

Análisis microeconómico con la familia de medidas de pobreza multidimensional de Alkire-Foster: una discusión abierta

Autores: Jimena Macció¹ y Ann Mitchell²

RESUMEN

La medida de Alkire-Foster (2011) es una de las más utilizadas para la medición multidimensional de la pobreza (PNUD y OPHI, 2019; CEPAL, 2014). Se trata de una familia de medidas construida a partir del recuento ponderado de privaciones a las que están sujetos los hogares, evaluados en un conjunto de dimensiones de su bienestar. El objetivo de este artículo es poner en discusión los posibles métodos de estimación microeconómica de los determinantes de la pobreza multidimensional calculada mediante el método de Alkire-Foster. Busca avanzar sobre cuestiones relevantes tales como ¿Pueden modelarse de manera simultánea los tres miembros de la familia Alkire-Foster mencionados? ¿Cómo se modela la censura en la proporción de privaciones producida al establecer el umbral de pobreza? ¿Hay dos procesos independientes que dan lugar al hecho de que el hogar sea pobre o no, y a la cantidad de privaciones que acumule por encima del umbral? Se propone una posible modelización a través del modelo beta inflado truncado (Pereira, Botter y Sandoval, 2013; Gray y Hernandez, 2018). A manera de ilustración se presenta la aplicación de esta técnica a la evaluación de impacto del programa de vivienda de emergencia de la ONG TECHO a partir de una medida multidimensional de privaciones acumuladas en dimensiones vinculadas a la vivienda, sobre la base de Mitchell, Macció y Mariño Fages (2019) y Mitchell y Macció (2018).

Palabras clave: Alkire-Foster, pobreza multidimensional, impacto, inflated-truncated beta regression

JEL: C51, I32

¹ Facultad de Ciencias Económicas, Pontificia Universidad Católica Argentina, jmaccio@uca.edu.ar. Docente en FCE-UBA.

² Facultad de Ciencias Económicas, Pontificia Universidad Católica Argentina, annmitchell@uca.edu.ar.

1. Introducción

La medida de Alkire-Foster (2011) es una de las más frecuentemente utilizadas para la medición multidimensional de la pobreza. Se trata de una familia de medidas que surge de la extensión multidimensional de los índices de Foster, Greer y Thorbecke (1984), construida a partir del recuento ponderado de privaciones a las que están sujetos los individuos u hogares, evaluados en un conjunto de dimensiones de su bienestar. El método propone la construcción de un indicador compuesto, que puede presentarse como un conteo (cantidad de privaciones) o como una proporción (proporción de indicadores en situación deficitaria sobre el total de indicadores considerados). Luego determina la aplicación de un umbral de pobreza multidimensional, denominado k : aquellos individuos u hogares que acumulen más de k privaciones serán considerados multidimensionalmente pobres. La medida resulta atractiva porque cumple con un conjunto de axiomas o propiedades deseables, y por ser flexible en adaptarse a diversas decisiones metodológicas que pueden tomar sus usuarios (cantidad y tipo de dimensiones, cantidad de indicadores en cada dimensión, indicadores de medición cuali o cuantitativos, ponderaciones intra e inter-dimensiones, umbrales de privación y umbral de pobreza).

El uso de esta metodología está ampliamente difundido para ejercicios de medición y de diagnóstico de pobreza multidimensional (PNUD y OPHI, 2019; CEPAL, 2014). De hecho, ha sido la metodología más elegida en América Latina para la aplicación de medidas oficiales de pobreza multidimensional (véanse Ministerio de Desarrollo Social de Chile, 2015; INEC, 2015; STPP y MINEC-DIGESTYC, 2015; MEF, MDF e INEC, 2017, entre otras). Sin embargo, su uso como variable dependiente en ejercicios explicativos está menos difundido. Las razones detrás de este hecho son tanto teóricas como prácticas (Alkire et al., 2015). Entre las primeras, fundamentalmente se encuentra la idea de que la medida en sí misma debería incorporar todas las dimensiones pertinentes, incluso algunas que podrían muchas veces entenderse como determinantes de la pobreza en sí mismos. Entre las segundas, la flexibilidad de la medida permite modelarla ya sea como un indicador de recuento, como una proporción, a través de una variable dicotómica cuando simplemente se consideran las situaciones de pobre/no pobre, o incluso modelar de manera conjunta la incidencia y la intensidad de la pobreza multidimensional.

