



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ECONOMÍA

Aproximaciones al estudio de la desigualdad en economía. Teorías, indicadores y aplicaciones a la distribución salarial por medio de contrafactuales.

TESIS

QUE PARA OBTENER EL GRADO ACADÉMICO DE

LICENCIADO EN ECONOMÍA

P R E S E N T A

OMAR RUIZ GUTIÉRREZ

TUTOR: DR. GABRIEL ALEJANDRO MENDOZA PICHARDO

COMITÉ TUTOR:

DR. ALEJANDRO ROGELIO ALVAREZ BEJAR
DRA. MA. DE LA LUZ BLANCA ARRIAGA LEMUS
DRA. XIMENA VALENTINA ECHENIQUE ROMERO
MTRA. ELENA SANDRA MARTÍNEZ AGUILAR



CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO.

2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

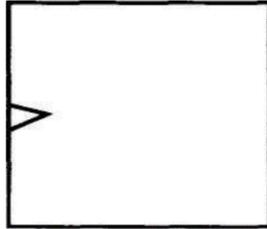
DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

“ 13 de febrero

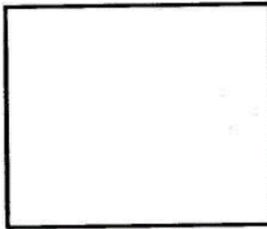
¿Qué hay detrás de la ventana?



Una estrella.

14 de febrero

¿Qué hay detrás de la ventana?



Una sábana extendida.

15 de febrero

¿Qué hay detrás de la ventana?



Roberto Bolaño, “Los detectives salvajes”.

Agradecimientos

Con este texto culmino un paso necesario de mi vida. Durante el trayecto, la gestación de ideas y el tiempo que me tomó escribirlo, he convivido con personas increíbles que me han permitido seguir creciendo de maneras que no imaginé.

En especial, quiero agradecer a mis padres por no dejar de transformarse y por apoyar cada una de mis decisiones. Los amo. Agradezco a mis hermanas Ilse y Abi, por estar al pendiente de mis avances y por siempre hacerme preguntas de economía.

A mis tíos Ruiz y Gutiérrez por darme mucho cariño y muchas ganas de mejorar cada día. Los quiero mucho.

A mis abuelos por darme tantas enseñanzas, consejos e historia. Muchas gracias.

A los grandes profesionales con quienes he tenido la fortuna de trabajar tanto en mi carrera profesional como académica. Especialmente, a la Dra. Y. Valles, el Dr. C. Serrano, el Dr. P. Peña, el Dr. G. Mendoza, el Dr. C. Ruiz por ser pilares en mi formación y desarrollo en etapas tempranas de mi carrera. A la Dra. L. Hernández, la Mtra. M. Perez-Tejada y el Dr. J. Ibarreche por ser grandes seres humanos y dejarme aprender un poco de ellos. Al Dr. Alejandro Álvarez, la Dra. María de la Luz Arriaga, la Mtra. Elena Martínez y la Dra. Ximena Echenique por aceptar ser sinodales de este trabajo. Me siento muy afortunado de contar con su supervisión en mi primer trabajo de investigación. ¡Muchas gracias!

A la extraordinaria institución que representa para mí la Universidad Nacional, de la cual he recibido incontables oportunidades y satisfacciones, y a la que debo los mayores esfuerzos para contribuir en su incansable labor de hacer de este país un lugar mejor.

Al cerrar el papel, no pienso sino en todas las experiencias que durante mis estudios me dio la Universidad: cafés, amigos, ratos bajo la lluvia, descubrimientos, sinsabores y una fuente inagotable de cambio y proyección. Gracias México, gracias Universidad por transformarme y darme la oportunidad de educarme y guiarme en tus aulas.

Agradezco a la DGAPA por la beca concedida para participar en el proyecto PAPIIT IN307611 “Tendencias y ciclos económicos: teoría y aplicación empírica en países en vías de desarrollo 1980-2010”

ÍNDICE

| | |
|---|----|
| Introducción..... | 1 |
| I. Capítulo 1. Aproximaciones desde la teoría social al fenómeno de la desigualdad..... | 3 |
| I.1 La desigualdad | 3 |
| I.1.1 La desigualdad en la teoría social..... | 4 |
| I.1.1.1 Algunas visiones liberales de la desigualdad..... | 5 |
| I.1.1.1.1 Rousseau..... | 5 |
| I.1.1.1.2 John Rawls..... | 8 |
| I.1.1.2 La visión materialista de la desigualdad..... | 10 |
| I.1.1.2.1 Friedrich Engels y Karl Marx..... | 11 |
| I.2 Conclusiones..... | 14 |
| II. Capítulo 2. La medición de la desigualdad..... | 16 |
| II.1 Las medidas de la desigualdad..... | 16 |
| II.1.1 Medidas de tendencia central..... | 16 |
| II.1.2 La varianza | 19 |
| II.1.3 Coeficiente de variación..... | 21 |
| II.1.4 La curva de Lorenz y el coeficiente de Gini..... | 24 |
| II.1.5 Medidas surgidas de la teoría del bienestar..... | 27 |
| II.1.6 La entropía, el índice de Theil y las medidas surgidas de la teoría de la información..... | 28 |

| | |
|--|----|
| II.2 Propiedades de los indicadores de desigualdad..... | 35 |
| II.3 Perspectivas empíricas..... | 38 |
| II.3.1 La evolución de la desigualdad en el mundo..... | 38 |
| II.3.2 La desigualdad en México..... | 40 |
| II.4 Conclusiones sobre los hechos e indicadores de la desigualdad..... | 42 |
| | |
| III. Capítulo 3. La desigualdad salarial en México y sus determinantes..... | 43 |
| III.1 La desigualdad salarial como una faceta de la desigualdad social..... | 43 |
| III.2 Revisión de la literatura..... | 44 |
| III.2.1 Habilidades y desigualdad de ingresos y salarios..... | 44 |
| III.2.2 Crecimiento económico y educación..... | 45 |
| III.2.3 Edad y escolaridad..... | 46 |
| III.2.4 Calidad de la educación y factores que inciden en su disponibilidad: perspectivas lingüístico-culturales..... | 47 |
| III.2.5 Habilidades cognitivas y no cognitivas..... | 49 |
| III.3 Análisis causal y contrafactuales..... | 51 |
| III.4 Regresión cuantílica para el estudio de distribuciones..... | 55 |
| III.5 Inferencia del escenario contrafactual..... | 56 |
| III.6 Conclusiones..... | 59 |

| | |
|--|-----|
| IV. Capítulo 4. Análisis empírico y resultados..... | 61 |
| IV.1 Datos sobre educación y salarios en México..... | 61 |
| IV.2 Ingreso por trabajo y definición de salario..... | 64 |
| IV.3 Estadísticos de desigualdad para la distribución del salario por hora..... | 67 |
| IV.4 Definición de las variables a partir de la ENIGH. Análisis descriptivo..... | 70 |
| IV.5 Estimación del escenario observado por medio de regresión cuantílica..... | 78 |
| IV.6 Inferencia de dos escenarios no observados..... | 85 |
| IV.7 Conclusiones..... | 97 |
| Comentarios finales..... | 98 |
| Bibliografía | 101 |
| | |
| Anexo1.Límites de $I_{ext}^{\alpha,\kappa}$ cuando $\alpha \rightarrow 0$ y $\alpha \rightarrow 1$ | 107 |
| Anexo 2. La regla de L'Hôpital..... | 112 |
| Anexo3. Incidencia de bilingüismo en población hablante de lenguas indígenas, por lengua indígena. ENIGH 2016..... | 113 |
| Anexo 4. Resultado de las pruebas de significancia estadística de los coeficientes del escenario observado en cuantiles seleccionados..... | 116 |
| Anexo 5. Correlación de las variables empleadas en la regresión por cuantiles..... | 118 |

Índice de Cuadros y Gráficas.

| | |
|---|----|
| Cuadro II.1 Ejemplo de cálculos basados en medidas de tendencia central..... | 22 |
| Cuadro II.2 Impactos de variaciones en los índices basados en medidas de tendencia central.... | 24 |
| Cuadro II.3 Entropía en dos sistemas..... | 31 |
| Cuadro IV.1 Composición del ingreso corriente trimestral por fuente de ingreso por deciles de personas, ENIGH 2016..... | 62 |
| Cuadro IV.2 Contribución del ingreso corriente trimestral por persona, por deciles de personas, ENIGH 2016..... | 64 |
| Cuadro IV.3 Estimación del salario promedio por deciles de personas..... | 65 |
| Cuadro IV.4 Estimación del salario promedio en deciles superiores | 66 |
| Cuadro IV.5 Indicadores de desigualdad para la distribución del salario por hora | 69 |
| Cuadro IV.6 Incidencia de discapacidad entre la población de estudio, por tipos y causas reportadas. ENIGH 2016..... | 72 |
| Cuadro IV.7 Incidencia de paternidad o maternidad en la población de estudio. ENIGH 2016..... | 73 |
| Cuadro IV.8 Incidencia de bilingüismo por lengua indígena hablada, ENIGH 2016..... | 74 |
| Cuadro IV.9 Incidencia de problemas de salud incapacitante en el año de la entrevista..... | 75 |
| Cuadro IV.10 Estimación del promedio del índice de hacinamiento por tamaño de localidad, ENIGH 2016..... | 76 |
| Cuadro IV.11 Construcción de años de escolaridad..... | 77 |
| Cuadro IV.12 Resultados de la regresión por cuantiles..... | 80 |
| Cuadro IV.13 Efectos de la política en índices de desigualdad. 0.1%-99.9%..... | 91 |
| Cuadro IV.14 Efectos de la política en índices de desigualdad. 4.9%-94.9%..... | 93 |

| | |
|--|----|
| Cuadro IV.15 Efectos de la segunda política en índices de desigualdad..... | 96 |
| Gráfica II.1 Ejemplo de distribuciones polarizadas y centradas..... | 23 |
| Gráfica II.2 Curva de Lorenz para ingresos por salarios.I..... | 25 |
| Gráfica II.3 Curva de Lorenz para ingresos por salarios.II. | 26 |
| Gráfica IV.1 Participación de ingresos por trabajo, transferencias y estimación de alquiler en el total del decil por deciles de personas, ENIGH 2016..... | 63 |
| Gráfica IV.2 Salario por hora promedio para la población de 15 a 64 años, ENIGH 2016..... | 66 |
| Gráfica IV.3 Distribución del salario por hora..... | 67 |
| Gráfica IV.4 Distribución del logaritmo del salario por hora..... | 68 |
| Gráfica IV.5 Estructura etaria de la población objeto de estudio, por sexo | 70 |
| Gráfica IV.6 Distribución del índice de hacinamiento..... | 76 |
| Gráfica IV.7 Distribución de años de escolaridad..... | 78 |
| Gráfica IV.8 Distribución condicional del salario en logaritmos respecto de la educación..... | 79 |
| Gráfica IV.9 Coeficientes estimados por percentil. Escenario observado..... | 81 |
| Gráfica IV.10 Proporción de personas que declararon ser indígenas, por años de escolaridad... | 85 |
| Gráfica IV.11 Distribución de años de escolaridad por edad..... | 86 |
| Gráfica IV.12 Distribución observada y contrafactual de años de escolaridad con una política para población con auto adscripción indígena..... | 88 |
| Gráfica IV.13 Densidad observada y contrafactual del salario por hora..... | 89 |
| Gráfica IV.14 Migración por cuantiles. 0.1%-99.9%. Primera política..... | 90 |
| Gráfica IV.15 Migración por cuantiles. 4.9%-94.9%. Primera política..... | 93 |

| | |
|---|----|
| Gráfica IV.16 Distribución observada y contrafactual de años de escolaridad con una política sobre escolaridad ideal..... | 94 |
| Gráfica IV.17 Densidad observada y contrafactual del salario por hora..... | 94 |
| Gráfica IV.18 Migración por cuantiles. Segunda política..... | 95 |

Introducción

La desigualdad como un problema social ha generado gran interés durante los últimos años, debido a la expansión de los valores democráticos en diferentes países y a los efectos del comercio global en diferentes regiones del planeta. La economía ha sido el principal campo en que este problema se manifiesta de manera más evidente pues la concentración surgida de los procesos de expansión económica necesariamente obliga a plantear diversas preguntas: ¿por qué países similares hace algunos siglos tienen niveles de vida o trayectorias dispares actualmente? ¿por qué algunas personas tienen una posición social o económica tan dispar respecto al resto?, ¿qué determina la posición social, económica o incluso política de un país o un individuo? La economía como ciencia social ha buscado responder el primer tipo de preguntas por medio de la teoría del desarrollo económico y del crecimiento económico. Sin embargo, el segundo y tercer tipo de preguntas generalmente escapan de su interés. No obstante, el problema de la desigualdad es subyacente a diversos problemas económicos como la pobreza y el bienestar, entre otros. Diversos teóricos sociales han planteado explicaciones que compiten por dar sentido al problema y de hecho han influido de manera sustancial en algunos planteamientos dentro de la ciencia económica.

El presente trabajo de investigación busca, en primera instancia, rescatar dos visiones que han tenido relevancia en los tratamientos teóricos de la desigualdad en economía: el liberalismo y el materialismo histórico. Este acercamiento al problema de la desigualdad desde la teoría social, que se realiza en el primer capítulo, se justifica en tanto permite identificar autores clave que han moldeado la manera en que se piensa la desigualdad no solo en economía, sino como un problema social ante el cual cada visión ha desarrollado diagnósticos específicos. Estas dos explicaciones permitirán contextualizar el problema central de la desigualdad como un problema de elección ética y empírica de quien la observa. Por ello, ambos tratamientos teóricos son revisados con el interés de mostrar sus límites y conocer sus planteamientos. De ellos se extraerán únicamente intuiciones que serán puestas a prueba por medio de los métodos más comunes en los estudios de la desigualdad en economía.

El objetivo último de esta investigación es analizar el efecto del nivel educativo de las personas, así como de políticas sobre la educación en la distribución de ingresos salariales en México por medio del empleo de una técnica contrafactual. Para ello una vez analizadas las dos reflexiones de la desigualdad desde la teoría social, se procede a analizar las medidas estadísticas empleadas para

medir la desigualdad así como las técnicas econométricas necesarias para realizar un análisis contrafactual.

En el segundo capítulo, se revisan la construcción e interpretación de los principales índices de desigualdad pues constituyen la principal herramienta analítica para dimensionar numéricamente la desigualdad que se plantea medir. Asimismo, se analizan algunos hechos importantes de la desigualdad a nivel global y nacional que se han estudiado empíricamente con el doble objetivo de observar el funcionamiento de los índices y obtener precedentes para el estudio de la desigualdad en los salarios en México.

En el tercer capítulo se plantea el método de estimación econométrica por cuantiles para aproximar los efectos de algunas variables seleccionadas a partir de las intuiciones rescatadas en el primer capítulo. El objetivo es definir cómo la estructura del salario responde a vínculos causales con las variables elegidas. Por ello, se plantea el problema de la causalidad en economía y se explica el acercamiento contrafactual como un enfoque para mitigar las limitaciones metodológicas en esta disciplina y poder obtener efectos causales a partir de una técnica estadística propuesta por los investigadores Victor Chernozhukov del MIT, Iván Fernández-Val de la Universidad de Boston y Blaise Melly de la Universidad de St. Gallen.

En el cuarto capítulo se muestran los resultados de la implementación del método de estimación por cuantiles sobre la distribución salarial obtenida por medio de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2016 realizada por el Instituto Nacional de Estadística, Geografía e Informática. Asimismo, se analiza a partir de la metodología contrafactual, el impacto de dos políticas elegidas sobre la distribución de años de escolaridad en los niveles de desigualdad salarial, así como la efectividad comparada entre una y otra, como una respuesta a la necesidad de llevar a cabo medidas que busquen mejorar el ingreso y el nivel de desigualdad salarial en México. Finalmente, en el apartado de conclusiones se hace un recuento del trabajo y se incluyen algunas reflexiones.

I. Capítulo 1. Aproximaciones desde la teoría social al fenómeno de la desigualdad.

En este capítulo se estudia la desigualdad desde dos tipos de planteamientos surgidos de la teoría social. Primeramente, se definen las características del fenómeno a partir del planteamiento del liberalismo de Jean-Jacques Rousseau y John Rawls. Lo propio sigue para el caso del planteamiento del materialismo histórico a partir de la exposición hecha por Abramovic Leontiev de la teoría de Karl Marx y Friedrich Engels. El objetivo de este capítulo es presentar estos planteamientos como antecedentes intuitivos para la contextualización teórica de algunas ideas actuales sobre la desigualdad, las cuales comúnmente están ancladas explícita o implícitamente, a alguno de estos dos tipos de planteamientos.

I.1 La desigualdad

La afirmación de que al menos dos entidades o individuos son desiguales surge de la comparación de alguna o varias características comunes entre ellos. Un planteamiento de este tipo es también un proceso de conocimiento que depende de dos ideas. Por una parte, se asume que es posible dicha comparación, entre otras razones, porque se cree que entre las dos entidades (o más) existe algo común que las asemeja, es decir, se concibe una idea de igualdad entre ellas. Por otra parte, se cree que esta comparación es constatable con la realidad, es decir, una afirmación de igualdad puede constatarse empíricamente. Así, a partir de la constatación de una concepción de igualdad determinada se conoce el mundo y se descubren las similitudes y desavenencias en él (Rousseau, 1755 :153).

Amartya Sen, uno de los expertos modernos en temas de desigualdad, reconoce que la idea de la desigualdad está ligada a un planteamiento del tipo: “todos los hombres son iguales en x aspecto”. Para él, la aprobación de una idea igualitaria basada en un criterio implica el rechazo de una idea igualitaria bajo otro criterio, es decir, a la justificación de algún tipo de desigualdad. Por ejemplo, creer que todos debemos acceder al mismo ingreso mensual es incongruente con una aprobación de igualdad en el plano del bienestar (Sen, 1995: 29). Esta contradicción exhibe la multiplicidad de conceptos de igualdad posibles, los cuales son resultado de diferentes criterios y creencias. La elección de la variable (o criterio) que se utiliza como referencia (Sen llama dichas variables “focales”), para una idea de igualdad constituye una voluntad teórica que solo puede justificarse en el plano ético (Sen, 1995: 30-32).

En suma, plantear una idea sobre la desigualdad es también plantear una hipótesis sobre una creencia formada por una noción moral, ética o empírica de quien la propone, la cual será sujeta a su respectiva justificación ética y constatación empírica. Es pues necesario reconocer estos aspectos como inherentes al estudio de la desigualdad. Con este objetivo en mente, se busca explorar dos planteamientos relevantes respecto a la desigualdad como referentes teóricos y poder así contextualizar los estudios de desigualdad en México.

I.1.1 La desigualdad en la teoría social

A continuación, se plantean dos enfoques teóricos relacionados a la desigualdad que, aunque no son los únicos¹, considero que son útiles para discernir, por un lado, las perspectivas éticas o morales inherentes a cualquier planteamiento igualitario, y, por el otro, los límites concretos que supone su ejecución en la realidad. Ambos planteamientos, se advierte, se reconocen como esquemas hipotéticos que permiten dar una explicación al fenómeno de la desigualdad. Su valor teórico es justamente el de aportar visiones intuitivas que ayudan a dar sentido a las desigualdades humanas, planteando también acciones ideales dados sus diagnósticos.

En este contexto, se propone pensar el problema de la desigualdad en una dicotomía no excluyente, como resultado de: a) una condición natural de los hombres, o b) una condición histórica relacionada al devenir material de la sociedad. El primer enfoque se puede apreciar directamente en el tratamiento de la desigualdad, planteado por Jean-Jacques Rousseau en su *“Discurso sobre el origen de la desigualdad”* y después de un modo distinto en el trabajo de John Rawls, *“A Theory of Justice”* publicado en 1971. El segundo enfoque, el materialista, se puede observar en las visiones de Karl Marx depositadas en *“El Capital”* y que son revisadas por Abramovich Leontiev en su libro *“El capital de Carlos Marx y la igualdad”* publicado en español en 1971.

La justificación ética de ambas posibilidades trasciende los límites de este trabajo, sin embargo, algo de sus planteamientos morales puede vislumbrarse a medida que se expone su entramado. Lo que destaca de ambas propuestas, además, es que a pesar de que cada una concluye con esquemas de acción distintos ante la desigualdad, las dos coinciden en que este fenómeno se legitima con la propiedad privada. En el planteamiento liberal de Rawls, sin embargo, la desigualdad en este

¹ Por ejemplo, el planteamiento de John Locke sobre el derecho natural de todo individuo y el planteamiento de Amartya Sen sobre la igualdad de libertades no serán tratados aquí.

aspecto podría llegar a ser tolerable bajo ciertas condiciones. Esta conclusión común, considero, es precedente de cualquier elaboración teórica que utilice la desigualdad social en su esquema explicativo y por tanto será una de las variables a analizar empíricamente en la tercera sección, a la par de algunos otros conceptos que se desprenderán de estos planteamientos.

I.1.1.1 Algunas visiones liberales de la desigualdad.

La economía ortodoxa no plantea una visión normativa respecto a la desigualdad, salvo cuando esta se plantea bajo el criterio de optimalidad paretiana. A juicio del autor, el problema normativo se convierte en un problema filosófico que las diversas corrientes de filosofía política han tratado abundantemente (Bunge, 2009: 25-26; Kymlicka, 1995: 11-17). Incluso Amartya Sen, un economista centrado en el estudio de la desigualdad se apoya en elaboraciones modernas de la teoría de la justicia de Rawls para plantear conceptos aplicables en economía.

Es por esto que se parte de dos visiones filosóficas del problema de la desigualdad desde el liberalismo igualitario. En específico, se parte de las ideas detalladas por Rousseau en su “*Discurso sobre el origen de la desigualdad*” para exponer la conexión de su planteamiento con visiones liberales más modernas como la de John Rawls, la cual ha modificado el discurso liberal a partir de una variación al planteamiento contractualista de Rousseau².

I.1.1.1.1 Rousseau

Rousseau plantea la existencia de dos tipos de desigualdad en los hombres: la desigualdad física y la desigualdad política o moral (Rousseau, 1755). Por un lado, la desigualdad física está dada naturalmente por los atributos físicos que se distribuyen desigualmente entre los hombres. A su vez, la desigualdad política o moral surge de las instituciones sociales que legitiman el derecho a la propiedad y a la riqueza, privilegiando solo a algunos individuos de la sociedad.

El objetivo de Rousseau es esclarecer, a partir de la hipótesis de un estado natural del hombre, los orígenes de las desigualdades morales (aquellas surgidas de las instituciones). Para ello plantea dos visiones del hombre: un hombre civilizado y un hombre salvaje. Para Rousseau, el hombre civilizado es aquel ser humano que gracias a su capacidad de resistir las operaciones que le impone

² El trabajo de Luiselli (2008) brindó a esta investigación una luz muy especial para comprender la influencia inequívoca del planteamiento contractualista de Rousseau en John Rawls.

la naturaleza, es capaz de ejercer su libertad y, al utilizarla para perfeccionar sus hábitos, logra alejarse de su estado natural (aquel en el que sobrevive sin perfeccionarse). El hombre salvaje, en cambio, reacciona a la naturaleza y coexiste en ambientes de menor complejidad, aunque logra satisfacer sus necesidades por medio de los atributos con los que cuenta, principalmente la fuerza física.

Estos dos tipos de hombres se relacionan y mientras algunos individuos perfeccionan sus hábitos, otros permanecen más cercanos a su estado natural, por ello surgen las desigualdades. Sin embargo, ¿qué obliga al hombre a alejarse de su estado natural? Para Rousseau, el estado natural del hombre se altera cuando el más fuerte trabaja más, cuando el más hábil ahorra más su trabajo o el sembrador de trigo requiere más hierro: cuando surge un deseo de obtener del otro el fruto de su trabajo, generalmente por medio de la fuerza. La comodidad conquistada o ganada por los hábitos de perfeccionamiento corrompe al hombre y constituye un alejamiento de la igualdad natural. Dicho alejamiento se originaría por la mirada que se lanzan entre sí los individuos y que se conjuga con el deseo de ser mirado lo que deriva en deseo de estimación pública basada en el logro individual.

Este proceso se habría apoyado en la existencia de la propiedad privada y en el trabajo como una necesidad de satisfacción humana que supera las necesidades de supervivencia (Rousseau, 1775: 170). La desigualdad de posesiones se habría extendido y se legitimaría para, entonces, permanecer entre los hombres y afectar la suerte de los particulares (Rousseau, 1775: 174). Aquí podemos observar una hipótesis interesante: las desigualdades morales surgidas de este modo afectarían no solo la vida de los hombres sino también la de su descendencia, suponiendo una condición intergeneracional.

Por lo tanto, las desigualdades físicas, que se acompañan del deseo de trabajar más de lo necesario y del perfeccionamiento de los hábitos humanos, se preservan mediante el derecho de propiedad - el catalizador de su perpetuidad, dice Rousseau. Este derecho supone una regla de justicia en la que se reparte el fruto del trabajo según cuanta propiedad se posee. Por lo que esta regla de justicia habría dado pie a la constitución de un poder supremo que sancionara el caos surgido de la confrontación entre el desposeído y el poseedor, dicha figura sería el Estado.

Es así que, en las sociedades con derechos de propiedad, se observarían al menos cuatro diferentes tipos de desigualdad: la desigualdad de riqueza, la de nobleza o rango, la de poder y la de mérito personal. Para Rousseau, el mérito personal engendra el resto de las desigualdades y la desigualdad de riqueza sería la consecuencia última de las demás desigualdades pues la forma monetaria de la desigualdad es de fácil transmisión y es capaz de proveer inmediatamente el bienestar a quien la posee³ (Rousseau, 1775: 186).

Los conflictos que surgen entre particulares corrompen también a la sociedad y el despotismo aparece como la última consecuencia de la desigualdad, incluso en presencia del Estado. Rousseau plantea este escenario como un círculo que se cierra, pues la desigualdad extrema vuelve a igualar a los hombres en un nuevo estado de naturaleza en el que la ley del más fuerte se vuelve a imponer ante ellos, como la naturaleza se imponía al hombre salvaje: se vislumbra como resultado del exceso de corrupción de los hombres poseedores y la opresión de los desposeídos. En un trabajo posterior “*El contrato social*” (Rousseau, 1762), la afamada frase rousseauiana que inicia dicho trabajo abarcaría este y otros aspectos de las libertades humanas⁴:

“*L’homme est né libre et partout il est dans les fers*”

“Man is born free and everywhere he is in chains”

“El hombre ha nacido libre, y sin embargo, vive en todas partes entre cadenas”

En otras palabras, Rousseau considera que el origen último de las desigualdades es la corrupción del hombre pues:

“El salvaje vive en él mismo, el hombre sociable siempre fuera de sí no sabe vivir más que en la opinión de los otros de cuyo juicio extrae el sentimiento de su propia existencia” (Rousseau, 1775: 188)

Rousseau concluye con una idea tajante. La extrema desigualdad, el despotismo, es prueba de que la mayor parte de nuestros males son *nuestra* propia obra y a la que nos lanzamos al evitar la

³ Todas estas desigualdades forman prejuicios que Rousseau considera contrarios a la razón, la virtud y la felicidad.

⁴ El autor agradece muy profundamente al profesor Joshua Dienstag el permitirle vislumbrar a partir de sus clases este y otros problemas filosóficos en los pensadores políticos europeos.

manera de vivir prescrita por la naturaleza. Las desigualdades que pasan por naturales en realidad son obra del hábito y de la adopción de un género de vida posible en la sociedad.

El planteamiento de Rousseau se apoya en un estado natural hipotético que es débil teóricamente pues no tiene una constatación empírica. En todo caso, funciona como un referente explicativo -y hasta cierto punto ético, que permite comprender una postura en la que el gobierno surge para preservar el poder de los ciudadanos y donde este se ancla en la legitimación de la propiedad como base de los derechos y obligaciones. Además, aporta la noción de que por medio de la propiedad privada se mantiene la desigualdad y que esta es una elección posible de los hombres cuya última consecuencia es el despotismo. Esta es la conclusión ciertamente breve de su visión de las desigualdades cuya relevancia se hará evidente en el tercer capítulo de esta investigación.

I.1.1.1.2 John Rawls

En 1971 con “*A Theory of Justice*”, John Rawls plantea un esquema explicativo similar al estado natural rousseauiano en el que los hombres y mujeres eligen, como en el contrato social de Rousseau, una serie de principios que serían aplicables a todo ser humano y que por tanto harían justa la constitución de ciertas desigualdades. Su propuesta tiene como objetivo plantear “un equilibrio reflexivo entre las intuiciones teóricas sobre la justicia, previas a su trabajo, y el utilitarismo que imponía nociones de justicia limitadas” (Kymlicka 1995: 63).

Rawls visualiza un estado inicial, llamado “posición original”, en el que los individuos abstraídos de todo conocimiento relacionado a su posición social, raza (sic), religión, posesiones económicas, concepción del bien o tendencias psíquicas, entre otras, erigirían una serie de principios que les asegurarían la posibilidad de una vida humana. Esta abstracción pone a los hombres bajo un velo de ignorancia, el cual asegura que nadie tiene una ventaja sobre el otro al momento de decidir qué principios guiarán a la sociedad (Kymlicka, 1995: 63-91 y Luiselli, 2008: 9). De este modo no se privilegiaría a una clase social o económica en particular. De hecho, como apunta Luiselli (2008) el hombre de Rawls actuaría como un individuo racional, el cual persigue sus propios intereses y

que, al atarse a este velo de ignorancia, escogería principios que resultarían justos para todos⁵ (Luiselli, 2008: 10):

Los principios elegidos de esta forma serían dos (Kymlicka, 1995: 66-68; Luiselli 2008: 12-14):

a) Principio de libertad: “Cada persona tiene el mismo derecho irrevocable a un esquema plenamente adecuado de libertades básicas iguales, compatible con un esquema similar de libertades para todos” (Luiselli, 2008: 12).

Aquí, Rawls demuestra su vertiente liberal al defender las libertades básicas de derecho al voto, la presentación a un cargo, la libre circulación, entre otros (Kymlicka, 1995: 67)

b) Principio de diferencia y de oportunidades equitativas

La estructura de las desigualdades económicas y sociales debe ser congruente con un sistema en el que redunden:

b.1 Mayores beneficios para los menos aventajados según un principio de ahorro justo (principio de diferencia o maximin- maximización de las ventajas para los menos favorecidos).

b.2 Los cargos y funciones públicas son asequibles para todos bajo la igualdad de oportunidades (principio de oportunidades equitativas).

Kymlicka (1995: 66) considera que la concepción de justicia que defiende Rawls es la siguiente: Todos los bienes primarios sociales como la libertad, la oportunidad, los ingresos y la riqueza, así como los fundamentos de la propia estima deben estar distribuidos de modo igual a menos que una distribución desigual de estos bienes sea ventajosa para los más desfavorecidos.

Así, los principios propuestos por Rawls como los que *naturalmente* escogería un agente racional que está bajo un velo de ignorancia permitirían que las desigualdades morales entre los individuos que pueden estar originadas por condiciones físicas o morales, sean las mínimas y permitan una sociedad justa. El acceso a los bienes primarios (sociales y naturales) es el mejor para todos en

⁵ Rawls asume ciertas condicionantes para que ocurra un contrato social. Básicamente son aquellas que permiten una cooperación social como coexistencia física y temporal, igualdad de recursos y capacidades físicas y mentales, vulnerabilidad a la agresión y escasez moderada. Ver Luiselli (2008: 12).

este planteamiento. Los bienes primarios sociales serían distribuidos por las instituciones, dichos bienes son: el ingreso, la riqueza, las oportunidades, los poderes, los derechos y las libertades. Los bienes primarios naturales son aquellos dados por la naturaleza a los individuos como el vigor, la imaginación, la salud, la inteligencia y pueden también afectarse por las instituciones (Kymlicka, 1995: 78).

La teoría rawlsiana, sin embargo, está sujeta a diversas críticas que van desde lo acomodaticio de la posición inicial para asegurar los dos principios propuestos, hasta los problemas internos de la teoría (la compensación por las desigualdades naturales y la subvención de elecciones erróneas de las personas). Justamente estas fallas han dado pie también a mayores elaboraciones teóricas (Kymlicka, 1995: 85-91).

A pesar de estas críticas, la propuesta de Rawls es contundente respecto a las desigualdades tolerables en la sociedad y ha sido fundante para el tratamiento de diversos problemas contemporáneos, como el tratamiento de Luiselli (2008) a la migración internacional. En una sociedad rawlsiana se permiten únicamente aquellas desigualdades que no afectan a los más desventajados, es decir, solo aquellas desigualdades que beneficien a todos y que alimenten las diferencias socialmente útiles serán aceptadas. Esto abre el camino a una visión más fragmentada de la desigualdad en la que incluso la propiedad privada sería tolerable en tanto redunde en mayores ventajas a los menos favorecidos.

1.1.1.2 La visión materialista de la desigualdad.

Contraria a la visión liberal de la desigualdad, el marxismo considera que la igualdad es un concepto social que se determina según las condiciones de producción material, las relaciones sociales y la lucha de clases que prevalecen en una etapa determinada de desarrollo de la sociedad (Leontiev, 1971: 13)

Leontiev (1971) revisa las concepciones de la igualdad desde el materialismo histórico a partir del trabajo de Friedrich Engels “*Anti-Dürhing*” y Karl Marx “*El Capital*”. A continuación, se presentarán estas concepciones a la luz de Leontiev (1971) a fin de tener presentes estos planteamientos teóricos.

1.1.1.2.1 Friedrich Engels y Karl Marx

De acuerdo con Friedrich Engels, en las comunidades primitivas el trabajo igualitario y poco productivo favorecía concepciones igualitarias limitadas. Dada la imperiosa necesidad de dominar la naturaleza, el trabajo comunitario en el que todos recibían la misma cantidad de producto del trabajo, aseguraba la supervivencia de la especie. Esta era la razón principal para que existiera la idea de igualdad de consumo (a pesar de que reconoce que existía desigualdad entre los sexos surgida a partir de la división del trabajo) y no la socialización expresa de los medios de producción como preconiza el materialismo histórico (Leontiev, 1971:15-17). La propiedad privada y el paso de sociedades agrícolas a ganaderas, implicó la esclavitud de unas tribus conquistadas por otras. La comunidad se redujo entonces al conjunto de miembros de la sociedad que no eran esclavos, ni mujeres ni forasteros (Leontiev, 1971: 17). Es decir, la comunidad primitiva dio paso a una sociedad clasista basada en la explotación del hombre por el hombre.

La contradicción formada entre libres y esclavos sobre la que se basaba el régimen de producción del mundo antiguo preservaba la desigualdad política y económica por medio de la concepción de desigualdad natural entre los hombres, que habría surgido como resultado de estas condiciones de producción. Para Platón, por ejemplo, la sociedad se dividía en clases, una ocupada en la producción que vive en circunstancias comunes y la otra que vive en un medio ideal también común. (Leontiev, 1971: 18). Así, las clases sociales libres de la realización de trabajo productivo, por ejemplo, los guerreros, disfrutaban de un comunismo que se basaba en la posesión de esposas y propiedades. En general, este comunismo selectivo, relegaba la codicia y la propiedad privada a la clase libre. El derecho romano surgiría entonces, dice el materialismo, de la concepción de igualdad prevaleciente entre personas detentadoras de propiedades como un mecanismo para preservar su continuidad.

