cada grupo étnico por separado.

Tal como se muestra en el Cuadro 3, el promedio nacional estaba ocultando realidades distintas en cada grupo étnico. Luego de controlar por la presencia de restricciones materiales inmediatas, las mujeres quechuas y amazónicas aún reportan un riesgo de deserción casi 30% mayor que los hombres de sus respectivos grupos étnicos.

Cuadro 3: Efecto proporcional sobre la tasa de riesgo de la variable sexo

Variable explicativa	Cohorte 25-30	Cohorte 31-40	Cohorte 41-50	Cohorte 51-60
Mujer mestiza/blanca	0.94***	0.99	1.05***	1.17***
Mujer quechua/aymara	1.29***	1.28***	1.33***	1.40***
Mujer amazónica	1.28***	1.13*	1.14	1.33*
Mujer afrodescendiente	0.93	0.94	0.88	1.31

Notas: (***) significancia al 99%, (**) significancia al 95% y (*) significancia al 90%. Regresiones controladas por ámbito geográfico y educación del padre.

Modelo de riesgos no proporcionales

Tal como fue explicado anteriormente, el modelo desarrollado hasta ahora impone la restricción de que los regresores tienen un efecto proporcional sobre un riesgo base. Si evaluamos esta restricción, notaremos que la hipótesis de proporcionalidad se rechaza de manera global para todos los modelos y de manera individual para la mayoría de regresores (ver Anexo 2). Este resultado nos indica que el efecto que tiene la etnicidad sobre la probabilidad de deserción es distinto según sea la etapa del proceso educativo considerada por lo que cabe estimar un modelo más flexible que permita la interacción de los regresores con la función de riesgo base.

El modelo de riesgos no proporcionales explicado líneas arriba permite la estimación de los coeficientes asociados a cada regresor en cada uno de los elementos en los que se descompone el logaritmo de la función de riesgo acumulado (véase ecuación 9). Los resultados de estas estimaciones se muestran en el Anexo 3.

Para facilitar la interpretación y comparación con los resultados mostrados en el acápite anterior se utilizaron los coeficientes estimados para calcular las funciones de riesgo para cada categoría bajo estudio (las cuatro etnias y los dos sexos) manteniendo el resto de variables en la media de su distribución. Luego, se construyeron los riesgos relativos tomando al grupo mestizo/blanco como categoría de referencia para las etnias y al hombre como categoría de referencia para el sexo. Los resultados se muestran en los gráficos siguientes. En cada caso y para la cohorte más reciente¹⁶, se comparan: el riesgo relativo no proporcional, el riesgo relativo proporcional (estimado con el modelo discutido en el acápite anterior), y el valor unitario asociado a la igualdad de riesgo de deserción.

Los gráficos muestran que la evolución del riesgo relativo de deserción a lo largo del proceso educativo es marcadamente distinta entre las tres etnias consideradas. Para el caso quechua/aymara, se observa que el riesgo relativo registra su valor más elevado al inicio del proceso, decae durante la primera mitad de la primaria y comienza a crecer nuevamente para registrar otro pico hacia el final de la secundaria e inicio de la educación superior (años 11 y 12). En el caso de la etnia amazónica, el riesgo relativo también registra valores comparativamente altos durante el tránsito por la secundaria pero registra su valor más bajo al inicio del proceso educativo, lo que le da una forma de U invertida a la función de riesgo. Al igual que para el grupo quechua, la etnia afrodescendiente comienza el proceso educativo con un riesgo relativo elevado pero, a diferencia de las otras dos etnias, luego exhibe un comportamiento decreciente que se mantiene hasta el final del proceso educativo.

Este análisis nos ha permitido mostrar que, luego de aislar el efecto de restricciones relacionadas con la disponibilidad de recursos en el hogar y la oferta educativa, las dificultades que tienen los grupos minoritarios para progresar en el sistema educativo a un ritmo similar al del grupo blanco/mestizo se dan en diferentes etapas del proceso educativo. Los quechuas y afrodescendientes tienen mayores dificultades al inicio de la educación primaria. Para la etnia amazónica, en cambio, estas dificultades se concentran en el tránsito por la educación secundaria.

El modelo de riesgos no proporcionales no sólo nos ha permitido capturar esta heterogeneidad en el riesgo de deserción. También nos ha permitido reevaluar la magnitud

