



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE POSGRADO EN ECONOMÍA

FACULTAD DE ECONOMÍA ♦ DIVISIÓN DE ESTUDIOS DE POSGRADO

DESIGUALDAD DE OPORTUNIDADES EN MÉXICO: UNA MEDICIÓN
CON VARIABLES LATENTES.

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

Maestra en Economía

PRESENTA:

Flor Elisa Hernández Reyes

TUTOR:

Dr. Curtis Huffman Espinosa

Programa Universitario de Estudios del Desarrollo, UNAM

MIEMBROS DEL JURADO:

Dr. Héctor Nájera Catalán

Programa Universitario de Estudios del Desarrollo, UNAM

Dra. Iliana Yashine Arroyo

Programa Universitario de Estudios del Desarrollo, UNAM

Dra. Laura Vazquez Maggio

Facultad de Economía, UNAM

Dr. Juan Carlos Moreno Brid

Facultad de Economía, UNAM

Ciudad Universitaria, Cdmx, Mayo 2021



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.

A mi familia

Reconocimientos

Esta tesis representa una de las experiencias más satisfactorias que he tenido y no puedo más que expresar agradecimiento con cada una de las maravillosas personas que me han acompañado a lo largo de este camino. En primer lugar agradezco al Dr. Curtis Huffman Espinosa no solo por ser un gran guía y apoyo durante estos años sino también por ser una gran persona y fuente de inspiración desde el día en que lo conocí. Gracias por ser un asesor comprometido con sus estudiantes, por estar siempre dispuesto a ayudar y por todo lo enseñado durante estos años. A los integrantes de mi jurado, el Dr. Héctor Nájera, la Dra. Iliana Yaschine, la Dra. Laura Vazquez y el Dr. Juan Carlos Moreno Brid, gracias por haber compartido conmigo su experiencia, por haber dedicado tiempo a la revisión de mi trabajo y por sus valiosos comentarios, me siento muy honrada.

Durante toda la vida he sido afortunada de contar con el amor y el apoyo incondicional de la maravillosa familia que tengo, muchas gracias a mi papás Juan Hernández y Reina Reyes por ser un gran ejemplo y por motivarme a ser mejor cada día, gracias a mi hermana Abril Hernández, por ser la mejor amiga y compañera de vida. A mi mamá Mil, la abuelita más cariñosa y a Abraham Cruz por siempre creer en mí más de lo que a veces yo lo hago, por ser mi cómplice en la vida y por siempre estar a mi lado. A Eduardo Salinas por convertirse en un gran amigo y ser parte de la familia. Siempre estaré agradecida con ellos y es por eso que les dedico esta tesis.

A mis amigos de la maestría, David Flores, Eduardo Castro, Marvin Trejo, Omar Ruiz, Carolina Morgan, Dulce Fuentes, Marissa Pérez, Cosijoeza Zárate y Rocío Aguilar, gracias por las aventuras y días de estudio, porque sin su amistad y el gran equipo que formamos esto no hubiera sido posible. Gracias a Cruz Eduardo Vera por su amistad, apoyo y buenos consejos en la edición de esta tesis. A Andrea Larios por motivarme todos los días y ser una amiga inigualable. A Luis Ángel Monroy Gómez Franco porque sus trabajos inspiraron esta tesis, por su apoyo durante su revisión y al Dr. Carlos A. López por siempre estar disponible para dar palabras de apoyo y los mejores consejos de vida. Gracias al Posgrado de la Facultad de Economía y al Consejo Nacional de Ciencia y Tecnología por la beca otorgada durante toda la maestría.

Resumen

Las estimaciones de desigualdad de oportunidades dependen notablemente de la manera en que se miden y correlacionan los logros de vida y el origen social, fenómenos que no se observan directamente. Esto resulta en estimaciones atenuadas de desigualdad de oportunidades, lo que sugiere que han perdido intensidad y que podrían ser mayor de lo que serían de llevarse a cabo una medición adecuada. En este trabajo se construyen medidas confiables de logros de vida y de origen social a partir del análisis factorial confirmatorio y del análisis de confiabilidad. Se encontró conveniente medir origen social y logros de vida a partir de indicadores de bienestar económico y una correlación observada a nivel nacional de 0.49, la cual resulta en 0.59 después de desatenuarla. Es muy probable que las medidas de desigualdad de oportunidades revisadas en la bibliografía estén subestimadas por un 20 por ciento.

Palabras clave: logros de vida, origen social, coeficiente de correlación, confiabilidad, desatenuación y desigualdad de oportunidades.

Índice general

Índice de figuras	IX
Índice de tablas	XI
1. Introducción	1
2. Marco de referencia	5
2.1. Desigualdad de oportunidades y su delimitación conceptual	5
2.2. Desigualdad de oportunidades a través de su medición	6
3. Metodología	9
3.1. Teoría de la medición	9
3.1.1. Métodos	10
3.1.1.1. Análisis factorial confirmatorio y análisis de confiabilidad	10
3.1.1.2. Desatenuación del coeficiente de correlación	13
3.1.2. Variables utilizadas en la medición de desigualdad de oportunidades	15
3.1.2.1. Descripción de la base de datos	16
4. Resultados de la estimación	19
4.1. Análisis factorial confirmatorio y análisis de confiabilidad	19
4.1.1. Logros de vida	20
4.1.2. Origen social	21
4.2. Desatenuación de la correlación entre logros de vida y origen social . . .	23
5. Conclusiones	27
A. Anexo	31
A.1. Cuadros y esquemas adicionales/Código en R	31
Bibliografía	45

Índice de figuras

3.1. Variables utilizadas en la bibliografía para medir desigualdad de oportunidades	14
3.2. Regiones ESRU-EMOVI 2017	17
4.1. Ilustración Modelo I Logros de vida a través del bienestar económico actual	20
4.2. Ilustración Modelo II Origen social a través del bienestar económico a los 14 años	22
A.1. Modelo I Logros de vida medido a través del bienestar económico	31
A.2. Modelo II Origen social medido a través del bienestar económico	32
A.3. Modelo III Logros de vida medido a través del acceso a servicios públicos	33
A.4. Modelo IV Origen social medido a través del acceso a servicios públicos	34
A.5. Código en R. Desigualdad de oportunidades en México: una medición con variables latentes	37

Índice de tablas

3.1. Índices de ajuste en un modelo CFA	11
3.2. Variables utilizadas en la bibliografía para medir desigualdad de oportunidades	18
4.1. Estadísticos de ajuste y confiabilidad para modelos I y II	22
4.2. Correlación entre logros de vida y origen social: bienestar económico . .	25
A.1. Estadísticos descriptivos para medir logros de vida y origen social	35
A.2. Estadísticos de ajuste y confiabilidad para modelos alternativos (III y IV)	36

Introducción

Reducir las desigualdades sociales es uno de los retos más grandes que enfrentan nuestras sociedades. Fallar en la importante tarea de monitorear y controlar apropiadamente la desigualdad puede conducir fácilmente a catástrofes políticas, económicas y sociales (Alvaredo, Chancel, Piketty, Saez, y Zucman, 2018). Esto es particularmente acuciante en América Latina donde países como México figuran entre los más desiguales del mundo (Oxfam México, 2018).

Actualmente, el análisis empírico de la desigualdad gira en torno al concepto de igualdad de oportunidades. Este concepto emerge de un campo teórico y empírico que entretiene la idea general de que, al analizar la desigualdad económica, lo más importante es comprender y atender el origen de ésta. En el análisis empírico de los factores detrás de la desigualdad (i.e. los determinantes de los logros de vida), lo usual es distinguir entre aquellos factores asociados al origen social de los individuos, ajenos a su control (sexo, género, color de piel, contexto familiar, región de nacimiento, bienestar económico en el hogar de origen, suerte, talento, etc.), y otros atribuibles al individuo como el esfuerzo (Roemer, 1998).

De esta manera, el análisis empírico de la desigualdad de oportunidades distingue entre fuentes de desigualdad justas (de responsabilidad individual) e injustas (ajenas al individuo). En la medida en la que, en una sociedad, la exposición de los individuos a circunstancias más allá de su control, como su origen social, se encuentre más estrechamente ligada o correlacionada con variaciones en sus logros de vida, la sociedad en cuestión será considerada menos justa y la desigualdad existente debería eliminarse.

Por el contrario, en una sociedad justa, con igualdad de oportunidades, no habría de observarse correlación alguna entre las distribuciones empíricas de los logros de vida y el origen social de los individuos. Esto no quiere decir que en dicha sociedad no haya desigualdad, sino que ésta resulta éticamente aceptable al depender de factores de responsabilidad individual como el esfuerzo o el talento (Piketty, 2014). Esta desigualdad, dentro de ciertos límites razonables, se alinea con la idea de la utilidad social derivada

1. INTRODUCCIÓN

de incentivar el trabajo sobresaliente y es en esa medida necesaria.

Analizar adecuadamente la desigualdad de oportunidades requiere buenas mediciones tanto de los logros de vida como del origen social de los individuos. Por ello las preguntas de investigación que motivan esta tesis son las siguientes: ¿cómo se mide la desigualdad de oportunidades y qué dificultades o limitaciones se encuentran al medirla?, ¿cómo es posible corregir las limitaciones encontradas en la medición de desigualdad de oportunidades?.

Para estimar adecuadamente la desigualdad de oportunidades lo ideal sería contar con observaciones precisas de los logros de vida y del origen social. Sin embargo, en ciencias sociales la medición no es tan simple como en ciencias exactas debido a que lo que se quiere medir son atributos no físicos. No es posible observar a una persona y conocer con exactitud cuál fue su origen social cuáles fueron sus logros de vida, no existe un dispositivo o herramienta para medir esto. En lugar de ello tenemos que construir entidades teóricas conocidas como constructos o variables latentes a partir de indicadores observables que caractericen y se aproximen a lo que queremos medir. De no llevarse a cabo una medición adecuada de logros de vida y origen social, resultarán estimaciones que subestiman la verdadera desigualdad de oportunidades (Hausman, 2001).

La hipótesis que pone a prueba esta tesis es que una mala medición de los constructos que tienen que ver con desigualdad de oportunidades, logros de vida y origen social, conduce a estimaciones subestimadas que indican que hay menos desigualdad de lo que sugeriría otra medida y de lo que realmente existe. Esto limita seriamente la comparabilidad en la investigación sobre desigualdad (Donni, Rodríguez, y Dias, 2015; Wendelspiess Chavez Juarez, 2015). Por este motivo esta tesis busca contribuir a la bibliografía sobre la desigualdad de oportunidades mediante la selección de un conjunto de indicadores finito y confiable que nos permita medir mejor nuestros conceptos de interés: logros de vida y origen social y con ello estimar su correlación insesgada por error de medición.

Además, en esta tesis se discute y se aclara el verdadero problema de los análisis sobre desigualdad de oportunidades, que no es solo buscar mejores métodos de estimación, como típicamente se reduce, sino también y aún más importante llevar a cabo una medición adecuada de los conceptos involucrados. El hacer esto nos lleva a hacer la distinción entre el error de estimación y el error de medición, que es relevante porque de no hacerla y corregir únicamente el error de estimación, trivializar la medición y suponer que las variables utilizadas han sido medidas sin error, como comúnmente se hace, limitará nuestro análisis al no considerar el impacto del error de medición y nos llevará a buscar soluciones en lugares incorrectos (Bandalos, 2018; Revelle, W., y Condon, 2014).

La teoría de la confiabilidad, como parte de la teoría de la medición, será utilizada como punto de partida para que, junto con una apropiada técnica de agregación, sea

posible reducir el error de medición y el sesgo de desigualdad de oportunidades. Adicionalmente se presenta una discusión sistemática sobre la corrección de los coeficientes de correlación entre los constructos origen social y logros de vida. La corrección de la atenuación es el uso más importante de la teoría de la confiabilidad al permitir que se estime la correlación verdadera entre dos constructos, en este caso origen social y logros de vida, cuando estos son medidos perfectamente sin error de medición; lo cual requiere encontrar la confiabilidad de cada constructo por separado (Revelle, W., y Condon, 2014). Cuando un coeficiente de correlación está atenuado, éste ha perdido potencia o intensidad y sugiere que la correlación obtenida podría ser más importante de lo que sería si se llevara a cabo una medición adecuada de los constructos de interés. Por ello el desatenuar la relación entre logros de vida y origen social nos aproximará a la verdadera (insesgada) desigualdad de oportunidades en México y sus regiones durante el 2017 al ser el año más reciente para el cual se tienen datos disponibles de la Encuesta de Movilidad Social (ESRU-EMOVI). Para realizar lo anterior se propone un nuevo abordaje a la medición de la desigualdad de oportunidades. A través del análisis factorial confirmatorio, el análisis de confiabilidad y la desatenuación de la correlación entre los logros de vida y origen social se contribuye a la operacionalización de la desigualdad de oportunidades a partir del enfoque propuesto por Roemer (1998).