Por otra parte, Leontiev (1971) señala que en el desarrollo de las concepciones históricas de la igualdad identificadas por Engels se pueden identificar dos componentes: uno burgués y otro proletario, de acuerdo con los constructos marxistas de clases sociales.

En una primera etapa, los movimientos campesinos surgidos durante el feudalismo son entendidos por el materialismo como resultado de las reivindicaciones del discurso igualitario en ausencia de una burguesía organizada. De hecho, estos movimientos habrían tomado, a su vez, dos vertientes.

La primera es la del discurso plebeyo-cristiano que exigía la supresión de las clases sociales (imposible para la época, según Leontiev). La segunda defendía el incipiente discurso burgués, y abogaba por la igualdad jurídica frente a la clase eclesiástica (Leontiev, 1971: 25).

Eventualmente las, así llamadas, demandas burguesas se satisficieron y sus conquistas devinieron legislaciones llamadas a preservar los derechos humanos en general. Esto planteó la libertad y la igualdad como derechos naturales e inalienables y, eventualmente, la justicia tomó el lugar de la igualdad burguesa ante la ley (Leontiev, 1971: 33-37).

Hasta aquí la lectura materialista del concepto histórico de igualdad se comprende a partir de la evolución, si es que existe, de los modos de producción sociales. El derecho a la propiedad surgiría del desarrollo histórico de la concepción de igualdad burguesa que se trasladó de una demanda campesina a una ley, llamada natural por los partidarios de dichos discursos.

Lo más relevante de la visión materialista de la igualdad, empero, es la lectura anclada en la explicación de la sociedad capitalista según lo plantea Karl Marx en *“El Capital”*. Para Marx, el valor es la esencia del fenómeno de la mercancía en la sociedad capitalista. A partir del análisis de esta, plantea que el intercambio surge a partir de una concepción de igualdad de diferentes “valores de uso” y de “valores”, lo cual constituye, dice Marx, una contradicción profunda del sistema capitalista y que prevalece como la idea de igualdad en el modo capitalista de producción. Por un lado, un objeto puede ser valor de cambio sin por ello ser valor de uso, es decir, sin tener “valor”. Por otra parte, la abstracción del “valor” de una mercancía en su forma de “valor de cambio” es necesaria para igualar lo heterogéneo del valor de las mercancías, y hacer posible así el intercambio (Leontiev, 1971: 48-49).

El proceso de intercambio supone, entonces, una abstracción del producto de trabajo que debe igualar a un trabajo humano en general, es decir, la variedad de los tipos de trabajo concreto contenidos en las diversas mercancías se abstrae y solo se compara, en el proceso de intercambio, el trabajo humano abstracto (Leontiev, 1971: 50). La suma de este trabajo humano abstracto es el trabajo social, el cual en el caso de la sociedad capitalista es un trabajo individual agregado y no un trabajo socializado (Leontiev, 1971: 52).

Posteriormente, el dinero surgiría como la mercancía que da respuesta a la contradicción del sistema y constituye un equivalente radical frente al que todas las mercancías se equiparan como

valores de uso. Su aparición elimina los límites del intercambio de mercancías que, en su ausencia, se determinan por las necesidades individuales de los poseedores de mercancías deseosas de intercambiarlas. Sin embargo, aporta una complicación mayor pues todo valor contenido en una mercancía debe trasladarse a una noción de precio y dada la incongruencia en esta transformación, el intercambio puede ser injusto respecto a la idea de que se intercambian valores semejantes e iguales en un intercambio sin dinero. Es así como el dinero hace impersonal y hace prevalecer la desigualdad social pues esta se encubre en el intercambio.

El proceso de intercambio implica, además, dice Marx, una relación de dominación de la clase capitalista respecto a la clase trabajadora. En el proceso de producción y en el proceso de circulación pueden ubicarse dos fuentes de esta dominación que se agrava con la reproducción capitalista. En el primero, el valor producido se logra por medio de la valorización que hace el trabajo de los medios de producción. En el segundo, el intercambio de trabajo y salario que, según Marx, no se intercambian en su misma cuantía pues el valor producido y el salario no se igualan, constituye la muestra de la desigualdad inherente en el sistema.

La idea de igualdad que defiende el materialismo es una en la que los medios de producción son socializados y por tanto todos tienen acceso a ellos, sin necesidad de que existan las clases sociales. La igualdad planteada por las democracias liberales constituiría, en este contexto, la preservación de la desigualdad originada por el proceso de producción capitalista en el que la propiedad privada y los medios de producción son privados y el trabajo concreto se obliga a una igualdad ficticia para lograr el intercambio. Además, según el marxismo dicha ideología liberal buscaría plantear un concepto de justicia que no sería necesario si la sociedad aboliera los derechos de propiedad y si los medios de producción fueran socializados.

A pesar de la generalización que esta vertiente de la teoría social hace de las visiones liberales como surgidas de intereses clasistas legitimados mediante la ley liberal, no resulta muy claro cómo su propuesta de igualdad de acceso a los medios de producción y, eventualmente, a los recursos producidos podría ser compatible con una igualdad de acceso a la toma de decisiones de un país, si es que existiera un Estado socialista. Una crítica adicional lanzada por Kymlicka (1995) es respecto al funcionamiento real del acceso a los recursos escasos, y a los procesos estrictos en que la alienación del trabajo por una parte y la explotación económica, por la otra, pudieran realmente

superarse por medio del comunismo o la socialización de los medios de producción (Kymlicka, 1995: 179-218).

1.2 Conclusiones

En este capítulo se expusieron dos planteamientos aparentemente contrarios en cuanto a la acción social necesaria para asegurar alguna noción de justicia, que en ambos casos se acerca a la determinación de algún tipo de igualdad en un aspecto específico. A pesar de dicho antagonismo, en ambos prevalece un diagnóstico importante: las desigualdades surgidas de la naturaleza (condiciones físicas, habilidades o inteligencia) persisten por medio de las desigualdades de la propiedad, aunque en uno de estos planteamientos estas pueden ser tolerables si redundan en los máximos beneficios para los menos aventajados.

Rousseau imagina un estado social natural para comprender el funcionamiento y origen de las desigualdades sociales. Su conclusión es que las desigualdades se originan por la naturaleza y se legitiman por medio de los derechos de propiedad. El ser humano alentaría este proceso y sería el responsable de que estas desigualdades persistieran. Por su parte, la elaboración rawlsiana de la justicia plantea una posición inicial en la que todos los seres humanos al ser abstraídos de sus intereses más claros podrían elegir los principios de justicia sobre los que se apoyaría una vida en sociedad. Con ello, Rawls parece plantear una esencia humana que decide qué es y qué no es justo sin importar prejuicios físicos, mentales, sociales, entre otros. Así, la igualdad se plantea en cuanto al acceso a los bienes primarios que se conjuga con cierta libertad de actuar agravando las desigualdades que se juzgan inofensivas para todos, en especial para los menos aventajados. A pesar de sus limitaciones, esta teoría resulta innovadora al poner “prioridades lexicológicas”⁶ (Kymlicka, 1995: 66) sobre las decisiones de una igualdad social, las cuales fragmentan la visión patrimonial de la desigualdad y permiten visualizar diversas fuentes del problema.

La visión materialista pone en duda las nociones liberales de igualdad pues considera que todo derecho de propiedad es antagónico a la verdadera justicia que se alimenta de la abolición de las clases sociales. El derecho a la propiedad permite la profundización de las desigualdades del

⁶ Como lo señala Flores, I. (1999), John Rawls “propone reglas de prioridad en las que se debe respetar el orden serial o léxico de los tres principios propuestos” Flores (1999: 98). Esto significa que la prelación o importancia de cada principio va en orden decreciente de acuerdo con su orden de enunciación. Así, el primer principio rige sobre el segundo y así sucesivamente.

sistema capitalista y se legitima en el proceso de intercambio conformándose bajo la forma del dinero. Esta visión de la sociedad es relevante puesto que, a pesar de plantear una acción radical sobre el origen de la desigualdad, ubica su fuente en el derecho de propiedad y como surgida de la lucha entre las clases sociales.

Este capítulo permitió ubicar dos de las posiciones más relevantes en los análisis de la desigualdad social. En específico, la desigualdad de propiedad por un lado y por otro, la fragmentación de la visión rawlsiana que plantea diversos criterios que permiten una sociedad justa, entre los que figura el derecho a la propiedad privada pero también al voto y a la igualdad de oportunidades. En el segundo capítulo se analizarán algunas concepciones sobre tipos de desigualdad en la literatura económica, a la luz de las nociones empíricas y estadísticas de la desigualdad que se expondrán en el inicio de dicho capítulo. Contrastar dichas concepciones con las ideas propuestas en este apartado, en especial con el enfoque de la desigualdad de habilidades, permite integrar el conocimiento surgido en los estudios empíricos dentro de un flujo de ideas sobre las concepciones igualitarias. En el tercer capítulo se desarrollará un análisis de desigualdad salarial influido por la perspectiva liberal sobre las habilidades, se incluyen también algunas variables relativas a las condiciones de propiedad individual que, sin embargo, no logran dimensionar cabalmente el aspecto estructural que ese tipo de desigualdad representa. El análisis de la desigualdad de propiedades merece una atención similar que no ha sido posible brindar en este trabajo. Esa veta de investigación es también una aportación de este capítulo a futuros análisis sobre las desigualdades sociales.

II. Capítulo 2. La medición de la desigualdad.

Más allá de la elección de la variable sobre la que se elige plantear una hipótesis de desigualdad y los juicios morales surgidos de dicha elección, existen medidas desarrolladas por la estadística que permiten cuantificar en el tiempo o entre diferentes sistemas el grado de desigualdad (o de igualdad) existente en ellos. Sin embargo, no todas estas medidas reflejan de igual modo esta condición. Es decir, dependiendo de los planteamientos estadísticos y la información disponible para su análisis, cada una de estas medidas refleja de manera diferente el comportamiento de la desigualdad. En este capítulo se aborda primeramente el fundamento y el análisis de las propiedades de estos indicadores y posteriormente se observa su uso en algunos estudios empíricos enfocados a México y el mundo.

II.1 Las medidas de la desigualdad

Las medidas estadísticas más inmediatas retoman los momentos de la distribución estadística de la variable analizada (i.e. media, varianza, desviación estándar). Otras medidas más complejas son aquellas surgidas de ecuaciones de bienestar, como el índice de Dalton y Atkinson (Cowell, 2009: 39) y los índices surgidos del concepto de entropía y que han sido retomados desde la teoría de la información. A continuación, se plantean las definiciones de estos indicadores de desigualdad siguiendo los trabajos de Coulter (1989), Cowell (2009) y Cortés, F. y R. Rubalcava (1984), así como sus respectivas propiedades como indicadores de desigualdad. Concluyo con una revisión de las propiedades deseables establecidas por Cowell (2009). Comienzo, puntualmente con definiciones para las medidas de tendencia central para después poder definir las medidas de varianza y el coeficiente de variación. Procedo a deducir el índice de Gini, los índices surgidos de la teoría del bienestar y los índices surgidos de la teoría de la información. Finalmente, realizo un balance de las propiedades comparativas de cada uno de los indicadores propuestos.

II.1.1 Medidas de tendencia central

Algunas variables en economía pueden ser resultado de experimentos aleatorios, es decir, que son fruto del azar como, por ejemplo, la distribución de habilidades entre las personas. Si consideramos que una variable, llamémosla X , captura los resultados de un experimento, se pueden hacer al menos dos cosas. La primera es obtener de dicha variable los posibles valores que ha tomado (x_0, x_1, x_2, \dots), es decir, obtener el espacio muestral de la variable. En segundo lugar, se pueden obtener

las probabilidades de que la variable tome un determinado valor dentro de un espacio muestral finito. Dependiendo de si se trata de variables aleatorias o continuas, estas probabilidades deben cumplir dos axiomas, de acuerdo con (Greene, 1998: 56):

A. Variables discretas

Primer axioma: $0 \leq \Pr(X = x) \leq 1$; Segundo axioma: $\sum_x f(x_i) = 1$

B. Variables continuas

Primer axioma: $\Pr(a \leq x \leq b) = \int_a^b f(x)dx \geq 0$; Segundo axioma: $\int_{-\infty}^{+\infty} f(x)dx = 1$

Donde $f(x)$ y $f(x_i)$ son las funciones de distribución de la variable x para los casos variable y discreto, respectivamente. Cuando estos axiomas se cumplen, se puede explícitamente establecer una primera relación entre el espacio muestral y la probabilidad de un cierto evento. La función de distribución (F.Dist.) se define del siguiente modo:

Sea x una variable aleatoria, la probabilidad de que x sea menor o igual a un valor a se representa como $F(a)$. Así, $F(x)$ es la función de distribución (F.Dist.) de dicha variable (Greene, 1998: 56).

La F.Dist. permite, además, contemplar una relación entre la probabilidad de un evento y los valores que puede tomar una variable dependiendo si estos son discretos o continuos:

A. Variables discretas:

Con un espacio muestral finito y ordenado con valores discretos un evento x tiene la siguiente F.Dist:

$$F(x_i) = \sum_{X \leq x} f(x_i) = \Pr(X \leq x)$$

y se sigue que

$$f(x_i) = F(x_i) - F(x_{i-1})$$

la cual es la función de densidad de probabilidad (f.den.) de la variable y se puede determinar a partir de su respectiva F.Dist. Esta función refleja la probabilidad de que un valor específico sea el tomado por la variable analizada.

B. Variables continuas

Para la F.Dist.:

$$F(x) = \int_{-\infty}^x f(x)dx$$

Donde $f(x)$ representa la función de densidad f.den de la variable x. Por lo que también, la f.den. se puede definir como:

$$f(x) = \frac{dF(x)}{dx}$$

Si la F.Dist. cumple con que a) $0 \leq F(x) \leq 1$; b) Si $x > y$, $F(x) > F(y)$; c) $F(+\infty) = 1$; y d) $F(-\infty) = 0$, la relación entre el espacio muestral y las probabilidades de la variable son válidas para determinar los tres conceptos siguientes:

Valor esperado o media

Se define del siguiente modo, para el caso discreto y continuo, respectivamente:

$$\mu = E[x] = \sum_x xf(x)$$

$$\mu = E[x] = \int xf(x)$$

El cual informa el valor que caracteriza a la variable. Es decir, resume en un valor específico las probabilidades y los valores posibles de una variable.

Cowell (2009) establece una definición similar y añade la posibilidad de utilizar una transformación logarítmica de la variable analizada para el cálculo de la media geométrica:

| Nombre | Símbolo | Distribución | |
|------------------------------------|-----------|---|------------------------------------|
| | | Discreta | Continua |
| Media Aritmética. (Ec. II.1) | \bar{x} | $\frac{1}{n} \sum_{i=1}^n x_i$ | $\int x dF$ |
| Media Geométrica. (Ec. II.2) | x^* | $\exp\left(\frac{1}{n} \sum_{i=1}^n \log(x_i)\right)$ | $\exp\left(\int \log(x) dF\right)$ |

Fuente: Cowell (2009: 150-151).

Mediana

Se define como el valor m tal que $\Pr(X \leq m) \geq \frac{1}{2}$ y $\Pr(X \geq m) \geq \frac{1}{2}$. Es decir, es el valor que se encuentra justo al centro de la F.Dist., por lo que informa de una frontera a partir de la cual ambas partes tienen igual probabilidad de ocurrir.

Moda

Es el valor de la variable para el cual la f.den., $f(x)$, obtiene su valor máximo. Esto es equivalente a encontrar el valor más frecuente o con mayor probabilidad.

II.1.2 La varianza

Una vez consideradas las medidas de tendencia central, es posible plantear algunos indicadores más complejos que aportan información relevante de la variable analizada. En primer lugar, se define a la varianza como el promedio del cuadrado de la diferencia de las observaciones respecto al promedio de la distribución. Formalmente se define como:

$$(Ec. II.3) \quad V = \frac{1}{n} \sum_{i=1}^n [x_i - \bar{x}]^2$$

Donde, n es el tamaño de la población estudiada. Esta medida es sensible a la dimensión de la variable, es decir, refleja si esta se encuentra en miles, millones o en números menores a uno. Sin embargo, la desviación estándar permite una reducción del problema pues es la raíz cuadrada de este indicador $d = \sqrt{V}$.

En el contexto del estudio de la desigualdad del ingreso, Cowell (2009) plantea dos versiones de la varianza. La primera consiste en utilizar las diferencias logarítmicas de cada observación respecto a la media aritmética del ingreso \bar{x} resultando la varianza logarítmica v . La segunda se calcula respecto a la media geométrica del ingreso x^* resultando la varianza de los logaritmos del ingreso v_1 :

$$(Ec. II.4) \quad v = \frac{1}{n} \sum_{i=1}^n \left[\log \left(\frac{x_i}{\bar{x}} \right) \right]^2$$

$$(Ec. II.5) \quad v_1 = \frac{1}{n} \sum_{i=1}^n \left[\log \left(\frac{x_i}{x^*} \right) \right]^2$$

Al transformar los niveles de las variables analizadas en logaritmos, se debe prestar particular atención a que esta transformación se permite únicamente para variables con valores superiores a cero, lo cual es razonable para la mayor parte de las variables en economía.

Los indicadores de varianza V , v y v_1 permiten capturar la distribución de una variable en un momento determinado y pueden ser de utilidad para reportar cambios en su distribución al compararse temporalmente. Sin embargo, también dependen de la escala de la variable utilizada por lo que en el caso del indicador V , la comparación entre variables distintas (por ejemplo, años de escolaridad e ingreso mensual) no sería válida sin reconocer la diferencia de escalas. La propuesta de Cowell para indicadores en logaritmos reduce este problema y permite una comparación más certera entre variables, sin embargo, preserva un problema adicional que afecta a los tres indicadores. Dicho problema consiste en que los valores demasiado alejados de la media pueden incidir en la medida de varianza seleccionada sin importar si estos se encuentran a la izquierda (por debajo) de la media o bien si se encuentran a la derecha (por encima) de la media. Piénsese el caso de una distribución del ingreso relativamente equitativa, pero con algunas pocas personas con ingresos casi nulos, el indicador de la varianza crecerá y podría coincidir a una distribución donde el mismo número de personas estuvieran a la misma distancia del promedio, pero por encima (pocas personas con ingresos exorbitantes). Esto significa que este indicador aporta poca información sobre si la desigualdad surge de puntos extraordinarios por arriba o por debajo de la media, por lo que el origen de un cambio en dicha variable sería difícil de identificar.

En la Gráfica II.1 pueden observarse dos ejemplos de distribuciones, el caso polarizado y el centrado. En dichas distribuciones el término centrado o polarizado hace referencia a la distancia media de cada observación respecto al promedio. Nótese cómo la transformación logarítmica relativiza las distancias entre los valores en niveles y permite visualizar en un rango más corto su comportamiento y cómo los valores más pequeños se dimensionan relativamente más respecto a los más grandes. Además, en el Cuadro II.1 se puede observar el cálculo de los indicadores de varianza de dichas distribuciones. Se incluye también el cálculo del coeficiente de variación, que se define a continuación.

II.1.3 Coeficiente de variación

Un último indicador de desigualdad basado en las medidas de tendencia central se obtiene al expresar la raíz de la varianza (desviación estándar) como proporción de la media. Este indicador se denomina coeficiente de variación:

$$(Ec.II.6) \quad c = \frac{\sqrt{V}}{\bar{y}}$$

El coeficiente de variación, c , informa sobre la concentración de la distribución respecto al valor medio, en términos porcentuales y elimina el problema de la dimensión de la variable analizada, que se presentaba en el caso de los índices de varianza. Como indicador de desigualdad, se utiliza la siguiente interpretación:

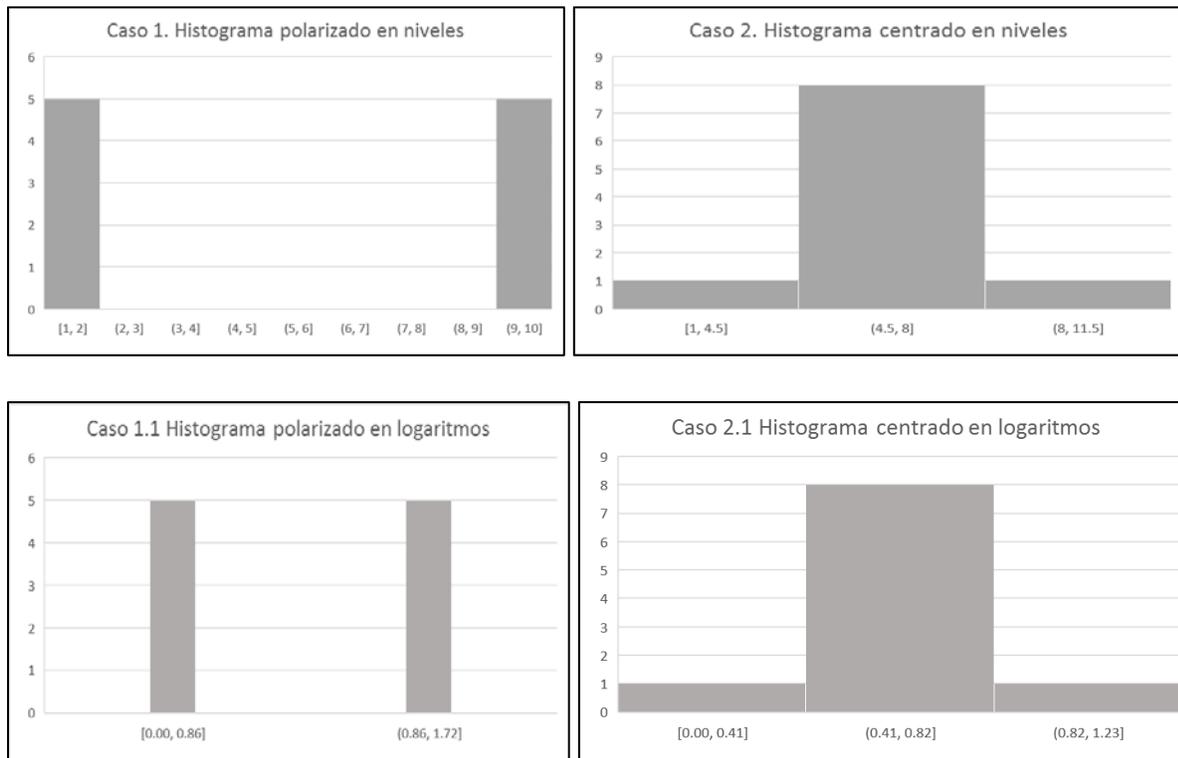
Un coeficiente de variación bajo, tendiente a cero, es congruente con una distribución más igualitaria. Una distribución menos igualitaria tenderá a tener un coeficiente con un valor muy grande a medida que la varianza de la distribución sea muy grande en relación al promedio. Es necesario, sin embargo, que la media de la distribución sea diferente de cero para que el indicador exista, lo cual es congruente en la mayoría de las variables en economía.

Aunque este índice es comparable entre variables, mantiene la otra limitación de los índices de varianza al tratar igualmente los valores por encima y por debajo del promedio. Es decir, no aporta información sobre si la distribución está concentrada en un polo en específico.

Cuadro II.1 Ejemplo de cálculos basados en medidas de tendencia central

| Caso 1: Desigualdad polarizada | | Caso 2: Desigualdad centrada | |
|---------------------------------------|-------------|-------------------------------------|-------------|
| x_i | $\log(x_i)$ | x_i | $\log(x_i)$ |
| 10 | 1.00 | 10 | 1.00 |
| 10 | 1.00 | 5 | 0.70 |
| 10 | 1.00 | 5 | 0.70 |
| 10 | 1.00 | 5 | 0.70 |
| 10 | 1.00 | 5 | 0.70 |
| 1 | 0.00 | 5 | 0.70 |
| 1 | 0.00 | 5 | 0.70 |
| 1 | 0.00 | 5 | 0.70 |
| 1 | 0.00 | 5 | 0.70 |
| 1 | 0.00 | 1 | 0.00 |
| V | 20.25 | V | 4.09 |
| υ | 0.65 | υ | 0.06 |
| υ_1 | 0.25 | υ_1 | 0.06 |
| μ | 5.50 | μ | 5.10 |
| $\log(\mu)$ | 0.74 | $\log(\mu)$ | 0.71 |
| c | 3.68 | c | 0.80 |

Gráfica II.1 Ejemplo de distribuciones polarizadas y centradas.



Como una forma de resumen de los cuatro indicadores hasta ahora presentados, en el Cuadro II.2 se muestra el impacto que tendría la transformación de una variable analizada cuando se le suma un valor a homogéneamente a cada observación, o bien el impacto de una multiplicación homogénea por dicho valor a en el cálculo de cada medida. También se observan las restricciones de la variable para realizar el cálculo.

Cuadro II.2 Impactos de variaciones en los índices basados en medidas de tendencia central

| Nombre | Definición | Restricción | Impacto cuando x_i sufre una: | |
|--------|---|-----------------------|---------------------------------|---|
| | | | Multiplic. por a | Suma de a |
| V | $\frac{1}{n} \sum_{i=1}^n [x_i - \bar{x}]^2$ | $n > 0$ | $a^2 V$ | Ninguno |
| C | $\frac{\sqrt{V}}{\bar{x}}$ | media no igual a cero | Ninguno | $\frac{\sqrt{V}}{a + \bar{x}}$ |
| v | $\frac{1}{n} \sum_{i=1}^n \left[\log \left(\frac{x_i}{\bar{x}} \right) \right]^2$ | $n > 0$ | Ninguno | $\frac{1}{n} \sum_{i=1}^n \left[\log \left(\frac{x_i + a}{\bar{x} + a} \right) \right]^2$ |
| v_1 | $\frac{1}{n} \sum_{i=1}^n \left[\log \left(\frac{x_i}{x^*} \right) \right]^2$ | $n > 0$ | Ninguno | $\frac{1}{n} \sum_{i=1}^n \left[\log(x_i + a) - \sum \log(x_i + a) / n \right]$ |

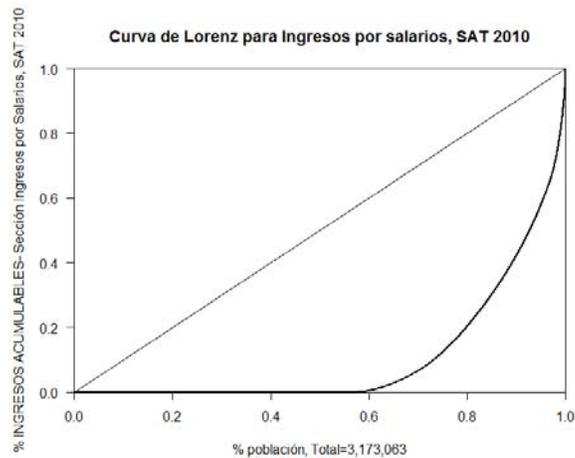
Nótese que la restricción del coeficiente de variación puede evitarse al sumar a toda la distribución un valor (digamos 1) lo que permitiría su cálculo.

II.1.4 La curva de Lorenz y el coeficiente de Gini

Max Otto Lorenz planteó en 1905 pensar la concentración de la riqueza a partir de identificar cuánta riqueza acumula cierto porcentaje de personas en un país. Así, la curva que se trazaría en un espacio donde en el eje horizontal se presenta la proporción de la población acumulada de un país (cuyo máximo es 1 o 100%) y en el eje vertical la riqueza acumulada por cada porcentaje acumulado de población (con máximo 1 o 100%), reflejaría las condiciones de desigualdad de riqueza poseída dentro de un país determinado. Esta curva conocida como curva de Lorenz es ampliamente usada siguiendo la idea planteada por su creador. Usualmente esta línea se contrasta con un ideal igualitario planteado también por Lorenz (1905: 217) donde a medida que se contabiliza la población de un país, la riqueza se acumularía *pari passu*, es decir, el ideal en que todos tienen la misma cantidad de riqueza. Esta curva ideal se observa en las gráficas de la curva de Lorenz como una línea recta que conecta los puntos de 0% de población y de ingreso). Como

con el correspondiente punto de 100% de población e ingreso. Como un ejemplo empírico de este enfoque, se muestra en la Gráfica II.2 la curva de Lorenz que se desprende de 3,173,063 declaraciones fiscales de personas físicas contenidas en la base de datos publicada el 16 de Agosto de 2017 por el Servicio de Administración Tributaria de México. El rubro específico empleado fue la declaración de ingresos por salario en 2010. En el eje horizontal se observa la acumulación de declaraciones de personas físicas. y en el eje vertical los ingresos acumulados por cada declaración:

Gráfica II.2 Curva de Lorenz para ingresos por salarios. I.



Fuente: Elaboración propia con base en datos abiertos de SAT (2010)

Esta gráfica informa un aspecto sustancial de la desigualdad en el universo de declaraciones en 2010. Solo a partir de poco menos del 60% de las declaraciones de personas físicas, comienza a despegar la acumulación de ingresos por salarios. Esto se explica porque únicamente 1,339,112 de las observaciones en 2010 cuentan con valores positivos en este rubro. Es de considerar que la Ley de Impuesto a la Renta aplicable en 2010 considera además de los ingresos por salarios, aquellos originados por arrendamientos, ganancias por enajenación de bienes, entre otros, como susceptibles de reglas especiales para la declaración fiscal.

Al tomar únicamente las observaciones con ingresos por salario positivos, el universo es únicamente aquellas declaraciones que contuvieron montos de Ingresos por Salarios. La curva resultante es muy diferente (ver Gráfica II.3): un poco más igualitaria, en tanto que se acerca a la línea recta diagonal.

Gráfica II.3 Curva de Lorenz para ingresos por salarios.II.



Fuente: Elaboración propia con base en datos abiertos de SAT (2010)

El enfoque de Lorenz permitió, algunos años después, al estadístico italiano Corrado Gini en 1912 plantear una medida que reflejara el promedio de las distancias absolutas entre cada una de las observaciones respecto al ingreso total de la población (Cowell, 2009: 26) y que se equipara al área comprendida entre el ideal igualitario de Lorenz (línea recta diagonal) y la curva definida por la distribución del ingreso (curva de Lorenz). Si bien Gini planteó originalmente alrededor de 13 índices en su trabajo de 1912 (Ceriani y Verme, 2012), es el índice que a continuación se presenta el más ampliamente usado como indicador de desigualdad, el cual es llamado índice de Gini en honor a su creador⁷:

$$(Ec. II.7) \quad G = \frac{1}{2n^2 \bar{x}} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|; \forall i \neq j$$

Esta medida tiene una propiedad cardinal importante: representa menor desigualdad cuando se aproxima a 0, es decir, cuando en promedio la distancia o valor absoluto entre el ingreso de cada pareja de individuos es nula. En términos de la curva de Lorenz, la menor desigualdad se observa cuando la curva empírica y la de igualdad plena son prácticamente las mismas. Por el contrario, el índice de Gini obtiene su valor máximo cuando todos los individuos, salvo uno, no tienen ingreso

⁷ Existen múltiples expresiones del índice de Gini, sin embargo, se considera que esta expresión a pesar de ser poco intuitiva es la más sintética para su enunciación para los fines de este trabajo. Para una muestra de variantes de esta expresión, se recomienda revisar Yitzhaki S. y E. Schechtman (2013).

y solamente uno lo obtiene todo, es decir, cuando la curva es realmente solo una línea súper exponencial que acumula todo el ingreso (o riqueza) en el último porcentaje de población acumulada. En este caso, mientras n , el total de individuos de la economía, se aproxime a un valor muy grande, el máximo tenderá a 1, siendo $(n-1)/n$ el valor máximo para una población dada.

II.1.5 Medidas surgidas de la teoría del bienestar

Un grupo de medidas de desigualdad surge de los axiomas de la teoría del bienestar directamente desde la economía. Los índices que componen este grupo capturan los niveles de desigualdad del ingreso desde una perspectiva normativa (Cowell, 2005: 39-52). Presentamos, siguiendo el trabajo de Cowell (2005), los índices de Atkinson y Dalton. Específicamente, estos índices de desigualdad permiten, a partir de un nivel de aversión a la desigualdad⁸, establecer cómo se combinan el nivel de utilidad (o ingreso) promedio y la utilidad (o ingreso) ideales para establecer una medida de desigualdad.

De este modo, para un nivel del parámetro de aversión a la desigualdad ε se define el índice de Dalton como:

$$(Ec. II.8) \quad D_{\varepsilon} = 1 - \frac{\frac{1}{n} \sum_{i=1}^n [x_i^{1-\varepsilon} - 1]}{\bar{x}_i^{1-\varepsilon} - 1}$$

En el que la fracción considera la relación entre la utilidad promedio de un país dado un nivel de aversión a la desigualdad (ε) y el nivel de utilidad de un nivel de ingreso totalmente equitativo ($\bar{x}_i^{1-\varepsilon} - 1$). Esto significa que, si esta razón es igual a 1, la utilidad promedio del país es la misma que la generada de una distribución del ingreso equitativa. Por lo que el índice de Dalton será igual a 0 y el nivel de desigualdad es el deseado dada la función de utilidad social.

⁸ Para mayor detalle sobre el rol del nivel de aversión a la desigualdad ε , se recomienda revisar el apartado de Social Welfare Functions en Cowell (2009). En él, se puede observar cómo este parámetro afecta el comportamiento de las funciones de utilidad, pues representa la elasticidad del peso que tiene un individuo en el nivel de utilidad del país según su nivel de ingreso. Así, el incremento del ingreso de una persona con ingresos altos tendrá un menor impacto en el bienestar social que el incremento de una persona con niveles de pobreza extrema a medida que ε disminuye.

A su vez, Atkinson plantea un índice muy similar, el cual compara, en vez de utilidades, el ingreso equivalente (el cual genera el nivel de utilidad promedio) respecto a la media del ingreso (pensado como el ingreso equitativo), dado un nivel de aversión a la desigualdad (ε):

$$(Ec. II.9) \quad A_{\varepsilon} = 1 - \left[\frac{1}{n} \sum_{i=1}^n \left[\frac{x_i}{\bar{x}} \right]^{1-\varepsilon} \right]^{1-\varepsilon}$$

Nuevamente, este índice cambia su definición según la aversión a la desigualdad que se busque emplear.

La ventaja de estas definiciones de desigualdad es que permiten ubicar los cambios del índice dadas las propiedades de la función de utilidad empleada. Es decir, por medio de la función de utilidad elegida se capturan los juicios morales respecto a la organización social según el orden de preferencias sobre los posibles estados deseables. Una desventaja de estos índices es que recaen en definiciones estrictas de las funciones de utilidad social y por tanto de un criterio *a priori* sobre el comportamiento agregado de los individuos, que se afecta también por el nivel de aversión a la desigualdad (ε).

II.1.6 La entropía, el índice de Theil y las medidas surgidas de la teoría de la información.