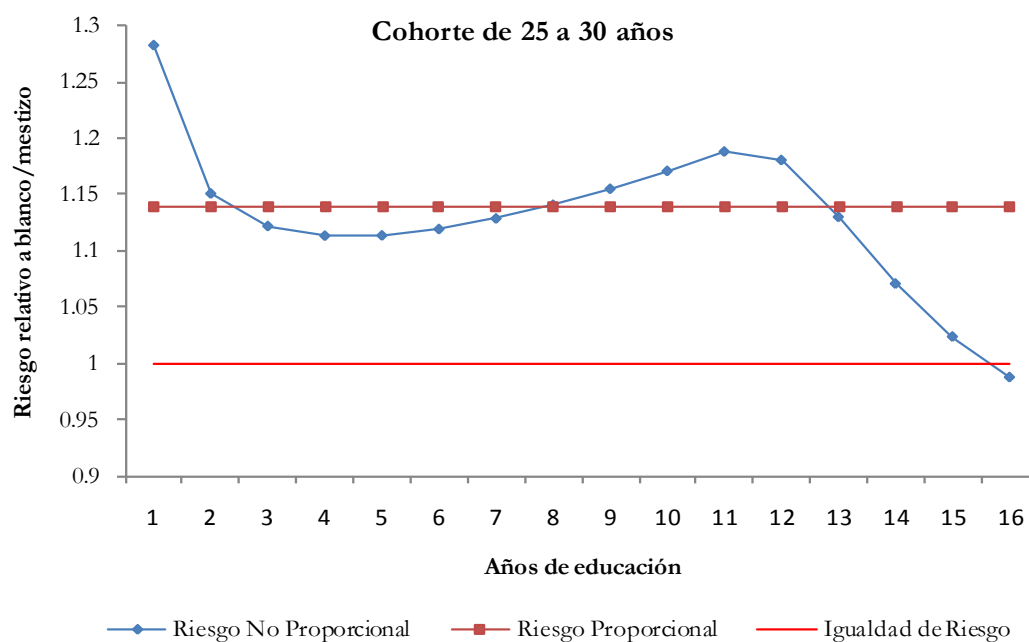
¹⁶ Los resultados para la cohorte más antigua se muestran en el Anexo 4. Los principales resultados respecto a la evolución en el tiempo de las tasas de riesgo relativo se mantienen.

de estos riesgos respecto a los resultados discutidos en el acápite anterior. Si tomamos como referencia el valor más alto de cada función, confirmamos que la etnia quechua es la que menos riesgo relativo presenta (un máximo de 26% más que el grupo blanco/mestizo al iniciar la educación primaria). A diferencia de los resultados obtenidos con el modelo de riesgo proporcional, sin embargo, ahora observamos una diferencia importante entre los riesgos de las etnias amazónica y afrodescendiente. Los primeros tienen un riesgo que es, como máximo, 50% mayor que el del grupo de referencia (cercano al riesgo relativo proporcional discutido en el acápite anterior). El grupo afrodescendiente, en cambio, exhibe más del doble de probabilidad de desertar que el grupo blanco/mestizo al inicio de la primaria. Esto sitúa al grupo afrodescendiente en clara desventaja respecto al resto de etnias. El riesgo proporcional estimado para este grupo en el acápite anterior (42% mayor que el grupo blanco/mestizo) subestimaba la brecha que existe entre esta etnia y el grupo blanco/mestizo durante toda la educación básica y, en especial, durante los primeros años de la instrucción primaria.

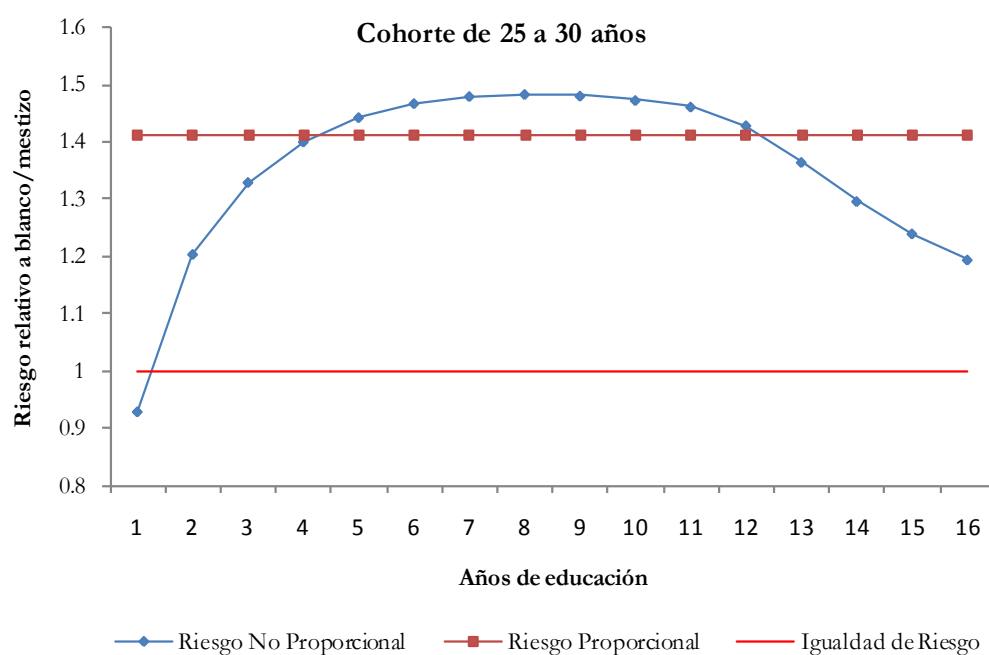
Con respecto a las diferencias según sexo, si bien el modelo de riesgos proporcionales muestra que no existe diferencias en la cohorte más joven, el Gráfico 5 revela que el sexo femenino exhibe un riesgo significativamente mayor que el sexo masculino durante toda la educación primaria (primeros 6 años). Al igual que en el caso del grupo afrodescendiente, el supuesto de proporcionalidad utilizado para evaluar el riesgo relativo en el modelo anterior nos ofrecía un resultado distorsionado sobre las dificultades que tienen las mujeres para avanzar a través de la educación básica a un ritmo similar que los hombres.