El objetivo de este artículo es poner en consideración un conjunto de opciones metodológicas para la estimación microeconómica de los determinantes de la pobreza multidimensional cuando la variable dependiente es una medida multidimensional de pobreza construida mediante el método de Alkire-Foster. En primer lugar, se presenta el método de Alkire-Foster y su familia de medidas: recuento (H), promedio de privaciones de los pobres (A) y recuento ajustado ($M0$). Busca avanzar sobre cuestiones relevantes tales como ¿Se puede suponer que hay dos procesos independientes que dan lugar, por un lado, al hecho de que el hogar sea pobre o no, y por el otro, a la cantidad de privaciones adicionales que acumule por encima del umbral? ¿Pueden modelarse de manera simultánea los tres miembros de la familia Alkire-Foster mencionados? ¿Cómo se modela la censura en la proporción de privaciones producida al establecer el umbral de pobreza? ¿En qué medida resulta interesante analizar la contribución de las variables explicativas a la variabilidad del indicador de pobreza multidimensional? Para procurar responder estas preguntas, entre otras, en el documento se propone una posible modelización a través del modelo beta inflado truncado (véanse Pereira, Botter y Sandoval, 2013; Gray y Hernandez, 2018) entre otros). Asimismo, se presenta la aplicación de esta técnica a la evaluación de impacto del programa de vivienda de emergencia de la ONG TECHO (www.techo.org) a partir de una medida multidimensional de privaciones acumuladas en dimensiones

vinculadas a la vivienda, sobre la base de Mitchell, Macció y Mariño Fages (2019) y Mitchell y Macció (2018).

El documento se desarrolla de la siguiente manera. En la primera sección se presenta y describe brevemente el método de Alkire-Foster (2011). En la segunda sección se presentan las consideraciones necesarias para el uso de esta medida en análisis explicativos. Con el objeto de realizar una ilustración de los métodos, se presenta la evaluación de impacto del programa de vivienda de emergencia de la organización TECHO, incluyendo una descripción del programa, del diseño de la evaluación y la construcción de la medida Alkire-Foster empleada. La siguiente sección describe las alternativas econométricas disponibles para el modelaje de la familia de medidas. La sección seis se ocupa de los resultados. Finalmente, se presentan algunas consideraciones que plantean posibles respuestas a los interrogantes enunciados.

2. Método de Alkire-Foster para la medición de pobreza multidimensional

La medición multidimensional de la pobreza ha logrado una aceptación generalizada como forma de complementar las mediciones tradicionales de pobreza a través del ingreso. El método de Alkire-Foster es uno de los más difundidos, tanto en términos académicos como en cuanto a sus aplicaciones como medidas nacionales de pobreza, particularmente en América Latina. Para medir la pobreza de manera multidimensional es necesario definir varios aspectos constitutivos de la medida. El más significativo es el de las dimensiones que comprenderá la medida, en función de la concepción que se tenga del bienestar subyacente (Duclos, 2006)³.

El método de Alkire-Foster se emplea para la medición de pobreza sobre la base del enfoque de recuento, es decir, calculando el número de privaciones que sufren las personas en un conjunto de dimensiones de sus vidas. Está comprendido por dos etapas: la etapa de la identificación y la etapa de la agregación

La identificación consiste en poder reconocer quién es pobre, sobre la base de un conjunto de criterios. En primer lugar, es necesario elegir las dimensiones de análisis, y operacionalizarlas empleando indicadores para cada dimensión. Debe seleccionarse un umbral para cada indicador, que expresa el nivel mínimo que es necesario lograr para no ser considerado privado en ese aspecto. Asimismo, se definen ponderaciones intra e inter-dimensiones, que permiten expresar la importancia relativa de cada aspecto en la medida global de pobreza. Sobre la base de estos criterios generales, se calcula para cada hogar (o cada individuo) la suma ponderada de privaciones, c_i . Este puede ser expresado como el recuento de privaciones, que yace dentro del intervalo $[0, K]$, siendo K la cantidad total de indicadores, o bien puede construirse una proporción de privaciones al dividir la cantidad de privaciones por el total posible (de manera tal que c_i se desarrolle en el intervalo $[0,1]$).