Además de los momentos de la distribución de una variable, es posible plantear índices útiles para el estudio de la desigualdad a partir del concepto de entropía. Este concepto fue originalmente planteado en la física, después fue empleado en la teoría de la información por Claude Shannon en 1948 y recientemente ha encontrado aplicaciones también en el estudio de la desigualdad económica. Debido a que no es un concepto de fácil definición, a continuación, se mencionan algunos ejemplos en cada una de estas disciplinas para facilitar su comprensión. Más adelante se establece su formulación a partir de la teoría de la información, así como el índice de Theil derivado a partir de dicho concepto para la desigualdad económica, también se ofrece la generalización de medidas entrópicas planteada por Cowell (2005), la cual es ilustrativa del poder de este concepto para la generación de índices de desigualdad.

El concepto de entropía fue primeramente empleado en la física donde se utiliza como una medida del caos que existe en un sistema termodinámico. Herrera (2016) propone el siguiente ejemplo para su comprensión: al encender un cerillo se genera humo debido a la combustión; después dicho humo compuesto por partículas se difumina en el aire. La forma que adopta el humo al difuminarse refleja el caos existente en el sistema en que las moléculas de aire y de humo interactúan hacia un equilibrio. Al caos de un sistema como este, se le denomina entropía. En el caso específico del humo, la forma que adopta es difusa y fugaz, pues sus moléculas ocupan todos los espacios posibles sin un orden definido por lo que desaparece de nuestra vista. Esto ocurre debido a que cada espacio en el aire tiene la misma probabilidad de ser ocupado por el humo. De hecho, la aleatoriedad y uniformidad de probabilidades que presenta la interacción del humo con el aire es lo que caracteriza a un sistema caótico con alta entropía, en el que no existe un orden definido. En sentido contrario, la entropía presente en la interacción de las partículas de un bloque de agua congelada con el aire es menor, pues solo con la sucesiva aplicación de energía ese cuerpo podrá aumentar su entropía y fundirse en el entorno.

En el caso de la teoría de la información, la entropía no se plantea en el contexto de un equilibrio térmico sino en el de estructuras de datos y canales de transmisión. La teoría de la información busca dar soluciones a problemas como el que supone la transmisión óptima de volúmenes de información por medio de canales de transmisión dados. El tipo de soluciones que plantea, sin embargo, ha encontrado diversas aplicaciones en física, matemáticas, ingeniería, ciencias de la computación y economía y se considera ya un campo que se intercepta con dichas áreas (Cover y Thomas, 2006: 1).

La entropía en la teoría de la información establece el valor mínimo de información (usualmente expresado en *bits*) que permitiría describir un mensaje sin que todo éste sea transmitido, dada la estructura de la información que contiene. Es decir, dada la distribución de probabilidad de los valores que toma cada carácter en el mensaje, la entropía determina cuántos caracteres son los mínimos necesarios para describirlo, lo cual refleja la estructura y el caos de la información a transmitir.

Así, información con una estructura bien definida requerirá un menor número de caracteres para ser transmitida. En la teoría de la información se asume que una cadena de información puede ser expresada en *bits* (0 ó 1), por lo que cada posición en dicha cadena responde a una pregunta

dicotómica: ¿es cero o uno? Una cadena con el número óptimo de caracteres permitiría brindar la información óptima respecto a la estructura de los datos que contiene.

En otras palabras, la medida de entropía permite en esta disciplina identificar cuántas posiciones de información (o preguntas) son suficientes para describir una variable, dadas las probabilidades de ocurrencia de cada posible combinación de caracteres.

Cover y Thomas (2006: 13) plantean la siguiente definición:

La entropía $H(X)$ de una variable aleatoria discreta X se define como:

$$H(X) = -\sum_{x \in X} p(x) \log_2 p(x)$$

También llamada $H(p)$. Nótese que, debido al planteamiento de información en bits, el logaritmo de esta ecuación se encuentra en base 2. De hecho, es posible plantear este mismo concepto en términos de logaritmo natural, lo que implicaría la transmisión de *nats*.

Esta medida tiene la característica de no depender de los valores de la variable sino de su distribución de probabilidad, por lo que los efectos de traslación (suma de un valor) o multiplicación no la afectan.

Para su mayor comprensión, supongamos un caso sencillo desde este enfoque en el que en una oficina se recaba información binaria sobre los usuarios que ingresan a ella. La primera posición describiría el sexo de la persona (1=mujer, 0=hombre) y la segunda su contratación (empleado=1, no empleado=0).

Por tanto, se tienen cuatro posibles combinaciones de cadenas de bits generados por el sistema (00, 01, 10, 11). Ahora supongamos que, en dos puntos del día, se observan los siguientes escenarios. En el primer punto, la oficina recibe a casi todas sus empleadas, a muy pocos empleados y prácticamente a nadie del público, mientras que en el segundo punto se reciben todo tipo de personas debido a que se abre al público. El Cuadro II.3 resume las probabilidades de cada evento en ambos puntos del día y la aplicación de la medida de entropía (se utiliza la convención de que $-\log(p_i)(p_i) = 0$ cuando $p_i = 0$):

Cuadro II.3 Entropía en dos sistemas.

| Valores | p_i | | $-\log(p_i)(p_i)$ | |
|---------|---------|-----------------|-------------------|-------------|
| | Punto 1 | Punto 2 | Punto 1 | Punto 2 |
| 00 | 0.00 | 0.25 | 0.00 | 0.50 |
| 01 | 0.10 | 0.25 | 0.33 | 0.50 |
| 10 | 0.00 | 0.25 | 0.00 | 0.50 |
| 11 | 0.90 | 0.25 | 0.14 | 0.50 |
| | | Entropía | 0.47 | 2.00 |

Con esta medida es posible comparar qué tanta información genera el flujo de personas en dos puntos del día, dada la estructura de sus probabilidades. Recordemos que la medida de entropía nos indica las posiciones mínimas a transmitir para dar información del sistema. Tomando el punto 1, nótese que la entropía es incluso menor a 1 posición, es decir, respondiendo a la pregunta ¿es hombre o mujer, (0 o 1 en la primera posición)? es posible saber que en el primer punto del día se trata de empleados o empleadas, lo cual se anticipa por la concentración de probabilidades en los valores con segundo dígito 1. En el segundo caso, debido a que las probabilidades son homogéneas a través de todas las opciones, es necesario recabar la información de las dos posiciones (o hacer las dos preguntas) dada la distribución de las probabilidades. El segundo punto del día es más caótico respecto al primero y requiere de mayor número de bits (o respuestas) para ser descrito (incluso en este caso, informando sobre ambos caracteres no sería posible saber qué perfil específico tienen los visitantes).

El paralelismo con el uso del concepto en física se observa en que, en la teoría de la información, una alta entropía ocurre en sistemas con información compleja y con las mismas probabilidades de ocurrir. De hecho, un sistema con probabilidades homogéneas sería de máxima entropía o caótico, justo como el segundo punto del día en el ejemplo recién analizado.

El concepto de entropía refleja tan efectivamente las distribuciones de probabilidad que el profesor de las Universidades de Utrecht y Chicago, Henri Theil, propuso en 1967 su aplicación también en economía. Planteó pensar la distribución del ingreso como una distribución de probabilidades

por lo que también se podría obtener un índice entrópico de la distribución del ingreso. Así, cada probabilidad p_i en el planteamiento de la teoría de la información, pasa a pensarse como la participación del ingreso de la persona i en el ingreso nacional, s_i . También sería válido plantear esta participación como sigue:

$$s_i = \frac{x_i}{n\bar{x}}$$

Donde \bar{x} es el ingreso medio. Además, se cumple que $\sum_{i=1}^n s_i = 1$. No se admiten probabilidades negativas.

De este modo, la entropía del ingreso es el cálculo siguiente:

$$\text{entropía del ingreso} = -\sum_{i=1}^n s_i \log(s_i)$$

Sin embargo, por sí sola esta medida no permite relativizar la posición de un nivel de entropía determinado respecto al ideal de igualdad plena (máxima entropía).

Por ello, Theil sugirió el siguiente índice en el que se resta del nivel de máxima entropía de la distribución del ingreso (con participaciones homogéneas o sistema caótico), el nivel de entropía real de dicho sistema:

$$(Ec. II.10) \quad T = -\sum_{i=1}^n \frac{1}{n} \log\left(\frac{1}{n}\right) + \sum_{i=1}^n s_i \log(s_i)$$

Este índice se conoce como el índice de Theil y se observa directamente su relación con el concepto de entropía presente en las disciplinas mencionadas más arriba. Nótese que la máxima entropía queda descrita con el primer operador de la suma y que es dependiente del tamaño de la población (n). En este operador se supone una distribución igualitaria del ingreso. A su vez, el segundo operador suma plantea la inclusión de la entropía del ingreso según la distribución real del mismo.

Este mismo índice se puede expresar de las maneras siguientes:

$$T = \sum_{i=1}^n s_i \left[\log(s_i) - \log\left(\frac{1}{n}\right) \right] = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{\bar{x}} \log\left(\frac{x_i}{\bar{x}}\right)$$

Recordemos que $s_i = \frac{x_i}{n\bar{x}}$.

La medida de Theil resume la información que se desprende de la función de distribución de una variable por lo que permite realizar comparaciones entre diferentes variables sin importar su escala. Su valor máximo es $\log(n)$ (máxima desigualdad) y su valor mínimo es 0 (igualdad plena). Aunque finalmente el valor máximo depende del tamaño de la población estudiada, es posible realizar comparaciones entre variables con más o menos la misma cantidad de observaciones (incluso es posible dividir el índice por su valor máximo y obtener un índice alternativo, independiente del tamaño de la muestra, si se desea). Este indicador tiene la ventaja de que su cálculo no es un promedio de desviaciones cuadradas respecto al promedio y por tanto el impacto de transferencias entre los miembros de la sociedad tienen igual importancia tanto si ocurren entre individuos cerca o lejos del promedio, a diferencia de los indicadores de la varianza con logaritmos, que consideran mayores impactos para cambios en la zona inferior de la distribución y subrepresentan aquellos dados en la parte superior (debido a la naturaleza de la transformación logarítmica). Sin embargo, esta sensibilidad ante los cambios en la distribución se puede definir cuando se selecciona un valor específico en una clase de índices, de la cual el índice propuesto por Theil es una instancia.

A. Clase extendida de índices entrópicos susceptible a descomposición

Es importante recalcar que la medida de entropía propuesta por Theil y Cover y Thomas (2006: 13) es un caso particular de una función de valuación $h(p_i)$ en una medida del tipo $p \cdot h(p_i)$ que cumple totalmente con los siguientes requisitos necesarios para que la medición refleje el concepto de entropía:

- a. Valuación nula de la certidumbre (máxima desigualdad) $h(1) = 0$
- b. Valuación decreciente de las probabilidades: Si $p > p'$, entonces $h(p) < h(p')$
- c. Posibilidad de sumar eventos independientes: $h(pp') = h(p) + h(p')$

Esto implica la selección de la función $h(p) = -\log(p)$ la cual es la única capaz de funcionar como medición de entropía para los enfoques anteriores.

Sin embargo, la elección de esta función $h(p)$ es una entre varias posibles si no se toman en cuenta las restricciones mencionadas. Cowell (2005: 8) plantea una clase más general de índices donde utiliza un parámetro α que representa la sensibilidad que tendrá el índice a los cambios de probabilidad (transferencias de ingreso) en zonas específicas de la distribución. Si α es pequeño, el índice será más sensible a cambios en la zona de ingreso más bajo de la distribución y si α es grande será más sensible a cambios en la zona de ingresos superiores. Además, esta clase incluye un parámetro κ que representa un factor de traslación de la distribución (es decir una constante que se suma a cada valor).

Considerando que el factor de traslación determina el valor de α , es decir, $\alpha = \gamma + \beta\kappa$ y que $\alpha \in (-\infty, 1)$, $\kappa \in (0, \infty)$, $\gamma \in \mathbb{R}$ y $\beta \in \mathbb{R}_+$, Cowell (2005:8) plantea la *clase extendida intermedia susceptible a descomposición* la cual toma la forma siguiente:

$$(Ec. II.11) \quad I_{ext}^{\alpha, \kappa}(F) = \frac{1 + \kappa^2}{\alpha(\kappa)^2 - \alpha(\kappa)} \int \left[\left[\frac{x + \kappa}{\mu(F) + \kappa} \right]^{\alpha(\kappa)} - 1 \right] dF(x)$$

Donde $\mu(F)$ es la aplicación de la media aritmética a la distribución F .

Se anticipa, a partir de su nombre, que esta clase cumple con diversas exigencias teóricas deseables para medir la desigualdad:

1. Es una clase extendida que abarca diversos índices que dependen del valor asignado a cada parámetro y por tanto permite entender la elección de la sensibilidad a cambios redistributivos (o de probabilidad).
2. A diferencia del índice de Gini y los índices surgidos de la teoría del bienestar, esta clase de índices siempre puede descomponerse por grupos excluyentes que caractericen a una población (por ejemplo, zonas geográficas, sexo, entre otros atributos) (Cowell, 2005: 4-6).
3. Todos estos índices capturan el efecto de la redistribución entre individuos sin importar su ubicación en la distribución (aunque pueden generarse índices con un énfasis deseado en una parte de la misma).

4. Es independiente de la escala de la variable y de transformaciones que pueda sufrir esta (como traslaciones o multiplicaciones) aunque es sensible en algunos casos a los valores de la variable (ingresos negativos) (Cowell, 2005:6).

El poder de esta clase recae justamente en que dos casos particulares (cuando $\alpha(0) = 0$ y $\alpha(0) = 1$) son índices útiles planteados por Theil, los cuales preservan las características de esta clase.

Para el caso en que $\alpha(0) = 0$, se obtiene el siguiente indicador denominado también Desviación Logarítmica Promedio (*Mean Logarithmic Deviation, MLD*).

$$(Ec. II.12) \quad \lim_{\alpha \rightarrow 0} I_{ext}^{\alpha, \kappa=0}(F) = \frac{1}{\alpha^2 - \alpha} \int \left[\left[\frac{x}{\mu(F)} \right]^\alpha - 1 \right] dF(x) = - \int \left[\log \left[\frac{x}{\mu(F)} \right] \right] dF(x)$$

Para el caso en que $\alpha(0) = 1$, se obtiene el índice de Theil mostrado más arriba:

$$(Ec. II.13) \quad \lim_{\alpha \rightarrow 1} I_{ext}^{\alpha, \kappa=0}(F) = \frac{1}{\alpha^2 - \alpha} \int \left[\left[\frac{x}{\mu(F)} \right]^\alpha - 1 \right] dF(x) = \int \frac{x}{\mu(F)} \log \left[\frac{x}{\mu(F)} \right] dF(x)$$

La realización de estos límites requiere del uso de la regla de L'Hôpital para su cálculo. En el Anexo 1 se puede consultar la obtención de estos indicadores a partir de la fórmula aquí planteada. Bellú y Liberati (2006) consideran el uso de la regla de L'Hôpital en ambos casos, aunque a partir de una fórmula expresada en términos discretos.

II.2 Propiedades de los indicadores de desigualdad.

Una vez definidos estos indicadores y habiendo conocido sus limitantes y aciertos como indicadores de desigualdad, vale la pena remitirse a los requisitos intuitivos que plantea Cowell (2009) para la elección de una medida de desigualdad. A continuación, se enumeran y analizan los cinco requisitos definidos por dicho autor.

1) Principio débil de transferencias

Una medida de desigualdad adecuada debería reaccionar ante la transferencia de ingreso entre dos individuos. De modo que, si existieran dos individuos con una distancia positiva entre su respectivo ingreso, y si aquel individuo con mayor ingreso transfiere una cantidad proporcional de

dicha distancia al individuo de menor ingreso, el índice deberá mostrar una baja en el nivel de desigualdad, todo lo demás constante. Este requisito se denomina principio débil de transferencia. Los índices de Varianza (V), el coeficiente de variación, los coeficientes de Gini y los índices de Theil, Akitson y Dalton para aversiones positivas cumplen con este requisito (Cowell 2009: 61).

2) Independencia de escala

Bajo el supuesto de que los ingresos de los individuos crecen de manera uniforme en el tiempo y, por tanto, sus respectivas proporciones se mantienen, el índice asociado a los dos estados de la economía permanece invariante. Este requisito lo infringen los índices de Dalton y todos los índices de varianza.

3) Principio de tamaño

La desigualdad debe mantenerse invariante ante cambios en el tamaño de la población. Así, si dos economías con distribuciones del ingreso idénticas e ingreso total idénticos se mezclan, el índice debe ser el mismo antes y después de dicha mezcla. Este principio lo cumplen prácticamente todos los índices mencionados, salvo directamente el índice de Theil, pues refleja un caso particular en el que la desigualdad aumenta levemente a medida que crece la población, sin embargo, el aumento marginal es irrelevante en tanto este es cada vez menor con poblaciones más grandes, por lo que este principio se considera satisfecho por este indicador Cowell (2009: 62)

4) Posibilidad de descomposición

Al plantear una medida de desigualdad, es deseable que esta refleje los cambios en la desigualdad que ocurren entre y dentro de los grupos que componen una economía. Así, en su definición se podría establecer una relación funcional entre el nivel general de desigualdad y el existente entre y dentro de sus respectivos grupos. Cowell (2009) discute las propiedades intrínsecas de los diversos índices y concluye que el coeficiente de Gini no permite dicha descomposición para poblaciones no ordenadas mientras que las medidas surgidas de la ecuación generalizada (I_{ext}^G) sí lo permiten.

5) Principio fuerte de transferencia.

Toda medida de desigualdad que satisface simultáneamente la propiedad débil de transferencias, la descomposición, la independencia de la escala y la independencia de población (los cuatro criterios previos) debe ser expresable ya sea del modo

$$E_{\theta} = \frac{1}{\theta^2 - \theta} \left[\frac{1}{n} \sum_{i=1}^n \left[\frac{x_i}{\bar{x}} \right]^{\theta} - 1 \right]$$

O como una transformación ordinalmente equivalente de esta ecuación (E_{θ}) donde θ es un parámetro que puede tener cualquier valor real. Nótese que esta definición es justamente equivalente a la *clase de índices entrópicos extendida susceptible a descomposición* cuando $k = 0$ y se normaliza con n . Esta clase se denomina comúnmente como medidas entrópicas generalizadas $GE(k)$, Cowell (2009: 65).

El cumplimiento del principio fuerte de transferencias garantiza que un índice puede ser generado sin necesidad de recurrir a ecuaciones de bienestar, es decir, únicamente determinando un nivel de sensibilidad, α o θ , se puede obtener un índice deseable. Por otra parte, para efectos de análisis empírico, un índice de este estilo tiene propiedades cardinales⁹ deseables (está acotado entre cero y uno), aunque estas propiedades varían según el valor α o θ elegido.

Elección de medidas de desigualdad.

A partir del análisis de los indicadores más usuales relativos a la desigualdad, es posible comprender sus respectivas propiedades y elegir el indicador adecuado para su utilización como medida de desigualdad. En especial, los índices surgidos de la clase de medidas entrópicas generalizadas y particularmente el índice de Theil, cumplen con los requisitos deseables en una medida de desigualdad según lo plantea Cowell (2009), le siguen el coeficiente de Gini (aunque

⁹ La cardinalidad deseada se refiere a que el conjunto de valores que puede tomar una familia de índices E_{θ} tienen una lectura específica. Así, un valor de 1 significaría mayor o menor desigualdad dependiendo del índice que se trate. A su vez, la ordinalidad se refiere a que cada valor entre los posibles que tome un índice debe reflejar un orden coherente, de modo que dichos valores puedan representar una serie con un orden. De modo que si 0 es menor desigualdad y 1 mayor desigualdad, un valor en medio de ambos no puede representar un valor de desigualdad menor al que representa 0 o mayor al que representa 1.

este falla si el análisis se piensa realizar en términos de grupos sin orden dentro de una economía), después los índices de Atkinson y Dalton (por su dependencia a criterios extra estadísticos en su formulación) y finalmente las medidas de varianza y coeficiente de variación, que dependen de la escala y el promedio, aspectos que imposibilitan su adaptabilidad ante diferentes variables o cambios redistributivos.

II.3 Perspectivas empíricas

II.3.1 La evolución de la desigualdad en el mundo.

A continuación, se ilustrará por medio de algunos estudios de la literatura económica los usos de variables focales, de índices y métodos más comunes para el estudio de la desigualdad. Esta rama de la literatura económica es diversa en sus conceptos y métodos de medición. En el plano empírico, resalta el trabajo de Atkinson, Piketty y Saez (2011) quienes construyeron diversas series con los ingresos tributarios de veinte países para periodos tan largos que inician en el siglo XIX. Su análisis se basó en el ingreso gravable como variable focal, la cual mide los recursos recibidos debido al trabajo realizado y al capital poseído. Para cada país calcularon el porcentaje de participación del ingreso del 1% de la población en cada año disponible en el total nacional. Este es el indicador considerado como medida de desigualdad, aunque en casos necesarios proponen un índice surgido de esta noción llamado el coeficiente Pareto inverso que refleja la concentración en la parte más alta de la distribución. Encuentran que durante la primera mitad del siglo XX, la desigualdad medida con dichos indicadores cayó en cinco de los países de estudio para los que hay información, en gran medida gracias a las guerras y a la depresión económica.

Durante los últimos treinta años del siglo XX, en cambio, señalan que la desigualdad ha ido en aumento en prácticamente todos los países. Una parte de este aumento en la desigualdad se debe al crecimiento de la proporción del ingreso surgida de los salarios respecto a la proporción de los ingresos por capital. En términos geográficos, la mayor parte del crecimiento en la desigualdad se debe a la expansión de esta en Estados Unidos para los años posteriores a 1970. Debido a la naturaleza de la información no se puede determinar si este aumento está ligado a la evolución de las innovaciones en tecnología o si se debe al desempeño en otros sectores. Sin embargo, resalta la bifurcación en las experiencias dentro de cada grupo de estudio durante el siglo XX.

Para los países anglosajones los autores observan un comportamiento en forma de U: el ingreso recibido por el 1% más rico cae desde 1920 hasta los años 70 y posteriormente emprende un alza que ha persistido durante los últimos años. Para los países de Europa continental y Japón este indicador comienza con altos niveles al inicio del siglo y declina rápidamente después de 1945 para permanecer relativamente constante, esto exhibe un comportamiento en forma de L. Los países nórdicos parecen experimentar un proceso similar al de Europa continental, salvo que su decrecimiento repentino comenzó en la primera década del siglo XX para permanecer con un decrecimiento leve durante prácticamente todo el siglo (solo en los últimos años se ha observado un repunte en los niveles de ingreso recibido por el 1% más rico en estos países). Finalmente, los países en desarrollo, entre los que se encuentra solo uno latinoamericano – Argentina, muestran un comportamiento en forma de U, aunque más errático. Este último grupo vio disminuir los ingresos del top a partir de 1945 sin un mínimo muy definido. Solo a partir de los años 80 se observa un crecimiento en la desigualdad para todos los países de este grupo.

Esta evidencia coincide con el movimiento de la porción *within* (dentro del país *i*) de los índices calculados por Sala-i-Martin (2002). Aunque a nivel global este autor encuentra una disminución de la pobreza y de la desigualdad entre países, surgida principalmente por el progreso económico en China. La variable focal que utiliza es el ingreso personal deducido a partir del PIB nacional de 125 países de la base Heston, Summers y Aten (2001) para el periodo de 1970-1998.

Como caso arquetípico, Weil (2013) muestra la evolución del coeficiente de Gini para Estados Unidos en el periodo 1947-2000. Este coeficiente mantiene la tendencia mostrada por Atkinson et al (2011): la desigualdad ha incrementado consistentemente desde 1970 después de una tendencia a la baja desde antes de 1947. A su vez, este autor provee algunas explicaciones de este movimiento: el cambio tecnológico adverso al trabajo no calificado, el incremento del ingreso de personas educadas de países desarrollados dada la apertura comercial y el comportamiento de ingresos exorbitantes para personas con altas capacidades (superestrellas en una actividad específica).

Algunas lecturas relacionadas al análisis institucional de la desigualdad ligan la prevalencia de instituciones democráticas liberales al pasado colonial de cada nación Acemoglu, D., S. Johnson y J. A. Robinson (2001), o al funcionamiento efectivo de las instituciones como soluciones eficientes a las demandas civilizatorias North (1991). Una visión común desde esta perspectiva es

que los estados latinoamericanos tienen una tendencia a la profundización de las desigualdades por medio de las instituciones: para López-Calva y Lustig (2010), la desigualdad en Latinoamérica es resultado de la captura del Estado por parte de las elites, de las imperfecciones del mercado, de la desigualdad de oportunidades, de la segmentación de los mercados de trabajo y de la discriminación racial (sic) y de género, López-Calva y Lustig (2010: 2). Estos hechos son relevantes como precedentes empírico-históricos del estudio de la desigualdad en el caso de México.

II.3.2 La desigualdad en México.

Los principales análisis empíricos de la desigualdad en México muestran una tendencia a su mejora a partir de la primera década del siglo XXI. Sin embargo, los niveles de desigualdad aún son altos debido a las tendencias previas en el país. Lustig y Szekely (1997) identifican que en el periodo de 1950 a 1968 el factor que explicaba los altos niveles de desigualdad era la propiedad de la tierra. Más tarde, este proceso se vio influido por la educación de la población, Cortés (2013) identifica una reducción de la desigualdad medida con el índice de Gini debida a un crecimiento de los deciles intermedios entre 1963 y 1968. A partir de 1984 y hasta 1994 Lustig y Szekely (1997) consideran que el aumento en los niveles de desigualdad se originó por el aumento de la desigualdad salarial surgido de la transición hacia un régimen de liberalización comercial que condujo a una mayor demanda y premio del trabajo calificado. Sin embargo, esta tendencia comienza a revertirse dada la sobreoferta de trabajo calificado.

Cortés (2013) muestra dos elementos importantes a considerar en las tendencias de desigualdad en México. Por una parte, existe un movimiento cíclico del índice de Gini en los últimos cincuenta años y por otra parte, el efecto de las crisis parece coincidir con un proceso de “equidad por empobrecimiento” en el que los deciles superiores ven afectados sus ingresos y por tanto a pesar de que el impacto a la baja del ingreso es generalizado en todos los deciles, los individuos de los primeros muestran una mayor participación en el total lo cual tiene el efecto de una menor desigualdad debida a la mayor caída del ingreso en el decil superior. A partir de 2002 la desigualdad ha mantenido una tendencia a la baja. Una explicación de esta última tendencia es la disminución del premio del trabajo calificado versus el trabajo no calificado. A pesar de esta reducción, México mantiene altos niveles de desigualdad en comparación con los otros países de América Latina.

Otro rasgo a contemplar en la evolución del nivel de desigualdad en México, tiene que ver con la política social enfocada en las transferencias condicionadas y en la expansión de la cobertura educativa que se refleja en un nivel de educación más alto en los jefes de hogares pobres, PNUD (2006). Esto ha permitido reducir la pobreza agrícola rural tornando el problema más concentrado en las zonas urbanas y en el sector comercial.

A nivel regional, Székely, M., L. F. López-Calva, A. Meléndez, E. Rascón y L. Rodríguez-Chamussy (2007) proponen un índice de Theil susceptible de descomposición regional a partir de la distribución del ingreso siguiendo un proceso de imputación de dos etapas. Identifican que el estado más desigual de México es el de Querétaro, le siguen Chiapas, Oaxaca y Guerrero, aunque el Estado con mayores niveles de pobreza es el Estado de Puebla. Por otra parte, los estados con menores niveles de desigualdad son los estados de Baja California Sur, Baja California y Chihuahua. El nivel nacional, sin embargo, se determina en mayor medida por los niveles de desigualdad del Estado de México y de Ciudad de México, debido a su volumen poblacional. Resalta que no existe ningún estado al sur del país con bajos niveles de desigualdad, aunque Quintana Roo cuenta con un nivel promedio. Los estados con menores niveles de desigualdad son los del centro y norte del país.

A nivel municipal, este mismo estudio muestra que el 10% de todos los municipios cuentan con niveles altos de desigualdad, aunque el 75% tiene niveles bajos. Esto significa que algunos municipios preservan niveles de desigualdad muy importantes al interior de ellos, el municipio de San Miguel de Allende resalta como el más desigual del país. Por otra parte, la desigualdad intermunicipal añade un componente importante a la desigualdad nacional: 32.6% de la desigualdad total se explica por la desigualdad imperante entre municipios. En ausencia de esta desigualdad, los niveles totales no serían tan severos. Parte de la desigualdad entre municipios puede comprenderse por las reglas de transferencias de recursos federales que consideran el tamaño de la población, pero no necesariamente la necesidad de cada estado o municipio en términos de inversión para el desarrollo humano. De hecho, Torres (2010) realiza un análisis de los componentes del Índice de Desarrollo Humano a nivel municipal y muestra que la mayor parte de la desigualdad entre municipios depende en primer lugar del componente del ingreso y en segundo lugar del nivel educativo. El estudio PNUD (2006) muestra que el gasto social no necesariamente sigue una regla de justicia distributiva sino una fórmula en la que se considera el

tamaño de la población y no las carencias específicas de cada entidad. Es probable que parte de la desigualdad intermunicipal sea resultado del enfoque de esta política fiscal.

II.4 Conclusiones sobre los hechos y medidas de la desigualdad.

La desigualdad económica puede capturarse por medio de diferentes indicadores y cada uno aporta diferente información respecto al fenómeno. A nivel global la desigualdad ha crecido en la mayor parte de los países anglosajones, como lo muestran Atkinson, Piketty y Saez (2011). Esto puede explicarse por las recientes innovaciones en los sectores de tecnología que sustituyen diversos puestos de trabajo no calificado. Sin embargo, con los indicadores propuestos por estos autores no es fácil identificar las fuentes sectoriales de esta variación. A nivel nacional, diversos estudios apuntan a una condición institucional inicial relacionada a la concentración de propiedades que gradualmente se ha paliado con los aumentos en el nivel educativo. Este es un rasgo muy importante que resaltar pues al parecer el principal mecanismo de movilidad social en la actualidad es la educación y la política social en general. La innovación tecnológica, sin embargo, parece agravar esta tendencia a nivel global y puede hacer parecer contraproducente el logro de mejores niveles de educación si se evalúa únicamente su dimensión como capital humano. Es importante también resaltar que hay una dimensión regional de la desigualdad en México pues entre estados y entre municipios existen desigualdades importantes que podrían atenderse desde una perspectiva fiscal. La inercia a la desigualdad entre municipios permite que la desigualdad nacional sea más profunda de modo que podría coaccionar y retrasar las ganancias sociales de un mayor nivel educativo en México.

III. Capítulo 3. La desigualdad salarial en México y sus determinantes.

Hasta este punto se ha tratado el tema de la desigualdad tanto desde el punto de vista teórico-intuitivo como metodológico. Como se vio en los capítulos anteriores, las ideas sobre la desigualdad son múltiples y los índices planteados pueden ser calculados para cada variable de interés según las propiedades específicas de la información y del índice a calcular. A la luz de dichos precedentes, en este capítulo se plantea investigar si el grado de desigualdad salarial en México depende de la formación educativa de las personas, para probar las intuiciones vertidas en el capítulo 1 relacionadas a las habilidades individuales como causas de la desigualdad. Con dicho fin, se analizarán a detalle algunas variables que potencialmente determinan el nivel salarial de acuerdo con algunos estudios académicos para posteriormente pasar a su utilización en un modelo de regresión de la distribución de salarios en México. Dada la naturaleza de la pregunta causal a analizar, se presenta el enfoque contrafactual empleado en la literatura económica, el cual constituye un medio para identificar vínculos causales a partir de información no experimental. La idea es comparar dos escenarios de la realidad (uno observado y otro no observado) y, a partir de medir la variación de una variable en ambos casos, identificar el efecto causal de interés. El primero de dichos estados corresponde al caso observado, el cual será analizado por medio de un modelo de regresión por cuantiles dadas sus propiedades para capturar los efectos específicos de los regresores en diferentes puntos de la distribución de los salarios. El segundo es un escenario hipotético (o contrafactual) el cual no es observado y por ello requiere ser inferido. Esto es posible por medio del método de inferencia de escenarios contrafactuales, propuesto por Chernozhukov, Fernández-Val y Melly (2013) el cual será empleado en esta investigación. La ventaja de este método es que las estimaciones obtenidas son consistentes para calcular la mayoría de los estadísticos de desigualdad vistos en el capítulo 2. De este modo, será posible derivar conclusiones causales del análisis de regresión y dar una respuesta a la pregunta del vínculo entre la educación y el salario que surge de la intuición planteada en los capítulos anteriores. Los detalles metodológicos, así como los resultados del modelo por cuantiles y el escenario contrafactual se presentan en el cuarto capítulo.

III.1 La desigualdad salarial como una faceta de la desigualdad social.

La principal manera en que las estructuras económicas pueden transformarse y así generar cambios sustanciales en el nivel de vida para la sociedad y para las personas es por medio del empleo de

las habilidades individuales en la producción (de bienes, servicios, o ideas), siempre y cuando los mecanismos de retribución al trabajo garanticen condiciones para la obtención de un salario que permita modificar la estructura de las desigualdades sociales como la desigualdad de propiedades o en capacidades físicas, políticas o aquellas entre grupos étnicos. Solo entonces, los mecanismos últimos que definirían la movilidad social y la capacidad de dinamismo de una sociedad serían aquellos ligados a la obtención del salario y no los mecanismos que permiten la preservación de desigualdades mediante la herencia, el rango o la desigualdad en el acceso al poder y la justicia. Esta intuición surgida del capítulo 1, guía la aplicación empírica pues refleja tanto la calidad de las decisiones individuales en relación con el esfuerzo, las limitaciones que enfrentan los individuos para adquirir más habilidades, así como las decisiones de política laboral y las rigideces de la provisión de educación que impactan finalmente en el nivel salarial. A continuación, nos centraremos en identificar cuáles son los determinantes del nivel salarial de las personas con el objetivo de poder aislar correctamente el vínculo de la educación con los salarios, el cual es el interés de este trabajo. Aproximaremos la identificación ubicando los factores que inciden en el nivel de educación, los factores micro y macroeconómicos que afectan su adquisición y disponibilidad, de modo que pueda controlarse en un modelo empírico la exogeneidad de estos factores y así poder identificar el vínculo educación salario de manera más efectiva. Consideraremos también algunas variables relevantes de acuerdo con las explicaciones de la desigualdad vistas en el capítulo 1.

III.2 Revisión de la literatura.

III.2.1 Habilidades y desigualdad de ingresos y salarios.

Como puede apreciarse la elección de la variable sobre la que nos hemos planteado la pregunta sobre la desigualdad, permite una aproximación al proceso descrito por Rousseau, en el que las personas buscan diferenciarse continuamente, ahondando las diferencias en habilidades y por tanto los niveles de desigualdad.

La capacidad humana de razonar en términos abstractos y poder solucionar los problemas que representa la civilización es un aspecto medular del progreso social y tecnológico. Por ello podríamos esperar que aquellas sociedades con mejores niveles de educación alcancen mayor progreso social comparadas con aquellas sociedades que no tienen altos niveles educativos. Es por

eso por lo que también a nivel individual, como lo planteó Rousseau, podríamos esperar que las personas con mayores habilidades alcancen mayores niveles de ingreso, comparadas entre sí nacional y globalmente.

III.2.2 Crecimiento económico y educación.