Gráfico 4: Riesgo relativo al grupo mestizo/blanco

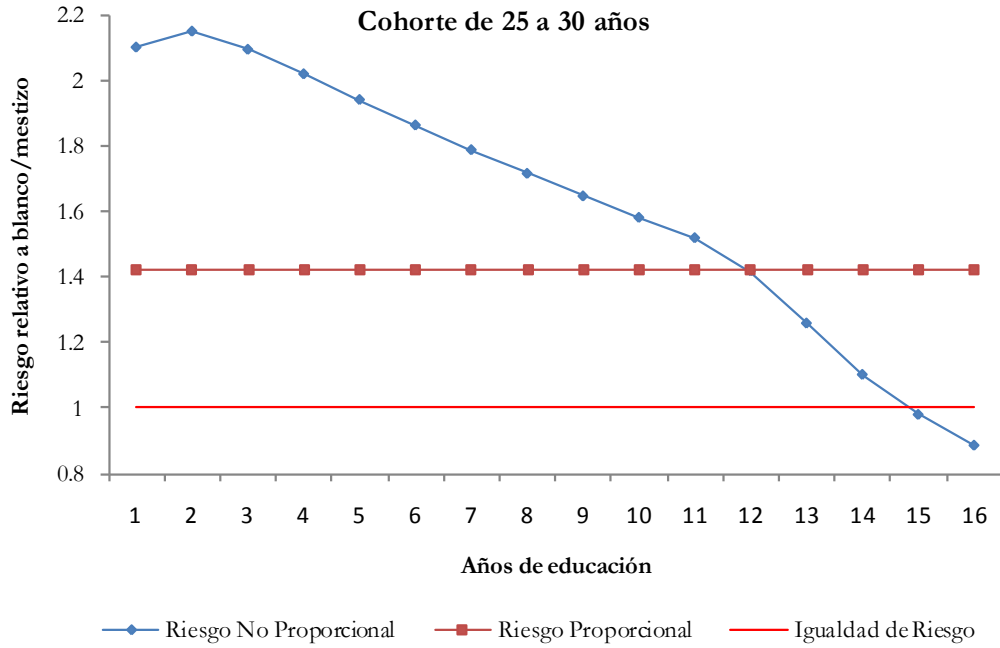
(A) Etnia Quechua



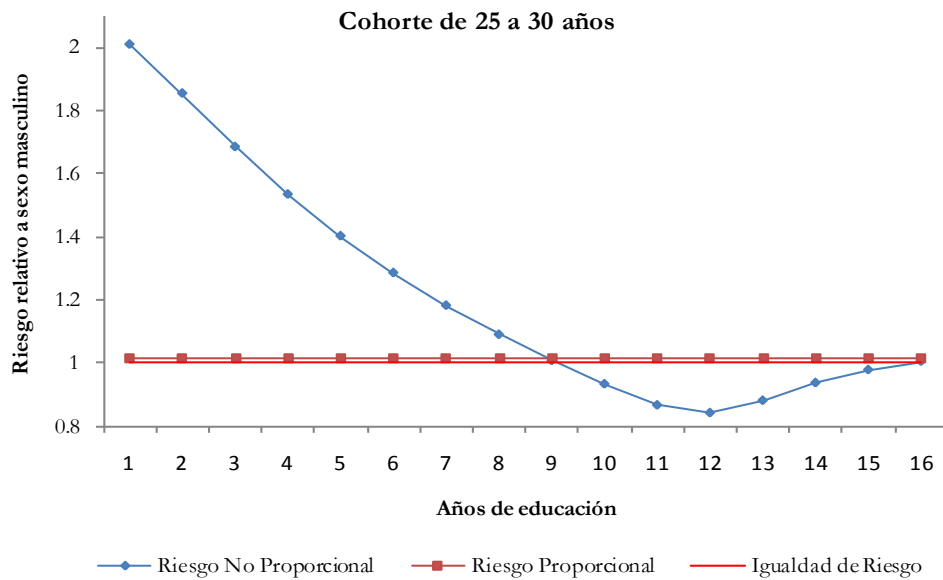
(B) Etnia Amazónica



(C) Etnia Afrodescendiente



**Gráfico 5: Riesgo estimado del sexo femenino
(Con respecto a sexo masculino)**



VI. CONCLUSIONES E IMPLICANCIAS DE POLITICA

En esta investigación se documentaron y analizaron las brechas étnicas y de sexo en el progreso educativo en los últimos 50 años en el Perú. Para esto, modelamos el progreso educativo (los años de educación acumulados) como una *duración* por lo que nuestra medida de “brecha” vino dada por el riesgo de abandonar el sistema educativo medido respecto al del grupo blanco/mestizo (para las comparaciones de etnia) y respecto al del hombre (para las comparaciones de sexo).