El desarrollo de la tesis se divide en los siguientes apartados. Después de esta breve introducción, el segundo apartado presenta el marco de referencia donde se describe el abordaje teórico y empírico de la igualdad de oportunidades como teoría de justicia distributiva, así como los desafíos presentes en su medición. En un tercer apartado se presentan los métodos de la teoría de la medición involucrados en nuestra propuesta metodológica y la base de datos a usar. En el siguiente apartado se discuten los resultados obtenidos y por último un apartado de comentarios a modo de conclusiones.

Marco de referencia

2.1. Desigualdad de oportunidades y su delimitación conceptual

La igualdad de oportunidades es un concepto de justicia social que ha sido trabajado por autores como John Rawls (1971), Ronald Dworkin (1981), Amartya Sen (1985) y John Roemer (1998). Este concepto parte de analizar los logros de vida de cada individuo como atribuibles a factores ajenos a su control (sexo, género, color de piel, contexto familiar, región de nacimiento, bienestar económico en el hogar de origen, suerte, talento, etc.), los cuales llamaremos origen social; así como a factores que sí controla como el esfuerzo individual (Roemer, 1998). Entre mayor sea la correlación del origen social con variaciones en sus logros de vida, la sociedad se considerará menos justa y existirá desigualdad de oportunidades (Hufe, Kanbur, y Peichl, 2018). En contraste, en una sociedad justa, con igualdad de oportunidades, la correlación entre las distribuciones empíricas de los logros de vida y el origen social no habría de observarse. Así la desigualdad existente puede considerarse como éticamente aceptable ya que depende de factores de responsabilidad individual como el esfuerzo.

Al medir el grado de desigualdad de una sociedad, normalmente se analiza la distribución que se observa en los logros de vida de los individuos. Esta distribución resulta de un proceso que involucra el origen social, y que es clave al analizar si las desigualdades son justas o injustas (Hufe, et al., 2018; Kanbur y Snell, 2019). Sin embargo, la no observabilidad directa de este proceso que generan los distintos resultados en los logros de vida de la población conduce generalmente a un análisis de desigualdad se queda tan solo en los logros de vida, por ejemplo, el nivel de ingresos, que son lo único observable (Kanbur y Snell, 2019).

Medir la desigualdad de oportunidades busca recuperar parte de ese proceso que genera los logros de vida al identificar qué proporción de éstos se explican por el origen

social (Hufe, et al.,2018). Para realizar este análisis, típicamente se supone que: i) el esfuerzo es independiente del origen social y ii) el origen social es observable. No obstante, la baja probabilidad de que estos supuestos se cumplan produce estimaciones de desigualdad de oportunidades atenuadas, es decir que lo que se estima es la mínima desigualdad con la que se cuenta (Soloaga y Wendelspiess, 2010). Para corregir esta atenuación entre la correlación de los logros de vida y el origen social, esta tesis se centrará en construir dos medidas confiables de logros de vida y el origen social como primer paso para aproximarnos a la estimación de desigualdad de oportunidades en México durante 2017.

2.2. Desigualdad de oportunidades a través de su medición

Al revisar el abordaje que se le da en la bibliografía al concepto de desigualdad de oportunidades cuando se estima la correlación entre el origen social y los logros de vida es posible distinguir una metodología clásica. Esta consiste en identificar la variación en el logro de vida atribuible al origen social a través de dos etapas. La primera de ellas estima una distribución de logros de vida (típicamente el ingreso) en un escenario justo, con igualdad de oportunidades. En la segunda se mide la divergencia entre dicha distribución y una distribución de logros de vida resultante de las condiciones actuales observadas, aquella en donde la correlación entre origen social y logros de vida es distinta de cero. La desviación de estos escenarios constituye un índice que mide el peso de la desigualdad injusta en una distribución de resultados dada (Kanbur y Wagstaff , 2016 en Hufe, et al., 2018). Dentro de los autores que toman este método como base para sus trabajos empíricos destacan Paes de Barros, Ferreira, Vega, y Chanduvi (2009) ; Bourguignon, Ferreira, y Menéndez (2007), Ferreira y Gignoux (2011), Ferreira y Meléndez (2012); Santos Ochoa (2016), entre otros.

Durante los últimos años diversos autores como Ferreira y Gignoux (2011), Wendelspiess Chavez Juarez (2015), Brunori, Peragine, y Serlenga (2019), Hufe et al., (2018) y otros han reconocido que el problema que enfrenta esta metodología es la no observabilidad u observabilidad parcial de no sólo el origen social, que está fuera del control del individuo, sino también de los logros de vida al ser fenómenos que no se observan directamente. Esto tiene como consecuencia que el tamaño de la correlación entre origen social y logros de vida resulte subestimada o atenuada y en estimaciones no confiables de desigualdad de oportunidades. Ferreira y Gignoux (2011) señalan que las estimaciones de desigualdad de oportunidades están subestimadas porque el excluir variables, en este caso de origen social, decrece el coeficiente de determinación de todo modelo. Adicionalmente, Wendelspiess Chavez Juarez (2015) discute que la medición imprecisa de las variables de logros de vida termina sesgando las estimaciones obtenidas, lo cual

limita la interpretación de los resultados.

A causa de la no observabilidad directa de los fenómenos de interés, al medir desigualdad de oportunidades nos enfrentamos al error de medición definido como la variación aleatoria de una distribución que produce una diferencia entre los valores reales y los observados (Loken y Gelman, 2017). No obstante el enfoque clásico reduce el problema al error de estimación y busca solucionarlo al producir mejores métodos de estimación, pero no se trata sólo de un problema de estimación, sino de un problema de medición en el cual el error de medición agrega ruido, incrementa la incertidumbre en los parámetros estimados (Loken y Gelman, 2017), y afecta la correlación observada entre los fenómenos subyacentes de interés, al hacer que ésta esté atenuada o sea menor a la que debería esperarse cuando se lleva a cabo una correcta medición de dichos fenómenos (Trafimow, 2016).

Para mejorar la medición de la correlación entre logros de vida y origen social, algunos autores proponen la estimación de modelos con clases latentes (Donni et al., 2015), simulaciones Monte Carlo (Lara Ibarra y Martínez Cruz, 2015), análisis factorial (Wendelspiess Chavez Juarez, 2015) y modelos de ecuaciones estructurales (Nogales Carvajal, 2016). En particular, Donni et al., (2015) y Wendelspiess Chavez Juarez (2015) insertan sus trabajos en la teoría de la medición, la cual es muy útil en el marco de la desigualdad de oportunidades porque permite no solo describir las relaciones entre las variables observadas, sino también aquellas no observadas. Bajo esta teoría, la selección correcta de indicadores disminuye el error de medición y hace más confiables los resultados obtenidos. Además, no se basa en los supuestos tradicionales que a menudo se violan en la práctica (relación lineal, distribución normal, homogeneidad), por lo que se está sujeto a un menor sesgo (Magidson y Vermunt, 2002) y con ello es posible mejorar las estimaciones mínimas o atenuadas indicadas por autores como Ferreira y Gignoux (2011).

Hasta este punto hemos visto que las estimaciones de igualdad de oportunidades dependen notablemente de la manera en que se miden los logros de vida y el origen social. Lo ideal sería contar con observaciones de estos dos fenómenos de interés, sin embargo, esto no es posible en la práctica porque tanto origen social como logros de vida no se observan directamente, lo que limita seriamente la interpretación y la comparabilidad en la investigación sobre desigualdad. En particular, esto tiene como resultado que la mayoría de las estimaciones realizadas representen límites inferiores que subestiman la verdadera (desatenuada) desigualdad de oportunidades. Si bien en la bibliografía revisada estos problemas se abordan desde el punto de vista de la estimación, esta tesis enfatiza que también es un problema de medición y por ello, a partir del enfoque propuesto por Roemer (1998), el objetivo de esta tesis es construir medidas confiables de logros de vida y de origen social que nos permitan conocer la correlación verdadera que existe entre logros de vida y origen social dando cuenta de los errores de medición con el fin de aproximarnos a una estimación insesgada de la desigualdad de oportunidades en México utilizando datos de 2017.

3.1. Teoría de la medición

Las contribuciones a la medición de desigualdad de oportunidades comparten la idea común de relacionar indicadores de logros de vida con indicadores de origen social. Sin embargo, una limitación clave de las estimaciones derivadas de métodos como mínimos cuadrados ordinarios o el análisis de regresión múltiple, utilizado típicamente en medición de desigualdad de oportunidades, es la suposición de que las variables usadas se han medido sin errores, es decir, son perfectamente confiables (más adelante se explicará a detalle el significado de confiabilidad). Dado que esta suposición rara vez ocurre en las ciencias sociales, las estimaciones suelen estar atenuadas a un grado desconocido (Brown, 2015) (4) ya que los fenómenos o constructos analizados no son directamente observados (Wendelspiess Chavez Juarez, 2015).

La medición como base de toda investigación científica representa un gran desafío sobre todo en las ciencias sociales puesto que los fenómenos de interés no son directamente observables, tal es el caso de fenómenos como la inteligencia, la ansiedad, la pobreza, el altruismo, etc. En el caso particular de esta tesis no es posible observar a una persona y conocer con exactitud su origen social y sus logros de vida, no existe un dispositivo o herramienta para medirlo. En lugar de ello tenemos que construir entidades teóricas conocidas como constructos, variables latentes, factores o variables inobservables, a partir de indicadores observables que den cuenta de características de cada constructo (Bandalos, 2018).

La teoría de la medición es un marco analítico que nos permite vincular la teoría con la experimentación al mismo tiempo que asegura que los resultados obtenidos sean confiables, válidos y comparables. La idea principal es que los indicadores observados son una manifestación de un fenómeno subyacente (Spearman, 1904; en Nájera Catalán y Gordon, 2019) y en esta tesis habremos de abordar la desigualdad de oportunidades en México bajo este enfoque. En particular se propone realizar un análisis factorial

3. METODOLOGÍA

confirmatorio, un análisis de confiabilidad y desatenuación de la correlación entre origen social y logros de vida, como una primera aproximación a una medida confiable de desigualdad de oportunidades en México para 2017. Estos métodos permiten estimar relaciones entre constructos no observables después de ajustar por error de medición (Brown, 2015).

3.1.1. Métodos

3.1.1.1. Análisis factorial confirmatorio y análisis de confiabilidad

El análisis factorial confirmatorio (CFA) es un tipo de modelo de ecuaciones estructurales (SEM) que nos permite identificar un grupo de indicadores que varían y covarían en función de fenómenos latentes (Brown, 2015). La especificación del CFA está fuertemente impulsada por la teoría o la evidencia de investigación previa, pues será ésta la que sirva de guía en la especificación y evaluación del modelo. Al realizar un CFA se debe especificar toda la estructura del modelo, es decir, qué variables se utilizan y en qué factor(es), qué factores están correlacionados (si los hay) y si alguna de las varianzas de error de medición debe estar correlacionada (Bandalos, 2018).

Para llevar a cabo un CFA se necesita en primer lugar especificar el modelo. Las relaciones entre los indicadores observados y los factores se pueden representar a través de la siguiente ecuación:

$$X_{iv} = \lambda_{v1}\xi_{1i} + \lambda_{v2}\xi_{2i} + \dots + \lambda_{vf}\xi_{fi} + \epsilon_{iv} \quad (3.1)$$

$$X_{iv} = \sum_{j=1}^f \lambda_{vj}\xi_{ji} + \epsilon_{iv} \quad (3.2)$$

Donde X_{iv} representa los indicadores observados con subíndice i para persona y v para indicador. El término λ_{vf} es el coeficiente de las cargas factoriales del indicador v en el factor f , las cuales representan la cantidad de varianza en el indicador que se debe a, o es explicada por el factor (cuanto más explique el factor la variabilidad del indicador menor será el error y será mayor la confiabilidad a nivel de indicador). ξ_{fi} es la puntuación de la persona i en el factor f y ϵ_{iv} es el error asociado a la persona i en el indicador v , este error es la varianza del indicador que no tiene que ver con la variación de los factores latentes y se le atribuye al error de medición. La ecuación 3.1 y la 3.2 representan la puntuación variable (X) de una sola persona, pero en CFA nos interesa explicar las relaciones entre todas las variables entre todos los individuos (Bandalos, 2018).

Una vez especificado el modelo se identificará si (1) hay al menos tres variables que miden cada factor, (2) cada variable carga en un solo factor, y (3) los errores de medición no están correlacionados. Aunque estas condiciones son suficientes para la identificación, los modelos pueden identificarse si no se cumplen una o más de las condiciones (Bollen 1989; en Bandalos, 2018). El ajuste del CFA estará determinado por el tamaño de los parámetros estimados (por ejemplo, la magnitud de las cargas factoriales) y en qué medida las relaciones observadas entre las variables reflejen el fenómeno establecido desde la teoría.