Si consideramos que el progreso social se puede equiparar al crecimiento económico, podemos tomar como válida la evidencia de que la educación tiene un efecto importante en la ruta de desarrollo de los países como lo mostró Barro (1991) para 98 naciones en el periodo 1960-1985. El autor evidenció que el nivel educativo del que parten las economías explica las diferencias en tasas de crecimiento durante el periodo, lo cual, además, permite validar la hipótesis de convergencia entre países rezagados y avanzados controlando por el nivel de inversión física.

Es posible también que aquellos países con más altos niveles de ingreso sean en realidad más capaces de destinar mayores recursos al sector educativo. Incluso, como lo plantea Hanushek, E. y Woessmann, L. (2010), los estudios de regresión sobre datos internacionales como Barro (1991), Sala-i-Martin, Doppelhofer y Miller (2004), entre otros, podrían sufrir problemas de causalidad inversa o de determinación. Esto quiere decir que no podría descartarse la posibilidad de que el nivel de educación sea resultado de mayores niveles de crecimiento o bien, que no se ha determinado si es el nivel de escolaridad promedio o el cambio en dichos niveles lo que predice el nivel de crecimiento económico.

Si el nivel educativo es relevante para establecer o reflejar una ruta de crecimiento, cabría pensar entonces en cuáles son los factores que inciden en la determinación del nivel educativo de un país. Bourguignon y Verdier (2000) plantean un modelo que explica los mecanismos económicos y políticos que podrían definir tanto el nivel de crecimiento como el nivel de desigualdad en los países dinámicamente. La variable que vincula ambos fenómenos en su modelo es el nivel educativo promedio de la población, el cual, en conjunto con la estructura de la desigualdad del ingreso inicial, el retorno individual y social de la educación y el costo individual y social de educar a las personas, establecen la configuración que predice si un país será autárquico o si podrá tomar alguna ruta democrática con altos niveles de crecimiento o algún otro equilibrio.

La complejidad del modelo de Bourguignon y Verdier (2000) dificulta trasladarlo a un estudio empírico debido a la ausencia de información comparable e histórica entre países, aunque podría

ser posible su aplicación al considerar el caso de la experiencia de estados o municipios dentro de un país con suficientes datos. La dinámica planteada por los autores refleja la complejidad de interrelaciones temporales entre las instituciones que proveen educación, los individuos, las condiciones de los mercados laborales y los costos inherentes a la actividad educativa. Este esquema analítico permite considerar que a partir de decidir el nivel de educación promedio, un país decide en realidad su ruta de desarrollo, en cuya determinación intervienen no solamente el gobierno, sino también las clases medias y las instituciones que afectan los costos y beneficios sociales y privados para la educación. La relevancia del modelo es que la decisión de educación puede estar influida por aspectos bien definidos y que pueden modificarse por medio de cambios sociales como la democratización y la ampliación de la clase media, la mejora en los rendimientos privados y sociales de educar y el costo-beneficio de educarse.

III.2.3 Edad y escolaridad.

La literatura de la economía laboral permite identificar cómo el nivel educativo, el entrenamiento en el trabajo o incluso los movimientos migratorios afectan los salarios. Un hecho conocido es que mayores niveles de educación se asocian con altos niveles de ingreso y mayor progreso tecnológico como lo documentan Katz y Autor (1999: 1467-1481) para el caso de Estados Unidos. A las diferencias de ingreso por nivel educativo, se les denomina comúnmente “retorno educativo” y se interpreta como el incremento en el ingreso percibido después de estudiar un año adicional o todo un ciclo de estudios adicional. Sin embargo, las diferencias salariales así medidas no siempre toman en cuenta las diferencias etarias dentro de cada grupo de escolaridad. Por ello, para Estados Unidos, Canadá y el Reino Unido, Card D. y T. Lemieux (2001) identificaron cómo la estructura de edades explica las diferencias en ingreso de las personas con educación universitaria respecto a graduados de preparatoria. Debido a que las generaciones más jóvenes tuvieron un logro educativo relativamente menor a aquél de las generaciones anteriores, los retornos educativos (ingreso promedio relativo) de la educación universitaria se incrementaron debido a la mejora del salario en las generaciones jóvenes que compiten con menos personas en comparación con generaciones anteriores y quienes además encontraron una mayor demanda respecto a otros grupos de escolaridad. Controlar por la edad les permitió aislar los efectos de la experiencia o escasez relativa por nivel educativo dentro de cada cohorte generacional.

III.2.4 Calidad de la educación y factores que inciden en su disponibilidad: perspectivas lingüístico-culturales.

El nivel educativo de un individuo o el que caracteriza a una sociedad no lo es todo, pues es importante reconocer también su calidad y efectividad para generar mayores vínculos entre las personas (sean en la esfera económica, intercultural o social). El debate sobre la calidad de la educación es muy amplio y sobrepasa el interés de esta investigación. Una discusión suficiente sobre este concepto puede encontrarse en Blanco (2011) quien aborda el problema de la definición de calidad educativa atendiendo a diferentes argumentos teóricos y concluye en que esta puede entenderse como la capacidad del sistema educativo para asegurarse que los estudiantes realmente aprendan los contenidos del programa de estudios vigente en sus escuelas. A partir de dicho concepto, propone un análisis sobre la desigualdad educativa en términos de aprendizajes (calidades). Describe cómo las escuelas y los estudiantes en México presentan rezagos y diferencias cuando se les compara internacionalmente (*versus* otros países de la OECD mediante pruebas estandarizadas), pero también cuando se comparan entre sí, es decir, nacionalmente en términos de aprendizajes.

Algunos determinantes de la desigualdad en calidad educativa o de aprendizajes en México según Blanco (2011) son: la diversidad sociocultural y regional de la población, la cual se relaciona con las diferencias en aprendizajes en primaria, dados los programas de estudio; el tamaño del sistema educativo frente a los sistemas de gestión centralizados; la dispersión de la población en localidades muy pequeñas que reciben servicios públicos insuficientes; la clase económica del alumno así como su nivel de “capital familiar” (recursos afectivos y de seguimiento escolar brindados en casa).

El análisis de Blanco apunta a la idea de que las diferencias culturales, étnicas y económicas entre la población, además de otros factores, explican cómo nuestro país decide la provisión diferenciada de bienes públicos entre los diversos grupos que lo componen.

De hecho, este es un aspecto trascendental en la literatura sobre el cambio gubernamental¹⁰, el deseo de redistribución y la desigualdad. Si bien existe un amplio debate sobre si la educación es

¹⁰ La literatura sobre el cambio gubernamental estudia los mecanismos que hacen posible las transiciones democráticas en los países, así como los factores que inciden en el deterioro de las instituciones.

o no un bien público como lo señala Daviet (2016), aquí se resalta el rol de esta actividad no como un bien sujeto únicamente a la lógica de mercado, sino como un proceso de adecuación al entorno que puede ser financiado por el gobierno para sus ciudadanos en virtud de que es un “bien para las personas”, pues habilita al ser humano para el intercambio.

En este sentido, Baldwin y Huber (2010) analizan el efecto que tiene el grado de variación cultural y económica entre los grupos étnicos de una muestra de países en el nivel de provisión de bienes públicos. Muestran que las mediciones tradicionales que capturan únicamente la desigualdad lingüística o la desigualdad económica exclusivamente no permiten aislar el vínculo complejo de las diferencias económicas y culturales con la provisión de bienes públicos. Es decir, el grado en que una sociedad es diversa culturalmente (no solo lingüísticamente) y económicamente entre los grupos étnicos que conforman cada país explican qué tanto los gobiernos proveen bienes esenciales (bienes públicos) para el intercambio en sociedad. Como lo mencionan los autores, la diversidad cultural comúnmente se relaciona con mayores costos de transacción y menores niveles de confianza entre la población por lo que constituye una barrera natural de intercambio, más aún cuando a dichas diferencias se añaden diferencias económicas entre los grupos étnicos.

En ese mismo sentido, en 2007 Robert Putnam de la Universidad de Harvard publicó un estudio donde analiza el efecto de la diversidad cultural en los condados estadounidenses medida con un índice Herfindahl¹¹ en la preferencia por políticas de redistribución. Muestra que en los condados con mayores grados de diversidad étnica se presenta una menor disposición a la cooperación entre la población, lo cual afecta el nivel de deseos de políticas redistributivas en el corto plazo. Es decir, las diferencias económicas y lingüístico-culturales pueden ser relevantes a nivel individual y grupal para explicar la disponibilidad de recursos y calidades para fomentar las capacidades dentro de los países, por lo que constituyen factores que inciden en el logro educativo de las personas.

¹¹ El índice Herfindahl fue propuesto por Orris Herfindahl y Albert Hirschman para estudiar la concentración de poder de mercado y es ampliamente usado como un índice de concentración, donde su valor máximo es 10,000 y se obtiene por medio de la siguiente ecuación $H = \sum_{i=1}^N s_i$ donde s_i representa la cuota de mercado de la empresa i y N el total de empresas del mercado.

III.2.5 Habilidades cognitivas y no cognitivas

Una manera alternativa de conceptualizar las capacidades adquiridas en la escuela más allá del aprendizaje o nivel educativo, consiste en dividir las capacidades que tiene un individuo para solucionar problemas en dos tipos. Por una parte, las habilidades cognitivas (solución de problemas abstractos, conocimientos específicos, resultados en pruebas de inteligencia, etc.) y por otra, las habilidades no cognitivas (personalidad y carácter, entre otras). La manera en que cada individuo adquiere cada tipo de habilidades ha sido estudiada extensamente en diversos países, encontrando evidencia de efectos relevantes en los ingresos salariales de los individuos.

En el caso de México, la principal encuesta que atiende la necesidad de información en torno a este acercamiento a las capacidades individuales fue diseñada por un consorcio de universidades mexicanas y estadounidenses para capturar en un diseño longitudinal, los principales rasgos de la población y, específicamente, los rasgos relacionados con las habilidades cognitivas. Dicho estudio fue denominado Encuesta Nacional sobre los Niveles de Vida de los Hogares ENNViH y se ha realizado tres veces: 2002, 2005-2006 y 2009-2012 con representatividad nacional, urbano-rural y regional.

Al igual que en varios estudios sobre las habilidades cognitivas como en Lindqvist y Vestman (2011) y Bütikofer y Peri (2017), la ENNViH consideró una medida de inteligencia por medio de pruebas de Raven¹² adaptadas para dos grupos de edad. También en otras encuestas para grupos de edades avanzadas como el Estudio Nacional de Salud y Envejecimiento en México, esta es la medida propuesta para medir las habilidades cognitivas. De manera similar, una entrevista con un psicólogo es la principal fuente de información sobre las habilidades no cognitivas (sociabilidad, persistencia, carácter, entre otros) para Lindqvist y Vestman (2011) y Bütikofer y Peri (2017) quienes estudian estas características en poblaciones militares en Suecia y Noruega, respectivamente. La información sobre las capacidades no cognitivas no fue capturada en la ENNViH por lo que no se tiene un panorama sobre estas habilidades en la población mexicana.

¹² Las pruebas de Raven son ampliamente empleadas para medir el estado cognitivo de las personas y no requieren que los participantes cuenten con conocimientos específicos para su diagnóstico. Es posible que el lector se haya enfrentado a este tipo de pruebas que se componen de figuras o patrones geométricos que deben ser completados por medio de habilidades lógicas, por ejemplo en el examen de admisión al bachillerato de la UNAM.

Entre los elementos estudiados a partir de la información de la ENNViH, la transmisión intergeneracional de las habilidades cognitivas es de suma importancia. Mayer-Foulkes y López (2006) encuentran diferencias en la adquisición de habilidades cognitivas entre los estratos sociales pues parece existir un vínculo entre el nivel económico y los resultados en las pruebas de inteligencia, el cual se vincula con problemas de desnutrición en edades tempranas, la persistencia en tareas complejas y el desarrollo mental de los individuos. Chacón y Peña (2012) muestran cómo las horas que un niño pasa con sus padres para estudiar, el entorno y los modelos de éxito que se encuentren disponibles para los niños, y el financiamiento educativo pueden determinar el deseo de seguir estudiando y por tanto afectar la percepción de la capacidad individual. Es muy relevante la perspectiva planteada por estos autores desde el lado de la demanda de educación pues permite identificar cómo prevalecen las desigualdades en educación a partir de las restricciones con que cuentan los mismos estudiantes y que podrían ser objeto de política pública.

Esta evidencia apunta a que la aleatoriedad de las habilidades humanas al nacer, planteada por Rousseau, ya no es un supuesto válido para estudiar las condiciones actuales de la desigualdad. La constante estratificación histórica implica que no todos nacen con las mismas posibilidades de tener éxito en el plano educativo y potencialmente en su ingreso salarial, por lo que interesa investigar cómo las habilidades con que cuentan las personas se traducen en diferentes niveles de ingreso y por tanto en diferentes niveles de desigualdad dada su calidad de vida. Una variable que indicaría la calidad de vida, incluso ya realizado el nivel educativo puede ser el hacinamiento en la vivienda.

En conjunto, la literatura analizada permitirá identificar con más certeza el vínculo entre los años de estudio y el salario obtenido. Diversos factores entran en juego, pues la disponibilidad de escuelas y la calidad de la educación se enmarca en una dinámica global de diferentes ritmos de crecimiento, así como en una dinámica específica de decisiones individuales y sociales respecto al desarrollo. Asimismo, la expansión de la clase media y la consecuente expansión de la educación pública reflejan decisiones históricas en relación con la estructura de las desigualdades. En específico, un vínculo por medio del cual estas desigualdades podrían preservarse incluso ya determinado el nivel educativo es por medio de las características sociales de los individuos, como su adscripción etnolingüística, disponibilidad de modelos o concepción de las capacidades individuales las cuales pueden estar influidas por la dinámica económica de su localidad.

Asimismo, las condiciones en que los individuos pueden ejercer sus capacidades según su calidad de vida puede ser un factor determinante de su nivel de salario.

Para poder llevar estas relaciones al plano empírico y poder establecer un modelo estadístico, a continuación, se aborda el enfoque contrafactual para la identificación del vínculo causal entre dos variables. Este enfoque es relevante debido a que permitirá identificar el vínculo causal de la educación en el nivel salarial de las personas, dadas las variables recién expuestas.

III.3 Análisis causal y contrafactuales.

Las ciencias sociales en pocos contextos pueden responder preguntas bajo situaciones experimentales. Encontrar este tipo de situaciones permitiría poder diseñar un experimento ideal de modo que fuera posible identificar los efectos de variaciones de una variable en los grupos de control y tratamiento de manera repetida. Con ello, sería posible definir puntualmente los efectos causales de dicha variación en los resultados de una variable.

Debido a que este capítulo está orientado a definir un vínculo causal para probar empíricamente las intuiciones planteadas en el primer capítulo (en qué grado las habilidades explican las desigualdades salariales), interesa entonces establecer una estrategia para capturar en un escenario no experimental, el efecto causal de la variación de una variable en otra variable dependiente.

La literatura económica de la evaluación de impacto tiene una propuesta para analizar el tipo de pregunta que se plantea en esta tesis a partir de datos no experimentales. En esa corriente de análisis se busca comprender si una determinada política tiene el efecto deseado en la variable de interés. Por ejemplo, si un programa de entrenamiento en el trabajo impacta o no en el salario de quienes lo culminan (Gertler, P., et al., 2011: 33). Definir que una política causa el efecto deseado no es sencillo pues no basta con comparar el ingreso antes y después del programa o el de aquellos que participaron en él con aquellos que no, pues puede ser que incluso si el programa no existiera la variable de interés se habría comportado de manera similar (Angrist y Pischke, 2009: 12). Es ahí donde las técnicas de la evaluación de impacto juegan un rol sustancial al permitir establecer empíricamente hasta qué punto un programa en particular tiene un impacto en los resultados de la variable de interés.

Dicho impacto también llamado efecto puede trasladarse al siguiente planteamiento matemático:

$$\alpha = (Y|P = 1) - (Y|P = 0)$$

Donde α representa el impacto o efecto de la política y Y la variable de interés (en nuestro caso el salario). De este modo, el impacto queda explicado en términos de la diferencia en el resultado de una persona o unidad de observación si se emplea la política ($P=1$) respecto del resultado de esa misma persona si la política no se hubiese ejercido ($P=0$). Como puede apreciarse en realidad se comparan dos estados de la realidad alternativos, pero solamente se puede observar uno de ellos. Si fuera posible observar ambas posibilidades, en el mismo punto del tiempo, con las mismas características y sin influencia externa, podría identificarse directamente que la causa del comportamiento de la variable respuesta es justamente el empleo de la política.

Al problema de identificar el resultado no observado de la variable de interés ($Y|P=0$) para un momento del tiempo idéntico al resultado observado, se denomina problema contrafactual (Gertler, P., et al, 2011: 35). Puesto de otra manera el problema contrafactual se define como el problema de identificar cuál habría sido el resultado al no aplicar una política determinada (o al aplicarla). Dado que por definición la observación del escenario contrafactual es implausible, el problema es aproximar su existencia. Esto es posible por medio de inferir a partir de los hechos cuál habría sido el resultado en circunstancias idénticas salvo si la política en efecto no hubiese sido aplicada. Existen diferentes métodos de estimación del escenario contrafactual para poder calcular el impacto o efecto de la política.

En la estimación interviene por una parte una estrategia de identificación del escenario observado y por otra una estrategia no experimental para conocer el resultado hipotético. Una de las estrategias no experimentales más comunes busca identificar personas o grupos que sean similares uno a otro en características bien definidas en el mismo punto en el tiempo, para después aplicar una política y observar las diferencias en los resultados obtenidos. Se deben cumplir tres requisitos para que la diferencia en los resultados de ambos grupos sea atribuible a la política. Primero, dado que uno de los grupos recibiría el estímulo de una política mientras el otro no, se requiere que las diferencias entre ambos grupos o personas no sean origen de sesgo en la estimación. En segundo lugar, ambos grupos o individuos al ser sometidos al mismo estímulo deben reaccionar de la misma manera. Tercero, ninguno de ambos grupos debe tener estímulos distintos a la aplicación de la política en cuestión.

En el caso particular que se tratará en esta investigación, se empleará una estrategia distinta para inferir el escenario no observado. Esta consiste en un método de inferencia del escenario contrafactual de la distribución del salario cuando se modifica la distribución del nivel educativo de la población a partir de la metodología de Chernozhukov, Fernández-Val y Melly (2013), la cual será expuesta más adelante con mayor detalle. Por ahora interesa identificar el escenario observado ($P=0$), es decir, estimar mediante un método estadístico, la relación observada y sin política alguna del salario con las habilidades, pero también incluyendo los controles o variables explicativas ya identificadas para aislar lo más posible la relación de interés. En el cuarto capítulo se aborda la fuente de información que se empleará en esta estimación, así como los rasgos a resaltar de las variables elegidas provenientes de dicha fuente.

A continuación, se abordará la definición del modelo ideal y después el método de estimación de dicho modelo para finalmente presentar el método de inferencia propuesto por Chernozhukov, Fernández-Val y Melly (2013). El valor de este método de inferencia es que permitirá comparar la distribución de los salarios y sus respectivos índices cuando se altera la distribución del nivel educativo o su vínculo causal con el ingreso. En el cuarto capítulo se mostrarán los resultados del modelo y la estimación del escenario contrafactual, así como el impacto de la educación en los niveles de desigualdad con y sin políticas que afecten la educación.

A. Modelo ideal

De acuerdo con la revisión de la literatura se identificaron las siguientes variables que a nivel individual pueden asociarse con el nivel de retribución salarial. Este grupo de variables fue complementado con determinantes potenciales como la condición de salud, las capacidades físicas y la condición de vivienda del individuo. Se espera que las relaciones aquí capturadas sean suficientes para aislar el impacto de las habilidades intelectuales que serán aproximadas por el nivel educativo. Este modelo será el medio para capturar el escenario observado y poder inferir bajo esta misma definición, la relación no observada. La función que explica el salario es la siguiente:

$$(Ec. III.1) \quad \text{Salario} = f(E, S, C, F, EL, S, Ca, H,)$$

Donde:

E =Edad

S= Sexo

C=Capacidades físicas

F=Condición familiar

EL=Condición etnolingüística

S=Condición de salud

Ca=Calidad de vida

H=Habilidades intelectuales (Nivel educativo)

Cabe aclarar que este modelo será nutrido a partir de la Encuesta Nacional de Ingresos y Gastos de los Hogares de 2016 que constituye la fuente idónea para la observación de las variables propuestas. La definición puntual de estas variables se realiza en el capítulo IV. Se considera que el grupo de variables explicativas está compuesto de buenos controles pues estas variables son fijas o ya se han determinado al momento en que se capturó el salario mensual en la fecha de entrevista por lo que ninguna de ellas podría estar influida por la variable que buscan explicar. Así, la condición familiar que reflejaría si el individuo es padre o madre no puede estar influida por el salario mensual más reciente, tampoco lo podría estar la condición de salud del entrevistado durante el año de observación o si es o no hablante de lengua indígena, o bien la calidad de vida medida según las características de hacinamiento en su hogar. Asimismo, algunas de ellas fueron consideradas en un estudio similar para la ENIGH 2014 (Alberto, 2017). Este autor utiliza la técnica de estimación cuantílica que se usará en el capítulo IV, aunque sus variables no están construidas como se propone aquí y no realiza el análisis contrafactual que también se presenta en el capítulo IV A continuación, se detalla el método de estimación elegido para observar la relación de estas variables. Este método permitirá capturar el escenario observado y será la referencia para la inferencia del escenario contrafactual.

III.4 Regresión cuantílica para el estudio de distribuciones.

El método de estimación más sencillo de implementar para relaciones empíricas es el de Mínimos Cuadrados Ordinarios (MCO), el cual requiere de una serie de supuestos para corroborar su pertinencia según los datos disponibles. Asimismo, dicho método se adecúa a modelos que buscan explicaciones relacionadas a la función de esperanza condicional en la que las variables explicativas o controles tienen el poder de predecir *en promedio* el nivel de la variable dependiente. Sin embargo, cuando la variable modelada cuenta con una distribución continua, es posible que la respuesta promedio no refleje de manera concreta la realidad de toda la distribución ante cambios de las variables explicativas. Por ello, el método de estimación por mínimos cuadrados no necesariamente permite capturar las relaciones de interés pues solo se centra en el efecto promedio. Debido a que el interés de esta investigación recae en conocer los efectos de la desigualdad salarial a partir de los niveles de educación, el enfoque de regresión por mínimos cuadrados no resulta totalmente adecuado.

Es posible plantear un método alternativo que surge del interés de conocer el comportamiento de la variable dependiente en diferentes puntos de su distribución, dado un conjunto de regresores. Este método se conoce como regresión cuantílica y fue desarrollado primeramente por Koenker y Basset (1978). Ellos plantearon la posibilidad de utilizar el método de regresión lineal para aproximar el comportamiento de la distribución por medio de la minimización de la suma ponderada de desviaciones respecto de la media, en contraste con la minimización de errores al cuadrado que se emplea en el método de mínimos cuadrados.

Para ello definen la función de cuantil condicional, la cual refleja la distribución condicional de la variable dado el conjunto de regresores X_i (Angrist y Pischke, 2009: 270):

$$Q_\tau(Y_i | X_i) = F_y^{-1}(\tau | X_i)$$

Donde $F_y(y | X_i)$ es la función de distribución acumulada de y condicional en X_i .

Para un cuantil τ determinado, se plantea la solución del siguiente problema de estimación para Q_τ :

$$Q_{\tau}(Y_i | X_i) = \arg \min_{q(X)} E[\rho_{\tau}(Y_i - q(X))]$$

Donde $\rho_{\tau}(u) = (\tau - 1(u \leq 0))u$ representa una función de pérdida que da diferentes pesos a términos negativos y positivos dependiendo del cuantil en cuestión, y $u = Y_i - q(X)$. Asimismo $q(X_i)$ representa una función de los estimadores que puede tomar muchas formas.

Generalmente se plantea que $q(X)$ sea un modelo lineal, de modo que sea posible estimar un vector de coeficientes para cada regresor, dada la función de pérdida relativa al cuantil:

$$\beta_{\tau} \equiv \arg \min_b E[\rho_{\tau}(Y_i - X'_i b)]$$

Lo cual requiere además que $U \sim U(0,1)$ ¹³ de manera independiente de $X_i \sim F_X$. Este problema se soluciona por medio de la programación lineal y estima un modelo lineal de Y_i usando la función de pérdida $\rho_{\tau}(u)$ para un nivel τ en particular. Esto es, por medio de la regresión cuantílica se estima la relación lineal en diferentes cuantiles de la distribución de la variable de interés respecto de los diferentes regresores o variables explicativas. No existe una solución general de los coeficientes de interés como en el caso del método de mínimos cuadrados ordinarios, pues para cada programa lineal dado un cuantil τ se plantea una solución dependiendo del número de variables explicativas. Afortunadamente, los paquetes estadísticos más comunes cuentan con programas que permiten realizar esta estimación.

III.5 Inferencia del escenario contrafactual.

A continuación, interesa describir la metodología que permitirá estimar el escenario no observado ($P=1$), el cual sucedería si una política sobre la distribución de habilidades fuera aplicada. Gracias a ese escenario sería posible identificar cuánta variación experimentaría el salario y sus índices de desigualdad en presencia de dicha política y permite dimensionar y dar respuesta al vínculo causal que se intenta responder en esta investigación.

En el artículo “*Inference on counterfactual distributions*”, Chernozhukov, Fernández-Val y Melly (2013) desarrollaron un método de inferencia de distribuciones contrafactuales que es consistente

¹³ Esto significa que el vector de errores debe distribuirse con media cero y desviación estándar 1.

para los principales modelos de regresión, incluida la regresión cuantílica. Su trabajo es una aportación sustancial a la literatura de la evaluación de impacto pues abarca diversos estudios previos bajo una misma teoría asintótica que además provee regiones de confianza para las funciones estimadas. La definición teórica y sus demostraciones es amplia y muy especializada por lo que esta no se tratará aquí. En cambio, se ahondará en el planteamiento de este método como una forma de identificar los efectos de una política en una variable de interés.

Inicialmente podemos retomar el problema contrafactual de identificar cuál sería el comportamiento de una variable de interés si a una población le es aplicada una política o tratamiento específico, dado que solo puede observarse un escenario. Este problema es trasladado por Chernozhukov, et al. (2013) a una formulación muy particular en la que descomponen el problema en dos posibilidades. Consideran que una política determinada puede tomar dos formas. Primero, la política o tratamiento puede tomar la forma de una transformación en la distribución de X_i (por ejemplo, transformar la distribución del nivel educativo, nivel de ingreso, etc.) manteniendo constante la relación estimada $Q_\tau(Y_i | X_i)$. O bien, una política puede tomar la forma de un cambio en la distribución condicional de la variable dependiente $Q_\tau(Y_i | X_i)$ respecto de las características poblacionales (cómo se relacionan las variables explicativas con la variable de interés), manteniendo constante la distribución de las variables explicativas, X_i .

Estas dos posibles formas de la política son excluyentes y por tanto el análisis propuesto por ellos solo asume que ocurre una sola. Es decir, que si se transforma la distribución de las características X_i no puede al mismo tiempo considerarse la modificación de la distribución condicional de la variable dependiente $Q_\tau(Y_i | X_i)$ para efectos de medir el impacto de la primera política.

Dado que en el problema del contrafactual solo se puede observar un estado de las cosas, es necesario identificar cuál es el escenario observado y el no observado. En general, llamaremos escenario observado a aquél antes de la política y del cual se cuenta con información. En cambio, el escenario no observado es aquél al que se le aplicará una política potencial y será inferido a partir de un experimento.

| Escenario | Escenario observado | Escenario no observado |
|--------------------|-----------------------|-------------------------------|
| Política | Antes | Después |
| Experimento | $Q_{\tau}(Y_i X_i)$ | $Q_{\tau_0}(Y_i g(X_{i0}))$ |
| | | $Q_{\tau_1}(Y_i X_i)$ |

En el escenario observado el experimento es la medición del vínculo de la variable dependiente con sus variables explicativas en nuestro caso, $Q_{\tau}(Y_i | X_i)$.

Consideremos que por medio del índice $j \in (0,1)$ se puede identificar ya sea un subconjunto de la población estudiada o la misma población en otro punto del tiempo. Con $j=0$ identificamos la subpoblación de referencia o escenario observado. Mientras que cuando $j=1$ se identifica una subpoblación que dará origen a la política o un punto en el tiempo del cual extraeremos el experimento para aplicar al grupo de referencia y estimar el escenario no observado. Por ejemplo, podemos emplear las características del grupo 1, $X_{i1} \sim F_{X_0}$ en la función estimada del grupo 0, Q_{τ_0} , o alguna transformación $g(\cdot)$ para las características del mismo grupo 0, $X_{i0} \sim F_{X_0}$. Sean por ejemplo dos subpoblaciones hombres ($j=1$) y mujeres ($j=0$), al hacer este tipo de experimentos se puede estimar un escenario contrafactual de la variable dependiente cuando se emplean las características de los hombres $X_{i1} \sim F_{X_1}$ en la función específica para las mujeres Q_{τ_0} , es decir, el experimento consistiría en estimar cómo sería la distribución del salario de las mujeres si éstas tuvieran la misma distribución de características que los hombres, $X_{i1} \sim F_{X_1}$. Un escenario alternativo sería emplear la función Q_{τ_1} específica de los hombres para estimar cómo sería la distribución salarial de las mujeres cuando el mercado reacciona a sus características $X_{i0} \sim F_{X_0}$ igual que lo hace con los hombres. Estos experimentos permitirían conocer el efecto de la discriminación salarial entre hombres y mujeres. En la notación anterior F_X representa para cada matriz de variables explicativas X_i la función de distribución marginal de sus p variables.

El resultado de estos experimentos permite calcular el efecto de una política en el promedio, la varianza, y en los estadísticos de Gini, además del efecto en modelos de regresión de diversos tipos, Chernozhukov, Fernández-Val y Melly (2013: 8).

En el próximo capítulo se mostrará el tipo de experimento a emplear, así como los resultados de las estimaciones usando el comando *counterfactual* en STATA¹⁴ con el cual la estimación de los efectos específicos a los experimentos puede realizarse con relativa sencillez. Dicho comando fue desarrollado por los autores del método y está disponible en dicha paquetería para su instalación. Además, en el siguiente capítulo también se presentarán los resultados del análisis del escenario observado, las pruebas pertinentes para corroborar la validez y la estimación del escenario contrafactual así como sus efectos en los indicadores de desigualdad.

III.6 Conclusiones

La desigualdad salarial constituye un caso particular de las desigualdades sociales y puede tener un impacto sustancial en la preservación de la estructura de clases en la sociedad. De hecho, esta desigualdad refleja tanto decisiones respecto al nivel educativo promedio de un país, como aquellas decisiones individuales e institucionales que afectan la retribución al trabajo. Por ello, el estudio de la desigualdad salarial puede ser abordado a partir de la distribución de la educación entre individuos. Sin embargo, para aislar este vínculo es necesario contemplar controles que permitan aislar el vínculo de interés. En este capítulo se definieron dichos factores y se planteó la necesidad de identificar las relaciones causales entre el nivel educativo y los salarios por medio de un análisis contrafactual. Dicho análisis fue expuesto como una alternativa plausible ante los datos no experimentales disponibles en las ciencias sociales. Se mostró del mismo modo, una estrategia de identificación para estimar una distribución contrafactual a partir de la metodología de Chernozhukov, Fernández-Val y Melly (2013). Los tipos de experimentos realizados con esta metodología permiten establecer relaciones causales entre las variables de interés y observar variaciones en índices y distribuciones estimadas, con lo cual se busca dar respuesta a la pregunta

¹⁴ El programa *counterfactual* puede ser descargado en la paquetería de STATA e instalado para su ejecución. Una ventaja respecto a *tteffects* es que este último programa no admite regresiones cuantílicas para aproximar los efectos promedio de las políticas mientras que *counterfactual* fue diseñado para este tipo de regresiones. Se recomienda visitar <https://sites.google.com/site/blaisemelly/home/computer-programs/inference-on-counterfactual-distributions> para mayores detalles sobre el funcionamiento del comando.

sobre si la desigualdad en educación es determinante de la desigualdad salarial en México. El interés de definir este modelo es que permitirá establecer escenarios hipotéticos para conocer cómo sería la desigualdad de ingresos en México si las habilidades o la educación fueran distribuidas de manera diferente. En el siguiente capítulo se abordarán los aspectos prácticos de esta estimación, así como los resultados y detalles metodológicos.

IV. Capítulo 4. Análisis empírico y resultados.

IV.1 Datos sobre educación y salarios en México.

Para realizar esta investigación se accedió a los microdatos de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2016. Esta Encuesta tiene un diseño muestral probabilístico, bietápico y por conglomerados que garantiza representatividad a nivel estatal y nacional. Tiene como objetivo observar el comportamiento del ingreso y del gasto de los hogares en México a partir de aplicar diversos cuestionarios, también captura las características de la población que los habita y qué equipamiento tienen las viviendas donde se encuentran, por lo que es una fuente apropiada de información para este estudio.

En esta encuesta se calcula el ingreso corriente promedio trimestral a partir de un cuestionario de ingreso mensual en cada uno de los seis meses anteriores a la entrevista. Este cuestionario es aplicado a toda persona mayor de 12 años y permite identificar a la población económicamente activa a partir de saber si realizó alguna actividad laboral durante el mes anterior a la entrevista. También captura las horas trabajadas a la semana durante el último mes en uno o varios empleos, así como el sector económico en que se desempeña, si su actividad laboral le generó ingresos como empleado subordinado o bien a partir de un negocio propio, otro trabajo u otra retribución en especie. Entre las variables sociodemográficas capturadas para cada integrante de la familia, se pregunta si la persona habla alguna lengua distinta del español, su auto adscripción étnica, si sufre o no uno o más tipos de discapacidad, así como las causas de esta, su nivel educativo, si asiste o no a la escuela, en qué grado y la ubicación geográfica y el tamaño de la localidad donde se encuentra el hogar, entre otras.