Encontramos que el riesgo de deserción de los grupos quechua, amazónico y afrodescendiente es mayor que el del grupo de blancos/mestizos. Luego de controlar por factores relacionados con la disponibilidad de recursos monetarios en el hogar y una oferta educativa suficiente, encontramos que estas restricciones materiales explican buena parte de la brecha entre quechuas y blancos/mestizos. Al otro lado del espectro, la brecha afrodescendiente tiene poco que ver con este tipo de restricciones y estaría más influida por factores estructurales relacionados con el contexto familiar y educativo de la persona y que no son directamente observables (la presencia de un entorno favorable para el desarrollo de habilidades básicas durante la infancia, preferencias y percepciones familiares respecto al retorno a la educación, mecanismos de discriminación en el proceso educativo, etc.).

Una comparación entre cohortes reveló que las brechas étnicas se han acentuado con el tiempo sin que esto implique mayores riesgos de deserción en términos absolutos. El progreso educativo ha mejorado en todas las etnias, sólo que ha mejorado significativamente más para el grupo blanco/mestizo. Si bien los promedios parecen sugerir que la brecha de sexo ha desaparecido en la última década, todavía subsisten diferencias importantes entre hombres y mujeres (a favor de los primeros) al interior de las etnias quechua y amazónica.

También empleamos una extensión al enfoque clásico de los modelos de duración que nos permitió analizar los efectos de la etnicidad y el sexo sobre la probabilidad de deserción en cada etapa del proceso educativo. Este enfoque de *riesgos no proporcionales* fue fundamental para encontrar que la etnia afrodescendiente se encuentra en clara desventaja respecto al resto de etnias con un elevado riesgo de deserción concentrado en los primeros años de la educación primaria, y atribuible a factores que van más allá de las restricciones monetarias y de oferta educativa. Asimismo, nos permitió encontrar que detrás de una brecha hombre-mujer prácticamente cerrada (en promedio), subsiste aún una diferencia importante entre

los riesgos de deserción (a favor de los hombres) concentrada en los primeros años de la educación básica.

Desde el punto de vista de política, nuestros resultados sugieren que el aumento en la cobertura educativa todavía no es una empresa acabada en el Perú. Si queremos reducir la brecha de logro educativo entre quechuas, amazónicos y blancos/mestizos, hay todavía un rol importante para las barreras materiales relacionadas con la disponibilidad de recursos en el hogar y una oferta educativa suficiente. Al respecto, las escuelas rurales unidocentes y multigrado son una prueba palpable y muy lamentable de la inequidad todavía presente en el sistema educativo de nuestro país. Nuestros resultados también sugieren que habrían importantes ganancias en eficiencia y equidad si las políticas específicas para aumentar la cobertura de estas poblaciones de origen indígena se enfocan en las niñas, las que exhiben riesgos de deserción significativamente mayores que los hombres en estos grupos étnicos.

Por el lado de las preguntas para futuras investigaciones, nuestros resultados sugieren que es necesario profundizar sobre los elementos del contexto familiar y educativo que están detrás de las brechas étnicas que subsisten luego de controlar por barreras materiales observables. Esto es particularmente relevante para entender porqué la etnia afrodescendiente se encuentra en clara desventaja respecto a la blanca/mestiza en su tránsito por los primeros años de la educación primaria.

REFERENCIAS

Abada, Teresa; Feng Hou y Bali Ram (2008). "Group Differences in Educational Attainment Among the Children of Immigrants". Analytical Studies Branch Research Paper Series N° 308

Anderson, Deborah y David Shapiro (1996). "Racial Differences in Access to High-Paying Jobs and the Wage Gap between Black and White Women". *Industrial and Labor Relations Review*, Vol. 49, No. 2

Benavides, Martín y Martín Valdivia (2004). "Metas del Milenio y la brecha étnica en el Perú". GRADE.

Bello, Álvaro y Martín Hopenhayn (2001). "Discriminación étnico-racial y xenofobia en América Latina y el Caribe". CEPAL, Políticas Sociales Serie N° 47.

Buchmann, Claudia; Anne McDaniel, Thomas DiPrete y Uri Shwed (2009). "The Black Gender Gap in Educational Attainment: Historical Trends and Racial Comparisons".

Cameron, Colin y Pravin Trivedi (2005). "Microeconometrics: Methods and Applications". New York: Cambridge University Press.

Castro, J. F. and G. Yamada (2010). "Las diferencias étnicas y de género en el acceso a la educación básica y superior en el Perú", mimeo, Universidad del Pacífico.

Cox, D. R. 1972. "Regression Models and Life Tables" *Journal of the Royal Statistical Society, Series B* 34:187-220.