Para evaluar la calidad del modelo es importante considerar la bondad de ajuste del mismo, pero sobre todo la interpretabilidad de las estimaciones resultantes. La bondad de ajuste está determinada por la capacidad de la medida y las partes estructurales de ésta. Un aspecto clave es la capacidad de los parámetros para reproducir las relaciones observadas entre los indicadores (Brown, 2015). Chi-cuadrada es la primera prueba de bondad de ajuste y como depende en gran medida del tamaño de la muestra es conveniente evaluar con otros índices de ajuste, tales como el índice de ajuste comparativo (CFI), el índice de Tucker-Lewis (TLI) y el error cuadrático medio de aproximación (RMSEA), todos ellos con valores entre 0 y 1. En términos generales, para que el modelo se considere apropiado, los valores de CFI y TLI deben ser mayores que 0.95 y los de RMSEA inferiores a .05. A continuación se presenta una tabla con los valores recomendados para aceptar el ajuste de un modelo CFA.

Tabla 3.1: Índices de ajuste en un modelo CFA

Índice de Ajuste	Valores para un buen ajuste del modelo
Chi-cuadrada	Aumenta con el tamaño de la muestra
Error cuadrático medio de aproximación (RMSEA)	≤ 0.05
Índice comparativo (CFI)	≥ 0.95 (entre más alto mejor)
Índice de Tucker- Lewis (TLI)	≥ 0.90 (CFI es siempre mayor que TLI)

Fuente: Bandalos, 2018 y UCLA, 2020.

Los modelos CFA también pueden utilizarse para obtener estimaciones de confiabilidad, por ello una vez realizada la estimación de las medidas tanto de origen social como de logros de vida se evaluará cada una en términos de la confiabilidad.

La confiabilidad se refiere a la consistencia interna que existe entre los indicadores que buscan medir algún constructo. Esta consistencia permite que se pueda distinguir y ordenar a la población y que el error de medición sea bajo (Bandalos, 2018; Nájera Catalán y Gordon, 2019; Revelle y Condon, 2019). Este concepto también se refiere a la homogeneidad de una medida, aquella en la que todos sus indicadores son causados

3. METODOLOGÍA

por el mismo fenómeno (Revelle y Zinbarg, 2009). Para medir confiabilidad, existen aquellos estadísticos que tienen que ver con la teoría clásica del test, alfa y beta, y el estadístico de variables latentes, omega. Estos estadísticos tienen valores que van del cero al uno y mientras más cercano a uno más confiable serán los puntajes de la medida. Además si estos indicadores se acercan a uno, entonces los indicadores en conjunto resultarán consistentes y se podrá llevar a cabo una correcta medición de la variable latente y su correcta clasificación de la población (Nájera Catalán y Gordon, 2019). A continuación se presentan cada uno de los estadísticos de confiabilidad.

Alpha α : el coeficiente alfa de Cronbach (Cronbach, 1951; en Revelle y Condon, 2019) supone que cada indicador mide el mismo constructo latente en la misma escala y con el mismo grado de precisión. Un valor alto de alfa nos indicaría que todos los indicadores son útiles para entender el constructo de interés; a medida que el valor de alfa disminuye, debemos considerar que quizá haya variables que estén interfiriendo en nuestra medición. Los valores de $\alpha < 0.7$ indican baja confiabilidad (Nájera, 2018; en Nájera Catalán y Gordon, 2019). Si bien es el estadístico más utilizado de consistencia interna de una prueba, los supuestos que considera son muy restrictivos lo cual es una condición insostenible en la práctica (Revelle y Zinbarg, 2009).

Beta β : este coeficiente estima el peor escenario (Revelle, 1979; en Revelle y Condon, 2019), lo cual pasa cuando se considera la menor correlación posible entre los indicadores del índice. Cuando $\beta < 0.5$ este es indicio de una baja consistencia interna y muestra que al menos uno de los indicadores no tiene que ver con el constructo latente que se quiere analizar (Nájera Catalán y Gordon, 2019).

Existen dos estimaciones alternativas de confiabilidad que tienen en cuenta la estructura jerárquica de las medidas, omega jerárquico ω_J y la confiabilidad total de una medida, omega total ω_T (McDonald, 1999; en Revelle y Condon, 2019). Estos estadísticos se enmarcan dentro de los modelos de variables latentes y normalmente corresponden al análisis factorial confirmatorio.

Omega total ω_T y omega jerárquico ω_J : omega es el mejor estimador de confiabilidad o techo de confiabilidad. Un omega alto indica que los puntajes o scores y el ordenamiento de la población es consistente, lo cual implica que el error de medición es bajo y los scores son confiables. Es deseable un valor de $\omega > 0.8$ para asegurar una clasificación consistente de la muestra (Nájera, 2018; en Nájera Catalán y Gordon, 2019).

El estadístico omega se basa en que la puntuación observada se puede descomponer en cuatro partes, un factor general, factores formados por grupos de indicado-

res (factores de grupo), factores específicos de cada indicador y variación aleatoria (Revelle y Condon, 2019). ω_J es importante cuando evaluamos la importancia y confiabilidad del factor general de una medida, mientras que el omega total ω_T , es un estimador de la confiabilidad total de los puntajes de una medida.

Diversos autores han comparado los estadísticos de omega total y omega jerárquico de McDonald con el alfa de Cronbach y la beta de Revelle y han concluido que la mejor estimación es omega (Zinbarg et al., 2005; Revelle y Zinbarg, 2009; en Revelle y Condon, 2019). Al ser el marco de variables latentes más potente y flexible, en esta tesis se pondrá mayor atención al estadístico omega total.

3.1.1.2. Desatenuación del coeficiente de correlación

Verificar la confiabilidad es el primer paso que proponemos en la medición de la desigualdad de oportunidades, esto es importante en la selección adecuada de un conjunto de indicadores al mismo tiempo que es necesario para atender y calcular el problema de atenuación de la correlación entre el origen social y los logros de vida. Cuando un coeficiente de correlación está atenuado, éste ha perdido potencia o intensidad y sugiere que la correlación obtenida podría ser más importante de lo que parece.

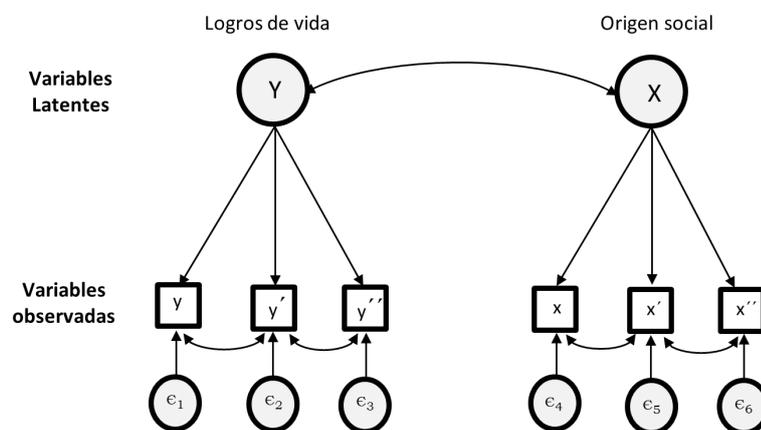
Consideremos una persona que realiza una prueba de conocimientos y obtiene una puntuación alta. El puntaje obtenido podría deberse a que la persona tiene un alto nivel de conocimiento y/o a cuestiones aleatorias como la buena suerte que la persona pudo haber tenido ese día. Si la misma persona realiza la prueba dos veces, bajo las mismas condiciones, y obtiene aproximadamente los mismos resultados, entonces diremos que los puntajes son confiables y esperaríamos que estos puntajes estuvieran altamente correlacionados. Entonces a menor aleatoriedad en los puntajes de las pruebas, mayor coeficiente de correlación y mayor confiabilidad. Por el contrario, a mayor aleatoriedad en los puntajes, menor coeficiente de correlación y menor confiabilidad. Esto último se conoce como el efecto de atenuación por el cual las magnitudes de los coeficientes de correlación obtenidos son más bajos (cerca de cero) de lo que se esperaría si las variables fueran perfectamente medidas (Trafimow, 2016).

La relación entre dos constructos se atenuará por el nivel de confiabilidad de una medida si este es bajo (Revelle, W., y Condon, 2014), por ello es importante asegurar un nivel de confiabilidad alto para cada una de nuestras medidas. En la figura 3.1 se describe de manera gráfica el modelo de variables latentes que se pretende realizar en la tesis, como podemos observar las variables latentes de interés están relacionadas entre sí y cada una con un conjunto de indicadores observados.

El desarrollo original de la teoría de la confiabilidad fue estimar la correlación de variables latentes en términos de correlaciones observadas corregidas por su confiabi-

3. METODOLOGÍA

Figura 3.1: Variables utilizadas en la bibliografía para medir desigualdad de oportunidades



Fuente: elaboración propia basada en el esquema de Revelle (2019)

alidad (Spearman, 1904). El efecto de atenuación se refiere a la magnitud en la cual los coeficientes de correlación son menores que los esperables si se midieran sin error alguno y tuviera perfecta confiabilidad (Trafimow, 2016). La correlación atenuada se calcula como sigue.

$$r_{XY} = R_{XY} \sqrt{\text{Confiabilidad}_X \text{Confiabilidad}_Y} \quad (3.3)$$

Donde r_{XY} es la correlación observada entre los dos constructos que en este caso son el origen social y los logros de vida, Confiabilidad_X es el coeficiente omega obtenido en el análisis del constructo “origen social” y Confiabilidad_Y es el coeficiente omega obtenido en el análisis del constructo “logros de vida”. Por su parte R_{XY} se refiere a la correlación verdadera o desatenuada que se esperaría si los resultados de ambos constructos fueran perfectamente confiables. Para calcular R_{XY} se tiene la siguiente ecuación.

$$R_{XY} = \frac{r_{XY}}{\sqrt{\text{Confiabilidad}_X \text{Confiabilidad}_Y}} \quad (3.4)$$

La correlación desatenuada entre dos constructos será el cociente entre su correlación observada y la raíz cuadrada del producto de sus respectivas confiabilidades. La corrección de la atenuación es uno de los usos más importantes de la teoría de la confiabilidad, ya que permite una estimación de la verdadera correlación entre dos

constructos cuando no hay una medición perfecta, sin errores. Al corregir la falta de fiabilidad de esta manera, podemos determinar las relaciones latentes subyacentes sin la distracción del error de medición (Revelle, W., y Condon, 2014).

Hasta este punto la estrategia metodológica planteada consiste en la construcción de dos medidas, una de logros de vida y otra de origen social, esto a través del análisis factorial confirmatorio. Adicionalmente para verificar el análisis realizado en términos de confiabilidad se analizará, a través de los estadísticos omega, alpha y beta, que tan homogéneos son los índices propuestos, es decir, en qué medida las variables consideradas nos explican, en conjunto, el origen social y los logros de vida, respectivamente. Asimismo, se realizará la estimación de la correlación desatenuada entre ambas medidas con el objetivo de conocer la correlación entre ambos constructos como una forma de aproximarnos a la estimación de desigualdad de oportunidades que existe en México y en sus distintas regiones durante 2017.

3.1.2. Variables utilizadas en la medición de desigualdad de oportunidades

Los enfoques revisados para medir la desigualdad de oportunidades varían en cuanto a las variables utilizadas tanto de origen social como de logros de vida. Por un lado, para logros de vida se utilizan el ingreso (Barros, Ferreira, Vega, y Chanduvi, 2009; Ferreira y Gignoux, 2011; Ferreira y Meléndez, 2012; Monroy-Gómez-Franco, Vélez-Grajales, y Yalonetzky, 2018), el logro educativo (Santos Ochoa, 2016), el bienestar económico medido a través número de bienes y servicios que posee el hogar (Wendelspiess Chavez Juarez, 2015), entre otros. Se identifican también indicadores multidimensionales que capturan dimensiones de bienestar material y la satisfacción de vida (Nogales Carvajal, 2016). En cuanto al origen social no sólo se consideran variables individuales y familiares, también se consideran variables de contexto territorial, éstas últimas son relevantes porque se ha demostrado que el origen social en el que nacen las personas es tan solo una parte de la lotería que deben sortear los individuos, pues el territorio en donde esto sucede es igual de importante (Monroy-Gómez-Franco y Corak 2019; en Centro de Estudios Espinosa Yglesias, 2019).

Si bien existen autores que incorporan variables de esfuerzo (Krafft y Assaad, 2015; Nogales Carvajal, 2016), con el fin de simplificar el ejercicio propuesto en esta tesis, no se medirá ésta variable y se supondrá que todos los individuos realizan el mismo esfuerzo. Nos concentraremos en medir el origen social y los logros de vida a partir del bienestar económico y del acceso a servicios públicos, ambos de manera retrospectiva y actual. Esto basado en lo propuesto por Wendelspiess Chavez Juarez (2015), Nogales Carvajal (2016), el Centro de Estudios Espinosa Yglesias, (2019b) y Soloaga y Wendelspiess, 2010). La tabla 3.2 describe a detalle las variables utilizadas en la bibliografía revisada, sin embargo, más adelante se presenta una sección de resultados que describe

3. METODOLOGÍA

específicamente las variables utilizadas en esta tesis.