A. Ingreso corriente trimestral por deciles de personas.

A partir de los microdatos publicados en INEGI para esta encuesta se obtuvieron las características de los residentes de cada hogar, con los cuales se pudo estimar por deciles de personas el ingreso corriente trimestral total en 2016, así como su descomposición que puede observarse en el Cuadro IV.1:

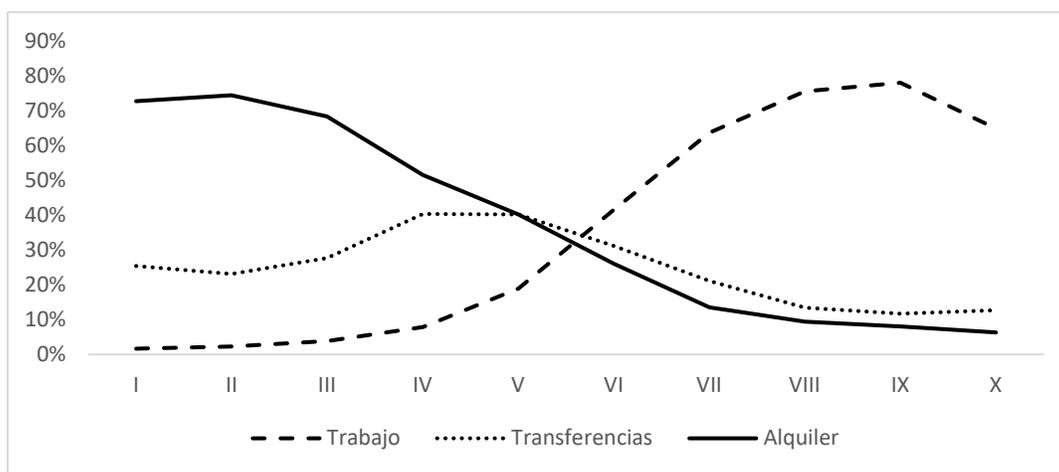
Cuadro IV.1 Composición del ingreso corriente trimestral por fuente de ingreso por deciles de personas, ENIGH 2016. Miles de pesos y porcentajes.

| Decil | % Ingreso corr. trim. por decil respecto del total | Edad media (años) | Componentes | | | | |
|--------------|--|-------------------|--------------------------|-------------------------|---------------------------------|-------------------------|--------------------|
| | | | % Ing. por trabajo decil | % Ing. por rentas decil | % Ing. por transferencias decil | % Estim. Alquiler decil | % Otros ing. decil |
| I | 0.2% | 15 | 1.7% | 0.0% | 25.4% | 72.8% | 0.1% |
| II | 0.6% | 16 | 2.3% | 0.0% | 23.1% | 74.5% | 0.1% |
| III | 1.0% | 19 | 3.9% | 0.1% | 27.7% | 68.4% | 0.0% |
| IV | 1.6% | 26 | 7.8% | 0.1% | 40.4% | 51.6% | 0.1% |
| V | 2.6% | 32 | 18.9% | 0.5% | 40.3% | 40.2% | 0.1% |
| VI | 4.7% | 38 | 41.7% | 0.9% | 31.2% | 26.1% | 0.2% |
| VII | 8.2% | 41 | 63.7% | 1.4% | 21.1% | 13.6% | 0.2% |
| VIII | 12.2% | 40 | 75.6% | 1.4% | 13.5% | 9.5% | 0.1% |
| IX | 18.3% | 41 | 78.1% | 2.1% | 11.7% | 8.0% | 0.1% |
| X | 50.6% | 44 | 65.0% | 15.9% | 12.7% | 6.3% | 0.1% |
| Total | 1,556,701,245.07 | 31 | 64.3% | 8.8% | 15.6% | 11.3% | 0.1% |

La estimación considera una población de 122,643,890 personas, a partir de 257,658 observaciones.
Fuente: Microdatos ENIGH 2016

Resalta de esta estimación el hecho de que los componentes más importantes del ingreso corriente en cada decil rivalizan entre la estimación del alquiler de la vivienda que, para deciles inferiores, consiste en hasta el 73% de su ingreso total y que disminuye para los deciles más altos a solo 6%, y el ingreso por trabajo que para deciles inferiores compone únicamente el 2% de su ingreso total, mientras que para el noveno decil este rubro compone el 78% del total (el décimo decil obtiene 16% de su ingreso por rentas de la propiedad por lo que el efecto de los ingresos por trabajo en el total de dicho decil se diluye). Esto puede apreciarse claramente en la Gráfica IV.1.

Gráfica IV.1 Participación de ingresos por trabajo, transferencias y estimación de alquiler en el total del decil por deciles de personas, ENIGH 2016.



Fuente: Estimación con microdatos de la ENIGH 2016.

A partir de este primer análisis, se pueden identificar dos aspectos importantes de la distribución de ingresos entre deciles de personas: los ingresos por trabajo reflejan la composición etaria de la población (pues los totales de ingreso se asocian con la edad promedio del decil: los niños pueden recibir una subvención importante en términos de alquiler de la vivienda como ingreso). Sin embargo, al restringir la observación a los deciles III en adelante donde los ingresos por trabajo podrían explicar una parte más importante del ingreso, se observa una curva creciente en la que los ingresos por trabajo expanden su predominio respecto de otras formas de ingreso.

Asimismo, en el Cuadro IV.2 puede apreciarse la composición del ingreso individual promedio del decil calculado como el total de cada tipo de ingreso entre la décima parte de la población. De nueva cuenta puede observarse la importancia de los ingresos por trabajo en los deciles superiores al V. De hecho en términos nominales, el ingreso corriente trimestral por trabajo se duplica pasando del decil V al VI. Asimismo, se observa la concentración del ingreso por rentas en el último decil, el cual prácticamente detenta el total de este tipo de ingreso. En la mayor parte de los casos, el ingreso promedio por tipo de ingreso se incrementa en más del doble pasando del decil IX al decil X. Esto refleja la estructura desigual de los ingresos sin importar el tipo de ingreso que interese estudiar.

Cuadro IV.2 Contribución del ingreso corriente trimestral por persona, por deciles de personas, ENIGH 2016. Miles de pesos.

| Decil | Ingreso Corriente Trimestral | Componentes | | | | |
|----------------|------------------------------|----------------------|---------------------|-----------------------------|-------------------------|----------------|
| | | Ingresos por trabajo | Ingresos por rentas | Ingresos por transferencias | Estimación del alquiler | Otros ingresos |
| I | 0.22 | 0.00 | 0.00 | 0.06 | 0.16 | 0.00 |
| II | 0.76 | 0.02 | 0.00 | 0.18 | 0.56 | 0.00 |
| III | 1.27 | 0.05 | 0.00 | 0.35 | 0.87 | 0.00 |
| IV | 2.00 | 0.16 | 0.00 | 0.81 | 1.03 | 0.00 |
| V | 3.27 | 0.62 | 0.01 | 1.32 | 1.32 | 0.00 |
| VI | 6.00 | 2.50 | 0.05 | 1.87 | 1.57 | 0.01 |
| VII | 10.38 | 6.61 | 0.15 | 2.19 | 1.41 | 0.02 |
| VIII | 15.51 | 11.73 | 0.21 | 2.09 | 1.47 | 0.01 |
| IX | 23.23 | 18.14 | 0.48 | 2.72 | 1.87 | 0.02 |
| X | 64.28 | 41.76 | 10.24 | 8.17 | 4.06 | 0.04 |
| Totales | 126.93 | 81.60 | 11.15 | 19.75 | 14.32 | 0.11 |

La estimación considera una población de 122,643,890 personas, a partir de 257,658 observaciones. Cada decil representa 12,264,389 personas.
Fuente: Microdatos ENIGH 2016

IV.2 Ingreso por trabajo y definición de salario.

Debido a que es el ingreso surgido del trabajo el interés de esta investigación, centraremos nuestra atención en identificar los factores que inciden en la determinación de su nivel y distribución por medio del cálculo del salario por hora percibido. En adelante se restringe el análisis al conjunto de personas no huéspedes o empleados domésticos (identificados por medio de la variable parentesco con valores 401-461 y 701-715) cuya edad está en el rango de 15 a 64 años y que declaró haber trabajado al menos 30 horas en la semana inmediata anterior a la entrevista. A este conjunto de personas, le fue calculado individualmente el salario promedio por hora de acuerdo con la definición empleada por Airola y Juhn (2005) quienes definen el salario por hora para el estudio de la desigualdad salarial en México a partir de una serie histórica de diversas ENIGH.

Dicha aproximación fue empleada en este trabajo debido a su consistencia con la estructura de información de la encuesta. Para cada persona en la muestra, se calculó la suma de su ingreso mensual por trabajo (ingresos por uno o varios trabajos subordinados, por negocios propios u otros trabajos percibidos en el mes inmediato anterior a la entrevista), así como el total de horas trabajadas en la semana en todos los trabajos reportados durante el mes en cuestión. Con esta información se obtuvo el salario por hora a nivel individuo dividiendo el ingreso por trabajo total mensual por las horas trabajadas a la semana multiplicadas por 4.33 (semanas promedio por mes):

$$\text{salario por hora} = \frac{\text{ing por trabajo}_{\text{mensual}}}{\text{horas trabajadas}_{\text{semana}} * 4.33}$$

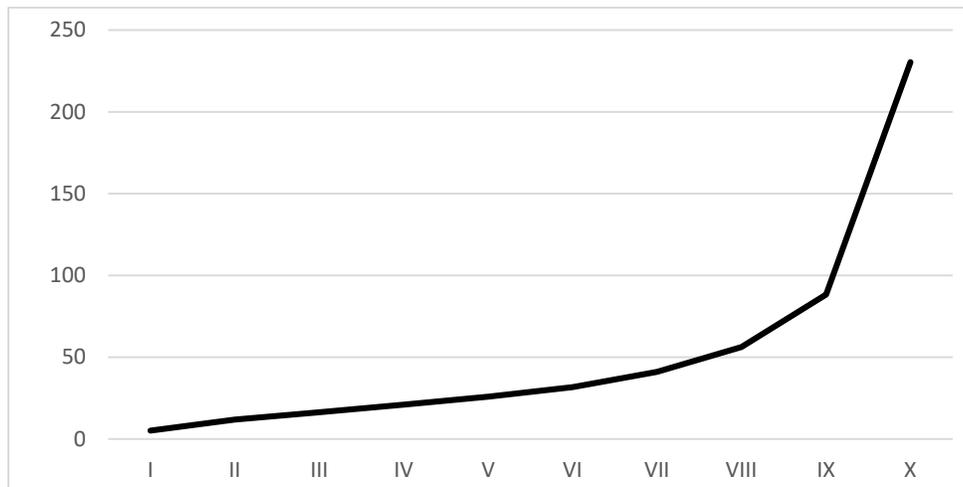
La distribución de esta variable se muestra a continuación en deciles de personas de salario por hora:

Cuadro IV.3 Estimación del salario por hora promedio por decil de personas entre 15 y 64 años que trabajaron al menos 30 horas a la semana. Pesos por hora.

| Estimación del salario por hora promedio por decil | | | | |
|---|--------------|--------------------|-----------------------|--------------|
| Decil | Media | Err. Están. | Int. 95% Conf. | |
| I | 5.14 | 0.04 | 5.05 | 5.23 |
| II | 11.84 | 0.03 | 11.79 | 11.89 |
| III | 16.45 | 0.02 | 16.41 | 16.49 |
| IV | 21.03 | 0.02 | 20.98 | 21.07 |
| V | 25.82 | 0.02 | 25.77 | 25.87 |
| VI | 31.81 | 0.04 | 31.73 | 31.89 |
| VII | 40.99 | 0.05 | 40.89 | 41.1 |
| VIII | 56.18 | 0.09 | 56 | 56.37 |
| IX | 88.23 | 0.23 | 87.78 | 88.69 |
| X | 230.39 | 3.45 | 223.63 | 237.15 |
| Total | 52.53 | 0.6 | 51.35 | 53.71 |
| Tamaño de la muestra: 41,385,988. | | | | |
| Fuente: Microdatos ENIGH 2016 | | | | |

La distribución del salario promedio por hora por decil puede apreciarse en la Gráfica IV.2:

Gráfica IV.2 Salario por hora promedio por decil para la población de 15 a 64 años, ENIGH 2016. Pesos.



Fuente: Microdatos ENIGH 2016.

Debido a que el salario promedio de los deciles superiores puede presentar alta variabilidad, se reporta a continuación el salario promedio en los percentiles 90 a 99:

Cuadro IV.4 Estimación del salario promedio en deciles superiores.

| Estimación del salario por hora promedio por percentiles superiores | | | | |
|--|--------------|--------------------|-----------------------|--------|
| Percentil | Media | Err. Están. | Int. 95% Conf. | |
| 90 | 112.63 | 0.11 | 112.41 | 112.85 |
| 91 | 119.90 | 0.13 | 119.64 | 120.15 |
| 92 | 128.85 | 0.16 | 128.54 | 129.16 |
| 93 | 140.60 | 0.19 | 140.23 | 140.97 |
| 94 | 153.62 | 0.20 | 153.23 | 154.01 |
| 95 | 168.98 | 0.25 | 168.49 | 169.47 |
| 96 | 190.20 | 0.38 | 189.46 | 190.95 |
| 97 | 217.00 | 0.47 | 216.07 | 217.93 |
| 98 | 255.26 | 0.80 | 253.70 | 256.82 |
| 99 | 319.19 | 1.46 | 316.33 | 322.06 |
| 100 | 598.94 | 19.43 | 560.86 | 637.03 |

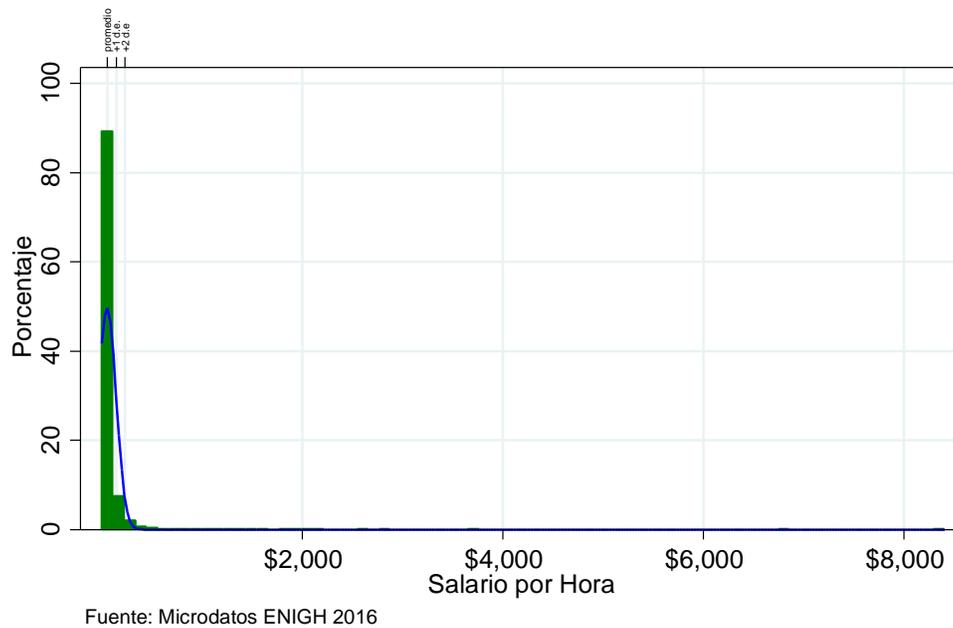
Tamaño de la muestra: 41,385,988.

Fuente: Microdatos ENIGH 2016

IV.3 Estadísticos de desigualdad para la distribución del salario por hora.

Al estimar la función de distribución del salario por hora, aplicando factores de expansión, observamos la estructura desigual de esta variable:

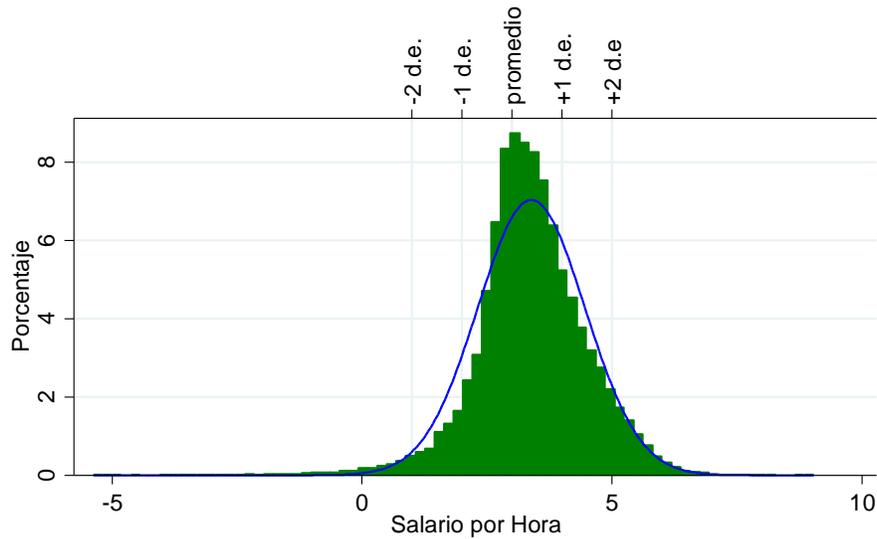
Gráfica IV.3 Distribución del salario por hora



Por lo que se aplica una transformación logarítmica para observar con más detalle su comportamiento¹⁵:

¹⁵ Esta misma transformación es aplicada por Alberto (2017) para el caso de la ENIGH 2014.

Gráfica IV.4 Distribución del logaritmo del salario por hora



Fuente: Microdatos ENIGH 2016

Se procedió a realizar el cálculo de los índices planteados en el primer capítulo para la distribución salarial. Las respectivas estimaciones se muestran en el Cuadro IV. 5 y consideran el diseño muestral de la ENIGH, en especial los factores de expansión¹⁶. Para los indicadores de varianza, varianza logarítmica v , varianza de los logaritmos v_l y coeficiente de variación se realizó el cálculo de manera rudimentaria, considerando la estimación de varianza, medias aritmética y geométrica para la distribución, después creando las variables al cuadrado y estimando el total para poder dividir entre el total de la población de estudio. Cabe aclarar que estos índices se realizaron para la distribución salarial sin transformación logarítmica debido a la potencial pérdida de datos al calcular logaritmos sobre valores negativos. Los índices de Gini, Atkinson, Theil y Desviación Logarítmica promedio fueron calculados por medio del comando *ineqdc* de STATA¹⁷.

¹⁶ Los factores de expansión constituyen una parte fundamental del diseño de la encuesta y por medio de ellos se puede conocer cuántas unidades poblacionales representa cada observación de la encuesta. Así, si una observación tiene asignado un valor de 200, eso significa que esa observación representa a 200 individuos de la población estimada.

¹⁷ Disponible en STATA para su instalación.

Cuadro IV.5 Indicadores de desigualdad para la distribución del salario por hora, ENIGH 2016.

| Indicador | Salario por hora |
|--|-------------------------|
| Media aritmética (Ec. II.1) | 52.53 |
| Media geométrica (Ec. II.2) | 29.62 |
| Varianza (Ec. II.3) | 7,887.70 |
| Varianza logarítmica (Ec. II.4) | 1.49 |
| Varianza de logaritmos (Ec. II.5) | 1.16 |
| Coefficiente de variación (Ec. II.6) | 1.69 |
| Coefficiente de Gini (Ec. II.7) | 0.55154 |
| Índice Atkinson, e=0.5 (Ec. II.9) | 0.25247 |
| Índice Atkinson, e=1 (Ec. II.9) | 0.43608 |
| Índice Atkinson, e=2 (Ec. II.9) | 0.79241 |
| Índice de Theil, GE(1) (Ec. II.13) | 0.60342 |
| Desviación Logarítmica promedio, GE(0) (Ec. II.12) | 0.57284 |
| 1/2 CV ² , GE(2) (Ec. II.11) | 1.42923 |
| Total de la población | 41,385,988 |

Fuente: Cálculos propios a partir de Microdatos de la ENIGH 2016.

Resalta de estos cálculos que todos muestran altos niveles de desigualdad salarial. El hecho de que la media geométrica disminuya a casi la mitad del cálculo de la media aritmética refleja la presencia de una cantidad importante de valores que son muy pequeños y que disminuyen el producto de todas las observaciones, sobre el que se calcula la raíz enésima. Asimismo, los indicadores de varianza y coeficiente de variación muestran que el salario promedio es poco representativo de la distribución total. El coeficiente de Gini muestra un mayor acercamiento hacia su valor máximo que hacia su valor mínimo, por lo que se considera que existe más desigualdad que igualdad en la distribución de estudio. Los índices de Atkinson muestran un crecimiento a medida que se ponderan con mayor fuerza los ingresos más bajos, es decir, a medida que crece, también lo hace la desigualdad reportada, por lo que los ingresos más bajos son la principal fuente de desigualdad en esta distribución. También los indicadores entrópicos generalizados, como el índice de Theil y la Desviación Logarítmica promedio reflejan niveles de desigualdad elevados pues ambos muestran un alejamiento del nivel de igualdad plena.

A continuación, se analiza la estructura de las variables explicativas elegidas y sus respectivas definiciones, para posteriormente observar cómo se relacionan con la distribución de interés.

Después de considerar el modelo propuesto, estos índices podrán ser empleados para corroborar la idea de que la educación determina el salario por hora de las personas.

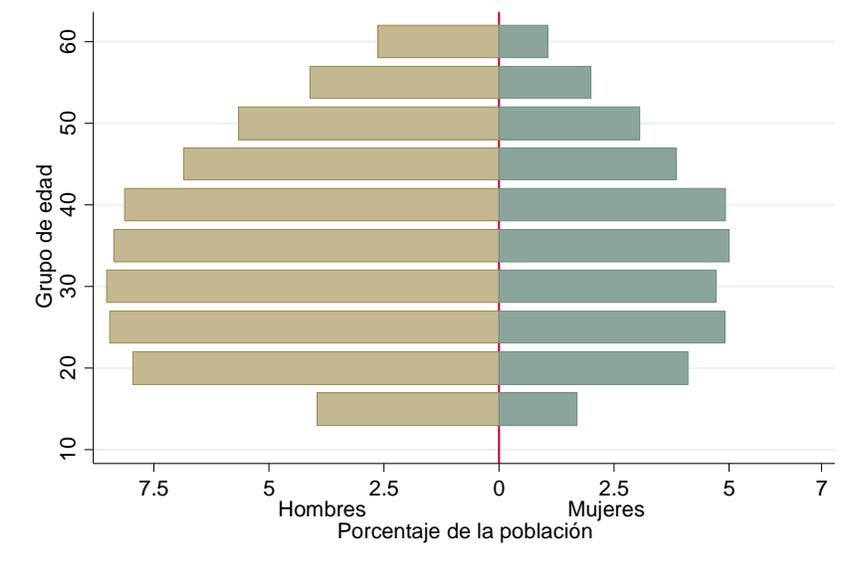
IV.4 Definición de las variables a partir de la ENIGH. Análisis descriptivo.

En el tercer capítulo se hizo la revisión de la literatura correspondiente y se concluyó que existirían 8 grupos de variables explicativas (Edad, Sexo, Habilidades intelectuales (Nivel educativo), Capacidades físicas, Condición familiar, Condición etnolingüística, Condición de salud y Calidad de vida, ver Ecuación III.1) por lo que en esta sección se definirá con mayor detalle la fuente de información empleada para cada variable, su definición, así como un resumen estadístico de la información que representan.

A. Edad

La variable edad fue obtenida de la tabla PERSONA de los microdatos de la ENIGH. La variable sexo fue transformada de modo que tuviera el valor 1 en el caso de hombres y 0 en el caso de que el individuo sea mujer. En la Gráfica IV.5 se muestra la composición etaria por sexo de la población objeto de estudio:

Gráfica IV.5 Estructura etaria de la población objeto de estudio, por sexo.



Fuente: Estimación propia con microdatos de la ENIGH 2016.

Considerando el factor de expansión, el 64.67% de la población en estudio son hombres cuya edad promedio es de 37 años. Asimismo, su salario promedio por hora es de 52.88 pesos que en promedio obtienen con 54 horas trabajadas a la semana. En cambio, las horas de trabajo promedio de las mujeres ascienden a 49 y su salario promedio es de 51.87. Ambas variables resultan relevantes en el análisis debido a que tanto la edad como el sexo representan características individuales que son consideradas en el mercado de trabajo y por tanto pueden incidir en el nivel salarial. Es de notar tanto la menor participación de las mujeres en la población de estudio, como en el tiempo de trabajo dedicado a obtener ingresos. Esto puede deberse a la división del trabajo en el hogar que en 2016 refleja aún rezagos en términos de igualdad por género.

B. Capacidades físicas

Las capacidades físicas de un individuo pueden medirse de diversas formas (presencia de enfermedades degenerativas, diabetes, problemas de alimentación, nutrición, etc.). Sin embargo y debido a que no se cuenta con una variable específica para este tipo de factores, se restringirá la observación a la presencia o no de al menos un tipo de discapacidad, pues es una de las variables registradas en los cuestionarios individuales. El Instituto Nacional de Estadística, Geografía e Informática pregunta si el individuo presenta al menos un tipo de discapacidad, así como sus causas. A partir de esta encuesta pudo determinarse que 3.8% de las personas de la población de estudio presentaron al menos un tipo de discapacidad. La composición de esta población por tipo y causa reportada se muestra en el Cuadro IV.6:

Cuadro IV.6 Incidencia de discapacidad entre la población de estudio, por tipos y causas reportadas. ENIGH 2016.

| Tipo de discapacidad | Causa | | | | | Total |
|---|----------------|----------------|----------------|----------------|---------------|------------------|
| | Enfermedad | Accidente | Nacimiento | Edad avanzada | Otra causa | |
| Caminar, moverse, subir o bajar | 341,280 | 302,620 | 38,228 | 69,759 | 40,057 | 791,944 |
| Ver, aun usando lentes | 267,423 | 71,635 | 70,957 | 89,894 | 36,097 | 536,006 |
| Oír, aun usando aparato auditivo | 46,985 | 43,416 | 20,239 | 17,186 | 12,393 | 140,219 |
| Hablar, comunicarse o conversar | 9,118 | 10,977 | 39,087 | 366 | 985 | 60,533 |
| Poner atención o aprender cosas sencillas | 4,820 | 6,124 | 9,081 | 2,541 | 2,963 | 25,529 |
| Tiene alguna limitación mental | 4,774 | 4,660 | 13,775 | 0 | 1,920 | 25,129 |
| Vestirse, bañarse o comer | 751 | 12,969 | 462 | 231 | 0 | 14,413 |
| Total | 675,151 | 452,401 | 191,829 | 179,977 | 94,415 | 1,593,773 |

Fuente: Elaboración propia a partir de Microdatos de la ENIGH 2016.

Resalta de este cuadro que la principal razón de discapacidad física proviene de discapacidad en movilidad debido a accidentes o enfermedades. Entre las causas de discapacidad surgidas exclusivamente de nacimiento se encuentra principalmente la discapacidad visual la cual también es consecuencia de la edad avanzada. Para efectos de esta investigación únicamente consideraremos la presencia o no de al menos un tipo de discapacidad por medio de una variable dicotómica llamada *discap* que tomará el valor de 1 cuando existe discapacidad en el individuo y 0 cuando no.

C. Condición familiar.

Dado el rol fundamental que implica la paternidad en la necesidad de subsistencia de los individuos y sus respectivas familias, se incluye un control por medio de una variable dicotómica que toma el valor de 1 cuando el individuo declara ser padre o madre. De la población objeto de estudio, el

61.92% reporta ser padre o madre. En el Cuadro IV.7 se presenta un detalle del comportamiento de la variable respecto a diferentes cohortes de edad. Resalta el incremento de la incidencia en los grupos de edad posteriores a los 25 años, donde el porcentaje aumenta a 51%.

Cuadro IV.7 Incidencia de paternidad o maternidad en la población de estudio. ENIGH 2016.

| Grupo de edad | Es padre o madre | | | | Total |
|---------------|-------------------|------------|-------------------|------------|-------------------|
| | No | | Sí | | |
| 15 a 20 | 2,126,526 | 91% | 214,560 | 9% | 2,341,086 |
| 20 a 25 | 3,394,804 | 68% | 1,598,988 | 32% | 4,993,792 |
| 25 a 30 | 2,716,041 | 49% | 2,823,444 | 51% | 5,539,485 |
| 30 a 35 | 1,727,521 | 32% | 3,751,967 | 68% | 5,479,488 |
| 35 a 40 | 1,268,471 | 23% | 4,264,324 | 77% | 5,532,795 |
| 40 a 45 | 1,091,934 | 20% | 4,314,051 | 80% | 5,405,985 |
| 45 a 50 | 965,173 | 22% | 3,464,254 | 78% | 4,429,427 |
| 50 a 55 | 987,700 | 27% | 2,619,146 | 73% | 3,606,846 |
| 55 a 60 | 873,200 | 35% | 1,654,901 | 65% | 2,528,101 |
| 60 a 64 | 607,983 | 40% | 921,000 | 60% | 1,528,983 |
| Total | 15,759,353 | 38% | 25,626,635 | 62% | 41,385,988 |

Fuente: Microdatos de la ENIGH 2016

D. Condición etnolingüística.

Como se mencionó más arriba, la condición etnolingüística interviene en la adquisición y disponibilidad de recursos públicos cuando las sociedades se encuentran altamente segmentadas por estos factores. Los estudios que analizan este tipo de efectos consideran alguna variable que capture la similitud entre las diversas lenguas que hablan los individuos, de modo que dos individuos que hablan lenguas totalmente disimiles como el japonés y el español reciban un valor más alto de disimilitud en comparación con aquellos que hablan español y catalán, por ejemplo, debido a que estas dos últimas comparten mayores ramas lingüísticas. Este es el caso de Baldwin y Huber (2010) quienes emplean una medida de similitud lingüística basándose en Fearon (2003). Debido a que en el caso de los datos de la ENIGH se reportan dos variables que reflejan esta información no se procedió a realizar una medida similar debido a la diversidad de lenguas habladas entre los individuos y a la alta incidencia de bilingüismo respecto al español (correlación de 0.97 con hablante de lengua indígena). Se consideró en cambio, una variable dicotómica para la presencia de lengua indígena (`lenguaindig`) que toma el valor de 1 cuando el individuo es

hablante de la lengua y 0 cuando no. Asimismo, se consideró una variable dicotómica con información de la autoadscripción étnica, que en el caso de la ENIGH considera únicamente dos opciones de adscripción (índigena o no). Sorprendentemente esta variable no tiene correlación con la incidencia de lengua indígena, por lo que se mantiene independiente y se considera en esta investigación por medio de una variable dicotómica que toma el valor de 1 cuando la adscripción es indígena y 0 cuando no. En el Cuadro IV.8 se muestra un detalle de las lenguas con mayor incidencia, así como la incidencia de bilingüismo respecto al español entre la población de estudio. Solamente los hablantes de Tzotzil y Tzeltal presentan tasas de bilingüismo menores al 90%. En el Anexo 3 puede consultarse la lista completa de Lenguas y su incidencia con el español.

Cuadro IV.8 Incidencia de bilingüismo por lengua indígena hablada, ENIGH 2016.

| LENGUA | CODIGO* | NO BILINGÜE | | BILINGÜE | | TOTAL |
|--------------------|---------|---------------|-----------|------------------|------------|------------------|
| | | | | | | |
| NÁHUATL | 1041 | 4,766 | 1% | 606,808 | 99% | 611,574 |
| MAYA | 922 | 4,529 | 1% | 313,442 | 99% | 317,971 |
| TZOTZIL | 934 | 34,229 | 19% | 143,675 | 81% | 177,904 |
| TZETZAL | 933 | 14,748 | 12% | 107,449 | 88% | 122,197 |
| OTOMÍ | 331 | 1,130 | 1% | 98,033 | 99% | 99,163 |
| TOTONACA | 711 | 3,036 | 3% | 92,852 | 97% | 95,888 |
| MIXTECO | 450 | 4,000 | 6% | 65,646 | 94% | 69,646 |
| ZAPOTECO DEL ISTMO | 422 | 630 | 1% | 64,390 | 99% | 65,020 |
| MAZTECO | 491 | 4,698 | 8% | 57,055 | 92% | 61,753 |
| MIXE | 811 | 2,513 | 4% | 58,316 | 96% | 60,829 |
| OTROS | | 18,129 | 3% | 622,791 | 97% | 640,920 |
| Totales | | 92,408 | 4% | 2,230,457 | 96% | 2,322,865 |

*Clasificación de Lenguas Indígenas, 2010

Fuente: Microdatos ENIGH 2016.

E. Condición de salud.

Para complementar el grupo de controles relativos a las capacidades físicas, se considera una variable adicional que refleja la salud del individuo como control para determinar su posible impacto en el salario percibido. Esta variable fue construida a partir de la pregunta sobre el año en que el individuo reportó haber tenido un dolor, malestar, enfermedad o accidente que le impidiera realizar sus actividades cotidianas (prob_anio). Para el caso de esta investigación, se generó a partir del año reportado una variable dicotómica (salud) que toma el valor de 1 cuando el año es 2016 y 0 en cualquier otro caso. Se considera únicamente los casos dentro del año en cuestión debido a

que el impacto de salud mayor a un año es potencialmente de menor relevancia para los ingresos percibidos durante el año del evento. Se estimó que solamente el 42.1% de la población presentó algún tipo de problema de este tipo durante 2016. En el Cuadro IV.9 se presenta la incidencia de problemas de salud en 2016 por grupos de edad:

Cuadro IV.9 Incidencia de problemas de salud incapacitante en el año de la entrevista.

| Grupo de edad | Problema de Salud en 2016 | | | | Total |
|---------------|---------------------------|------------|-------------------|------------|-------------------|
| | No | | Sí | | |
| 15 a 20 | 1,165,493 | 50% | 1,175,593 | 50% | 2,341,086 |
| 20 a 25 | 2,673,729 | 54% | 2,320,063 | 46% | 4,993,792 |
| 25 a 30 | 3,194,446 | 58% | 2,345,039 | 42% | 5,539,485 |
| 30 a 35 | 3,247,561 | 59% | 2,231,927 | 41% | 5,479,488 |
| 35 a 40 | 3,266,783 | 59% | 2,266,012 | 41% | 5,532,795 |
| 40 a 45 | 3,242,178 | 60% | 2,163,807 | 40% | 5,405,985 |
| 45 a 50 | 2,586,857 | 58% | 1,842,570 | 42% | 4,429,427 |
| 50 a 55 | 2,172,287 | 60% | 1,434,559 | 40% | 3,606,846 |
| 55 a 60 | 1,482,665 | 59% | 1,045,436 | 41% | 2,528,101 |
| 60 a 65 | 926,244 | 61% | 602,739 | 39% | 1,528,983 |
| Total | 23,958,243 | 58% | 17,427,745 | 42% | 41,385,988 |

Fuente: Microdatos de la ENIGH 2016

F. Calidad de vida.

Un grupo de variables de control relativo a la calidad de vida de los individuos debería tener en cuenta diversos factores que podrían incidir en el salario percibido. Entre dichos factores, de manera muy importante se encuentra el tiempo de traslado al centro de trabajo o bien, el tipo de transporte empleado o el tiempo de traslado a centros de esparcimiento o acceso a bienes públicos como parques, centros culturales, entre otros. Debido a que este tipo de variables no son capturadas directamente en la ENIGH, se procedió a incluir únicamente un control relativo a la calidad de vida dentro de los hogares, la cual se considera a partir del hacinamiento dentro de la vivienda. Esta variable se creó a partir de la información de cuántos habitantes hay en la vivienda del hogar (una vivienda puede tener uno o más hogares) y se complementó con el número de cuartos disponibles en dicha vivienda. De este modo, se calculó una razón para todos los individuos para conocer el número de personas por cuartos disponibles en su vivienda. La variable cuartos disponibles incluye la cocina pero excluye el baño o pasillos. Se considera que esta variable puede

incidir en el salario al afectar directamente la disponibilidad de espacio individual de la población en estudio.