Una vez que se conoce el recuento de privaciones de cada hogar (o individuo), es necesario definir un segundo umbral –el umbral de pobreza multidimensional, k . Este umbral es la proporción de privaciones ponderadas que un hogar debe experimentar para ser considerado multidimensionalmente pobre. Si un hogar acumula privaciones por encima de ese umbral ($c_i \geq k$), entonces es considerado

³ Se refiere a la elección del espacio de análisis. Entre las opciones se encuentra un análisis bienestarista, que estudia el bienestar en términos de la utilidad, o no bienestarista, que se basa en el enfoque de las necesidades o en el de las capacidades.

multidimensionalmente pobre. Adicionalmente, se puede construir un indicador censurado $c_i(k)$ que toma el valor de c_i si el hogar es multidimensionalmente pobre ($c_i \geq k$), o cero en caso contrario.

La etapa de agregación que plantea el método de Alkire-Foster es una extensión de la familia de medidas de pobreza unidimensional FGT (Foster, Greer y Thorbecke, 1984) para el caso de múltiples privaciones. En esta versión original, los autores proponían:

$$FGT_\alpha = \frac{1}{n} \sum_{i=1}^q \left(\frac{z - y_i}{z} \right)^\alpha \quad (1)$$

donde n es la cantidad total de hogares, q es la cantidad de hogares pobres, z es la línea de pobreza, y_i son los ingresos de los hogares pobres y α es un parámetro que puede tomar valores 0, 1 o 2. Cuando $\alpha = 0$, el FGT_0 representa la incidencia de la pobreza (proporción de hogares pobres sobre el total de hogares). Cuando $\alpha = 1$, el FGT_1 representa la brecha de pobreza promedio (la distancia promedio de los ingresos de los hogares pobres respecto de la línea de pobreza). Cuando $\alpha = 2$, el FGT_2 representa la severidad de la pobreza, dado que se toma en cuenta la distribución de los ingresos por debajo de la línea, penalizando en mayor medida aquellos ingresos que están alejados del umbral.

La extensión multidimensional que plantean Alkire y Foster (2011)⁴ se compone de:

- El índice de recuento multidimensional, H , que representa la proporción de hogares identificados como multidimensionalmente pobres, de acuerdo al umbral k :

$$H = \frac{q(k)}{n} \quad (2)$$

donde $q(k)$ es el número de hogares multidimensionalmente pobres y n es el número total de hogares.

- La intensidad de la pobreza multidimensional, A , se define a través de la proporción promedio de privaciones ponderadas que experimentan los hogares pobres

$$A = \sum_{i=1}^q \frac{c_i(k)}{q(k)} \quad (3)$$

Esta medida captura la intensidad de las privaciones que experimentan los hogares multidimensionalmente pobres. Se construye con el recuento censurado porque realiza foco sobre los hogares en situación de pobreza, ignorando aquellas privaciones que enfrentan los hogares no pobres.

⁴ La familia de Alkire-Foster cumple con un conjunto de axiomas o propiedades deseables (Bourgignon y Chakravarty, 2003; Santos, 2010).

- El índice de recuento ajustado, M_0 , se calcula como el producto de H y A:

$$M_0 = HA = \frac{q}{n} \frac{1}{q} \sum_{i=1}^q c_i(k) = \frac{1}{n} \sum_{i=1}^n c_i(k) \quad (4)$$

M_0 mide la suma de privaciones ponderadas que experimentan los hogares pobres, dividida por el máximo número de privaciones que podrían experimentar todos los hogares si fueran pobres y estuvieran privados en todas las dimensiones. M_0 se incrementa ya sea porque un nuevo hogar comienza a ser multidimensionalmente pobre o porque algún hogar que ya era pobre tiene una nueva privación en alguna dimensión evaluada. Es decir, M_0 satisface los axiomas de foco en la privación y foco en la pobreza: la medida de pobreza no cambia cuando existe un cambio en un logro de un hogar no pobre, y al mismo tiempo es sensible a la cantidad de privaciones de los pobres (Alkire et al., 2015). De esta manera, M_0 es un indicador de la incidencia de la pobreza multidimensional ajustado por –y por lo tanto sensible a– la intensidad de la pobreza.

Adicionalmente a estas medidas, el método de Alkire-Foster contempla el cálculo de

- M_1 : Considera la profundidad de las privaciones en cada dimensión. Es una versión multidimensional de la brecha de pobreza.
- M_2 : Es el índice de severidad de la pobreza, que toma en cuenta la desigualdad entre los pobres, priorizando a los más pobres.

Sin embargo, estos últimos dos miembros de la familia Alkire-Foster solo pueden calcularse en caso que la escala de medición de los indicadores sea al menos ordinal. Disponiendo de indicadores dicotómicos, no es posible construir una brecha entre el umbral de privación y diferentes grados de cumplimiento, lo cual invalida el cálculo de M_1 y M_2 .