Hauser, Robert (1993). "Trends in College Entry Among Whites, Blacks, and Hispanics." Pág. 61-104 en *Studies of Supply and Demand in Higher Education* (ed. Charles T. Clotfelter y Michael Rothschild). Chicago: University of Chicago Press.

Heckman, Lochner y Todd (2006). "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond". *Handbook of the Economics of Education*. Vol. 1. Chapter 7. Edited by Eric Hanushek and Finis Welch. Elsevier.

Maldonado, Stanislao y Vanessa Rios (2006). "Más allá de la igualdad de oportunidades: Desigualdad de ingresos, responsabilidad individual y movilidad social en el Perú". Consorcio de Investigación Económica y Social (CIES)

Montero, Ricardo y Gustavo Yamada (2011). "Raza, corrupción y acceso a servicios públicos en el Perú: ¿exclusión o discriminación?" Centro de Investigación de la Universidad del Pacífico. Junio 2011.

Ñopo, Hugo y Natalia Winder (2008). "Ethnicity and Human Capital Accumulation in Urban Mexico". Working Paper N° 660, Banco Interamericano de Desarrollo (BID).

Patachini, Eleonora y Yves Zenou (2006). "The Dynamics of the Racial Test Score Gap During the School Years in Britain". Research Institute of Industrial Economics

Roemer, John (1998) *Equality of Opportunity*. Harvard University Press. Cambridge.

Royston, Patrick (2001). “Flexible parametric alternatives to the Cox model, and more”. *The Stata Journal* 1, Number 1, pp. 1-28

Royston, Patrick (2004). “Flexible parametric alternatives to the Cox model: update”. *The Stata Journal* 4, Number 1, pp. 98-101.

Torero, Máximo; Jaime Saavedra, Hugo Ñopo y Javier Escobal (2002). “The Economics of Social Exclusion in Peru: An Invisible Wall?”. Grupo de Análisis para el Desarrollo

Yamada, Gustavo, Adriana Lizaaraburu y Katia Samanamud (2011). “Diferencias étnicas en el mercado laboral peruano: Un estudio comparativo de brechas de ingresos” Centro de Investigación de la Universidad del Pacífico, Junio 2011.

Anexo 1: Interpolación spline

Las técnicas de interpolación polinómica se utilizan para aproximar un conjunto de datos a una función. La interpolación spline pertenece a este tipo de técnicas, y es una de las más usadas para lograr el objetivo de construir una función con grandes oscilaciones. El objetivo de este tipo de funciones es construir espacios de funciones suavizadas y manejables. Así, la función spline está conformada por varios polinomios, cada uno definido sobre un sub-intervalo, que se unen entre sí obedeciendo a ciertas condiciones de continuidad. Estos polinomios pueden tomar cualquier grado; sin embargo, desarrollamos el grado cúbico ya que la metodología de riesgos no proporcionales propuesta por Royston y Parmar (2001) utiliza dicho grado.

Formalmente, se definen $n+1$ puntos o nodos: $t_0 < t_1 < \dots < t_n$. Sobre cada intervalo $[t_0, t_1], [t_1, t_2], \dots, [t_{n-1}, t_n]$, S está definido por un polinomio cúbico distinto. Luego la función spline cúbica es:

$$S(x) = \begin{cases} S_0(x) = S_0, & x \in [t_0, t_1) \\ S_1(x) = S_1, & x \in [t_1, t_2) \\ \cdot \\ \cdot \\ \cdot \\ S_{n-1}(x) = S_{n-1}, & x \in [t_{n-1}, t_n) \end{cases}$$

Los polinomios S_{i-1} y S_i interpolan el mismo valor en el punto t_i , es decir:

$$S_{i-1}(t_i) = y_i = S_i(t_i) \dots (1 \leq i \leq n-1)$$

Esta condición garantiza la continuidad de la función. Sin embargo, la solución al problema de interpolación necesita de dos condiciones adicionales: se cuenta con un total de $4n-2$ condiciones para determinar $4n$ coeficientes. En la metodología de Royston y Parmar (2001) se elige la interpolación spline cúbica natural, la cual consiste en asumir lo siguiente:

$$S''(t_0) = 0, S''(t_n) = 0$$

De esta manera, el sistema de ecuaciones puede ser resuelto con métodos numéricos.

Anexo 2: Prueba de Hipótesis de la proporcionalidad de cada variable j en el modelo de riesgos proporcionales (P-value)

Variable explicativa	Cohorte 25-30	Cohorte 31-40	Cohorte 41-50	Cohorte 51-60
Quechua	0.0165	0.0000	0.0003	0.1028
Amazónico	0.0098	0.4639	0.3246	0.8371
Mulato/negro	0.0005	0.0024	0.1724	0.0544
Sexo	0.0000	0.0000	0.0000	0.0000
Primaria padre	0.0000	0.0000	0.0000	0.0000
Secundaria padre	0.0000	0.0000	0.0000	0.0000
Superior padre	0.0000	0.0000	0.0000	0.0000
Global	0.0000	0.0000	0.0000	0.0000

Los resultados arrojan que existe evidencia estadística significativa de no proporcionalidad en el riesgo a nivel global ($p=0.000$) para todas las cohortes, y no proporcionalidad específica para la mayoría de variables.