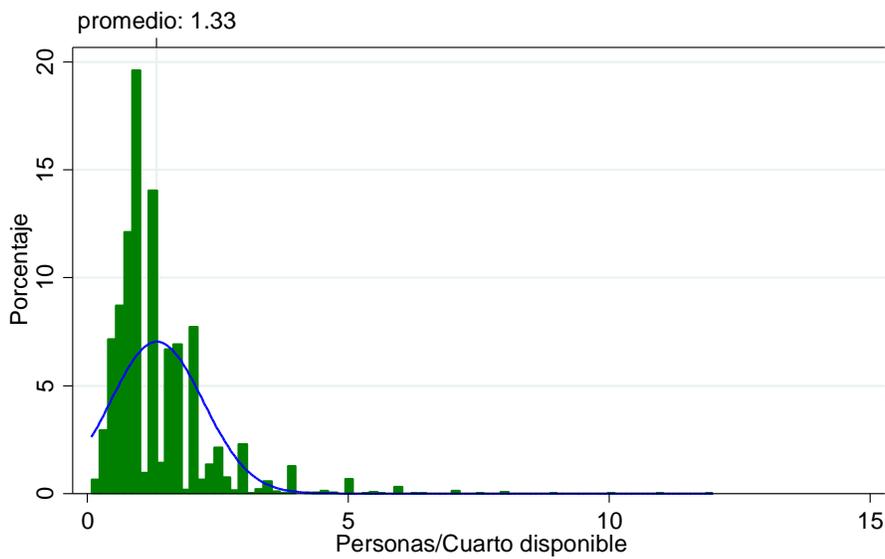
El resumen del número de personas por cuarto disponibles por tamaño de localidad aparece en el Cuadro IV.10. El tamaño de localidad también es considerado como un control en la investigación, por medio de crear una variable dicotómica para cada nivel de la variable original reportada en la ENIGH. También se incluye la Gráfica IV.6 donde se observa la distribución a nivel global, pues no parece haber diferencia entre las estimaciones por tamaño de localidad.

Cuadro IV.10 Estimación del promedio del índice de hacinamiento por tamaño de localidad, ENIGH 2016.

| Tamaño de localidad | Índice de hacinamiento promedio | Err. | Int. 95% Conf. | |
|---------------------------|---------------------------------|----------|----------------|----------|
| 100 mil o más hab. | 1.18310 | 0 .00891 | 1.165632 | 1.200581 |
| 15 mil a menos de 100 mil | 1.39126 | 0 .02305 | 1.346072 | 1.436452 |
| 2.5 mil a menos de 15 mil | 1.47793 | 0 .01871 | 1.441254 | 1.514611 |
| Menos de 2.5 mil | 1.60046 | 0 .01696 | 1.567221 | 1.633716 |

Fuente: Microdatos ENIGH 2016

Gráfica IV.6 Distribución del índice de hacinamiento.



Fuente: Microdatos ENIGH 2016

G. Habilidades intelectuales

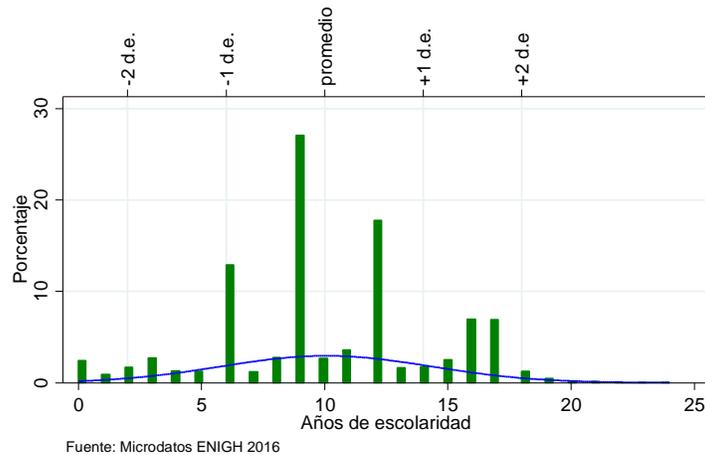
Como se comentó en los capítulos anteriores, las habilidades intelectuales no son observadas en la ENIGH, por lo que se considera en cambio, la distribución de los años de escolaridad como variable explicativa de interés. Esta variable fue nombrada “educación” y construida a partir de dos variables recogidas en la ENIGH a nivel individuo. En los cuestionarios individuales se pide al entrevistado mencionar su nivel máximo de estudios (nivelaprob) así como el grado máximo que aprobó (gradoaprob) dentro del nivel máximo cursado. Con esta información se logró definir cuántos años de escolaridad reportó cada individuo de acuerdo con el Cuadro IV.11:

Cuadro IV.11 Construcción de años de escolaridad, ENIGH 2016

| nivelaprob | Años de escolaridad |
|-------------------|--|
| 0 Ninguno | 0 |
| 1 Preescolar | 0 |
| 2 Primaria | gradoaprob |
| 3 Secundaria | 6+gradoaprob |
| 4 Preparatoria | 9+gradoaprob |
| 5 Normal | si ant_esc=Primaria, 6+gradoaprob, si ant_esc=Secundaria, 9+gradoaprob, si ant_esc=Preparatoria, 12+gradoaprob |
| 6 Carrera técnica | si ant_esc=Primaria, 6+gradoaprob, si ant_esc=Secundaria, 9+gradoaprob, si ant_esc=Preparatoria, 12+gradoaprob |
| 7 Profesional | 12+gradoaprob |
| 8 Maestría | 16+gradoaprob |
| 9 Doctorado | 18+gradoaprob |

A partir de la construcción mencionada, pudo estimarse la distribución de los años de escolaridad del total de la población. El nivel promedio de escolaridad asciende a 10 años, lo cual es equivalente a secundaria terminada y al menos un año de preparatoria. Además, como se muestra en la Gráfica IV.7, la variabilidad del nivel educativo es bastante alta. El caso de mayor frecuencia es el de secundaria concluida (años de escolaridad=9), con 11,206,834 personas cuya edad promedio es de 36 años. En cambio, 6,585,334 personas cuentan con estudios universitarios concluidos (años de escolaridad mayores o iguales a 16) y su edad promedio es de 38 años.

Gráfica IV.7 Distribución de años de escolaridad



Esta distribución de los años de escolaridad será el objeto de política para analizar el impacto que tiene de la educación en el ingreso salarial de los individuos. Más adelante se retoma el método a emplear para determinar su distribución estimada y contrafactual.

IV.5 Estimación del escenario observado por medio de regresión cuantílica.

A partir de las variables expuestas más arriba, en este apartado se presentan los resultados de la estimación del escenario observado por medio de una regresión cuantílica sobre los cuantiles 5 a 95. Para esta estimación retomamos el planteamiento del modelo ideal, el cual para efectos de esta investigación pasa a tomar la siguiente forma:

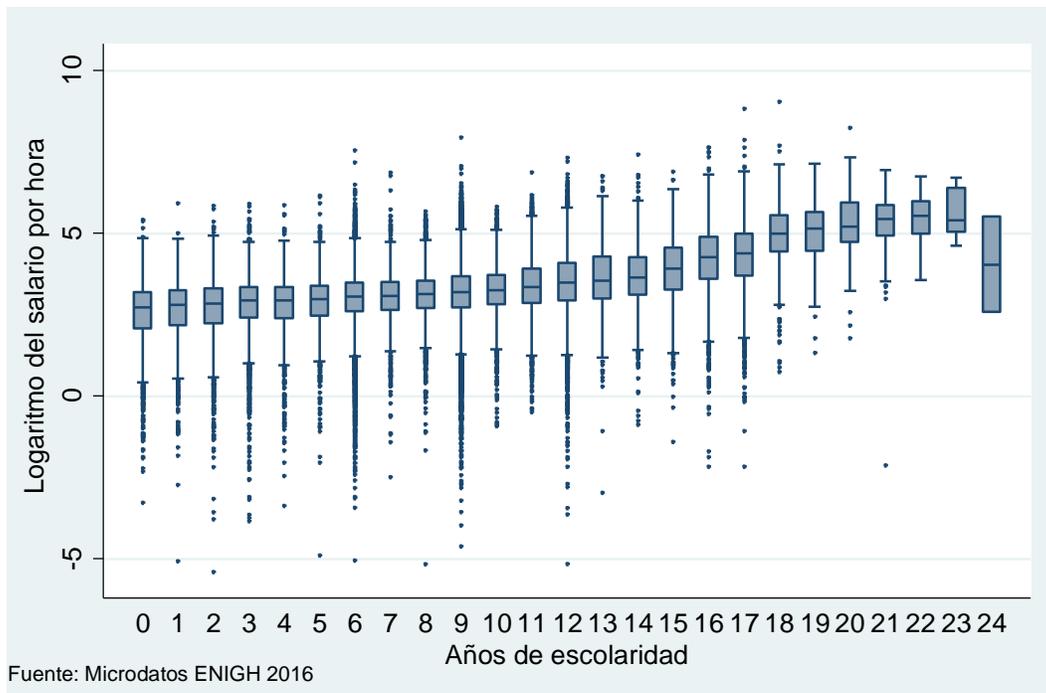
$$y = X\beta(\tau) + \varepsilon$$

Donde y es el vector de observaciones individuales del salario, X la matriz de variables explicativas y β el conjunto de coeficientes calculados para un cuantil τ específico de la distribución. El vector ε representa el error de estimación. Como se mencionó en el tercer capítulo, generalmente el modelo de regresión lineal estima este tipo de ecuaciones para encontrar el regresor asociado a la media condicional de la variable dependiente. Sin embargo, bajo el enfoque de regresión cuantílica, se busca identificar el coeficiente de cada variable dado el programa de optimización específico para cada cuantil. Este tipo de planteamientos resultan útiles cuando se presentan distribuciones continuas en la variable dependiente cuya estimación promedio puede infringir el supuesto de homocedasticidad del modelo de regresión lineal. Asimismo, la

estimación de los parámetros promedio en presencia de dispersiones desbalanceadas puede inhibir el poder explicativo del coeficiente obtenido (Davino, C., Furo, M. y Vistoso, D., 2014:29). Por medio de la regresión cuantílica, se obtienen, en cambio, estimadores específicos a las regiones de la distribución de interés.

Para puntualizar con mayor claridad el problema que soluciona la regresión cuantílica en la Gráfica IV.8 se puede apreciar el comportamiento del logaritmo del salario por hora respecto de los años de escolaridad. Se incluye gráficas de caja condicionales a cada valor de año de escolaridad, con los cuales se busca ilustrar cuántas observaciones quedan fuera del rango intercuantil condicional. Una estimación por medio de regresión lineal buscaría conectar el promedio condicional y por tanto sería poco explicativa de los diversos casos que se encuentran por debajo y por encima de dichos valores.

Gráfica IV.8 Distribución condicional del salario en logaritmos respecto de la educación



Para la estimación se empleó la transformación logarítmica del salario como variable dependiente. Las variables explicativas se mantienen en sus niveles originales, por lo que sus respectivos coeficientes se interpretan como aproximaciones al impacto porcentual de la variación de una unidad (o de un grupo a otro en el caso de las variables dummy) en la variable dependiente. La Ecuación IV.1 permite estimar el efecto de los regresores en la variable de interés:

$$\log(\text{salario}_i)_\tau = \alpha_\tau + \beta_{1\tau}\text{sexo} + \beta_{2\tau}\text{edad} + \beta_{3\tau}\text{discap} + \beta_{4\tau}\text{pad_mad} + \beta_{5\tau}\text{lenguaindig} + \beta_{6\tau}\text{etnia} + \beta_{7\tau}\text{educacion} + \beta_{8\tau}\text{salud} + \beta_{9\tau}\text{sitviv} + \beta_{10\tau}\text{tam_loc2} + \beta_{11\tau}\text{tam_loc3} + \beta_{12\tau}\text{tam_loc4} + \varepsilon_i$$

Ecuación (IV.1)

Nótese que los estimadores se asocian a un cuantil τ específico, por lo que se considera que los efectos de estas variables pueden ser distintos a lo largo de la distribución.

El escenario observado aproximado por este método de estimación se muestra en el Cuadro IV.12:

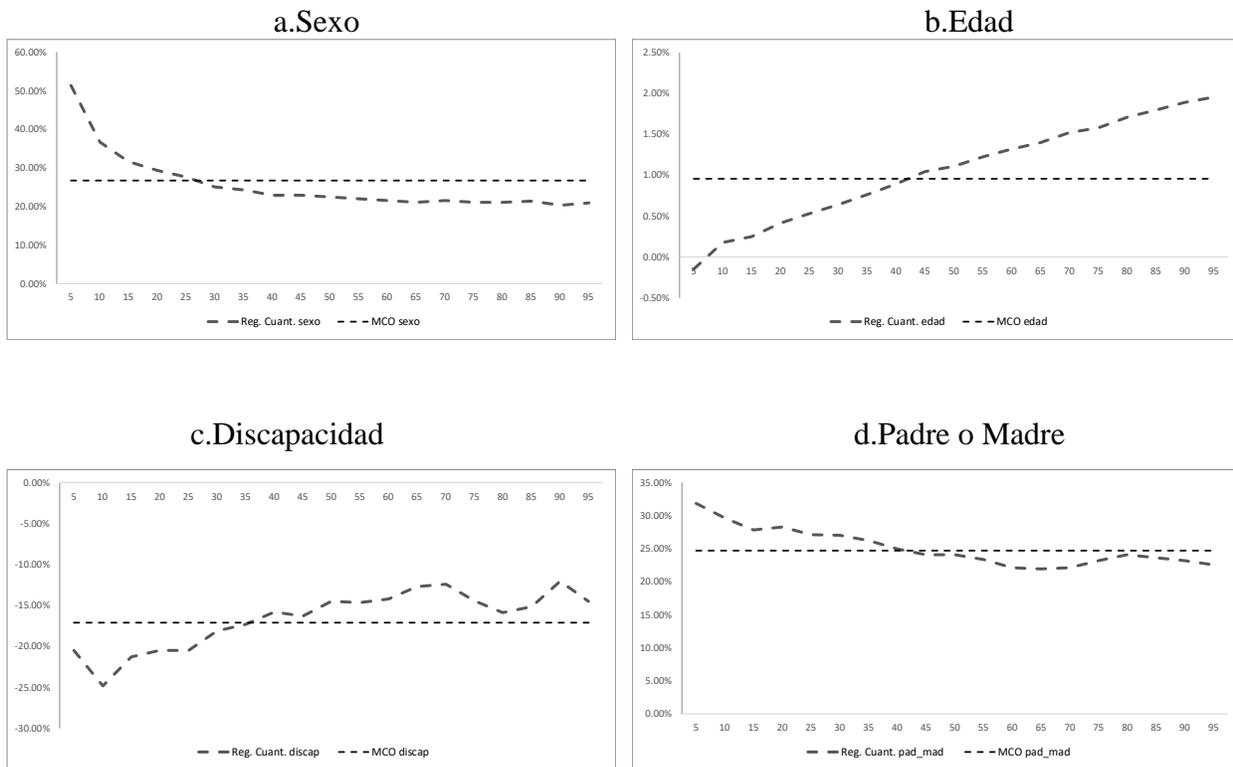
Cuadro IV.12 Resultados de la regresión por cuantiles.

| EFFECTO DE DIVERSOS FACTORES EN EL LOGARITMO DEL SALARIO POR HORA EN CUANTILES SELECCIONADOS Y EN LA MEDIA CONDICIONAL. | | | | | | |
|--|------------|------------|-------------|------------|-------------|------------|
| log(salario) | MCO | 0.1 | 0.25 | 0.5 | 0.75 | 0.9 |
| | (1) | (3) | (6) | (11) | (16) | (19) |
| sexo | 0.2378* | 0.3129* | 0.2448* | 0.2032* | 0.1918* | 0.1854* |
| edad | 0.0096* | 0.0018* | 0.0053* | 0.0110* | 0.0157* | 0.0187* |
| discap | -0.1876* | -0.2851* | -0.2287* | -0.1565* | -0.1560* | -0.1281* |
| pad_mad | 0.2209* | 0.2596* | 0.2401* | 0.2156* | 0.2088* | 0.2087* |
| lenguaindig | -0.3393* | -0.6667* | -0.4227* | -0.2568* | -0.1782* | -0.1600* |
| etnia | -0.1155* | -0.1262* | -0.1152* | -0.1098* | -0.0990* | -0.0925* |
| salud | -0.0336* | -0.0683* | -0.0411* | -0.0244* | -0.0167* | -0.0032* |
| sitviv | -0.1125* | -0.0898* | -0.0981* | -0.1067* | -0.1135* | -0.1216* |
| tam_loc2 | -0.1692* | -0.2377* | -0.1860* | -0.1558* | -0.1303* | -0.1114* |
| tam_loc3 | -0.3037* | -0.4080* | -0.3402* | -0.2804* | -0.2402* | -0.2147* |
| tam_loc4 | -0.4960* | -0.7852* | -0.5532* | -0.4124* | -0.3418* | -0.2884* |
| educacion | 0.1083* | 0.0894* | 0.0905* | 0.1042* | 0.1163* | 0.1229* |
| _cons | 2.0384* | 1.5218* | 1.8878* | 2.0455* | 2.2443* | 2.5240* |

*p<0.01

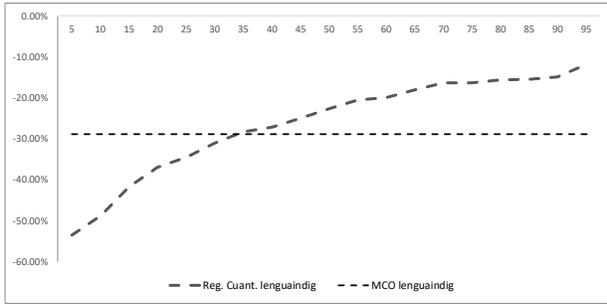
De acuerdo con la especificación del modelo, se determinó que existe un efecto diferenciado del género en el ingreso que se manifiesta de manera distinta en los cuantiles de la distribución. En general, los hombres del primer decil ganan en promedio 51.36%¹⁸ más que una mujer en dicho decil. Si bien este efecto desciende a medida que aumenta el ingreso, en ningún punto de la distribución las remuneraciones por hora al trabajo son iguales entre hombres y mujeres (ver panel Gráfica IV.9 a). Por medio del método de mínimos cuadrados ordinarios se estimó un efecto promedio de 26.85% de ventaja en el salario de un hombre respecto a una mujer. La estimación por medio de la regresión cuantílica brinda una noción más específica del efecto de las variables explicativas en cada percentil de la distribución. En la Gráfica IV.9, se presentan las gráficas de las estimaciones para cada regresor por decil. En todos los casos se grafica la estimación exacta de la variación del ingreso ante cambios unitarios (o de grupo) de cada nivel en la variable en cuestión.

Gráfica IV.9 Coeficientes estimados por percentil. Escenario observado.

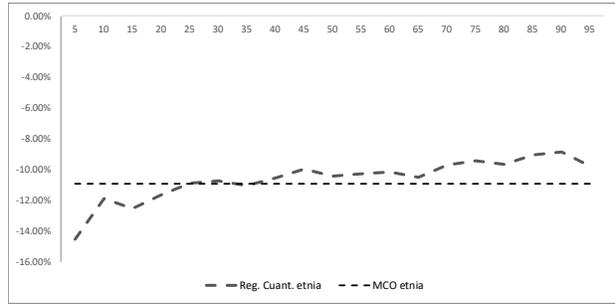


¹⁸ Después de aplicar la transformación $EXP(\hat{\beta}) - 1$ al coeficiente estimado. Esta transformación será empleada en el análisis de los resultados de la regresión a menos que se especifique lo contrario.

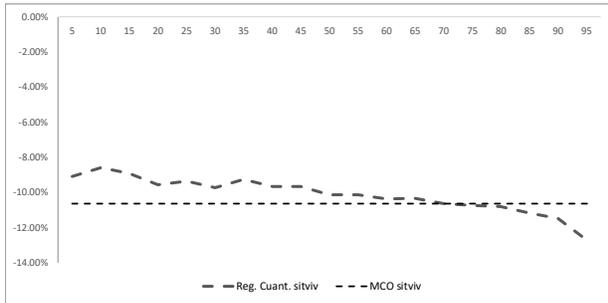
e. Lengua indígena



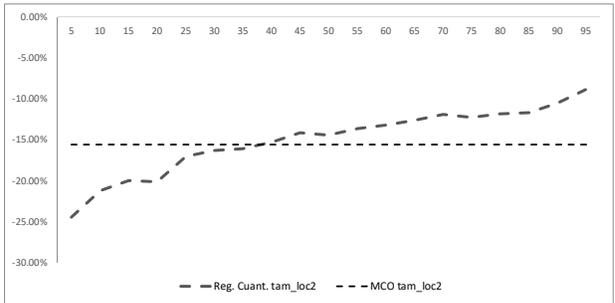
f. Etnia



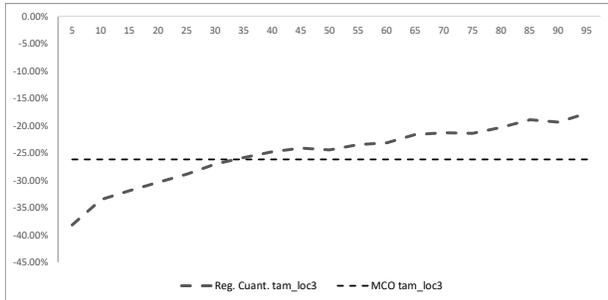
g. Índice de hacinamiento



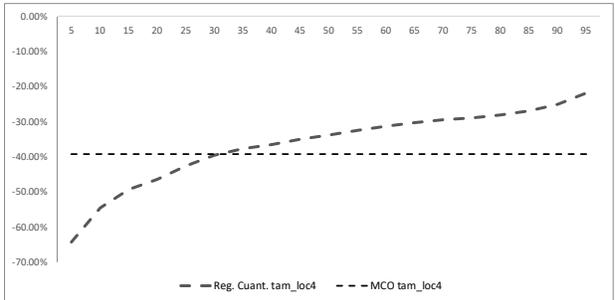
h. Localidad de 15 mil h. a menos de 100 mil



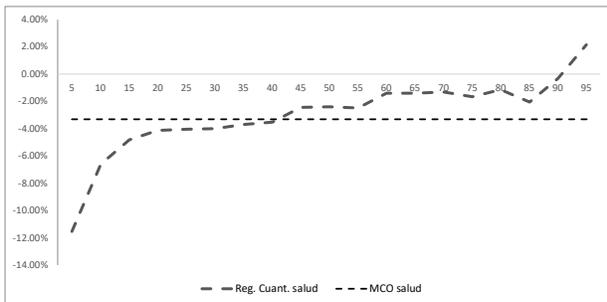
i. Localidad de 2.5 mil h. a menos de 15 mil



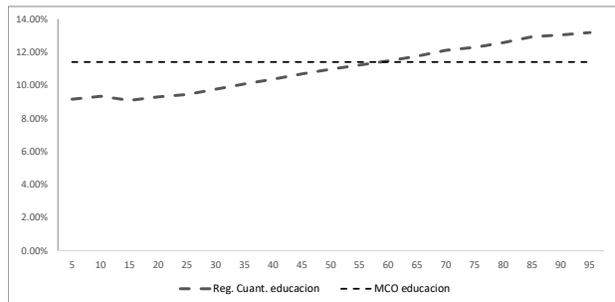
j. Localidad de menos de 2.5 mil h.



k. Problemas de salud en 2016



l. Educación



Los coeficientes estimados de los factores pueden agruparse en tres tipos. Aquellos que tienen efectos negativos, mixtos y positivos. En cada grupo, además se pueden distinguir subgrupos por el tipo de pendiente que tienen los resultados. Así, dentro del grupo de efectos negativos con pendiente positiva se encuentran los coeficientes de discapacidad, lengua indígena, todos los de localidades y etnia. Estos casos representan efectos adversos en el salario que agravan más fuertemente en los deciles más bajos. Así, por ejemplo, la presencia de al menos un tipo de discapacidad se traduce en un impacto de -20.49% en el percentil 5 y llega a tener su menor impacto en el percentil 90 donde éste se reduce a la mitad: -12.02%. La presencia de lengua indígena, a su vez, presenta un impacto que llega a ser -53.48% en los percentiles más bajos y se mantiene negativo solo hasta el -11.82% (equiparable con el impacto de una discapacidad física) en el percentil 95. El tamaño de localidad también tiene un impacto adverso y regresivo entre los percentiles de personas. Las localidades más pequeñas suponen una reducción del ingreso de las personas en los percentiles más bajos de -64.30% respecto a las localidades con más de 100 mil habitantes. Sin embargo, el efecto es de solo -21.9% para las personas en los percentiles más altos dentro de ese grupo de localidades (ver panel j.). Los efectos son sucesivamente menores a medida que la localidad es mayor (paneles h.-j.). La adscripción étnica también presenta un efecto adverso en el salario por hora percibido, aunque no logra anularse o generar un beneficio: su nivel mínimo de impacto es de -8.83% en el percentil 90. El comportamiento de este coeficiente con una pendiente cada vez más suavizada refleja una resistencia denominada comúnmente “techo de cristal” donde la presencia de la condición capturada por la variable constituye un impedimento constante para generar mayores ingresos, a pesar de que éste aumente por otras razones.

Dentro de los efectos negativos con pendiente negativa se ubica el índice de hacinamiento cuyo comportamiento es progresivo y afecta más gravemente el ingreso de las personas en los percentiles más elevados. Una persona adicional en cada cuarto de su vivienda impactaría en hasta -12.69% el ingreso salarial de una persona en los percentiles más altos, a pesar de que solo impactaría en -9.08% en los percentiles más bajos. Aquí, puede existir una interpretación inversa: una persona menos por cuarto puede reflejar hasta 12.69% más ingresos para los individuos en los deciles más altos.

Dentro del grupo de variables con efectos mixtos únicamente se identificó el caso de la variable que captura la presencia de problemas de salud en el año en curso, pues este factor actúa

adversamente sobre todo entre la población en los cuatro primeros deciles con una severidad de - 11.54% en el salario por hora percibido, respecto al beneficio que representa este factor en la población con más altos ingresos (+2.13% en el percentil 95). Para el resto de la distribución el efecto es relativamente constante, lo cual refleja cierta igualdad a través del nivel de ingreso en especial entre los percentiles 60 a 80.

Entre las variables que reflejan beneficios en términos salariales y que se presentan una pendiente negativa, resalta la presencia de paternidad o maternidad, la cual en promedio se incide en un incremento de 24.72% del salario por hora para toda la población, aunque hay variaciones importantes entre deciles, el efecto se mantiene positivo desde el incremento en 31.89% de las personas en los percentiles más bajos hasta el incremento de 22.55% en los percentiles más altos. Un caso similar es la presencia de género masculino, que genera mayores beneficios en los deciles más bajos, los cuales disminuyen en los deciles más altos, como se comentó más arriba.

Finalmente, el grupo de variables con efectos positivos con pendiente positiva se compone de la variable edad y educación. Específicamente este último grupo de variables refleja las oportunidades potenciales para el crecimiento salarial. La edad muestra un efecto positivo que genera mayores impactos en los deciles más elevados, es decir, que un año adicional de edad tiene mejores rendimientos entre mayores sean los ingresos percibidos. En el caso de la educación, un año adicional de estudios genera beneficios diferenciados entre la población, pues el impacto en el ingreso es de 9.19% entre la población del percentil más bajo frente al 13.21% de la población en el percentil más alto con un aumento pequeño pero constante a partir del percentil 15. La estimación promedio predice un impacto de 11.43% el cual es representativo del percentil 60. Si bien los mayores retornos se encuentran en la parte superior de la distribución, la diferencia entre ambos no es muy grande. El hecho de que la educación presente retornos positivos, permite justificar una política hipotética que si es bien focalizada, puede generar beneficios en términos de ingresos salariales a la población.

IV.6 Inferencia de dos escenarios no observados.

A partir del escenario observado fue posible identificar cómo se comporta la distribución del ingreso dadas algunas variables identificadas en la literatura. A continuación, se plantea observar el impacto en el ingreso y en la desigualdad salarial al aplicar un grupo de políticas totalmente hipotéticas a algunas variables explicativas de interés, manteniendo fijas las respuestas en el resto de las variables observadas.

Un primer caso de interés constituye el fomento del nivel educativo entre la población con autoadscripción étnica indígena, la cual constituye el 30% de la población en estudio. Si bien la presencia de esta condición afecta adversamente el salario percibido, por medio de aplicar una política de mejoramiento en el nivel educativo de esta población se espera observar mejoras en el nivel de ingreso en general y menor desigualdad medida por medio de los índices planteados observados. En la Gráfica IV.10 puede observarse cómo a medida que aumentan los niveles de escolaridad la proporción de personas indígenas disminuye fuertemente. Esto puede deberse al proceso educativo que transforma la cultura del individuo y por tanto altera la respuesta a la pregunta, o bien a que efectivamente existen menos personas con origen indígena en niveles de escolaridad mayores.

Gráfica IV.10 Proporción de personas que declararon ser indígenas, por años de escolaridad.

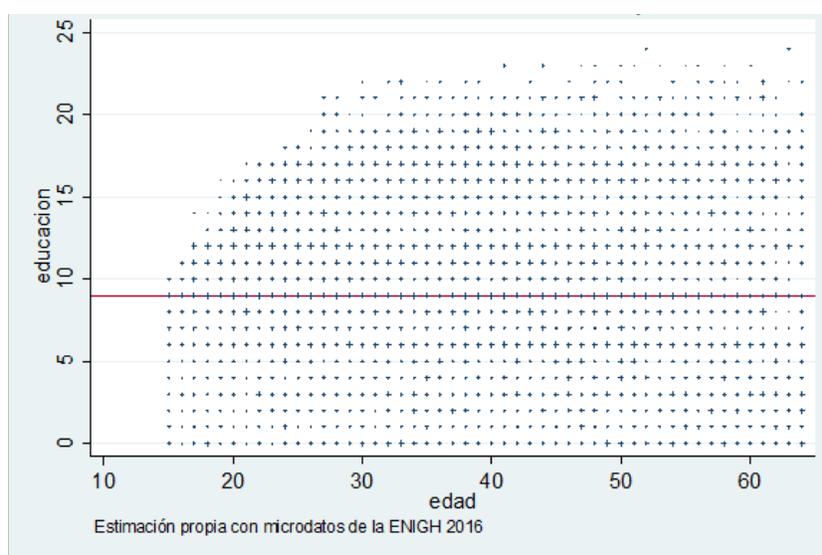


Fuente: Estimación propia a partir de microdatos de la ENIGH 2016.

Otro escenario de interés constituye la aplicación de una política de erradicación del logro educativo que sea inferior al mínimo esperado dada la edad biológica de la población. La

importancia de esta política puede apreciarse en la distribución que se presenta en la Gráfica IV.11. La línea roja muestra el límite mínimo de años de escolaridad ideal dado el mínimo de edad dentro de la población de estudio. Esto es, un individuo de 15 años idealmente debió haber cursado al menos la secundaria completa (años de escolaridad=9). Este es el nivel mínimo esperado de educación para personas de 15 años o más. Por ello, la política consistirá en buscar la erradicación de ese rezago histórico aumentando los niveles educativos para todos aquellos que no hayan terminado la secundaria, tengan más de 15 años y trabajen a cambio de un salario.

Gráfica IV.11 Distribución de años de escolaridad por edad.



A partir de los microdatos de la ENIGH se estima que 11,215,979 en la población de estudio que declaró trabajar al menos 30 horas durante la semana de entrevista no concluyó la formación educativa mínima ideal de la población. El total de la población de estudio es de 41,385,988 personas, por lo que esta población representa el 27.10% del total. El impacto esperado de una política hipotética de este estilo en la desigualdad es que esta podría generar mejoras sustanciales en el nivel de ingresos de la población y mejorar los niveles de desigualdad. La población susceptible de ser receptora de la política se encuentra distribuida por tamaño de localidad de manera muy polarizada pues casi dos tercios se encuentran o en poblaciones muy grandes o muy pequeñas. Los canales de acción para llevar a cabo la política pueden adecuarse de acuerdo con el tipo de localidad.

| Tamaño de localidad | Número de individuos elegibles | Porcentaje de la población elegible |
|----------------------------|---------------------------------------|--|
| 100 mil o más hab. | 3,837,698.00 | 34% |
| 15 mil a menos de 100 mil | 1,706,155.00 | 15% |
| 2.5 mil a menos de 15 mil | 1,875,626.00 | 17% |
| Menos de 2.5 mil | 3,796,500.00 | 34% |
| Total | 11,215,979.00 | 100% |

La política planteada en ambos casos consiste en mejorar el nivel educativo de cada individuo en 4 años. Es decir, se transformará la distribución de los años de escolaridad aumentando en el primer escenario la escolaridad observada en cuatro años únicamente para aquellos que mencionaron ser indígenas. Asimismo, para las personas con nivel educativo menor al mínimo esperado (educación < 9), se plantea una política que busque aumentar su nivel de escolaridad en 4 años. El resto de la distribución permanece invariante por lo que no se consideran los efectos dinámicos de la educación (cohortes que avanzan y acumulan niveles de escolaridad) ni tampoco los posibles efectos en los mercados laborales. Estos escenarios pueden pensarse también como hipotéticos en el sentido de que brindan una idea de cuál sería la realidad actualmente si las cosas hubiesen sido distintas de modo que las políticas planteadas fuesen la realidad observada.

A continuación, se muestran los resultados de estas políticas y la estimación del impacto en los niveles de salario por hora.

A. Resultados del primer escenario no observado.

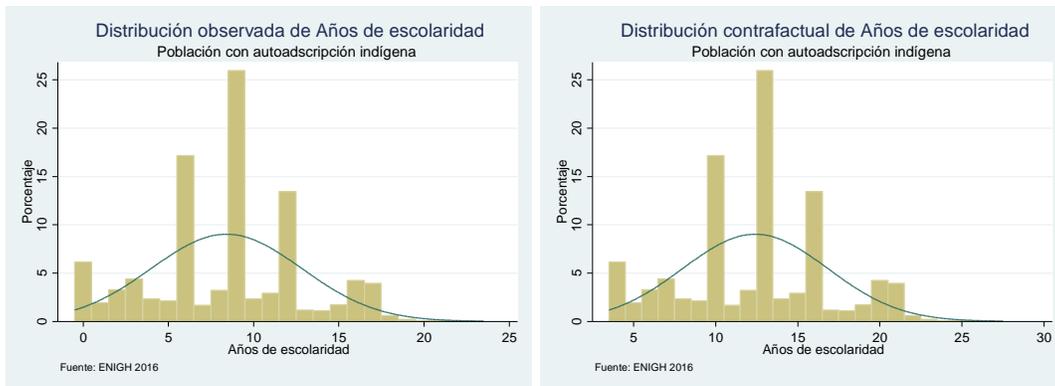
La formulación del primer escenario no observado consistió en modificar la estructura de la educación para la población con autoadscripción indígena en cuatro años adicionales, sin importar su nivel de salario o edad. Debido a que esta transformación consiste en un efecto de traslación en la distribución de años de escolaridad para el grupo de población indígena (ver Gráfica IV.12), se esperaba que la distribución del ingreso mejorase los niveles de desigualdad observados. La identificación de la distribución contrafactual del salario se obtiene a partir de aproximar en primer lugar la distribución observada del salario por hora, considerando los controles empleados en la estimación cuantílica vista en el apartado anterior $\hat{F}_{Y_o|X_o}$. Es decir, se estima la distribución condicional del salario por hora dados los diversos regresores para integrar respecto a una

estimación de la distribución contrafactual de las variables explicativas \hat{F}_{X_c} , la cual mantendrá idénticas las distribuciones del resto de las características salvo la distribución de los años de escolaridad. A continuación, se muestra la formulación específica a estimar, siguiendo a Chernozhukov, Fernández-Val y Melly (2013:11):

$$\hat{F}_{Y(o|c)}(\text{salario}) = \int \hat{F}_{Y_o|X_o}(y|x) d\hat{F}_{X_c}(x)$$

Donde el subíndice “o” remite a la distribución observada y el subíndice “c” a la distribución contrafactual. Para llevar a cabo las aproximaciones necesarias en el cálculo propuesto, se empleó el estimador basado en la regresión cuantílica definido en Chernozhukov, Fernández-Val y Melly (2013: 12) considerando alternativamente la exclusión del 1% y 5% por arriba y por debajo de la distribución. Ambas alternativas se muestran en la Gráfica IV.13. (paneles a. y b., respectivamente).