3.1.2.1. Descripción de la base de datos

La base de datos que se utilizará en este análisis de desigualdad de oportunidades es la Encuesta ESRU de movilidad social en México 2017 (ESRU-EMOVI 2017), la cual es realizada por el Centro de Estudios Espinosa Yglesias y financiada por la Fundación Espinosa Rugarcía (ESRU). La ESRU-EMOVI 2017, cuenta con 17,665 entrevistas y es representativa de hombres y mujeres entre 25 y 64 años a nivel nacional.

A diferencia de las versiones anteriores de la EMOVI, esta versión ofrece información representativa no solo a nivel nacional sino también para la Ciudad de México y cinco regiones del país: norte, norte-occidente, centro, centro-norte y sur (véase figura 3.2) (Centro de Estudios Espinosa Yglesias, 2019a). Asimismo, la cobertura temática de la encuesta incluye información contemporánea y retrospectiva sobre características sociodemográficas, contexto socioeconómico (riqueza y acceso a servicios), educación, ocupación y percepciones. En particular, estos indicadores no sólo nos permiten aproximarnos al origen social de los individuos, sino también, nos permite tener indicadores sobre percepciones materiales y no materiales que son útiles en la aproximación de los logros de vida.

Figura 3.2: Regiones ESRU-EMOVI 2017



Fuente: Centro de Estudios Espinosa Yglesias (2019a)

Nota: La región norte incluye a Baja California, Sonora, Chihuahua, Coahuila, Nuevo León y Tamaulipas; el norte-occidente comprende a Baja California Sur, Sinaloa, Nayarit, Durango y Zacatecas; el centro-norte considera a Jalisco, Aguascalientes, Colima, Michoacán y San Luis Potosí; el centro lo conforman Guanajuato, Querétaro, Hidalgo, Estado de México, Ciudad de México, Morelos, Tlaxcala y Puebla; y el sur incluye a Guerrero, Oaxaca, Chiapas, Veracruz, Tabasco, Campeche, Yucatán y Quintana Roo.

3. METODOLOGÍA

Tabla 3.2: Variables utilizadas en la bibliografía para medir desigualdad de oportunidades

Autor	Variables		
	Origen social	Logros de vida	País
Bourguignon, Ferreira y Menéndez (2007)	Regiones de nacimiento Raza Años de escolaridad de los padres Ocupación de los padres	Ingresos individuales	Brasil
Paes de Barros, et al. (2008)	Educación de los padres	Ingresos laborales	México, Brasil, Colombia
	Etnicidad	Ingresos del hogar	Ecuador, Panamá, Perú
	Lugar de nacimiento	Consumo del hogar	
Ferreira y Gignoux (2011)	Educación de los padres	Ingreso per cápita	Brasil, Colombia, Ecuador
	Ocupación de los padres	Gasto en consumo	Guatemala, Panamá, Perú
	Origen étnico		
	Región de nacimiento		
	Género		
Ferreira y Meléndez (2012)	Grupo étnico	Gasto per cápita	Colombia
	Educación de los padres	Ingreso Laboral	
	Lugar de nacimiento: zona urbana o zona rural		
Vélez-Grajales y Monroy-Gómez-Franco (2018)	Educación de los padres	Deciles de ingreso	México
Monroy-Gómez-Franco, et. al (2018)	Ocupación del padre		
	Origen étnico		
	Sexo		
	Comunidad urbana o rural		
	Color de piel		
Santos (2016)	Nivel educativo de los padres	Logro educativo	México
	Estatus ocupacional de los padres	Ingresos Laborales	
	Pertenencia étnica de los padres		
	Nivel socioeconómico del hogar en su juventud		
	Discapacidad		
Centro de Estudios Espinosa Yglesias (2019b)	Nivel educativo de los padres	Ingreso de los hogares	México
	Regiones		
Domii, et al. (2015)	Grupo socioeconómico parental		Gran Bretaña
	Educación de los padres		
	Capacidad cognitiva y no cognitiva en la infancia		
	Variables sobre la zona local		
	Porcentaje de desempleados		
	Inquilinos en viviendas sociales		
	Trabajadores no calificados		
Wendelspiess (2015)	Educación de los padres	Ingreso	México
	Condición socioeconómica a los 14 años	Número de bienes en el hogar	
	Origen étnico	Educación	
	Sexo		
Nogales (2016)	Género	Bienestar material	Bolivia
	Etnia	Artículos de la vivienda	
	Edad	Calidad de la vivienda	
	Antecedentes familiares	Habitantes en el hogar	
	Zona rural o urbana	Servicios básicos	
		Satisfacción de vida	
		Ingreso	
		Medio ambiente	
		Relaciones comunitarias	
		Afecto	
Soloaga y Wendelspiess (2010)	Educación del jefe/a del hogar	Primaria completa	México
	Acceso a servicios públicos en el hogar	Secundaria completa	
	Características del hogar en términos de activos	Preparatoria completa	
	Número de hermanos	Universidad completa	
		Ausencia de trabajo infantil	
Krafft, C., y Assaad, R. (2016)	Antecedentes familiares	Logro ocupacional	Egipto, Jordania
	Lugar de residencia	Edad del primer empleo	
	Género	Salario en el primer empleo	
		Creimiento del salario	
		Salario actual	
		Calidad del trabajo actual	
Saidi, A., y Hamdaoui, M. (2017)	Género	Acceso a servicios de salud	Túnez
	Lugar de residencia	Acceso a cuidados prenatales	
	Educación de los padres	Acceso a cuidados postnatales	
	Tamaño del hogar		
	Número de hijos por hogar		

Fuente: elaboración propia.

Resultados de la estimación

Esta sección presenta los hallazgos del análisis factorial confirmatorio llevado a cabo en la selección de variables para medir origen social y logros de vida. También se presentan los resultados del análisis de confiabilidad y de la desatenuación de la correlación entre logros de vida y el origen social como una aproximación a la desigualdad oportunidades en México durante 2017. Todos los análisis se realizaron con el software de análisis estadístico R (The R project for statistical computing) y la paquetería “semTools” (Jorgensen et al., 2020), “MBESS” (Kelley, 2020), “psych” (Revelle, 2020), “lavaan” (Rosseel, 2012) y “haven” (Miller, 2020). Las estimaciones podrán replicarse utilizando el código presentado en el Anexo (véase figura A.5).

4.1. Análisis factorial confirmatorio y análisis de confiabilidad

Se llevaron a cabo cuatro modelos diferentes, dos para medir logros de vida y dos para origen social. El primer par de modelos consideran variables que tienen que ver con el bienestar económico actual y el bienestar económico familiar cuando el entrevistado tenía 14 años. Este conjunto de indicadores incluye artículos del hogar. Los siguientes dos modelos incluyen variables de acceso actual y acceso retrospectivo a servicios públicos. Sin embargo, en este apartado de la tesis solo nos concentraremos en analizar el primer par de modelos, aquellos que miden el bienestar económico a través de artículos del hogar, puesto que estos tienen un mejor ajuste y mejores estadísticos de confiabilidad. Los resultados del segundo grupo de modelos se presentan únicamente en el Anexo (véase tabla A.2).

4.1.1. Logros de vida

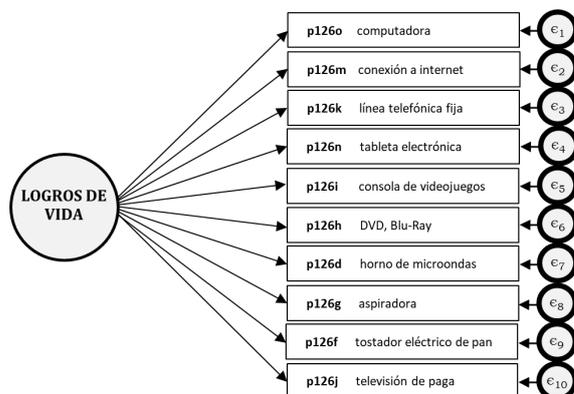
Modelo I: Logros de vida a través del bienestar económico actual

El primer modelo factorial confirmatorio considera lo propuesto por Wendelspiess Chavez Juarez (2015), Nogales Carvajal (2016) y el Centro de Estudios Espinosa Yglesias (2019b) quienes utilizan el bienestar económico actual de los entrevistados para aproximarse a sus logros de vida. El modelo presentado en esta tesis, representado en la figura 4.1, incorpora un conjunto de diez indicadores de artículos con los que cuenta la vivienda del entrevistado.

De acuerdo con los estadísticos de ajuste el modelo I tiene una bondad de ajuste buena, $\chi^2(23) = 222.583$, $CFI = 0.999$, $TLI = 0.998$ y $RMSEA = 0.022$. Al observar las cargas factoriales de este primer modelo (véase Anexo, figura A.1), lo primero que observamos es que la asociación entre los indicadores de artículos de la vivienda y la variable latente es aceptable pues estos indicadores se explican de un 50% a un 70% por los logros de vida. Los indicadores más relevantes en este primer modelo son la computadora y la conexión a internet, los cuales hoy en día son herramientas indispensables en el día a día de los individuos.

De acuerdo con la tabla 4.1, el estadístico de confiabilidad omega total $\omega_T = 0.82$, esto indica que la confiabilidad de los resultados de este modelo en conjunto es aceptable. Los estadísticos alpha (α) y beta (β) de 0.93 y 0.71, respectivamente, indican la consistencia interna de esta primera medida y que los indicadores por los que está conformada son homogéneos y útiles al explicar los logros de vida a través del bienestar económico actual de los individuos.

Figura 4.1: Ilustración Modelo I Logros de vida a través del bienestar económico actual



Fuente: elaboración propia con datos de ESRU-EMOVI 2017.)

A la par de la estimación de este primer modelo de logros de vida se llevó a cabo un modelo que incorpora indicadores de acceso a servicios públicos tomando como base lo propuesto por Donni, et al. (2015) y el Centro de Estudios Espinosa Yglesias (2019b) quienes utilizan variables referentes a la zona local donde vive el entrevistado y el acceso a servicios públicos, respectivamente.

De acuerdo con los estadísticos de ajuste este modelo alternativo de logros de vida tiene una bondad de ajuste aceptable, pero las cargas factoriales indicaron que este modelo no era el más adecuado porque la cantidad de varianza de los indicadores de acceso a servicios públicos es explicada por los logros de vida en solo un 50% y 60% (véase Anexo, figura A.3). Adicionalmente, se observa que el estadístico de confiabilidad omega total $\omega_T = 0.58$ lo cual fue decisivo para no considerarlo dentro del análisis pues esto implicaría utilizar una medida con alto error de medición y compuesta por indicadores que no son una buena aproximación para medir los logros de vida.

4.1.2. Origen social

Modelo II: origen social a través del bienestar económico a los 14 años

Este modelo para origen social incluyó indicadores del bienestar económico que tenía la familia del entrevistado cuando éste tenía 14 años. En específico este modelo está basado en lo propuesto por Soloaga y Wendelspiess (2010) y Centro de Estudios Espinosa Yglesias (2019b) quienes para medir origen social incorporan características del hogar de origen en términos de activos y servicios de la vivienda.

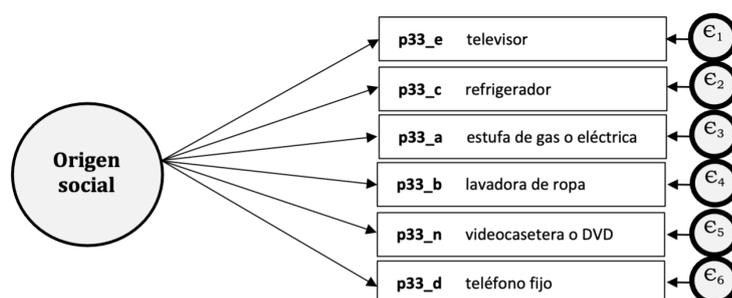
El modelo II de esta tesis, representado en la figura 4.2, incorpora un conjunto de seis indicadores de artículos de la vivienda de origen y presenta una bondad de ajuste aceptable, $\chi^2(7) = 309.86$, $CFI = 0.999$, $TLI = 0.998$ y $RMSEA = 0.044$. En cuanto a los estadísticos de confiabilidad, omega total $\omega_T = 0.842$, alpha (α) = 0.928 y beta (β) = 0.763 indican que la medida propuesta tiene consistencia interna y está compuesta por indicadores homogéneos. También sugiere que los indicadores incluidos miden la variable latente en casi la misma escala y con el mismo grado de precisión. Esto se corrobora con las cargas factoriales que indican que la cantidad de varianza de la mayoría de los indicadores es explicada del 60% al 70% por la variable latente origen social (véase Anexo, figura A.2). Dentro de los más relevantes se encuentran los siguientes artículos: televisor, estufa de gas, refrigerador y lavadora.