Si bien no es el único disponible, el método de Alkire-Foster es el más utilizado para la medición de pobreza multidimensional. En América Latina, 9 países han adoptado medidas oficiales de pobreza multidimensional, y en todos los casos -excepto uno (México)- aplican el método Alkire-Foster. Recientemente, el Programa de Naciones Unidas para el Desarrollo (PNUD) y la Oxford Poverty and Human Development Initiative (OPHI⁵) de la Universidad de Oxford, Reino Unido, han publicado el Índice de Pobreza Multidimensional Global para más de 100 países, usando esta misma metodología (PNUD y OPHI, 2019).

3. El uso del método Alkire-Foster en análisis explicativos

El uso del método AF también se encuentra difundido en el área académica, siendo profusos los estudios que lo utilizan, generalmente con objetivos descriptivos. Los análisis de tipo explicativos son menos frecuentes. En la literatura se suele usar una regresión logística o probit para analizar los determinantes de la pobreza unidimensional en donde la variable dependiente toma valor 1 si el hogar es pobre y valor 0, de lo contrario (Glewwe, 1990; Haughton et al., 1999; Haughton y Khandker, 2009). Las variables explicativas consideradas en estas regresiones pueden incluir características regionales o geográficas (por ejemplo, eventos climáticos, densidad poblacional o características institucionales

⁵ Sabina Alkire, quien desarrolló el método junto a James Foster, dirige este centro especializado en la medición de la pobreza multidimensional desde el enfoque de las capacidades.

y políticas), características comunitarias (acceso a servicios públicos o proximidad a mercados), características individuales (genero, edad, horas trabajadas, nivel educativo o estado de salud) o del hogar (tamaño, composición demográfica) (Haughton y Khandker, 2009).

Se ha adoptado una estrategia similar para analizar los determinantes de la pobreza multidimensional, tanto con la aplicación de la medida de Alkire-Foster (Alkire et al., 2015) como con otras medidas de privación multidimensional (Beduk, 2018). No obstante, en el análisis de la pobreza multidimensional existe un problema de endogeneidad potencial en la selección de las variables del “lado derecho”. Los mismos autores advierten acerca de esta circunstancia en relación a las dificultades propias que acarrea el método multidimensional:

al modelar la probabilidad de que un hogar ingrese en una situación de pobreza monetaria, varias variables relacionadas con la salud y la educación [por ejemplo], que no están integradas en la medida de pobreza monetaria, se usan como variables exógenas. En un caso multidimensional, estas variables exógenas pueden ser usadas directamente para construir la medida de pobreza, por lo cual los modelos de probabilidad están sujetos a problemas de endogeneidad potencial (Alkire et al., 2015, p.297, la traducción es propia)

Aun cuando esto sea cuidadosamente evitado en la especificación de los modelos de regresión, la dificultad principal que se plantea al modelar la familia de medidas de Alkire-Foster es la posibilidad de modelar de manera conjunta las tres medidas más básicas: H, A y M₀. Si se deseara estimar únicamente la incidencia de la pobreza multidimensional, H, bastaría con recurrir a una regresión logística, de la misma manera que en las estimaciones de pobreza unidimensional por ingresos, dado que la variable que da lugar a su construcción es una variable dicotómica que adquiere valor 1 en caso de que el hogar sea considerado multidimensionalmente pobre, y 0 en caso contrario. Sin embargo, esta opción no se ve enriquecida de la principal virtud del método AF: no es posible determinar en qué medida los factores tomados en cuenta “del lado derecho” tienen efecto sobre la intensidad de la pobreza (cantidad de privaciones por encima del umbral).

Tomando en cuenta este aspecto, podría modelarse el recuento de privaciones, c_i . En este caso, podría optarse por expresar el indicador como el recuento de privaciones, o bien dividir este recuento por el límite superior de la escala para conformar una proporción que se desarrolle en el intervalo [0,1]. Para los datos de recuento podría recurrirse a la estimación de un modelo de Poisson. Para los datos de proporción, podría recurrirse a un modelo fraccional. Alkire et al. (2015) sugieren emplear un Modelo Lineal General (GLM) con un link logit, cuyas características se describen en la próxima sección. Se descarta la distribución beta para este análisis, dado que solo contempla los valores dentro del intervalo abierto (0,1), si bien podría usarse el zero-one-inflated beta -ZOIB- para este propósito (Ferrari y Cribari-Neto 2004, Smithson y Verkuilen 2006). Sin embargo, modelar el recuento o la proporción de privaciones c_i deja de lado otro de los factores definitorios del método, que es el umbral de pobreza multidimensional.