Gráfica IV.12 Distribución observada y contrafactual de años de escolaridad con una política para población con auto adscripción indígena.



Esta estimación es posible por medio del comando *counterfactual* en STATA, que como ya fue explicado fue programado por Chernozhukov, Fernández-Val y Melly (2013). Se empleó la misma especificación que en el modelo de regresión por cuantiles visto en el apartado anterior para realizar el ajuste de la distribución observada.

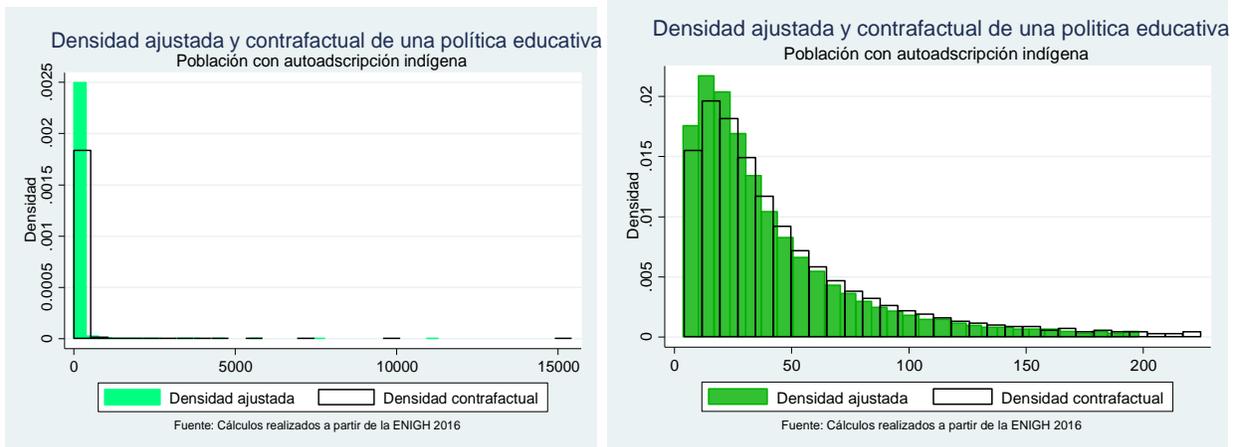
En un primer acercamiento se estimaron 998 cuantiles para los percentiles 0.1% a 99.9% y se obtuvieron la distribución ajustada y la distribución contrafactual al emplear como tratamiento el cambio en los niveles de escolaridad. En la Gráfica IV.13 se muestra el efecto que tiene la

implementación de esta política sobre el nivel educativo en la distribución salarial en los cuantiles 1 a 99 por ciento (panel a.) y 5 a 95 por ciento (panel b.). Es esta gráfica se muestra la función de densidad de la variable en su respectivo escenario en el eje vertical.

Gráfica IV.13 Densidad observada y contrafactual del salario por hora

a. 0.1%-99.9%

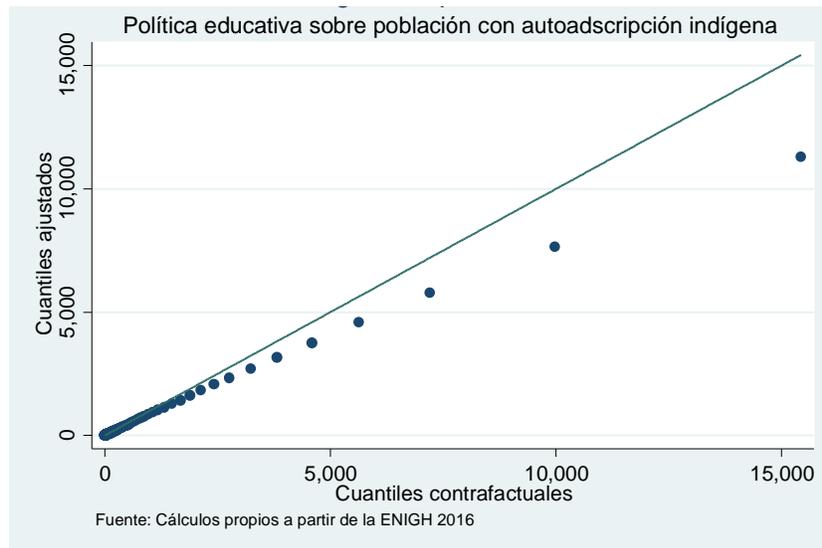
b. 4.9%-94.9%



Como puede apreciarse en el panel a. de la gráfica IV.13, la densidad ajustada al escenario observado mejora después de la implementación de la política en la primera parte de la distribución. Sin embargo, genera un efecto sustancial en polarizar los ingresos en la parte más alta de la misma, pues hace migrar más fuertemente los ingresos de la parte más alta de la distribución. Este mismo efecto, aunque más acotado se observa en el panel b. de la Gráfica IV.13 que considera exclusión del 5% más alto y más bajo de la distribución original. La migración por cuantiles que se muestra en la Gráfica IV.14 representa el movimiento de las observaciones que se evalúa según la pendiente que forman las observaciones. En el eje vertical se observa el cuantil inicial de cada observación y en el horizontal el cuantil final de la observación después de la política. Debido a que las observaciones con mayores ingresos se mueven más que las que tienen menos ingresos, esta migración informa sobre las regiones de la distribución más impactadas por la política implementada.

Gráfica IV.14 Migración por cuantiles. 0.1%-99.9%.

Primera política.



A.1 Efectos capturados en índices.

En términos de indicadores de desigualdad, el efecto solamente es comparable a partir de los indicadores asociados a la distribución ajustada y a la distribución contrafactual. Esto es así debido a que comparar con los datos puntuales surgidos de la ENIGH 2016 (Cuadro IV.5) no es válido en tanto las estimaciones obtenidas a partir de la metodología contrafactual consideran aproximaciones cuantílicas en 998 puntos de la distribución observada (ajustada) y contrafactual¹⁹, mientras que los cálculos del Cuadro IV.5 son más exactos pues fueron realizados a partir de la población de estudio considerando el diseño muestral de la ENIGH 2016. Por ello, los índices que se mostrarán para todos los ejercicios contrafactuales únicamente permitirán observar el efecto direccional de la política implementada (al aumentar o disminuir índices específicos) y por tanto permiten identificar su efectividad e intensidad en términos de niveles de desigualdad. Los índices que se muestran a continuación fueron obtenidos por medio del comando *ineqdeco*²⁰ en STATA aplicados a las distribuciones ajustada y contrafactual. En el Cuadro IV.13 aparece el efecto en la desigualdad para los cuantiles 0.1-99.9%.

¹⁹ El autor agradece especialmente a Blaise Melly por su apoyo en las estimaciones aquí realizadas. Sin su guía por medio de STATA list esta investigación habría enfrentado mayores dificultades.

²⁰ El comando *ineqdeco* fue desarrollado por Stephen Jenkins de la London School of Economics.

Cuadro IV.13 Efectos de la política en índices de desigualdad. 0.1%-99.9%

| Indicador | Ajustada | Contrafactual |
|--|-----------------|----------------------|
| Media aritmética | 102.7953 | 122.4918 |
| Media geométrica | 28.40864 | 32.28932 |
| Varianza | 293,354.8 | 494,606.70 |
| Coefficiente de variación | 5.2689 | 5.7415 |
| Coefficiente de Gini | 0.76648 | 0.778 |
| Índice Atkinson, e=0.5 | 0.53835 | 0.55727 |
| Índice Atkinson, e=1 | 0.72364 | 0.7364 |
| Índice Atkinson, e=2 | 0.99341 | 0.99289 |
| índice de Theil, GE(1) | 1.876 | 1.9969 |
| Desviación Logarítmica promedio, GE(0) | 1.28605 | 1.33331 |
| 1/2 CV ² , GE(2) | 13.86697 | 16.46573 |
| p90/p10 | 16.277 | 15.874 |
| p90/p50 | 3.965 | 3.977 |
| p10/p50 | 0.244 | 0.251 |
| p75/p25 | 3.83 | 3.81 |

Al realizar la medición de los principales índices, se observa que la media aritmética de la distribución ajustada no coincide con la media observada a partir de los microdatos de la ENIGH 2016. Esto puede deberse a que la distribución ajustada es capturada en 998 cuantiles específicos, considerando la especificación del modelo de regresión por cuantiles y a que la distribución ajustada puede sufrir de pocas observaciones en los extremos de la distribución observada. Sin embargo, el interés de realizar la medición recae en los efectos en términos de índices para la distribución ajustada y contrafactual. Las dos medias crecen, lo cual implica que la política tiene un efecto positivo en el nivel de salarios. Sin embargo, también crece la varianza de su distribución, así como el coeficiente de variación. Este efecto se debe a que un año adicional de educación genera mayores impactos salariales en las partes más altas de la distribución (ver Gráfica IV.9 panel 1.). El coeficiente de Gini y los dos primeros índices de Atkinson, así como el índice de Theil, la desviación logarítmica promedio y el índice entrópico GE(2) muestran un incremento sustancial en la desigualdad a partir de esta política. El índice de Atkinson cuando e=2 refleja una leve disminución debido al incremento de la media aritmética y al efecto positivo en los niveles de salario más bajos. Del mismo modo, los cuatro indicadores que miden la participación del percentil

“x” en el percentil “y” ilustran los efectos en diferentes puntos de la distribución. Estos indican que el efecto es benéfico en los niveles de ingreso más bajos pero casi inocuos en el rango de percentiles 25 y 90. Si bien hubo un efecto benéfico en los ingresos más bajos, la desigualdad aumenta sustancialmente debido al crecimiento en los niveles de salario más altos.

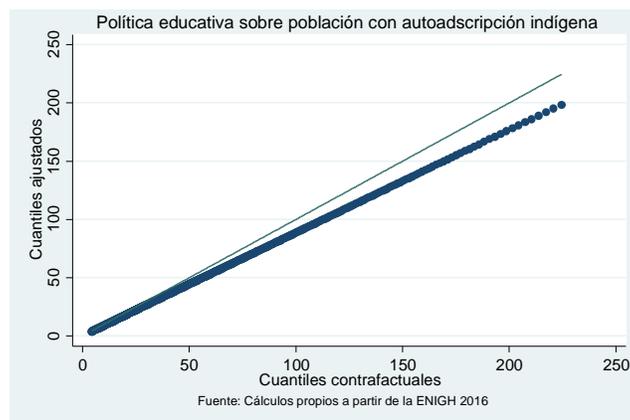
Al recortar el 5% más alto y más bajo de la distribución salarial, se realizó el ajuste para 900 posiciones desde el 4.9% y hasta el 94.9% de la distribución observada (Cuadro IV.14 y Gráfica IV.13 b.). Los resultados en términos de desigualdad fueron positivos pues mejoró el nivel de desigualdad de acuerdo con todas las medidas calculadas. Asimismo, el aumento en el salario por hora promedio y su varianza reflejan el mismo efecto surgido por la pendiente del coeficiente de años de escolaridad en los percentiles de la distribución. Es decir, que una política educativa que incida únicamente en la población indígena sin importar su nivel salarial podría generar efectos positivos en términos de salarios y niveles de desigualdad en los percentiles 4.9% a 94.9%. Potencialmente, un diseño de política más focalizado en la escolaridad de individuos con autoadcripción indígena y salario por hora debajo del promedio generará mayores beneficios en términos salariales y en disminución de la desigualdad pues el efecto regresivo de la educación sería anulado para individuos con salarios más altos. Esto puede observarse en la Gráfica IV.15 de migración, donde los cuantiles más altos tienen un mayor impacto que los más bajos después de implementar la política.

Cuadro IV.14 Efectos de la política en índices de desigualdad. 4.9%-94.9%

| Indicador | Ajustada | Contrafactual |
|--|----------|---------------|
| Media aritmética | 41.37303 | 46.74909 |
| Media geométrica | 29.34536 | 33.26346 |
| Varianza | 1,315.74 | 1,678.31 |
| Coefficiente de variación | 0.8767 | 0.8763 |
| Coefficiente de Gini | 0.43919 | 0.43842 |
| Índice Atkinson, e=0.5 | 0.15335 | 0.15252 |
| Índice Atkinson, e=1 | 0.29071 | 0.28847 |
| Índice Atkinson, e=2 | 0.50408 | 0.4983 |
| índice de Theil, GE(1) | 0.31882 | 0.31773 |
| Desviación Logarítmica promedio, GE(0) | 0.34349 | 0.34034 |
| 1/2 CV ² , GE(2) | 0.3839 | 0.38354 |
| p90/p10 | 9.609 | 9.439 |
| p90/p50 | 3.071 | 3.977 |
| p10/p50 | 0.32 | 0.326 |
| p75/p25 | 3.263 | 3.25 |

Gráfica IV.15 Migración por cuantiles. 4.9%-94.9%.

Primera política.

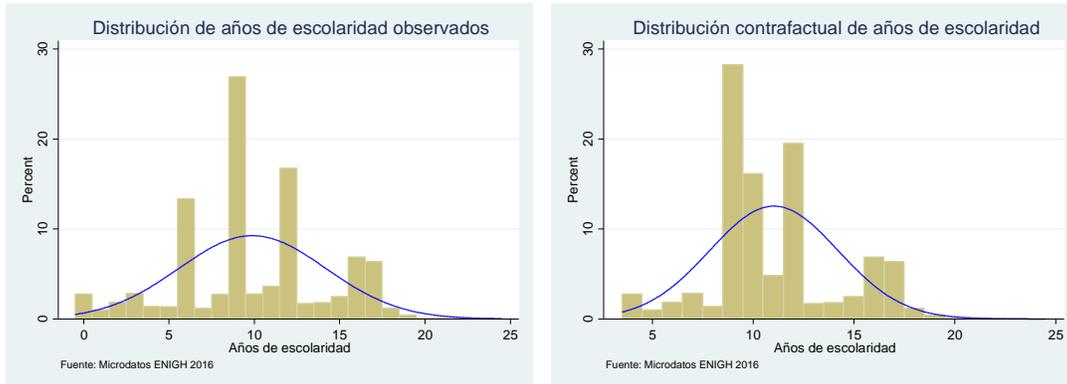


B. Resultados del segundo escenario no observado.

La segunda política consistió en transformar la distribución del nivel educativo para individuos que no tienen el nivel mínimo esperado, esto es, personas con años de escolaridad menor a los 9 años que un individuo con 15 años idealmente debe haber terminado. Esta transformación se muestra en la Gráfica IV.16. Mientras que la Gráfica IV.17 muestra la densidad ajustada y

contrafactual del salario para el 0.1% a 99.9% (panel a) y el 0.5% a 95% (panel b) de la distribución.

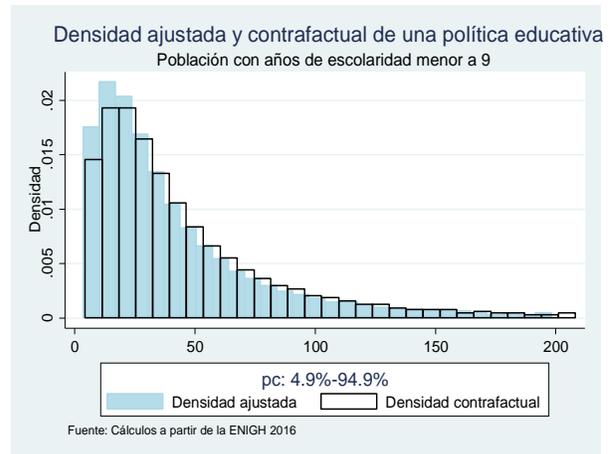
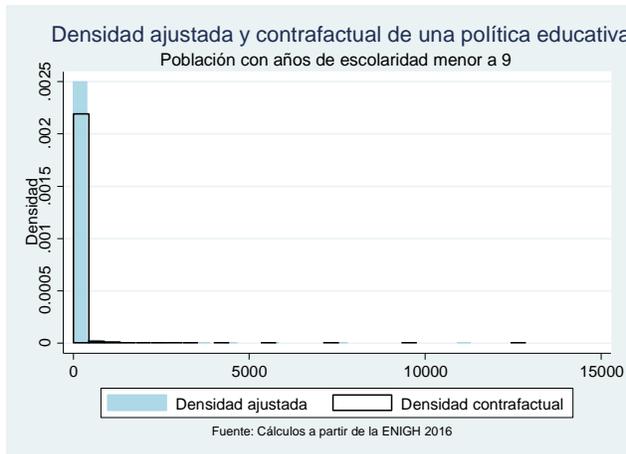
Gráfica IV.16 Distribución observada y contrafactual de años de escolaridad con una política sobre escolaridad ideal.



Gráfica IV.17 Densidad observada y contrafactual del salario por hora.

a. 0.1%-99.9%

b. 4.9%-94.9%



Como puede observarse en la Gráfica IV.17 panel a. la densidad ajustada mejora considerablemente a partir de la implementación de la política en los percentiles más bajos. A su vez, en la distribución acotada en el 5% más alto y más bajo de la distribución, el efecto del cambio en la distribución del nivel educativo ocurre sustancialmente en los percentiles inferiores con solo un efecto muy leve en los niveles de salario más elevados (Gráfica IV.17 b.).

B.1 Efectos capturados en índices.

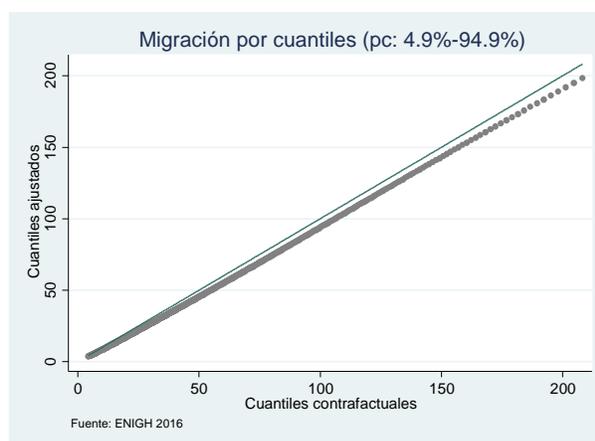
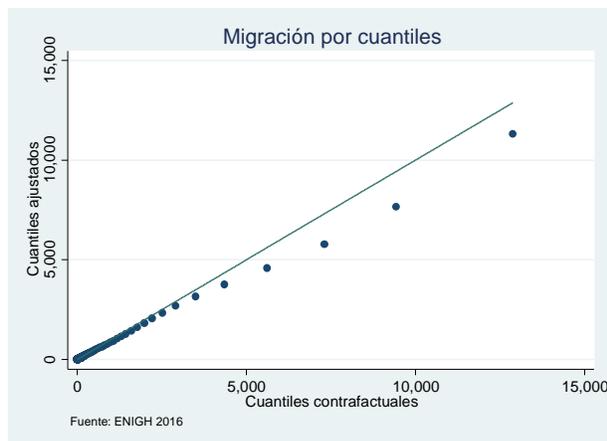
Para medir los efectos en términos de índices para esta segunda política, cabe recalcar que al igual que en el caso de la política para personas indígenas y por las razones que se explican allí, los índices no son comparables con los calculados en el Cuadro IV.4. Para el ajuste más amplio (panel a. en Cuadro IV.15) los impactos en índices fueron más discretos que en el caso de la primera política. La migración de cuantiles (Gráfica IV.18 a.) refleja una menor dispersión en los cuantiles superiores respecto al mismo ajuste para la primera política. Esto se refleja en los indicadores de desigualdad, los cuales presentan un efecto adverso de esta política, pero de menor intensidad respecto a la primera política. De hecho, los índices de Atkinson cuando $e=1$ y $e=2$ reflejan una disminución en términos de desigualdad debido al mayor progreso de los ingresos en percentiles más bajos en comparación con los percentiles más altos. También las participaciones de percentiles muestran una mejora de la distribución en la parte central e inferior. Los índices de Gini y los de la familia entrópica reflejan mayor desigualdad, salvo el caso de la Desviación Logarítmica Promedio, que muestra una disminución debida al efecto positivo de la política en las regiones cercanas a la media.

Gráfica IV.18 Migración por cuantiles.

Segunda política.

a. 0.1%-99.9%.

b. 4.9%-94.9%.



Cuadro IV.15 Efectos de la segunda política en índices de desigualdad.

| a. 0.1%-99.9% | | | b. 4.9%-94.9% | | |
|--|------------|-------------------|--|----------|---------------|
| Indicador | Ajustada | Contrafactua l | Indicador | Ajustada | Contrafactual |
| Media aritmética | 102.7953 | 115.9095 | Media aritmética | 41.37303 | 45.24827 |
| Media geométrica | 28.40864 | 32.05975 | Media geométrica | 29.34536 | 33.07337 |
| Varianza | 293,354.82 | 403,982.28 | Varianza | 1,315.74 | 1,430.57 |
| Coefficiente de variación | 5.2689412 | 5.483553 | Coefficiente de variación | 0.8767 | 0.8359 |
| Coefficiente de Gini | 0.76648 | 0.7679 | Coefficiente de Gini | 0.43919 | 0.42218 |
| Índice Atkinson, e=0.5 | 0.53835 | 0.54465 | Índice Atkinson, e=0.5 | 0.15335 | 0.14135 |
| Índice Atkinson, e=1 | 0.72364 | 0.72341 | Índice Atkinson, e=1 | 0.29071 | 0.26907 |
| Índice Atkinson, e=2 | 0.99341 | 0.99246 | Índice Atkinson, e=2 | 0.50408 | 0.47227 |
| índice de Theil, GE(1) | 1.876 | 1.93609 | índice de Theil, GE(1) | 0.31882 | 0.2932 |
| Desviación Logarítmica promedio, GE(0) | 1.28605 | 1.28521 | Desviación Logarítmica promedio, GE(0) | 0.34349 | 0.31344 |
| 1/2 CV ² , GE(2) | 13.86697 | 15.01963 | 1/2 CV ² , GE(2) | 0.3839 | 0.34897 |
| p90/p10 | 16.277 | 14.22 | p90/p10 | 9.609 | 8.569 |
| p90/p50 | 3.965 | 3.721 | p90/p50 | 3.071 | 2.915 |
| p10/p50 | 0.244 | 0.262 | p10/p50 | 0.32 | 0.34 |
| p75/p25 | 3.83 | 3.587 | p75/p25 | 3.263 | 3.081 |

Al considerar un ajuste que no involucre la porción 5% más alta y más baja, el efecto distributivo de la segunda política es contundentemente mejor que la primera política (panel b. Gráfica IV.17), pues mejora con más efectividad la distribución del ingreso, mejorando la porción central de la distribución y funcionando como un mecanismo posible de fomento de la clase media asalariada. Esto puede apreciarse claramente en el movimiento más moderado de la migración por cuantiles (panel b. Gráfica IV.18) que refleja un progreso relativamente equitativo a lo largo de la distribución de ingresos salariales. De hecho, debido a que hay más personas en la región inferior de la distribución que en la superior, se puede observar que esta política tiene un buen efecto sobre el ingreso y los niveles de desigualdad.

IV.7 Conclusiones

En este capítulo se implementó el análisis por medio de regresión por cuantiles a la distribución salarial de la población que trabajó al menos 30 horas en la semana de entrevista. Se controló por diversas variables que fueron identificadas en la literatura y que mostraron afectar de manera diferenciada en el nivel salarial según el percentil de la distribución. Así, condiciones físicas como la discapacidad o las enfermedades incapacitantes o el género juegan un rol adverso más intenso en las personas con salarios más bajos. En conjunto, las características sociodemográficas permitieron aislar más certeramente el nivel de educación, nuestra variable de interés, de modo que fuera posible aproximar el impacto de un año adicional de educación en el nivel salarial. A partir de dicha modelación, se planteó emplear un enfoque contrafactual para evaluar el impacto que tendrían dos políticas en los niveles de desigualdad y en el nivel salarial de las personas. Para ello, se empleó la metodología propuesta por Chernozhukov, Fernández-Val y Melly (2013) para estimar la distribución contrafactual y la distribución ajustada a los datos. Con dichas estimaciones fue posible observar con dos ajustes los impactos de dichas políticas y permitir vislumbrar algunos caminos posibles para la reducción de la desigualdad salarial en México. Los ajustes correspondientes fueron la exclusión de las porciones inferiores y superiores de la distribución en 1% y 5% y el resultado de los mismos refleja un rol fundamental de la educación en los niveles de crecimiento del ingreso salarial promedio y en los niveles de desigualdad. Si la educación es repartida de manera adversa en este sentido, puede tener efectos fundamentales en los procesos de desarrollo de México. Si bien el ejercicio contrafactual es totalmente hipotético y sujeto a diversos supuestos, su poder es que permite cuantificar y prever la dirección que una política potencial podría tener en el nivel de desigualdad económica, o bien, bajo especificaciones distintas, medir el impacto de la discriminación por género en las poblaciones de estudio.

Comentarios finales.

La teoría social es una fuente importante de intuiciones para los economistas. Si bien los argumentos aquí analizados desde el liberalismo y el materialismo podrían evaluarse como puramente hipotéticos, las conexiones latentes surgidas en las explicaciones de esta disciplina permitieron identificar en un primer momento, las causas más inmediatas que hacen surgir y prevalecer las desigualdades sociales. Como ciencia social, la economía ha sido influida en diversos aspectos por el problema que representa la desigualdad (económica, educativa, de oportunidades, entre otras). Si bien, este problema permanece como una problemática ética en el origen, las nociones materialista y liberal de la desigualdad, ambas relacionando causas tan claras como la desigualdad natural o de posesiones, animan a buscar algún vínculo científico que permita dar una mayor cabalidad a las explicaciones de este acuciante fenómeno.

Tal vez la economía no ha logrado dar una explicación definitiva de la desigualdad, pero los avances sustanciales que ha desarrollado la posicionan como una de las principales portadoras de las nociones sociales surgidas de la teoría social. Uno de sus muchos avances es el aparato metodológico que se ha conformado a partir de conceptos estadísticos como la medición de índices y la determinación de un enfoque que permita relativizarlos para cuanto se quiera decir respecto a la desigualdad. Este mismo aparato le ha permitido desarrollar diversas ramas de estudio que son ampliamente empleadas no solo en temas estrictos de la desigualdad social, sino también en aspectos más orientados al análisis de concentración de mercados y, en otras vertientes, de poder económico y político, entre otros. También ha permitido desarrollar enfoques empíricos sobre multiplicidad de aspectos que pueden ser evaluados y sintetizados por medio de la lupa de los índices planteados, permitiendo abarcar un mayor rango de hipótesis, así como establecer muestras más grandes de países y personas.

En esta investigación se buscó abordar el problema de la desigualdad salarial por medio de las explicaciones intuitivas vistas en el capítulo 1 y de los índices planteados en el capítulo 2. El enfoque adoptado consistió en una perspectiva metodológica que a la vez que permitiera dar una respuesta a la validez de las intuiciones identificadas, permitiera también realizar un análisis de las realidades contemporáneas en México. Por ello, se propuso tomar en cuenta las respuestas representativas de toda la población mexicana para medir cómo afectan las distribuciones de variables como la educación, la situación de vivienda, entre otras, a la distribución del salario por

hora por medio de una regresión por cuantiles. Este enfoque permitió capturar información relevante de una parte de la población que declaró ser económicamente activa y por tanto, resalta el rol de las variables propuestas en personas orientadas al trabajo y a la búsqueda de su bienestar. Para evaluar la importancia de las variables en esta relación, se propuso la realización de un análisis contrafactual a partir de los microdatos de la ENIGH 2016. Dicho análisis fue expuesto como una alternativa plausible ante los datos no experimentales surgidos de encuestas para derivar conclusiones causales y fue implementado a partir de la metodología de Chernozhukov, Fernández-Val y Melly (2013). Con ello, fue posible identificar la efectividad comparada de dos políticas que podrían subvertir la tendencia adversa de la polarización salarial por medio de mejorar la estructura educativa de la población trabajadora. Se logró identificar que una política focalizada en poblaciones rezagadas históricamente (individuos sin la educación mínima según el rango mínimo de edad) podría beneficiar más fuertemente la estructura de las desigualdades salariales, en comparación con una política orientada a disminuir la variación de los ingresos entre grupos étnicamente diversos. En especial, cualquier subsidio por medio de la educación permitiría reducir la tendencia a la polarización, siempre y cuando la población susceptible de tal política se encuentre por debajo del nivel salarial promedio. En poblaciones con ingresos superiores al promedio el impacto acentúa la desigualdad debido a la diferencia del rendimiento de un año adicional de educación en la población que recibe más por su trabajo, respecto a aquellos que reciben menos. Es muy posible que, de implementarse, una mejora de la distribución educativa mejore las condiciones no solo de aquellos beneficiados, sino de toda la sociedad que podría ver un mayor crecimiento de la clase media y por tanto mejorar la ruta de democratización y crecimiento económico. Este efecto aparentemente contrario a la racionalidad económica, que dictaría invertir más en las poblaciones con mayores rendimientos, refleja el conflicto inequívoco que representan los estudios de la desigualdad en economía. Los efectos de una política determinada son cada vez más predecibles por medio de la estadística, sin embargo, no lo son las intenciones de los diferentes tomadores de decisiones (a nivel individual y social) ni los acuerdos respecto al papel que se desea que juegue la educación en un país en vías de desarrollo como México. Un modelo que reconozca tanto los efectos dinámicos del mercado laboral, como una mayor cantidad de características respecto a los procesos de decisión social e individual de los años de escolaridad, permitiría dar una lectura más precisa del tamaño de los efectos esperados a partir de políticas educativas como las aquí propuestas. No obstante, el acercamiento aquí

propuesto busca ser una contribución a los debates fundamentales sobre el futuro del país pues la educación es uno de los mecanismos intergeneracionales que tiene impactos persistentes en los individuos y tal vez sea uno de los caminos más largos pero más efectivos para el progreso nacional.

Bibliografía

Acemoglu, D., S. Johnson y J. Robinson (2001) "The Colonial Origins of Comparative Development: An Empirical Investigation". *The American Economic Review*, 91, (5), 1369-1401.

Airola, J. y C. Juhn (2005) "Wage Inequality in Post-Reform Mexico". Institute for the Study of Labor, Discussion Paper No. 1525, 2-48.

Alberto, A. (2017) "Determinantes de los ingresos por trabajo: una aproximación por regresión cuantílica". Tesis de Maestría en Economía. Instituto de Investigaciones Económicas, UNAM.

Angrist, J. D. y J.-H. Pischke (2009) "Mostly Harmless Econometrics. An Empiricist's Companion". Princeton University Press.

Atkinson, A. B., T. Piketty, y E. Saez (2011), "Top incomes in the long run of history", *Journal of Economic Literature* 49 (1) , 3-71.

Baldwin, K. y J. Huber (2010) "Economic Versus Cultural Differences: Forms of Ethnic Diversity and Public Goods Provision". *American Political Science Review* 104, November ,(4).

Barro, R. (1991) "Economic Growth in a Cross Section of Countries". *The Quarterly Journal of Economics*, 106 , (2), 407-443.

Bellú, Lorenzo y P. Liberati (2006). "Describing income inequality. Theil Index and Entropy Class Indexes". Food and Agriculture Organization of the United Nations, EASYPol. Consultado en http://www.fao.org/docs/up/easypol/445/theil_index_051en.pdf el 1 de Agosto 2017.

Blanco, E. (2011) "Los límites de la escuela. Educación, desigualdad y aprendizajes en México". El Colegio de México. México D.F.

Bourguignon, F. y T. Verdier (2000) "Oligarchy, democracy, inequality and growth." *Journal of Development Economics*, 62 , 285-313.

Bunge, M. (2009). "Filosofía política. Solidaridad, cooperación y Democracia Integral". Editorial Gedisa. Barcelona, España.

Bütikofer, A. y G. Peri (2017) “The Effects of Cognitive and Noncognitive Skills on Migration Decisions” National Bureau of Economic Research WP No. 23877.

Card, D. y T. Lemieux (2001) “Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis”. National Bureau of Economic Research WP No. 7655.

Ceriani, L. y P. Verme (2012) “The origins of the Gini index: extracts from Variabilità e Mutabilità (1912) by Corrado Gini”. *Journal of Economic Inequality*, 10, 421-443.

Chacón, A. y P. Peña (2012) ” Cómo cambiar historias. Lo que podemos hacer los individuos, las empresas y las organizaciones sin fines de lucro por la educación en México” Fondo de Cultura Económica, México D.F.

Chernozhukov, V., I. Fernández-Val y B. Melly (2013) “Inference on counterfactual distributions”. ArXiv ID: 0904.0951 v6. url: arxiv.org/pdf/0904.0951.pdf

Cortés, F. y R. M. Rubalcava (1984) “Técnicas estadísticas para el estudio de la desigualdad social”. El Colegio de México-Facultad Latinoamericana de Ciencias Sociales.

Cortés, F. (2013) “Medio siglo de desigualdad en el ingreso en México”. *Economía UNAM* ,10 ,(29), 12-34.

Coulter, P. B. (1989) “Measuring Inequality. A Methodological Handbook”. Westview Press.

Cover, T. y J. A. Thomas (2006) “Elements of Information Theory”. Second Edition. John Wiley y Sons.

Cowell, F. A. (2003) “Theil, Inequality and the Structure of Income Distribution”. Distributional Analysis Research Programme at the London School of Economics and Political Science. Discussion Paper No. 67.