Este modelo es relevante de acuerdo con la bibliografía no solo de desigualdad de oportunidades sino también de movilidad social, en donde el bienestar económico determina la posibilidad de dar a los hijos oportunidades de acceso a bienes y servicios que mejoran su bienestar y mejores oportunidades en términos de educación y salud. A medida que el hogar de origen cuenta con un mejor bienestar económico entonces

4. RESULTADOS DE LA ESTIMACIÓN

mayor será la probabilidad de hacer frente a choques económicos negativos que afecten a los hogares (Centro de Estudios Espinosa Yglesias, 2019b).

Figura 4.2: Ilustración Modelo II Origen social a través del bienestar económico a los 14 años



Fuente: elaboración propia con datos de ESRU-EMOVI 2017

El modelo alternativo que se realizó para origen social considera indicadores de acceso a servicios públicos a los 14 años. La selección de los indicadores de este cuarto modelo se realizó con base a Soloaga y Wendelspiess (2010). Sin embargo, omega total $\omega_T = 0.704$ y al revisar que la varianza de los indicadores en su mayoría es explicada por el 50 % y 60 % del constructo origen social (véase Anexo, figura A.4). Parecería que el entorno en donde se desarrollan los individuos en sus primeras etapas de vida representa una gran ventaja o desventaja derivado del acceso desigual a servicios públicos, y esto a su vez puede ser determinante en los logros de vida de etapas posteriores (Centro de Estudios Espinosa Yglesias, 2019b). Sin embargo, los indicadores utilizados en este trabajo no resultan tan relevantes.

Tabla 4.1: Estadísticos de ajuste y confiabilidad para modelos I y II

	Ajuste del modelo						Confiabilidad		
	chisq	df	pvalue	CFI	TLI	RMSEA	Omega Total	Alpha	Beta
Logros de vida									
Modelo I bienestar económico	222.538	23	0.000	0.999	0.998	0.022	0.821	0.939	0.712
Origen social									
Modelo II Bienestar económico	309.861	9	0.000	0.999	0.998	0.044	0.842	0.928	0.763

Fuente: elaboración propia con datos de ESRU-EMOVI 2017.

De acuerdo con lo anterior los modelos más adecuados para medir logros de vida y

origen social son aquellos que miden estos conceptos a través del bienestar económico. Esto coincide con la bibliografía que nos dice que al medir logros de vida y origen social es conveniente utilizar indicadores de bienestar económico o riqueza del hogar de origen y el hogar actual, pues éste determina en un 46 % la desigualdad de oportunidades observada en México y el acceso a servicios públicos tan solo en un 11 % (Centro de Estudios Espinosa Yglesias, 2019b). Además, este es un resultado empírico de la base de datos pues la construcción y la disponibilidad de datos en la ESRU-EMOVI 2017 hace conveniente el uso de indicadores de bienestar económico.

4.2. Desatenuación de la correlación entre logros de vida y origen social

En este apartado se realiza la estimación de la correlación desatenuada que existe entre las medidas de origen social y los logros de vida basadas en el bienestar económico. Este análisis se realiza tanto a nivel nacional como para cinco regiones del país: norte, norte-occidente, centro, centro-norte y sur. La tabla 4.2 muestra por un lado la correlación atenuada que es la magnitud en la cual los coeficientes de correlación observados son menores que los esperados si se tuviera perfecta confiabilidad (Trafimow, 2016). Por otro lado, el mismo cuadro muestra los coeficientes de correlación desatenuada entre logros de vida y origen social como una propuesta para corregir los coeficientes de correlación atenuados y como una aproximación a la verdadera desigualdad de oportunidades en México. Asimismo se presentan los indicadores de confiabilidad calculados en el apartado anterior y utilizados para desatenuar las correlaciones observadas.

Para poder llevar a cabo comparaciones entre las distintas regiones de México, una vez establecida la estructura factorial de cada una de las medidas de interés, logros de vida y origen social, levantaremos el supuesto de invarianza en nuestras mediciones. Con ello supondremos que la comparación entre las cinco regiones tiene sentido pues cada indicador se relaciona de la misma manera con el constructo de interés en las diferentes regiones (grupos de interés). Diremos también que las diferencias regionales en la correlación estimada reflejan la verdadera heterogeneidad poblacional, es decir que, los cambios en las oportunidades de región a región deben reflejar cambios efectivos en los estándares de vida de la población analizada. Suponer invarianza significa que para cada uno de los diferentes grupos a analizar se mantendrá la misma estructura y significado tanto del constructo logros de vida y origen social (Putnick y Bornstein, 2016). Esto nos permite afirmar que la población de cada una de las regiones, entienden de la misma forma los factores latentes considerados en este trabajo, origen social y logros de vida. Este supuesto subyace a trabajos anteriores como los de Centro de Estudios Espinosa Yglesias (2019b), (Monroy-Gómez-Franco et al., 2018) y también subyace a mediciones de pobreza al comparar entre regiones una misma medida. En

4. RESULTADOS DE LA ESTIMACIÓN

las conclusiones se discute las implicaciones de este supuesto y las maneras de probarlo.

Al analizar las correlaciones entre logros de vida y origen social se considera que los contextos de ambos hogares son diferentes, por lo que al comparar de manera general la posición actual con respecto a la posición de origen obtenemos que el efecto de la atenuación entre logros de vida y el origen social es de 20 % pues dicho coeficiente pasa de 0.49 a 0.59 después de ser corregido, esto en el caso nacional. Este dato se asemeja al obtenido por Wendelspiess Chavez Juarez (2015) quien dice que el origen social, el cual está fuera del control del individuo representan alrededor del 40 % de la desigualdad total. También está en línea con lo encontrado por el Centro de Estudios Espinosa Yglesias (2019b) de 46 %.

Es importante advertir que ambas correlaciones, la observada o atenuada y la desatenuada, a lo largo de las diferentes regiones del país se observan diferenciadas, sobre todo estas se observan significativamente mayores en el sur respecto al resto del país. Con ello es posible concluir que, tal como lo sugiere la bibliografía revisada, las diferencias regionales importan y la población del sur se encuentra en desventaja con los del norte y las demás regiones estudiadas (Centro de Estudios Espinosa Yglesias, 2019b).

En la región norte es donde existe la menor correlación entre logros de vida y origen social, el coeficiente atenuado es de 0.39 y el desatenuado de 0.46, seguida de la región norte-occidente con coeficientes de correlación de 0.39 y 0.47. Por el contrario, las regiones en donde la exposición de los individuos a su origen social se encuentra más ligada o correlacionada con sus logros de vida, son la región centro y sur. Estas dos regiones cuentan con la mayor desigualdad de oportunidades en el país, el centro del país presenta coeficientes de correlación de 0.47 y 0.57 mientras que el sur de 0.51 y 0.62.

Los resultados diferenciados entre regiones coinciden con la bibliografía de desigualdad de oportunidades que estima en las regiones sur y centro del país una correlación de alrededor del 45 – 46 % y en las regiones del norte del 30 % (Centro de Estudios Espinosa Yglesias, 2019b). Estas diferencias confirman la heterogeneidad existente en México e indica que la región en donde se encuentran los individuos tiene un impacto diferenciado en los logros de vida de los mismos. Por tanto, el origen social en el que nacen y crecen los individuos es tan solo una parte de la lotería que deben sortear pues el territorio en donde esto sucede es igual de importante (Monroy-Gómez-Franco y Corak 2019; en Centro de Estudios Espinosa Yglesias, 2019b). Se tiene evidencia de que aproximadamente el 20 % de la desigualdad de oportunidades y hasta 60 % se explica por factores territoriales (Pereira y Soloaga 2016; en Centro de Estudios Espinosa Yglesias, 2019b).

De acuerdo con la tabla 4.2 las diferencias entre la correlación atenuada y desatenuada son mayores en la región centro norte, esta diferencia es del 23 %, seguido del centro y norte con una diferencia alrededor del 21 %, en el norte occidente del 20 %

Tabla 4.2: Correlación entre logros de vida y origen social: bienestar económico

	Atenuada	Desatenuada	Confiabilidad (Omega)	
Nacional	0.49	0.59		
Norte	0.39	0.46		
Norte-occidente	0.39	0.47	Logros de vida	Origen Social
Centro-Norte	0.46	0.57	0.82	0.84
Centro	0.47	0.57		
Sur	0.51	0.62		

Fuente: elaboración propia con datos de ESRU-EMOVI 2017.

y en el norte del 18%. Este incremento en la correlación depende de la correlación observada para cada una de las regiones y las dos confiabilidades calculadas para cada uno de los constructos. Dado que la confiabilidad permanece constante para cada una de los grupos analizados, como mayor sea la correlación observada o atenuada, mayor será la correlación desatenuada. Al corregir la falta de confiabilidad de esta manera, podemos determinar las relaciones latentes subyacentes sin la distracción del error de medición.

Con las diferencias entre las correlaciones a lo largo de las regiones analizadas, observamos que el acceso a bienes de la vivienda se refleja no solo en el nivel de cobertura de la población sino también en la distribución que existe a lo largo de las diferentes regiones del país, lo cual limita la igualdad de oportunidades y la movilidad social. El hecho de que el origen social y los logros de vida, medidos a través del bienestar económico, estén altamente correlacionados en las diferentes regiones del país, nos indica la doble lotería que tienen que enfrentar los individuos en cuanto al hogar donde se nace y el territorio donde esto ocurre. Esto representa un obstáculo para los logros de vida y el desarrollo de las capacidades y el talento individual, lo cual termina afectando el desarrollo de la sociedad en la que vivimos (Centro de Estudios Espinosa Yglesias, 2019b).

Conclusiones

El análisis empírico sobre desigualdad ha girado en torno al concepto de igualdad de oportunidades, concepto que enfatiza que al analizar la desigualdad económica lo más importante es comprender y atender el origen de ésta. John Roemer (1998) analiza los factores que originan la desigualdad y distingue entre aquellos ajenos al control de los individuos (sexo, género, lugar de nacimiento, contexto familiar, bienestar económico en el hogar de origen, etc.) y el esfuerzo individual como determinantes en los logros de vida. Así en la medida en la que la desigualdad económica se origine en factores ajenos al control del individuo como su origen social, la desigualdad será injusta y estaremos en desigualdad de oportunidades.

Al medir desigualdad de oportunidades tradicionalmente se relacionan variables de logros de vida con variables de origen social. Sin embargo, las estimaciones obtenidas generalmente están atenuadas debido a que los fenómenos analizados no son directamente observables (Wendelspiess Chavez Juarez, 2015). Esto representa limitaciones en el análisis y comparabilidad de la desigualdad de oportunidades. En esta tesis también se discute el que típicamente al estimar la desigualdad de oportunidades se reduzca el problema a buscar mejores métodos de estimación cuando el problema principal tiene que ver con la medición que muchas veces se trivializa al suponer que las variables utilizadas han sido medidas sin error. Corregir únicamente los errores de estimación, como comúnmente se hace, no nos dice nada del impacto del error de medición y poca información nos aporta si existe ruido en la medición.

Como propuesta para solucionar lo anterior se tomó como punto de partida la teoría de la medición y se llevó a cabo un análisis factorial confirmatorio, un análisis de confiabilidad para construir medidas confiables de logros de vida y origen social, y una desatenuación de los coeficientes de correlación entre los fenómenos de interés ya mencionados; esto para aproximarnos a una estimación insesgada de la desigualdad de oportunidades en México durante 2017. Realizar este tipo de análisis nos permite vincular la teoría con la experimentación para hacer observable lo inobservable al mismo tiempo que asegura resultados confiables, válidos y comparables.

5. CONCLUSIONES

La base de datos que se utilizó en este análisis de desigualdad de oportunidades es la Encuesta ESRU de movilidad social en México 2017 (ESRU-EMOVI 2017) y la selección de las variables a utilizar para origen social y logros de vida estuvo en función del análisis factorial confirmatorio realizado para cada uno de nuestros constructos de interés y de la bibliografía revisada.

Se llevaron a cabo cuatro modelos diferentes. La construcción del primer par de modelos para medir logros de vida y origen social, se basó en lo propuesto por el Centro de Estudios Espinosa Yglesias (2019b), Nogales Carvajal (2016) y Wendelspiess Chavez Juarez (2015). Estos modelos consideran variables de bienestar económico actual y familiar cuando el entrevistado tenía 14 años y cada uno de ellos incluye un conjunto indicadores de artículos de la vivienda del entrevistado. Los otros dos modelos incorporan variables de acceso actual y retrospectivo a servicios públicos tomando como base lo propuesto por el Centro de Estudios Espinosa Yglesias (2019b), Donni et al. (2015) y Soloaga y Wendelspiess (2010) quienes utilizan variables referentes al acceso a servicios públicos.