Anexo 3: Resultados de las estimaciones del modelo de riesgo no proporcional

3.1: Resultados de la regresión para la cohorte de 25 a 30 años

Variable	Coefficiente	Std. Err.	z	P>z
s0				
Quechua (raza=quechua)	-0.064	0.037	-1.71	0.088
Amazónico (raza=amazónico)	0.138	0.086	1.60	0.110
Afrodesc. (raza=afrodescendiente)	-0.282	0.119	-2.37	0.018
Área Rural	-0.729	0.041	-17.75	0.000
Educación primaria	0.369	0.038	9.63	0.000
Educación secundaria	0.995	0.072	13.92	0.000
Educación superior	1.449	0.169	8.57	0.000
Sexo (mujer=1)	-0.422	0.034	-12.45	0.000
Constante	3.033	0.052	58.49	0.000
s1				
Quechua (raza=quechua)	-0.014	0.013	-1.09	0.275
Amazónico (raza=amazónico)	0.069	0.033	2.10	0.036
Afrodesc. (raza=afrodescendiente)	0.091	0.042	2.15	0.032
Área Rural	0.226	0.013	17.51	0.000
Educación primaria	-0.087	0.014	-6.17	0.000
Educación secundaria	-0.280	0.021	-13.17	0.000
Educación superior	-0.466	0.045	-10.33	0.000
Sexo (mujer=1)	0.036	0.011	3.14	0.002
Constante	-0.276	0.017	-15.99	0.000
s2				
Quechua (raza=quechua)	0.026	0.008	3.21	0.001
Amazónico (raza=amazónico)	0.018	0.020	0.90	0.367
Afrodesc. (raza=afrodescendiente)	0.040	0.025	1.58	0.114
Área Rural	-0.034	0.007	-4.99	0.000
Educación primaria	-0.010	0.010	-1.01	0.314
Educación secundaria	-0.004	0.012	-0.33	0.742
Educación superior	-0.170	0.016	-10.46	0.000
Sexo (mujer=1)	-0.046	0.006	-7.89	0.000
Constante	0.005	0.011	0.47	0.636
Xb				
Quechua (raza=quechua)	0.283	0.094	3.02	0.002
Amazónico (raza=amazónico)	0.008	0.211	0.04	0.971
Afrodesc. (raza=afrodescendiente)	1.125	0.296	3.80	0.000
Área Rural	2.425	0.104	23.31	0.000
Educación primaria	-1.309	0.093	-14.03	0.000
Educación secundaria	-3.498	0.181	-19.32	0.000
Educación superior	-5.645	0.433	-13.04	0.000
Sexo (mujer=1)	1.078	0.086	12.59	0.000
Constante	-7.270	0.129	-56.16	0.000
Número de observaciones	28,636			
Prob > chi2	0.000			

3.2: Resultados de la regresión para la cohorte de 31 a 40 años

Variable	Coefficiente	Std. Err.	z	P>z
s0				
Quechua (raza=quechua)	-0.100	0.027	-3.75	0.000
Amazónico (raza=amazónico)	0.213	0.074	2.88	0.004
Afrodesc. (raza=afrodescendiente)	-0.097	0.095	-1.03	0.304
Área Rural	-0.486	0.031	-15.71	0.000
Educación primaria	0.385	0.027	14.06	0.000
Educación secundaria	0.777	0.069	11.31	0.000
Educación superior	1.591	0.303	5.25	0.000
Sexo (mujer=1)	-0.356	0.026	-13.96	0.000
Constante	2.575	0.038	68.21	0.000
s1				
Quechua (raza=quechua)	-0.001	0.011	-0.07	0.945
Amazónico (raza=amazónico)	0.048	0.033	1.46	0.145
Afrodesc. (raza=afrodescendiente)	0.021	0.041	0.51	0.609
Área Rural	0.183	0.013	14.51	0.000
Educación primaria	-0.047	0.012	-3.95	0.000
Educación secundaria	-0.306	0.024	-12.67	0.000
Educación superior	-0.338	0.123	-2.74	0.006
Sexo (mujer=1)	0.034	0.011	3.06	0.002
Constante	-0.202	0.016	-12.91	0.000
s2				
Quechua (raza=quechua)	0.025	0.007	3.43	0.001
Amazónico (raza=amazónico)	0.037	0.020	1.86	0.063
Afrodesc. (raza=afrodescendiente)	0.030	0.028	1.10	0.273
Área Rural	0.014	0.007	1.95	0.051
Educación primaria	-0.019	0.008	-2.24	0.025
Educación secundaria	-0.013	0.014	-0.97	0.334
Educación superior	-0.293	0.034	-8.65	0.000
Sexo (mujer=1)	-0.032	0.006	-4.96	0.000
Constante	-0.098	0.010	-9.99	0.000
Xb				
Quechua (raza=quechua)	0.373	0.065	5.71	0.000
Amazónico (raza=amazónico)	-0.257	0.179	-1.44	0.151
Afrodesc. (raza=afrodescendiente)	0.501	0.231	2.17	0.030
Área Rural	1.837	0.077	23.99	0.000
Educación primaria	-1.446	0.066	-22.01	0.000
Educación secundaria	-3.099	0.172	-17.98	0.000
Educación superior	-6.179	0.736	-8.39	0.000
Sexo (mujer=1)	0.956	0.063	15.21	0.000
Constante	-6.032	0.092	-65.25	0.000
Número de observaciones	35,846			
Prob > chi2	0.000			