Cowell, F. A. (2005) “Theil, Inequality Indices and Decomposition”. Society for the Study of Economic Inequality ECINEQ. Working Paper No. 2005-01. Disponible en <http://www.ecineq.org/milano/wp/ecineq2005-01.pdf>

- Cowell, F. A. (2009) "Measuring Inequality". London School of Economics and Political Science Papers. Disponible en http://darp.lse.ac.uk/papersDB/Cowell_measuringinequality3.pdf
- Davino, C., Furo, M. y Vistoso, D. ,(2014) "Quantile Regression. Theory and Applications". Wiley series in probability and statistics.
- Daviet, B. (2016) "Revisiting the Role of Education as a Public Good". Education Research and Foresight Working Papers. United Nations Educational, Scientific and Cultural Organization, July.
- Engels, F. (1878) "La revolución de la ciencia de Eugenio Dühring (Anti-Dühring)". Disponible en <https://www.marxists.org/espanol/m-e/1870s/anti-duhring/index.htm>
- Fearon, J. (2003) "Ethnic and Cultural Diversity by Country" *Journal of Economic Growth* ,8 (2), 195-222.
- Flores, I. (1999) "El liberalismo igualitario de John Rawls" *Cuestiones Constitucionales. Revista Mexicana de Derecho Constitucional*,1 (3) jul-dic, 85-122.
- Gertler, P., S. Martinez, P. Premand, L. Rawlings y C. Vermeersch (2011) "Impact Evaluation in Practice". The International Bank for Reconstruction and Development y The World Bank.
- Greene, W.H. (1998) "Análisis econométrico". Tercera edición. Prentice Hall.
- Hanushek, E. y L. Woessmann (2010) "Education and Economic Growth". International Encyclopedia of Education , 2,245-252.
- Herrera, G. (2016) "Varianza: Entropía y la arquitectura de Frank Gehry". Revista Este País. <http://www.estepais.com/articulo.php?id=799&t=varianza-entropia-y-la-arquitectura-de-frank-gehry>
- Heston, A., R. Summers y B. Aten (2001) "Penn World Table Version 6.0". Center for International Comparisons de la Universidad de Pennsylvania.
- Katz, L. y D. Autor (1999) "Changes in the Wage Structure and Earnings Inequality". Handbook of Labor Economics, 3, Ch. 26. 1463-1555.

- Kymlicka, W. (1995) "Filosofía política contemporánea: Una introducción". Editorial Ariel S.A.
- Leontiev, Liev A. (1971) "El capital de Carlos Marx y la Igualdad". Ediciones de Cultura Popular S.A.
- Lindqvist, E. y R. Vestman (2011) "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment". *American Economic Journal: Applied Economics* (3) pp- 101-128.
- Lopez-Calva, L. F. y N. C. Lustig (2010) "Declining Inequality in Latin America. A Decade of Progress?" United Nations Development Programme, New York Brookings Institution Press.
- Lorenz, M.O. (1905) "Methods of Measuring the Concentration of Wealth". Publications of the American Statistical Association, 9, (70), 209-219.
- Luiselli, V. (2008) "La teoría de la justicia como equidad frente a las exigencias de la justicia global: Una crítica a la teoría de Rawls desde el punto de vista de la migración". Tesis de Licenciatura en Filosofía, Facultad de Filosofía y Letras UNAM.
- Lustig, N. y M. Szekely (1997) "México: Evolución económica, pobreza y desigualdad". Banco Interamericano de Desarrollo. Documento de investigación Diciembre 1997. Washington, D.C.
- Mayer-Foulkes, D. y M.F. López (2006) "Transmisión intergeneracional de habilidades cognitivas por niveles socioeconómicos: una aplicación de modelos de swiching" Centro de Investigación y Docencia Económicas, DT No. 378.
- Marx, K. (1867) *El Capital: Crítica de la Economía Política*. Fondo de Cultura Económica, cuarta edición 2014. México. D.F.
- North, D. (1991) "Institutions". *The Journal of Economic Perspectives*, 5, (1) , 97-112.
- PNUD (2006) "México y la agenda pendiente: ¿Por qué empezar por la desigualdad?" Oficina del Informe sobre Desarrollo Humano en México.
- Putnam, R. (2007) "E Pluribus Unum: Diversity and Community in Twenty-First Century". The 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies* 30 (2), 137-174.

- Piskunov, N. (2015) “Cálculo diferencial e integral”. Limusa, Ciudad de México, México.
- Rawls, J. (1971) Teoría de la justicia. Fondo de Cultura Económica, segunda edición 1995. México, D.F.
- Rousseau, J.-J. (1755) “Discurso sobre el origen de la desigualdad”. Porrúa, vigésima edición 2016. México, D.F.
- Rousseau, J.-J. (1762) “El Contrato Social”. Porrúa, vigésima edición 2016. México, D.F.
- Sala-i-Martin, X. (2002) “The Disturbing “Rise” of Global Income Inequality”. National Bureau of Economic Research Working Paper No. 8904.
- Sala-i-Martin, X., G. Doppelhofer y R. Miller (2004) “Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach”. *American Economic Review* 94 (4), 813-835.
- Sen, A. (1995) “Nuevo Examen de la Desigualdad”. Editorial Alianza Editorial. Madrid, España
- Sen, A. (1997) “La Desigualdad Económica”. Fondo de Cultura Económica. Ciudad de México, México.
- Székely, M., L. F. López-Calva, A. Meléndez, E. Rascón y L. Rodríguez-Chamussy (2007) “Poniendo a la pobreza de ingresos y a la desigualdad en el mapa de México”. *Economía Mexicana Nueva Época* XVI (2).
- Torres, D. (2010) “Estudios de la pobreza y la desigualdad en México 2000-2006. Una revisión”. *Economía Informa* no. 365 pp. 54-73
- Urquieta J. E. y A. E. Paredes (2011) “Formation of Cognitive Abilities in Mexican Children (An Empirical Approach)”. Artículo presentado el Jueves 8 de Septiembre de 2011 en la Conferencia Internacional del 70 aniversario de la Revista Investigación Económica en México D.F. Texto disponible en: <http://www.depfe.unam.mx/70aniversario-ie/paredes-urquieta.pdf>
- Weil, D. (2013) “Economic Growth”. Third Edition. Pearson.

Wooldridge, J. M. (2009) “Introducción a la econometría: Un enfoque moderno”. 4ª edición. Cengage Learning.

Yitzhaki, S. y E. Schchtman (2013) “The Gini Methodology: A Primer on a Statistical Methodology”, Springer Series in Statistics 272.

Fuentes Estadísticas.

ENIGH (2016) Microdatos correspondientes a la Nueva serie 2016 de la Encuesta Nacional de Ingresos y Gastos de los Hogares. Consultados en

<http://www.beta.inegi.org.mx/proyectos/enchogares/regulares/enigh/nc/2016/default.html>

SAT (2010) Datos anonimizados (Art. Décimo Noveno Transitorio de la Ley de Ingresos Fiscales 2017). Declaraciones anuales de personas físicas (anonimizadas) 2010. Consultado en

http://www.sat.gob.mx/cifras_sat/Paginas/inicio.html

Anexo1. Límites de $I_{ext}^{\alpha, \kappa}$ cuando $\alpha \rightarrow 0$ y $\alpha \rightarrow 1$.

Caso 1. $\alpha \rightarrow 0$

$$\begin{aligned}
 \lim_{\alpha \rightarrow 0} \frac{\int \left[\left(\frac{x}{\mu} \right)^\alpha - 1 \right] dx}{\alpha^2 - \alpha} &= \\
 &= \lim_{\alpha \rightarrow 0} \frac{\frac{x \left(\frac{x}{\mu} \right)^\alpha}{\alpha + 1} - x}{\alpha^2 - \alpha} \\
 &= \lim_{\alpha \rightarrow 0} \frac{x \left(-1 - \alpha + \left(\frac{x}{\mu} \right)^\alpha \right)}{(\alpha + 1)(\alpha^2 - \alpha)} \\
 &= \left(\lim_{\alpha \rightarrow 0} \frac{1}{\alpha + 1} \right) \left(\lim_{\alpha \rightarrow 0} \frac{x \left(-1 - \alpha + \left(\frac{x}{\mu} \right)^\alpha \right)}{(\alpha^2 - \alpha)} \right) \\
 &= x \left(\lim_{\alpha \rightarrow 0} \frac{\left(-1 + \left(\frac{x}{\mu} \right)^\alpha \right)}{(\alpha^2 - \alpha)} - \lim_{\alpha \rightarrow 0} \frac{\alpha}{\alpha(\alpha - 1)} \right) \\
 &= x \left(\lim_{\alpha \rightarrow 0} \frac{\left(-1 + \left(\frac{x}{\mu} \right)^\alpha \right)}{(\alpha^2 - \alpha)} - \lim_{\alpha \rightarrow 0} \frac{1}{\alpha - 1} \right) \\
 &= x \left(\lim_{\alpha \rightarrow 0} \frac{\left(-1 + \left(\frac{x}{\mu} \right)^\alpha \right)}{(\alpha^2 - \alpha)} + 1 \right)
 \end{aligned}$$

Debido a que se obtiene un límite con forma indeterminada 0/0, se procede a aplicar la regla de L'Hôpital (Revisar Anexo 2).

Al obtener la derivada respecto a α de los factores del límite:

$$\lim_{\alpha \rightarrow 0} \frac{\frac{\partial}{\partial \alpha} \left(\left(\frac{x}{\mu} \right)^\alpha - 1 \right)}{\frac{\partial}{\partial \alpha} (\alpha^2 - \alpha)} = \lim_{\alpha \rightarrow 0} \frac{\left(\frac{x}{\mu} \right)^\alpha \log \left(\frac{x}{\mu} \right)}{2\alpha - 1}$$

Por lo que,

$$\begin{aligned} \lim_{\alpha \rightarrow 0} \frac{\int \left[\left(\frac{x}{\mu} \right)^\alpha - 1 \right] dx}{\alpha^2 - \alpha} &= \\ &= x \left(\lim_{\alpha \rightarrow 0} \frac{\left(\frac{x}{\mu} \right)^\alpha \log \left(\frac{x}{\mu} \right)}{2\alpha - 1} + 1 \right) \\ &= x \left(-\log \left(\frac{x}{\mu} \right) + 1 \right) \end{aligned}$$

El cual, expresado como una integral respecto a x, es:

$$\frac{\partial x \left(-\log \left(\frac{x}{\mu} \right) + 1 \right)}{\partial x} = x \left(\frac{\partial \left(-\log \left(\frac{x}{\mu} \right) \right)}{\partial x} \right) + -\log \left(\frac{x}{\mu} \right) + 1$$

$$\text{Si } \frac{\partial \log(u)}{\partial x} = \frac{\partial \log(u)}{\partial u} \frac{\partial u}{\partial x} \text{ y } u = \left(\frac{x}{\mu} \right), \frac{\partial -\log(u)}{\partial u} \frac{\partial u}{\partial x} = - \left(\frac{1}{\frac{x}{\mu}} \right) \left(\frac{1}{\mu} \right) = -\frac{1}{x}$$

$$\frac{\partial x \left(-\log \left(\frac{x}{\mu} \right) + 1 \right)}{\partial x} = -1 + -\log \left(\frac{x}{\mu} \right) + 1$$

Finalmente:

$$I_{ext}^{\alpha \rightarrow 0, \kappa=0} = \int -\log \left(\frac{x}{\mu} \right) dx = - \int \log(x) - \log(\mu) dx$$

Caso 2. $\alpha \rightarrow 1$

$$\begin{aligned} \lim_{\alpha \rightarrow 1} \frac{\int \left[\left(\frac{x}{\mu} \right)^\alpha - 1 \right] dx}{\alpha^2 - \alpha} &= \\ &= \lim_{\alpha \rightarrow 1} \frac{x \left(\frac{x}{\mu} \right)^\alpha - x}{\alpha^2 - \alpha} \end{aligned}$$

$$= \lim_{\alpha \rightarrow 1} \frac{x \left(-1 - \alpha + \left(\frac{x}{\mu} \right)^\alpha \right)}{(\alpha + 1)(\alpha^2 - \alpha)}$$

Lo que implica un resultado que no existe. Al aplicar la regla de l'Hôpital el límite se identifica:

$$= \frac{\lim_{\alpha \rightarrow 1} \frac{\partial}{\partial \alpha} x \left(\frac{\left(\frac{x}{\mu} \right)^\alpha}{\alpha + 1} - 1 \right)}{\lim_{\alpha \rightarrow 1} \frac{\partial}{\partial \alpha} (\alpha^2 - \alpha)}$$

Dado que

$$\frac{\partial}{\partial \alpha} x \left(\frac{\left(\frac{x}{\mu} \right)^\alpha}{\alpha + 1} - 1 \right) = x \left(\frac{(\alpha + 1) \left(\frac{x}{\mu} \right)^\alpha \log \left(\frac{x}{\mu} \right) - \left(\frac{x}{\mu} \right)^\alpha}{(\alpha + 1)^2} \right)$$

$$\frac{\partial}{\partial \alpha} (\alpha^2 - \alpha) = 2\alpha - 1$$

Entonces,

$$= \frac{\lim_{\alpha \rightarrow 1} \frac{\partial}{\partial \alpha} x \left(\frac{\left(\frac{x}{\mu} \right)^\alpha}{\alpha + 1} - 1 \right)}{\lim_{\alpha \rightarrow 1} \frac{\partial}{\partial \alpha} (\alpha^2 - \alpha)}$$

$$= \lim_{\alpha \rightarrow 1} \frac{x \left(\frac{(\alpha + 1) \left(\frac{x}{\mu} \right)^\alpha \log \left(\frac{x}{\mu} \right) - \left(\frac{x}{\mu} \right)^\alpha}{(\alpha + 1)^2} \right)}{2\alpha - 1}$$

$$\begin{aligned}
&= \lim_{\alpha \rightarrow 1} x \left(\frac{(\alpha + 1) \left(\frac{x}{\mu}\right)^\alpha \log\left(\frac{x}{\mu}\right) - \left(\frac{x}{\mu}\right)^\alpha}{(\alpha + 1)^2 (2\alpha - 1)} \right) \\
&= \lim_{\alpha \rightarrow 1} x \left(\frac{(\alpha + 1) \left(\frac{x}{\mu}\right)^\alpha \log\left(\frac{x}{\mu}\right) - \left(\frac{x}{\mu}\right)^\alpha}{(\alpha + 1)^2 (2\alpha - 1)} \right) \\
&= x \left(\frac{(2) \left(\frac{x}{\mu}\right) \log\left(\frac{x}{\mu}\right) - \left(\frac{x}{\mu}\right)}{4} \right)
\end{aligned}$$

Lo cual, re expresado como una integral de x, obtiene la siguiente forma:

$$\begin{aligned}
&= \frac{\partial}{\partial x} x \left(\frac{(2) \left(\frac{x}{\mu}\right) \log\left(\frac{x}{\mu}\right) - \left(\frac{x}{\mu}\right)}{4} \right) \\
&= \frac{1}{4} \frac{\partial}{\partial x} x \left((2) \left(\frac{x}{\mu}\right) \log\left(\frac{x}{\mu}\right) - \left(\frac{x}{\mu}\right) \right) \\
&= \frac{1}{4} \frac{\partial}{\partial x} x \left(\left(\frac{x}{\mu}\right) \left((2) \log\left(\frac{x}{\mu}\right) - 1 \right) \right) \\
&= \frac{1}{4} \left[x \left(\frac{\partial}{\partial x} \left(\left(\frac{x}{\mu}\right) \left((2) \log\left(\frac{x}{\mu}\right) - 1 \right) \right) \right) + \left(\frac{x}{\mu}\right) \left((2) \log\left(\frac{x}{\mu}\right) - 1 \right) \right] \\
&= \frac{1}{4} \left[x \left(\left(\left(\frac{x}{\mu}\right) \frac{\partial}{\partial x} \left((2) \log\left(\frac{x}{\mu}\right) - 1 \right) + \frac{\left((2) \log\left(\frac{x}{\mu}\right) - 1 \right)}{\mu} \right) \right) + \left(\frac{x}{\mu}\right) \left((2) \log\left(\frac{x}{\mu}\right) - 1 \right) \right]
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4} \left[x \left(\left(\left(\frac{x}{\mu} \right) \left(2 \frac{\mu}{x \mu} + \frac{(2) \log \left(\frac{x}{\mu} \right) - 1}{\mu} \right) \right) \right) + \left(\frac{x}{\mu} \right) \left((2) \log \left(\frac{x}{\mu} \right) - 1 \right) \right] \\
&= \frac{1}{4} \left[x \left(\left(2 \left(\frac{1}{\mu} \right) + \frac{(2) \log \left(\frac{x}{\mu} \right) - 1}{\mu} \right) \right) + \left(\frac{x}{\mu} \right) \left((2) \log \left(\frac{x}{\mu} \right) - 1 \right) \right] \\
&= \frac{1}{4} \left[\left(\frac{2x}{\mu} \right) + \frac{(2x) \log \left(\frac{x}{\mu} \right) - x}{\mu} + \left(\frac{x}{\mu} \right) \left((2) \log \left(\frac{x}{\mu} \right) - 1 \right) \right] \\
&= \frac{1}{4} \left[\left(\frac{2x}{\mu} \right) - \frac{x}{\mu} + \frac{(2x) \log \left(\frac{x}{\mu} \right)}{\mu} + \frac{(2x) \log \left(\frac{x}{\mu} \right)}{\mu} - \left(\frac{x}{\mu} \right) \right] \\
&= \frac{1}{4} \left[\frac{(4x) \log \left(\frac{x}{\mu} \right)}{\mu} \right] \\
&= \int \left(\frac{x}{\mu} \right) \log \left(\frac{x}{\mu} \right) dx
\end{aligned}$$

Anexo 2. La regla de L'Hôpital

De acuerdo con Piskunov (2015: 150-151) el cálculo de límites cuando se presentan indeterminaciones del tipo $0/0$, es posible si se trata de un cociente de dos funciones $f(x)$ y $\varphi(x)$ y además cada una cumple con ser diferenciable en un intervalo dado $[a,b]$

Si en un punto $x=a$ ambas funciones se anulan y esto da origen a la indeterminación $0/0$, y además para todo otro valor a la razón está definida, es posible obtener el límite de esta razón cuando $x \rightarrow a$.

El teorema o regla de L'Hôpital es el siguiente:

Si $\lim_{x \rightarrow 0} f(x) = 0$ y $\lim_{x \rightarrow 0} g(x) = 0$,

Entonces $\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow 0} \frac{f'(x)}{g'(x)}$

Anexo 3. Incidencia de bilingüismo en población hablante de lenguas indígenas, por lengua indígena. ENIGH 2016.

| LENGUA | CODIGO* | NO BILINGÜE | | BILINGÜE | | TOTAL |
|-------------------------------|----------------|--------------------|-----|-----------------|------|--------------|
| NÁHUATL | 1041 | 4,766 | 1% | 606,808 | 99% | 611,574 |
| MAYA | 922 | 4,529 | 1% | 313,442 | 99% | 317,971 |
| TZOTZIL | 934 | 34,229 | 19% | 143,675 | 81% | 177,904 |
| TZETZAL | 933 | 14,748 | 12% | 107,449 | 88% | 122,197 |
| OTOMÍ | 331 | 1,130 | 1% | 98,033 | 99% | 99,163 |
| TOTONACA | 711 | 3,036 | 3% | 92,852 | 97% | 95,888 |
| MIXTECO | 450 | 4,000 | 6% | 65,646 | 94% | 69,646 |
| ZAPOTECO DEL ISTMO | 422 | 630 | 1% | 64,390 | 99% | 65,020 |
| MAZTECO | 491 | 4,698 | 8% | 57,055 | 92% | 61,753 |
| MIXE | 811 | 2,513 | 4% | 58,316 | 96% | 60,829 |
| MIXTECO DE LA COSTA | 451 | 1,084 | 2% | 49,258 | 98% | 50,342 |
| CHOL | 931 | 4,785 | 10% | 42,021 | 90% | 46,806 |
| ZAPOTECO | 400 | 195 | 0% | 44,935 | 100% | 45,130 |
| MAZAHUA | 332 | 0 | 0% | 31,399 | 100% | 31,399 |
| TLAPANECO | 611 | 0 | 0% | 30,649 | 100% | 30,649 |
| POPOLUCA DE LA SIERRA | 821 | 0 | 0% | 29,473 | 100% | 29,473 |
| AMUZGO DE GUERRERO | 482 | 4,869 | 20% | 19,731 | 80% | 24,600 |
| TARAHUMARA | 1021 | 394 | 2% | 23,661 | 98% | 24,055 |
| CHONTAL DE TABASCO | 932 | 0 | 0% | 23,737 | 100% | 23,737 |
| PURÉPECHA | 1111 | 640 | 3% | 21,642 | 97% | 22,282 |
| ZAPOTECO VALLISTA | 421 | 0 | 0% | 21,648 | 100% | 21,648 |
| MAYO | 1022 | 350 | 2% | 20,952 | 98% | 21,302 |
| TOJOLABAL | 935 | 0 | 0% | 20,687 | 100% | 20,687 |
| MIXTECO DE LA MIXTECA BAJA | 453 | 1,257 | 6% | 18,608 | 94% | 19,865 |
| ZAPOTECO SUREÑO | 433 | 1,386 | 7% | 17,777 | 93% | 19,163 |
| HUASTECO | 911 | 0 | 0% | 17,458 | 100% | 17,458 |
| MIXTECO DE LA MIXTECA ALTA | 452 | 40 | 0% | 17,221 | 100% | 17,261 |
| CUICATECO | 461 | 416 | 3% | 15,969 | 97% | 16,385 |
| CHATINO | 441 | 1,557 | 10% | 13,717 | 90% | 15,274 |
| POPOLOCA | 494 | 0 | 0% | 13,753 | 100% | 13,753 |

| | | | | | | |
|------------------------------|------|-----|----|--------|------|--------|
| AMUZGO DE OAXACA | 483 | 0 | 0% | 13,629 | 100% | 13,629 |
| ZAPOTECO DE IXTLÁN | 411 | 0 | 0% | 12,707 | 100% | 12,707 |
| POPOLUCA | 800 | 0 | 0% | 11,254 | 100% | 11,254 |
| CHINANTECO DE OJITLÁN | 211 | 0 | 0% | 9,035 | 100% | 9,035 |
| PAME | 311 | 0 | 0% | 8,958 | 100% | 8,958 |
| ZOQUE | 823 | 0 | 0% | 8,510 | 100% | 8,510 |
| CORA | 1031 | 656 | 8% | 7,691 | 92% | 8,347 |
| HUAVE | 511 | 0 | 0% | 5,911 | 100% | 5,911 |
| CHINANTECO | 200 | 0 | 0% | 4,807 | 100% | 4,807 |
| MAME | 941 | 0 | 0% | 4,429 | 100% | 4,429 |
| TRIQUI | 471 | 0 | 0% | 4,353 | 100% | 4,353 |
| HUICHOL | 1032 | 176 | 4% | 4,140 | 96% | 4,316 |
| NO ESPECIFICADA | 9999 | 0 | 0% | 4,177 | 100% | 4,177 |
| CHICHIMECA | | | | | | |
| JONAZ | 321 | 0 | 0% | 4,035 | 100% | 4,035 |
| PIMA | 1011 | 324 | 9% | 3,204 | 91% | 3,528 |
| ZAPOTECO DE CUIXTLA | 431 | 0 | 0% | 3,375 | 100% | 3,375 |
| CHINANTECO DE VALLE NACIONAL | 232 | 0 | 0% | 3,300 | 100% | 3,300 |
| YAQUI | 1023 | 0 | 0% | 2,664 | 100% | 2,664 |
| CHONTAL DE OAXACA | 131 | 0 | 0% | 1,895 | 100% | 1,895 |
| POPOLUCA DE OLUTA | 812 | 0 | 0% | 1,218 | 100% | 1,218 |
| CHINANTECO DE USILA | 212 | 0 | 0% | 1,083 | 100% | 1,083 |
| MIXTECO DE LA ZONA MAZATECA | 454 | 0 | 0% | 970 | 100% | 970 |
| GUARIJÓ | 1024 | 0 | 0% | 910 | 100% | 910 |
| AMUZGO | 481 | 0 | 0% | 728 | 100% | 728 |
| CHUJ | 936 | 0 | 0% | 699 | 100% | 699 |
| TEPEHUANO DE DURANGO | 1015 | 0 | 0% | 622 | 100% | 622 |
| LACANDÓN | 921 | 0 | 0% | 580 | 100% | 580 |
| PÁPAGO | 1012 | 0 | 0% | 524 | 100% | 524 |
| JACALTECO | 971 | 0 | 0% | 519 | 100% | 519 |
| CHINANTECO DE LALANA | 241 | 0 | 0% | 512 | 100% | 512 |
| CHONTAL | 1311 | 0 | 0% | 405 | 100% | 405 |

| | | | | | | |
|---------------------------|------|---------------|-----------|------------------|------------|------------------|
| KANJOBAL | 961 | 0 | 0% | 311 | 100% | 311 |
| KEKCHI | 981 | 0 | 0% | 292 | 100% | 292 |
| POPOLUCA DE TEXISTEPEC | 822 | 0 | 0% | 261 | 100% | 261 |
| IXCATECO | 493 | 0 | 0% | 229 | 100% | 229 |
| TEPEHUANO | 1013 | 0 | 0% | 164 | 100% | 164 |
| IXIL | 942 | 0 | 0% | 147 | 100% | 147 |
| CAKCHIQUEL | 972 | 0 | 0% | 147 | 100% | 147 |
| KUMIAI | 115 | 0 | 0% | 63 | 100% | 63 |
| CHIAPANECO | 5024 | 0 | 0% | 37 | 100% | 37 |
| Totales | | 92,408 | 4% | 2,230,457 | 96% | 2,322,865 |

*Clasificación de Lenguas Indígenas, 2010

Anexo 4. Resultado de las pruebas de significancia estadística de los coeficientes del escenario observado en cuantiles seleccionados.

| Hipótesis | Etnia | Educacion | Salud |
|-----------------|---|--|--|
| Coef_05=Coef_95 | [q05]etnia - [q95]etnia = 0 F(1, 85251) = 4.29 Prob > F = 0.0382 | [q05]educacion - [q95]educacion = 0 F(1, 85251) = 202.26 Prob > F = 0.0000 | [q05]salud - [q95]salud = 0 F(1, 85251) = 13.58 Prob > F = 0.0002 |
| Coef_15=Coef_85 | [q15]etnia - [q85]etnia = 0 F(1, 85251) = 9.64 Prob > F = 0.0019 | [q15]educacion - [q85]educacion = 0 F(1, 85251) = 546.83 Prob > F = 0.0000 | [q15]salud - [q85]salud = 0 F(1, 85251) = 9.06 Prob > F = 0.0026 |
| Coef_25=Coef_75 | [q25]etnia - [q75]etnia = 0 F(1, 85251) = 5.64 Prob > F = 0.0175 | [q25]educacion - [q75]educacion = 0 F(1, 85251) = 533.69 Prob > F = 0.0000 | [q25]salud - [q75]salud = 0 F(1, 85251) = 3.55 Prob > F = 0.0595 |
| Coef_50=Coef_90 | [q50]etnia - [q90]etnia = 0 F(1, 85251) = 7.57 Prob > F = 0.0059 | [q50]educacion - [q90]educacion = 0 F(1, 85251) = 550.66 Prob > F = 0.0000 | [q50]salud - [q90]salud = 0 F(1, 85251) = 0.02 Prob > F = 0.8769 |
| Coef_10=Coef_50 | - [q10]etnia + [q50]etnia = 0 F(1, 85251) = 2.02 Prob > F = 0.1551 | - [q10]educacion + [q50]educacion = 0 F(1, 85251) = 54.80 Prob > F = 0.0000 | - [q10]salud + [q50]salud = 0 F(1, 85251) = 16.02 Prob > F = 0.0001 |

| Hipótesis | SitViv | Pad_Mad | LenguaIndig |
|-----------------|---|--|---|
| Coef_05=Coef_95 | [q05]sitviv - [q95]sitviv = 0 F(1, 85251) = 7.78 Prob > F = 0.0053 | [q05]pad_mad - [q95]pad_mad = 0 F(1, 85251) = 22.52 Prob > F = 0.0000 | [q05]lenguaindig - [q95]lenguaindig = 0 F(1, 85251) = 145.37 Prob > F = 0.0000 |
| Coef_15=Coef_85 | [q15]sitviv - [q85]sitviv = 0 F(1, 85251) = 11.69 Prob > F = 0.0006 | [q15]pad_mad - [q85]pad_mad = 0 F(1, 85251) = 31.04 Prob > F = 0.0000 | [q15]lenguaindig - [q85]lenguaindig = 0 F(1, 85251) = 255.34 Prob > F = 0.0000 |
| Coef_25=Coef_75 | [q25]sitviv - [q75]sitviv = 0 F(1, 85251) = 7.33 Prob > F = 0.0068 | [q25]pad_mad - [q75]pad_mad = 0 F(1, 85251) = 15.10 Prob > F = 0.0001 | [q25]lenguaindig - [q75]lenguaindig = 0 F(1, 85251) = 127.31 Prob > F = 0.0000 |
| Coef_50=Coef_90 | [q50]sitviv - [q90]sitviv = 0 F(1, 85251) = 20.55 Prob > F = 0.0000 | [q50]pad_mad - [q90]pad_mad = 0 F(1, 85251) = 0.86 Prob > F = 0.3540 | [q50]lenguaindig - [q90]lenguaindig = 0 F(1, 85251) = 27.12 Prob > F = 0.0000 |
| Coef_10=Coef_50 | - [q10]sitviv + [q50]sitviv = 0 F(1, 85251) = 1.92 Prob > F = 0.1657 | - [q10]pad_mad + [q50]pad_mad = 0 F(1, 85251) = 44.83 Prob > F = 0.0000 | - [q10]lenguaindig + [q50]lenguaindig = 0 F(1, 85251) = 191.47 Prob > F = 0.0000 |

| Hipótesis | Tam_loc2 | Tam_loc3 | Tam_loc4 |
|-----------------|--|--|---|
| Coef_05=Coef_95 | [q05]tam_loc2 - [q95]tam_loc2 = 0 F(1, 85251) = 16.16 Prob > F = 0.0001 | [q05]tam_loc3 - [q95]tam_loc3 = 0 F(1, 85251) = 75.75 Prob > F = 0.0000 | [q05]tam_loc4 - [q95]tam_loc4 = 0 F(1, 85251) = 263.80 Prob > F = 0.0000 |
| Coef_15=Coef_85 | [q15]tam_loc2 - [q85]tam_loc2 = 0 F(1, 85251) = 25.94 Prob > F = 0.0000 | [q15]tam_loc3 - [q85]tam_loc3 = 0 F(1, 85251) = 131.79 Prob > F = 0.0000 | [q15]tam_loc4 - [q85]tam_loc4 = 0 F(1, 85251) = 418.07 Prob > F = 0.0000 |
| Coef_25=Coef_75 | [q25]tam_loc2 - [q75]tam_loc2 = 0 F(1, 85251) = 24.03 Prob > F = 0.0000 | [q25]tam_loc3 - [q75]tam_loc3 = 0 F(1, 85251) = 86.05 Prob > F = 0.0000 | [q25]tam_loc4 - [q75]tam_loc4 = 0 F(1, 85251) = 226.73 Prob > F = 0.0000 |
| Coef_50=Coef_90 | [q50]tam_loc2 - [q90]tam_loc2 = 0 F(1, 85251) = 16.91 Prob > F = 0.0000 | [q50]tam_loc3 - [q90]tam_loc3 = 0 F(1, 85251) = 34.79 Prob > F = 0.0000 | [q50]tam_loc4 - [q90]tam_loc4 = 0 F(1, 85251) = 44.70 Prob > F = 0.0000 |
| Coef_10=Coef_50 | - [q10]tam_loc2 + [q50]tam_loc2 = 0 F(1, 85251) = 26.16 Prob > F = 0.0000 | - [q10]tam_loc3 + [q50]tam_loc3 = 0 F(1, 85251) = 42.16 Prob > F = 0.0000 | - [q10]tam_loc4 + [q50]tam_loc4 = 0 F(1, 85251) = 183.24 Prob > F = 0.0000 |

| Hipótesis | Sexo | Edad | Discap |
|-----------------|--|---|--|
| Coef_05=Coef_95 | [q05]sexo - [q95]sexo = 0 F(1, 85251) = 68.57 Prob > F = 0.0000 | [q05]edad - [q95]edad = 0 F(1, 85251) = 481.85 Prob > F = 0.0000 | [q05]discap - [q95]discap = 0 F(1, 85251) = 5.22 Prob > F = 0.0223 |
| Coef_15=Coef_85 | [q15]sexo - [q85]sexo = 0 F(1, 85251) = 40.18 Prob > F = 0.0000 | [q15]edad - [q85]edad = 0 F(1, 85251) = 1082.78 Prob > F = 0.0000 | [q15]discap - [q85]discap = 0 F(1, 85251) = 17.04 Prob > F = 0.0000 |
| Coef_25=Coef_75 | [q25]sexo - [q75]sexo = 0 F(1, 85251) = 21.63 Prob > F = 0.0000 | [q25]edad - [q75]edad = 0 F(1, 85251) = 626.14 Prob > F = 0.0000 | [q25]discap - [q75]discap = 0 F(1, 85251) = 10.20 Prob > F = 0.0014 |
| Coef_50=Coef_90 | [q50]sexo - [q90]sexo = 0 F(1, 85251) = 1.12 Prob > F = 0.2896 | [q50]edad - [q90]edad = 0 F(1, 85251) = 300.70 Prob > F = 0.0000 | [q50]discap - [q90]discap = 0 F(1, 85251) = 6.00 Prob > F = 0.0143 |
| Coef_10=Coef_50 | - [q10]sexo + [q50]sexo = 0 F(1, 85251) = 62.16 Prob > F = 0.0000 | - [q10]edad + [q50]edad = 0 F(1, 85251) = 663.90 Prob > F = 0.0000 | - [q10]discap + [q50]discap = 0 F(1, 85251) = 22.10 Prob > F = 0.0000 |

Anexo 5. Correlación de las variables empleadas en la regresión por cuantiles.

Coeficientes de correlación de Pearson. Regresores de la estimación por cuantiles.

| N = 41,385,988 | sexo | edad | discap | pad_ma d | lengindi g | etnia | educacio n | salu d | sitvi v | tam_loc 2 | tam_loc 3 | tam_loc4 |
|-------------------|--------|--------|--------|-------------|---------------|--------|---------------|-----------|------------|--------------|--------------|----------|
| sexo | 1 | * | * | * | * | * | * | * | * | * | * | * |
| edad | 0.004 | 1 | * | * | * | * | * | * | * | * | * | * |
| discap | 0.009 | 0.141 | 1 | * | * | * | * | * | * | * | * | * |
| pad_mad | -0.047 | 0.291 | 0.000 | 1 | * | * | * | * | * | * | * | * |
| lenguaindi g | 0.045 | 0.026 | 0.007 | 0.044 | 1 | * | * | * | * | * | * | * |
| etnia | 0.025 | 0.001 | 0.035 | 0.025 | 0.344 | 1 | * | * | * | * | * | * |
| educacion | -0.105 | -0.178 | -0.104 | -0.116 | -0.179 | -0.169 | 1 | * | * | * | * | * |
| salud | -0.075 | -0.042 | 0.042 | -0.005 | 0.001 | 0.034 | -0.039 | 1 | * | * | * | * |
| sitviv | 0.060 | -0.176 | -0.004 | 0.189 | 0.140 | 0.125 | -0.285 | 0.032 | 1 | * | * | * |
| tam_loc2 | -0.008 | -0.017 | 0.016 | -0.001 | -0.024 | 0.027 | -0.006 | 0.027 | 0.027 | 1 | * | * |
| tam_loc3 | 0.018 | -0.006 | 0.011 | 0.025 | 0.060 | 0.071 | -0.088 | 0.008 | 0.064 | -0.164 | 1 | * |
| tam_loc4 | 0.094 | -0.031 | 0.006 | 0.030 | 0.172 | 0.174 | -0.278 | 0.034 | 0.146 | -0.204 | -0.190 | 1 |

Nota: Todas las correlaciones fueron significativas al 99.99% de confianza.