Resultó que de estos cuatro modelos, los más convenientes para analizar desigualdad de oportunidades son aquellos que tienen que ver con el bienestar económico, ya que los indicadores seleccionados son homogéneos y atienden a la misma fuente de varianza. Además, los resultados obtenidos son consistentes y con bajo error de medición, lo cual nos permite hacer un análisis adecuado tanto de logros de vida como de origen social. Por el contrario, aquellos modelos construidos a partir de indicadores de acceso a servicios públicos presentaron bajos niveles de confiabilidad, lo cual sugeriría una mala discriminación y una pobre aproximación al fenómeno que queremos analizar, por lo que se decidió descartarlos.

Si bien se revisó que existe una amplia gama de indicadores para aproximarse a los logros de vida y origen social, esta decisión coincidió con el hecho de que una parte de la bibliografía de desigualdad de oportunidades apunta a que lo más conveniente al medir logros de vida y origen social es utilizar indicadores de bienestar económico o riqueza del hogar de origen y el hogar actual, pues éste determina en un 46 % la desigualdad de oportunidades observada en México mientras que el acceso a servicios públicos tan solo un 11 % (Centro de Estudios Espinosa Yglesias, 2019b). Además, esta elección también es un resultado empírico de la base de datos pues la construcción y la disponibilidad de datos en la ESRU-EMOVI 2017 hace conveniente el uso de indicadores de bienestar económico más que de otro tipo.

Una vez construidas las medidas de origen social y logros de vida se llevó a cabo la estimación de la correlación atenuada y desatenuada que existe entre los dos pares de medidas mencionadas. Este análisis se realizó a nivel nacional, así como para cinco regiones del país: norte, norte-occidente, centro, centro-norte y sur. Para poder llevar

a cabo el análisis entre regiones se supuso invarianza factorial en la medición. Poner a prueba este último supuesto queda pendiente para trabajos posteriores. Suponer invarianza implica que las propiedades de medida de cada uno de los instrumentos o los indicadores que lo conforman, son independientes de las características de los grupos o regiones que analiza este trabajo, pero no del constructo que el instrumento está midiendo, es decir, logros de vida y origen social.

En relación con las medidas de los logros de vida y el origen social a través del bienestar económico actual y retrospectivo, la correlación atenuada entre el origen social y los logros de vida es de 0.49 y después de corregir por confiabilidad es de 0.59 a nivel nacional. Esto nos indica que es muy probable que las medidas de desigualdad de oportunidades revisadas en la bibliografía estén subestimadas por un 20%. En la región norte es donde existe la menor correlación y ésta va en un rango del 0.39 al 0.46, mientras que las regiones en donde la exposición de los individuos a su origen social se encuentra más ligada o correlacionada con sus logros de vida, son la región centro y sur, con coeficientes de correlación del 0.47 al 0.57 y del 0.51 al 0.62 respectivamente. Esta heterogeneidad en la población a través de las distintas regiones analizadas es un claro indicador de que los factores territoriales son relevantes al medir la desigualdad de oportunidades, ya que condicionan las posibilidades de la población sobre todo en la región sur del país donde se observa una mayor correlación. También apuntan a la importancia de seguir profundizando en el estudio de la desigualdad de oportunidades pues solo conociendo bien el fenómeno se podrán llevar a cabo políticas económicas correctas que promuevan una igualdad de oportunidades en la población.

Nos encontramos con limitantes como la falta de datos para algunos indicadores que hubiera sido útil incorporar con respecto a la educación y el empleo de los entrevistados. Otra limitación tiene que ver con la encuesta y la información retrospectiva que contiene, ya que, al ser información del pasado, muchas veces puede existir dificultad para recordar por parte de los encuestados y esto de alguna manera conlleva a resultados no tan acertados. Finalmente nos encontramos con que muchos de los indicadores no eran los más adecuados para medir nuestros constructos de interés contrario a lo revisado en la bibliografía, como lo fue el caso del acceso a servicios del hogar.

A partir de la investigación desarrollada, surgen algunas posibilidades de investigaciones futuras. En primer lugar y como ya se dijo, es importante llevar a cabo un análisis detallado de invarianza, sobre todo si los resultados se utilizan para el diagnóstico y tratamiento del fenómeno en cuestión así como en la formulación de políticas públicas. Probar invarianza factorial es un proceso importante para demostrar que el funcionamiento de nuestro instrumento de medición es adecuado, en particular, que los indicadores utilizados son independientes de las características de los grupos comparados, pero no del constructo de interés.

Adicionalmente se pretende en trabajos posteriores realizar un índice más amplio

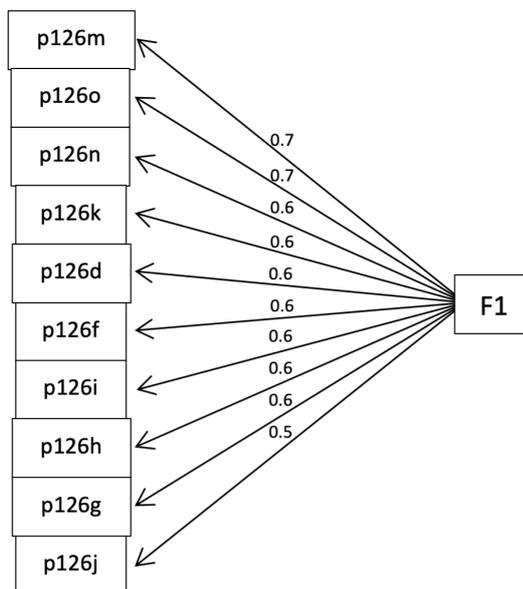
5. CONCLUSIONES

que incorpore diversas dimensiones tanto para origen social como para logros de vida, de modo que se cuente con información sobre educación y empleo, acceso a servicios de salud, etc. En resumen, hace falta llevar a cabo más medidas y más estudios en subpoblaciones antes de proponer una medida de bienestar económico plausible.

La manera en la que hemos estado midiendo desigualdad de oportunidades no ha sido la más adecuada, pero podemos hacerlo mejor si consideramos un conjunto de indicadores que realmente mida lo que queremos medir. Si logramos separar la señal del ruido, lo cual es un desafío en todas las mediciones, si no trivializamos la medición y recordamos la importancia de ésta. Hacer que nuestras medidas sean confiables aumentará la validez de nuestros análisis. Recordemos que no solo existe un problema de estimación sino también de medición y que muchas veces es este primer paso al que tenemos que poner más atención. A lo largo de este trabajo se buscó enfatizar en la importancia de medir bien y se incluyeron alternativas para lograrlo, por ejemplo, el análisis factorial confirmatorio, el análisis de confiabilidad y la desatenuación por confiabilidad de los coeficientes de correlación.

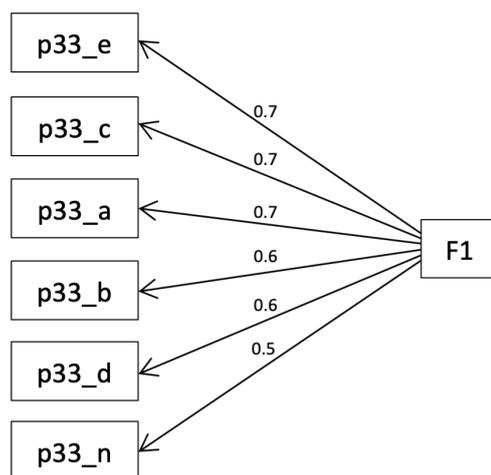
A.1. Cuadros y esquemas adicionales/Código en R

Figura A.1: Modelo I Logros de vida medido a través del bienestar económico



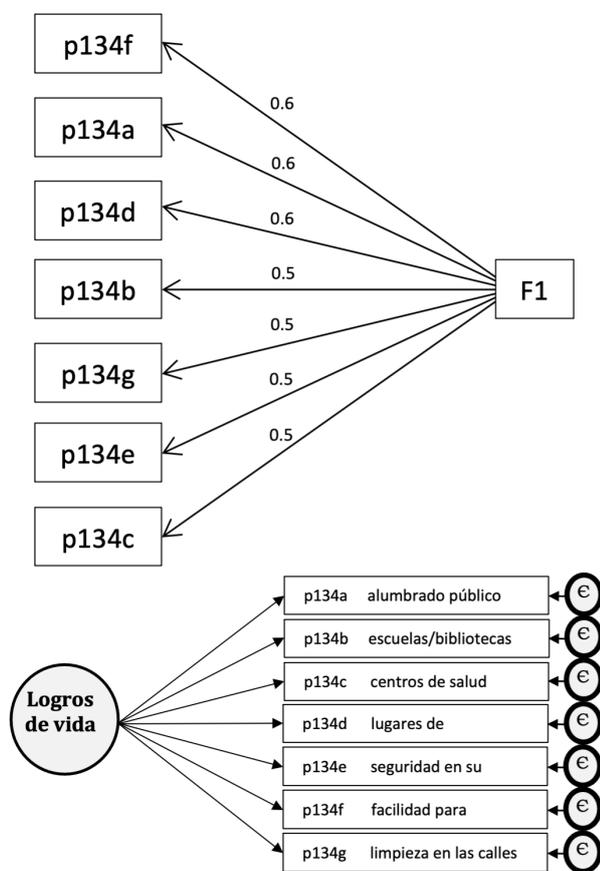
Fuente: elaboración propia con datos de ESRU-EMOVI 2017

Figura A.2: Modelo II Origen social medido a través del bienestar económico



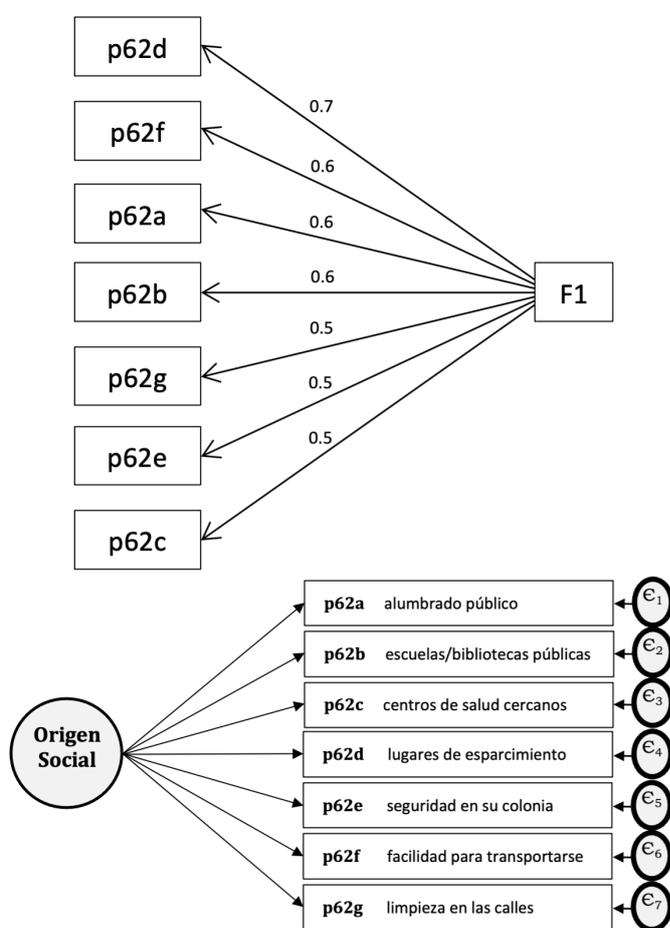
Fuente: elaboración propia con datos de ESRU-EMOVI 2017

Figura A.3: Modelo III Logros de vida medido a través del acceso a servicios públicos



Fuente: elaboración propia con datos de ESRU-EMOVI 2017

Figura A.4: Modelo IV Origen social medido a través del acceso a servicios públicos



Fuente: elaboración propia con datos de ESRU-EMOVI 2017

Tabla A.1: Estadísticos descriptivos para medir logros de vida y origen social

Variable	Etiqueta	Observaciones	Media	Desviación Estándar	Min.	Max
Modelos de bienestar económico (artículos de vivienda)						
Logros de vida						
p126o	computadora	17,665	1.676	0.468	1	2
p126m	conexión a internet	17,665	1.572	0.495	1	2
p126k	línea telefonica fija	17,665	1.617	0.486	1	2
p126n	tableta electrónica	17,665	1.771	0.420	1	2
p126i	consola de videojuegos	17,665	1.822	0.383	1	2
p126h	DVD, Blu-Ray	17,665	1.582	0.493	1	2
p126d	horno de microondas	17,665	1.479	0.500	1	2
p126g	aspiradora	17,665	1.861	0.346	1	2
p126f	tostador eléctrico de pan	17,665	1.783	0.412	1	2
p126j	televisión de paga	17,665	1.477	0.500	1	2
Origen Social (14 años del entrevistado)						
p33.e	televisor	17,665	1.422	0.780	1	8
p33.c	refrigerador	17,665	1.484	0.782	1	8
p33.a	estufa de gas o electrica	17,665	1.411	0.768	1	8
p33.b	lavadora de ropa	17,665	1.734	0.823	1	8
p33.n	videocasetera o DVD	17,665	1.896	0.890	1	8
p33.d	teléfono fijo	17,665	1.853	0.795	1	8
Modelos de acceso a servicios públicos						
Logros de vida						
p134a	alumbrado público	17,665	1.209	0.622	1	8
p134b	escuelas/bibliotecas públicas	17,665	1.150	0.673	1	8
p134c	centros de salud cercanos	17,665	1.273	0.721	1	8
p134d	lugares de esparcimiento cercanos	17,665	1.244	0.631	1	8
p134e	seguridad en su colonia	17,665	1.552	0.795	1	8
p134f	facilidad para transportarse	17,665	1.127	0.565	1	8
p134g	limpieza en las calles	17,665 1.447	0.720	1	8	
Origen Social (14 años del entrevistado)						
p62a	alumbrado público	17,665	1.459	0.787	1	8
p62b	escuelas/bibliotecas públicas cercanas	17,665	1.330	0.738	1	8
p62c	centros de salud cercanos	17,665	1.547	0.885	1	8
p62d	lugares de esparcimiento cercano	17,665	1.518	0.763	1	8
p62e	seguridad en la colonia	17,665	1.549	0.961	1	8
p62f	facilidad para transportarse	17,665	1.428	0.771	1	8
p62g	limpieza en las calles	17,665	1.520	0.930	1	8

Fuente: elaboración propia con datos de ESRU-EMOVI 2017.