3.3: Resultados de la regresión para la cohorte de 41 a 50 años

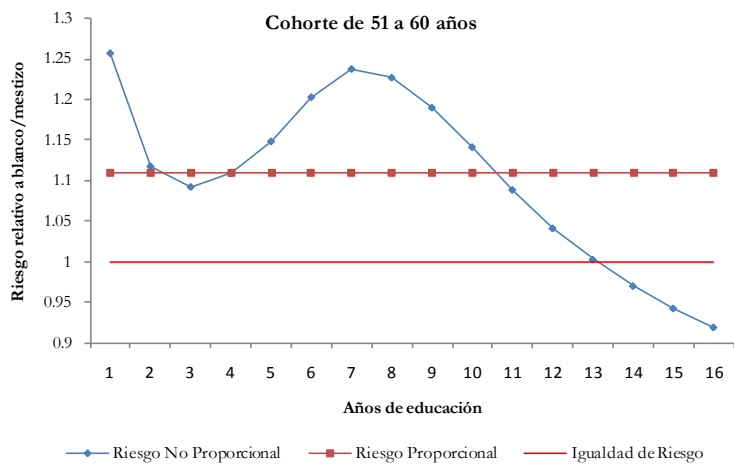
Variable	Coefficiente	Std. Err.	z	P>z
s0				
Quechua (etnia=quechua)	-0.088	0.026	-3.41	0.001
Amazónico (etnia=amazónico)	0.112	0.075	1.50	0.134
Afrodesc. (etnia=afrodescendiente)	-0.065	0.096	-0.67	0.502
Área Rural	-0.353	0.029	-12.11	0.000
Educación primaria	0.300	0.027	11.28	0.000
Educación secundaria	0.654	0.084	7.77	0.000
Educación superior	1.684	0.452	3.73	0.000
Sexo (mujer=1)	-0.357	0.025	-14.39	0.000
Constante	2.249	0.036	63.33	0.000
s1				
Quechua (etnia=quechua)	0.018	0.012	1.51	0.131
Amazónico (etnia=amazónico)	0.038	0.036	1.04	0.298
Afrodesc. (etnia=afrodescendiente)	0.013	0.045	0.28	0.780
Área Rural	0.170	0.013	12.88	0.000
Educación primaria	-0.072	0.013	-5.72	0.000
Educación secundaria	-0.317	0.035	-8.99	0.000
Educación superior	-0.257	0.214	-1.20	0.229
Sexo (mujer=1)	0.046	0.012	3.92	0.000
Constante	-0.176	0.016	-10.80	0.000
s2				
Quechua (etnia=quechua)	0.023	0.007	3.26	0.001
Amazónico (etnia=amazónico)	0.027	0.021	1.30	0.194
Afrodesc. (etnia=afrodescendiente)	0.019	0.029	0.66	0.509
Área Rural	0.063	0.007	8.98	0.000
Educación primaria	-0.036	0.008	-4.75	0.000
Educación secundaria	-0.115	0.016	-7.12	0.000
Educación superior	-0.327	0.049	-6.67	0.000
Sexo (mujer=1)	-0.038	0.006	-5.94	0.000
Constante	-0.081	0.009	-8.79	0.000
Xb				
Quechua (etnia=quechua)	0.299	0.062	4.83	0.000
Amazónico (etnia=amazónico)	-0.040	0.177	-0.23	0.820
Afrodesc. (etnia=afrodescendiente)	0.477	0.230	2.07	0.038
Área Rural	1.484	0.071	20.90	0.000
Educación primaria	-1.333	0.063	-21.26	0.000
Educación secundaria	-3.042	0.208	-14.64	0.000
Educación superior	-6.403	1.062	-6.03	0.000
Sexo (mujer=1)	1.014	0.060	16.89	0.000
Constante	-5.097	0.085	-59.65	0.000
Número de observaciones	25,909			
Prob > chi2	0.000			