Tabla A.2: Estadísticos de ajuste y confiabilidad para modelos alternativos (III y IV)

	Ajuste del modelo						Confiabilidad		
	chisq	df	pvalue	CFI	TLI	RMSEA	Omega Total	Alpha	Beta
Logros de vida									
Modelo III Acceso a servicios públicos	66.015	11	0.000	0.998	0.996	0.017	0.584	0.822	0.679
Origen social									
Modelo IV Acceso a servicios públicos	131.053	11	0.000	0.998	0.997	0.025	0.761	0.886	0.704

Fuente: elaboración propia con datos de ESRU-EMOVI 2017.

Figura A.5: Código en R. Desigualdad de oportunidades en México: una medición con variables latentes

```

library(psych)
library(lavaan)
library(haven)
library(semTools)
library(MBESS)

ESRU_EMOVI_2017_Entrevistado <- read_dta("~/Desktop/TESIS/3 BASES DATOS ESRU-EMOVI 2017/ESRU-E
MOVI 2017 Entrevistado.dta")

# ANALISIS FACTORIAL CONFIRMATORIO Y CONFIABILIDAD
#####
#####
#####-LOGROS DE VIDA#####
#####-
#####
#####
#Modelo I Logros de vida: bienestar economico (articulos del hogar)
#####
#####
# Se incluyen algunas correlaciones residuales porque resultaron ligeramente altas usando solo
el modelo de medicion
Logrosdevidal<- '
# measurement model
lv =~ p126o + p126m + p126k + p126n + p126i + p126h + p126d + p126g + p126f + p126j
# Residual Correlations
#Enterteinment
p126i ~~ p126h + p126j
p126h ~~ p126j
#Appliances
p126d ~~ p126g + p126f
p126g ~~ p126f
#TICS
p126o ~~ p126m + p126k + p126n
p126m ~~ p126k + p126n
p126k ~~ p126n
'

lv.cfa1 <- cfa(Logrosdevidal, data=ESRU_EMOVI_2017_Entrevistado,ordered=c("p126o", "p126m", "p
126k", "p126n", "p126i", "p126h", "p126d", "p126g", "p126f", "p126j"), std.lv=TRUE)
summary(lv.cfa1, fit.measures=TRUE)
residuals(lv.cfa1, type="cor")

#Confiabilidad

```

A. ANEXO

```
#Se calcula alpha y beta
x1.lv.rel<-reliability(lv.cfa1)
x1.lv.rel
x1.lv.omega3 <- x1.lv.rel[4,1]
x1.lv.omega3
logvid1<-as.data.frame(ESRU_EMOVI_2017_Entrevistado[c("p126o", "p126m", "p126k", "p126n", "p126i", "p126h", "p126d", "p126g", "p126f", "p126j")])
#Alpha 0.9393122
#Omega 0.8211323

#Se calcula beta
beta_c1<-iclust(logvid1)
beta_c1$beta
beta_c1
fa.diagram(beta_c1,n.factors=3)
#Beta 0.7122219

#Estadisticos de ajuste del modelo
chisq1<-fitmeasures(lv.cfa1,fit.measures = c("chisq","df","pvalue"))
relfit1<-fitmeasures(lv.cfa1,fit.measures = c("tli","cfi"))
rmsea1<-fitmeasures(lv.cfa1,fit.measures = c("rmsea","rmsea.ci.lower","rmsea.ci.upper","rmsea.pvalue"))
chisq1
relfit1
rmsea1
# chisq      df  pvalue
#222.583  23.000   0.000

# tli   cfi
#0.998 0.999

#rmsea rmsea.ci.lower rmsea.ci.upper  rmsea.pvalue
#0.022          0.020          0.025          1.000

#se crea un objeto para guardar los valores de cada una de las variable latente
x1.lv <- lavPredict(lv.cfa1, type="lv")
View(x1.lv)

#####
#####
#Modelo alternativo logros de vida: acceso a servicios publicos
#####
#####
Logrosdevida2<- 'lv =~ p134a + p134b + p134c + p134d + p134e + p134f +p134g
                #SERVICIOS
                p134a ~~ p134e + p134g
                p134e ~~ p134g
                ,

lv.cfa2 <- cfa(Logrosdevida2, data=ESRU_EMOVI_2017_Entrevistado,ordered=c("p134a", "p134b", "p134c", "p134d", "p134e","p134f","p134g"), std.lv=TRUE)
#Se cambia todo a 1/0 para generar el raw-index como la combinacion lineal de las x
lv.cfa2 <- cfa(Logrosdevida2, data=replace(x<-ESRU_EMOVI_2017_Entrevistado[, c("p134a", "p134b", "p134c", "p134d", "p134e", "p134f", "p134g")], x!=1, 0)
```

A.1 Cuadros y esquemas adicionales/Código en R

```
      ,ordered=c("p134a", "p134b", "p134c", "p134d", "p134e","p134f","p134g"), std.lv
=TRUE)
summary(lv.cfa2, fit.measures=TRUE)
residuals(lv.cfa2, type="cor")

#Confiabilidad
#Se calcula alpha y beta
x2.lv.rel<-reliability(lv.cfa2)
x2.lv.rel
#alpha 0.8112641
x2.lv.omega3 <- x2.lv.rel[4,1]
x2.lv.omega3
#omega 0.5394895

#Se calcula beta
logvid2<-as.data.frame(ESRU_EMOVI_2017_Entrevistado[c("p134a", "p134b", "p134c", "p134d", "p13
4e","p134f","p134g")])
#beta
beta_c2<-iclust(logvid2)
beta_c2$beta
beta_c2
fa.diagram(beta_c2,n.factors=3)
#0.6793877

#Estadisticos de ajuste del modelo
chisq2<-fitmeasures(lv.cfa2,fit.measures = c("chisq","df","pvalue"))
relfit2<-fitmeasures(lv.cfa2,fit.measures = c("tli","cfi"))
rmsea2<-fitmeasures(lv.cfa2,fit.measures = c("rmsea","rmsea.ci.lower","rmsea.ci.upper","rmsea.
pvalue"))
chisq2
relfit2
rmsea2
# chisq    df pvalue
#66.015 11.000 0.000

#tli    cfi
#0.996 0.998

#rmsea rmsea.ci.lower rmsea.ci.upper  rmsea.pvalue
#0.017          0.013          0.021          1.000

#se crea un objeto para guardar los valores de cada una de las variable latente
x2.lv <- lavPredict(lv.cfa2)
View(x2.lv)

#####
#####
#####-ORIGEN SOCIAL#####
#####
#####
#####
#Modelo II origen social: bienestar economico (articulos del hogar)
#####
```

A. ANEXO

```
origensocial3<- 'os =~ p33_e + p33_c + p33_a + p33_b + p33_n + p33_d '
```

```
os.cfa3 <- cfa(origensocial3, data=ESRU_EMOVI_2017_Entrevistado,ordered=c("p33_e", "p33_c", "p33_a", "p33_b", "p33_n", "p33_d"), std.lv=TRUE)
summary(os.cfa3, fit.measures=TRUE)
residuals(os.cfa3, type="cor")
```

```
#Confiabilidad
#Se calcula alpha y omega
reliability(os.cfa3)
x3.is.rel<-reliability(os.cfa3)
#Alpha 0.9276959
x3.is.omega3 <- x3.is.rel[4,1]
#Omega 0.841705
```

```
origsoc3<-as.data.frame(ESRU_EMOVI_2017_Entrevistado[c("p33_e", "p33_c", "p33_a", "p33_b", "p33_n", "p33_d")])
#Se calcula beta
beta_c3<-iclust(origsoc3)
beta_c3$beta
beta_c3
fa.diagram(beta_c3,n.factors=3)
#Beta 0.763169
```

```
#Estadisticos de ajuste del modelo
chisq3<-fitmeasures(os.cfa3,fit.measures = c("chisq","df","pvalue"))
relfit3<-fitmeasures(os.cfa3,fit.measures = c("tli","cfi"))
rmsea3<-fitmeasures(os.cfa3,fit.measures = c("rmsea","rmsea.ci.lower","rmsea.ci.upper","rmsea.pvalue"))
chisq3
relfit3
rmsea3
#chisq      df  pvalue
#309.861    9.000    0.000
```

```
#tli    cfi
#0.998  0.999
```

```
#rmsea  rmsea.ci.lower  rmsea.ci.upper  rmsea.pvalue
#0.044          0.039          0.048          0.995
```

```
#se crea un objeto para guardar los valores de cada una de las variable latente
x3.is <- lavPredict(os.cfa3, type="lv")
View(x3.is)
```

```
#####
#####
#Modelo alternativo origen social: acceso a servicios publicos
#####
#####
origensocial4<- 'os =~ p62a + p62b + p62c + p62d + p62e + p62f +p62g
#SERVICIOS
p62a ~~ p62e + p62g
```

A.1 Cuadros y esquemas adicionales/Código en R

```
      p62e ~~ p62g
      ,

os.cfa4 <- cfa(origensocial4, data=replace(x<-ESRU_EMOVI_2017_Entrevistado[, c("p62a", "p62b",
  "p62c", "p62d", "p62e","p62f","p62g")], x!=1, 0)
      ,ordered=c("p62a", "p62b", "p62c", "p62d", "p62e","p62f","p62g"), std.lv=TRUE)
summary(os.cfa4, fit.measures=TRUE)
residuals(os.cfa4, type="cor")

#Confiabilidad
#Se calcula alpha y omega
reliability(os.cfa4)
x4.is.rel<-reliability(os.cfa4)
x4.is.omega3 <- x4.is.rel[4,1]
#Alpha  0.8860616
#Omega  0.760556

origsoc34<-as.data.frame(ESRU_EMOVI_2017_Entrevistado[c("p62a", "p62b", "p62c", "p62d", "p62e"
,"p62f","p62g")])
#Se calcula beta
beta_c4<-iclust(origsoc34)
beta_c4$beta
beta_c4
fa.diagram(beta_c4,n.factors=3)
#Beta 0.7036738

##Estadisticos de ajuste del modelo
chisq4<-fitmeasures(os.cfa4,fit.measures = c("chisq","df","pvalue"))
relfit4<-fitmeasures(os.cfa4,fit.measures = c("tli","cfi"))
rmsea4<-fitmeasures(os.cfa4,fit.measures = c("rmsea","rmsea.ci.lower","rmsea.ci.upper","rmsea.
pvalue"))
chisq4
relfit4
rmsea4
#chisq      df  pvalue
#131.053  11.000   0.000

#tli   cfi
#0.997 0.998

#rmsea rmsea.ci.lower rmsea.ci.upper  rmsea.pvalue
#0.025          0.021          0.029          1.000

#se crea un objeto para guardar los valores de cada una de las variable latente
x4.is <- lavPredict(os.cfa4)
View(x4.is)

#####
#####
#####-CORRELACION#####
#####
#####
#####
#####
#En un solo modelo para Modelo I y II (LV1 Y OS3)
```