3.4: Resultados de la regresión para la cohorte de 41 a 50 años

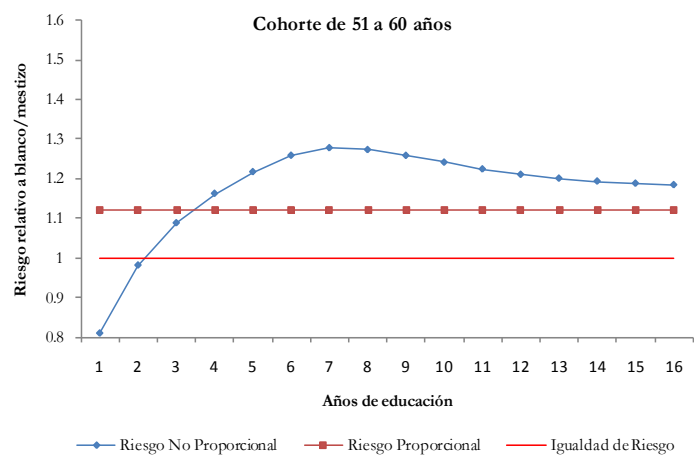
Variable	Coefficiente	Std. Err.	z	P>z
s0				
Quechua (etnia=quechua)	-0.068	0.027	-2.54	0.011
Amazónico (etnia=amazónico)	0.189	0.091	2.08	0.038
Afrodesc. (etnia=afrodescendiente)	-0.112	0.095	-1.18	0.237
Área Rural	-0.223	0.029	-7.64	0.000
Educación primaria	0.304	0.028	10.67	0.000
Educación secundaria	0.459	0.088	5.24	0.000
Educación superior	1.142	0.385	2.96	0.003
Sexo (mujer=1)	-0.207	0.026	-8.10	0.000
Constante	1.819	0.035	52.20	0.000
s1				
Quechua (etnia=quechua)	-0.014	0.014	-1.02	0.308
Amazónico (etnia=amazónico)	0.038	0.048	0.79	0.428
Afrodesc. (etnia=afrodescendiente)	0.021	0.050	0.42	0.678
Área Rural	0.125	0.014	8.69	0.000
Educación primaria	-0.085	0.015	-5.83	0.000
Educación secundaria	-0.321	0.040	-8.09	0.000
Educación superior	-0.394	0.192	-2.05	0.040
Sexo (mujer=1)	0.055	0.013	4.22	0.000
Constante	-0.063	0.017	-3.61	0.000
s2				
Quechua (etnia=quechua)	0.035	0.008	4.41	0.000
Amazónico (etnia=amazónico)	0.011	0.027	0.39	0.694
Afrodesc. (etnia=afrodescendiente)	-0.010	0.032	-0.33	0.744
Área Rural	0.050	0.008	6.48	0.000
Educación primaria	-0.034	0.009	-3.98	0.000
Educación secundaria	-0.134	0.018	-7.59	0.000
Educación superior	-0.349	0.049	-7.13	0.000
Sexo (mujer=1)	0.010	0.007	1.41	0.160
Constante	-0.101	0.010	-10.20	0.000
Xb				
Quechua (etnia=quechua)	0.273	0.063	4.36	0.000
Amazónico (etnia=amazónico)	-0.272	0.208	-1.30	0.193
Afrodesc. (etnia=afrodescendiente)	0.597	0.219	2.73	0.006
Área Rural	1.051	0.069	15.21	0.000
Educación primaria	-1.392	0.065	-21.33	0.000
Educación secundaria	-2.650	0.214	-12.39	0.000
Educación superior	-5.286	0.892	-5.92	0.000
Sexo (mujer=1)	0.754	0.060	12.49	0.000
Constante	-3.832	0.082	-47.00	0.000
Número de observaciones	15,821			
Prob > chi2	0.000			

Anexo 4: Riesgo relativo al grupo mestizo/blanco

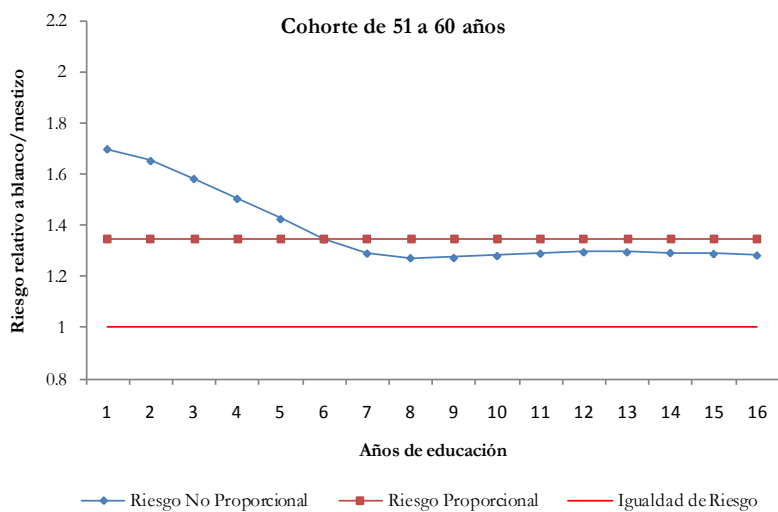
(A) Etnia Quechua



(B) Etnia Amazónica



(C) Etnia Afrodescendiente



Anexo 5: Riesgo estimado del sexo femenino (Con respecto a sexo masculino)