A. ANEXO

```
cor.desat<- '
# measurement model
lv =~ p126o + p126m + p126k + p126n + p126i + p126h + p126d + p126g + p126f + p126j
# Residual Correlations
#Enterteinment
p126i ~~ p126h + p126j
p126h ~~ p126j
#Appliances
p126d ~~ p126g + p126f
p126g ~~ p126f
#TICS
p126o ~~ p126m + p126k + p126n
p126m ~~ p126k + p126n
p126k ~~ p126n
os =~ p33_e + p33_c + p33_a + p33_b + p33_n + p33_d
lv~~os
'

lv1.os3<- cfa(cor.desat, data=ESRU_EMOVI_2017_Entrevistado,ordered=c("p126o", "p126m", "p126k",
, "p126n", "p126i", "p126h", "p126d", "p126g", "p126f", "p126j","p33_e", "p33_c", "p33_a", "p3
3_b", "p33_n", "p33_d"), std.lv=TRUE)
lv1.os3
summary(lv1.os3, fit.measures=TRUE)
residuals(lv1.os3, type="cor")
#La correlacion entre lv y os es la que buscamos. Como es la correlacion entre variables laten
tes es la correlacion desatenuada.
#lv ~~ os 0.590

#####
#####

#Modelos I y III
#correlacion observada
Logrosdevida1<- c("p126o", "p126m", "p126k", "p126n", "p126i", "p126h", "p126d", "p126g", "p12
6f", "p126j")
origensocial3<- c("p33_e", "p33_c", "p33_a", "p33_b", "p33_n", "p33_d")

#Se genera el raw score como una combinacion lineal simple de los items
LV1=rowSums(replace(x<-ESRU_EMOVI_2017_Entrevistado[,Logrosdevida1], x!=1, 0))
OS3=rowSums(replace(x<-ESRU_EMOVI_2017_Entrevistado[,origensocial3], x!=1, 0))

#CORRELACIONES ATENUADAS "EMP?RICAS" (ANTES DE CORREGIR LOS COEFICIENTES USANDO OMEGA "EMPPIRI
CO", omega3)
# Correlaci?n desatenuada con la f?rmula USANDO OMEGA "EMP?RICA" (omega3)
LV1.OS3<-cor(matrix(c(LV1,OS3),ncol=2))
LV1.OS3

#           [,1]      [,2]
#[1,] 1.0000000 0.4897565
#[2,] 0.4897565 1.0000000

# f?jate que aqu? usamos las omegas de las parciales, aunque podr?amos usar la del modelo comp
leto que son iguales
```

A.1 Cuadros y esquemas adicionales/Código en R

```
#LV1.OS3.rel<-c(lv1.os3.rel["omega3","lv"], lv1.os3.rel["omega3","os"])
LV1.OS3.rel<-c(x1.lv.omega3,x3.is.omega3)
#RELIABILITIES EN LA DIAGONAL PRINCIPAL, CORRELACION ATENUADA EN LA DIAGONAL DE ABAJO, DESATEN
UADA EN LA DE ARRIBA
round(correct.cor(LV1.OS3,LV1.OS3.rel),2) #F?JATE QUE ESTE ES EL MISMO RESULTADO QUE OBTENEMOS
DEL MODELO COMPLETO (QUE ES IGUAL A LA COVARIANZA)
#      [,1] [,2]
#[1,] 0.82 0.59
#[2,] 0.49 0.84

#####
#####
#####Análisis por regiones#####
#####-
#####
#####

# Correlaci?n desatenuada con la f?rmula USANDO OMEGA "EMP?RICA" (omega3)
#En las matriz resultante se encuentran los estadisticos de confiabilidad en la diagonal princ
ipal, la correlacion atenuada en la diagonal de abajo y la desatenuada en la de arriba

LV1.OS3<-cor(matrix(c(LV1,OS3),ncol=2))
LV1.OS3.rel<-c(x1.lv.omega3,x3.is.omega3)
round(correct.cor(LV1.OS3,LV1.OS3.rel),2)
#Se obtiene el mismo resultado que obtenemos del modelo completo

#[,1] [,2]
#[1,] 0.82 0.59
#[2,] 0.49 0.84

#Se estiman modelos para cada region a partir del modelo general ya ajustado y se usan las mis
mas confiabilidades obtenidas anteriormente (omega3)
ESRU_EMOVI_2017_Entrevistado$region
#Para la regi?n norte
LV1.OS3.N<-cor(subset(matrix(c(LV1,OS3),ncol=2),ESRU_EMOVI_2017_Entrevistado$region==1))
round(correct.cor(LV1.OS3.N,LV1.OS3.rel),2)
#      [,1] [,2]
#[1,] 0.82 0.46
#[2,] 0.39 0.84

#Para la regi?n Sur
LV1.OS3.S<-cor(subset(matrix(c(LV1,OS3),ncol=2),ESRU_EMOVI_2017_Entrevistado$region==5))
round(correct.cor(LV1.OS3.S,LV1.OS3.rel),2)
#[,1] [,2]
#[1,] 0.82 0.62
#[2,] 0.51 0.84

#Para la regi??n Norte.Occidente
LV1.OS3.NO<-cor(subset(matrix(c(LV1,OS3),ncol=2),ESRU_EMOVI_2017_Entrevistado$region==2))
round(correct.cor(LV1.OS3.NO,LV1.OS3.rel),2)
#[,1] [,2]
#[1,] 0.82 0.47
```

```
#[2,] 0.39 0.84

#Para la regi??n Centro.Norte
LV1.OS3.CN<-cor(subset(matrix(c(LV1,OS3),ncol=2),ESRU_EMOVI_2017_Entrevistado$region==3))
round(correct.cor(LV1.OS3.CN,LV1.OS3.rel),2)
#      [,1] [,2]
#[1,] 0.82 0.55
#[2,] 0.46 0.84

#Para la regi??n Centro
LV1.OS3.C<-cor(subset(matrix(c(LV1,OS3),ncol=2),ESRU_EMOVI_2017_Entrevistado$region==4))
round(correct.cor(LV1.OS3.C,LV1.OS3.rel),2)
#      [,1] [,2]
#[1,] 0.82 0.57
#[2,] 0.47 0.84

library(psych)
library(lavaan)
library(haven)
library(semTools)
library(MBESS)

citation(package = "psych", lib.loc = NULL)
citation(package = "haven", lib.loc = NULL)
citation(package = "lavaan", lib.loc = NULL)
citation(package = "semTools", lib.loc = NULL)
citation(package = "MBESS", lib.loc = NULL)
```

Bibliografía

- [1] Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2018). *World inequality report 2018*. Belknap Press.
- [2] Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford Publications.
- [3] Bourguignon, F., Ferreira, F. H. G., and Menéndez, M. (2007). Inequality of Opportunity in Brazil: A Corrigendum. *Review of Income & Wealth*, 59(3):551–555.
- [4] Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications. 9
- [5] Brunori, P., Peragine, V., and Serlenga, L. (2019). Upward and downward bias when measuring inequality of opportunity. *Social Choice and Welfare VO - 52*, (4):635.
- [6] Centro de Estudios Espinosa Yglesias (2019a). ESRU-EMOVI 2017.
- [7] Centro de Estudios Espinosa Yglesias (2019b). Informe de movilidad social en México 2019: hacia la igualdad regional de oportunidades.
- [8] Corak, M. (2016). Economic Theory and Practical Lessons for Measuring Equality of Opportunities.
- [9] Donni, P. L., Rodríguez, J. G., and Dias, P. R. (2015). Empirical definition of social types in the analysis of inequality of opportunity: a latent classes approach. *Social Choice and Welfare*, 44(3):673–701.
- [10] Dworkin, R. (1981). What is equality? Part 1: Equality of welfare. *Philosophy & public affairs*, pages 185–246.
- [11] Ferreira, F. H. G. and Gignoux, J. (2011). The Measurement of Inequality of Inequality of Opportunity: Theory and an Application to Latin America.
- [12] Ferreira, F. H. G. and Meléndez, M. (2012). Desigualdad de Resultados y Oportunidades en Colombia: 1997-2010.

BIBLIOGRAFÍA

- [13] Hausman, J. (2001). Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left. *The Journal of Economic Perspectives*, 15(4):57–67.
- [14] Hufe, P., Kanbur, R., and Peichl, A. (2018). Measuring Unfair Inequality: Reconciling Equality of Opportunity and Freedom from Poverty.
- [15] Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., and Quick, C. (2020). semTools: Useful tools for structural equation modeling. *R package version 0.5-1*.
- [16] Kanbur, R. and Snell, A. (2019). Inequality Indices as Tests of Fairness. *Economic Journal VO - 129*, (621):2216.
- [17] Kelley, K. (2020). MBESS: The mbess r package. *R package version*, 4(3).
- [18] Krafft, C. and Assaad, R. (2015). Inequality of Opportunity in the Labor Market for Higher Education Graduates in Egypt and Jordan.
- [19] Lara Ibarra, G. and Martinez Cruz, A. L. (2015). Exploring the sources of downward bias in measuring inequality of opportunity.
- [20] Loken, E. and Gelman, A. (2017). Measurement error and the replication crisis. *Science (New York, N.Y.)*, 355(6325):584–585.
- [21] Magidson, J. and Vermunt, J. K. (2002). A nontechnical introduction to latent class models. *Statistical Innovations white paper*, 1:15.
- [22] Miller, H. W. y. E. (2020). Haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files.
- [23] Monroy-Gómez-Franco, L., Vélez-Grajales, R., and Yalonetzky, G. (2018). Layers of inequality: Social mobility, inequality of opportunity and skin colour in Mexico. *Documento de trabajo*, (3).
- [24] Nájera Catalán, H. E. and Gordon, D. (2019). The Importance of Reliability and Construct Validity in Multidimensional Poverty Measurement: An Illustration Using the Multidimensional Poverty Index for Latin America (MPI-LA). *Journal of Development Studies*, 56(9):1763–1783.
- [25] Nogales Carvajal, C. R. (2016). Capabilities, opportunities and public policies: analyses through Simultaneous Equation Models with latent variables.
- [26] Oxfam México (2018). México justo: propuestas de políticas públicas para combatir la desigualdad.
- [27] Paes de Barros, R., Ferreira, F. H. G., Vega, J. R. M., and Chanduvi, J. S. (2009). *Measuring Inequality of Opportunities in Latin America and the Caribbean*. The World Bank.

- [28] Peragine, V. (2004). Measuring and implementing equality of opportunity for income. *Social Choice and Welfare*, 22(1):187.
- [29] Piketty, T. (2014). Capital in the Twenty-First Century: a multidimensional approach to the history of capital and social classes. *The British journal of sociology*, 65(4):736–747.
- [30] Rawls, J. (1971). A Theory of Justice, By John Rawls, Cambridge, Massachusetts, Harvard University Press, 1971. 607 pp. *Theology Today*, 30(3):306–312.
- [31] Revelle, W. (2013). Using R and the psych package to find ω . *Computer Software*. <http://personality-project.org/r/psych/HowTo/omega.tutorial/omega.html#x1-150005.1>.
- [32] Revelle, W. (2020). psych: Procedures for personality and psychological research.
- [33] Revelle, W. and Condon, D. M. (2019). Reliability from α to ω : A tutorial. *Psychological Assessment*, 31(12):1395–1411.
- [34] Revelle, W. and Zinbarg, R. E. (2009). COEFFICIENTS ALPHA, BETA, OMEGA, AND THE GLB : COMMENTS ON SIJTSMA. *Psychometrika*, 74(1):145–154.
- [35] Revelle, W., y Condon, D. M. (2014). Reliability. *Personality-project*.
- [36] Roemer, J. E. (1998). Equality of Opportunity.
- [37] Saidi, A. and Hamdaoui, M. (2017). On measuring and decomposing inequality of opportunity in access to health services among Tunisian children: a new approach for public policy. *Health and Quality of Life Outcomes*, 15(1):1–29.
- [38] Santos Ochoa, J. J. (2016). Desigualdad de oportunidades en educación e ingresos laborales en México.
- [39] Sen, A. (1985). ¿Cual es el camino del desarrollo?(Quel est le changement vers le développement). *Comercio Exterior de México Mexico*, 35(10):939–949.
- [40] Soloaga, I. and Wendelpiess, F. (2010). Desigualdad de Oportunidades: aplicaciones al caso de México. *Movilidad social en México. México, Fundación Espinosa Yglesias*.
- [41] Spearman, C. (1904). The proof and measurement of association between two things. *International journal of epidemiology*, 39(5):1137–1150.
- [42] Trafimow, D. (2016). The attenuation of correlation coefficients: a statistical literacy issue. *Teaching Statistics*, 38(1):25–28.
- [43] UCLA (2020). Confirmatory Factor Analysis (CFA) in R with Lavaan, Institute for digital research and education. *Statistical Consulting*.
- [44] Wendelspiess Chavez Juarez, F. (2015). Measuring Inequality of Opportunity with Latent Variables. *Journal of Human Development and Capabilities*, 16(1):106–121.