Para poder aproximarnos a la modelización de M₀, es necesario tomar en cuenta de manera simultánea ambos aspectos: el umbral y la cantidad de privaciones por encima del umbral. La distribución de la variable que se origina al hacerlo tiene características atípicas. En primer lugar, un exceso de ceros, que se genera al recodificar como 0 todos los valores de c_i de los hogares no pobres. En segundo lugar, este mismo procedimiento produce un conjunto de valores censurados, puesto que todos aquellos valores por debajo del umbral de pobreza han sido codificados como cero. Luego una

distribución, ya sea de proporciones o de recuento, probablemente sesgada hacia la derecha por la propia naturaleza de la variable (mayor cantidad de hogares con menor acumulación de privaciones). Finalmente, la posibilidad (pero no la certeza) de existencia de valores observados para la unidad (1), correspondiente a hogares que sufren privaciones en todas las dimensiones analizadas.

Aun cuando encontremos la forma de modelar una variable con estas características, también es necesario considerar el proceso generador de estos datos: ¿Se puede suponer que el proceso que genera que un hogar no sea pobre (probabilidad de ser o no ser multidimensionalmente pobre) es independiente del proceso que está detrás de la acumulación de privaciones por encima del umbral?

Ramalho y Vidiyal Da Silva (2009) analizan el supuesto de independencia entre las partes y sugieren que la decisión de estimar el modelo en una o dos partes depende de la interpretación de los ceros observados. En el caso del método Alkire-Foster, la variable dependiente tiene ceros reales (hogares que no sufren ninguna privación) y ceros contruidos por la decisión de establecer el umbral de pobreza multidimensional al momento de construir la medida de Alkire-Foster. El supuesto que estamos tomando en la definición de la medida es que la acumulación de privaciones por debajo del umbral establecido es -en términos prácticos- igual a cero, lo cual implica suponer que una persona que tiene menos de k privaciones, estaría en la misma situación si tuviera cero privaciones. En términos de los axiomas de las medidas multidimensionales, este proceso implica tener un foco en la pobreza: obtener una medida que es ciega a las privaciones de los hogares que no son multidimensionalmente pobres. Este es, efectivamente, el propósito del umbral en la medida Alkire-Foster. Modelar este umbral es ser consecuente con el diseño de la medida.

Este supuesto implica la independencia entre el hecho de no ser multidimensionalmente pobre (que la variable adquiera valor cero, el límite inferior) y el hecho de acumular privaciones por encima del umbral. Cada parte del modelo estimado busca explicar la proporción de privaciones en distintos tramos de la distribución de c_i : el primero corresponde al extremo inferior de la distribución, en tanto que el segundo corresponde al sector superior. Si entendemos a c_i como una medida agregada de privaciones, de la misma manera que el ingreso es una medida proxy del bienestar, puede ser razonable argumentar que los factores que explican la variación en la parte superior de la distribución sean diferentes de los que explican la variación en la parte baja, particularmente la explicación de si un hogar puede o no salir de la situación de acumulación crítica de privaciones.

En segundo lugar, podemos entender a c_i no como un índice sino como un conjunto compuesto de privaciones. Nuevamente buscando la analogía en términos de los ingresos, se podría argumentar la existencia de modelos diferentes que explican diferentes tipos de ingresos: el ingreso salarial en la base de la distribución y el ingreso por rentas con mayor relevancia en la parte superior de la distribución.

En relación con este procedimiento, también podría suponerse que el proceso que genera la proporción de privaciones es el mismo, sin importar cuántas privaciones tenga la persona (un proceso generador de pobreza en múltiples dimensiones). Ahora bien, cuando se procede operativamente aplicando el umbral de Alkire-Foster, se produce una censura en la medida. Al provocar esa censura, se vuelve necesario modelar un modelo de dos partes: uno para el exceso de ceros y otro para la proporción de privaciones más allá de ese umbral. En este escenario, el modelo de dos partes no se explica por la “naturaleza” de la variable, sino por el método que se usa para estimarla.

En la literatura se han presentado distintos argumentos tanto teóricos como empíricos, para justificar el supuesto de independencia de procesos. Beduk (2018), quien emplea un índice de recuento para

estudiar la privación material en 25 países de la Unión Europea, desarrolla argumentos y realiza pruebas estadísticas para demostrar que tener cero privaciones es un fenómeno cualitativamente diferente a tener privaciones por encima de un umbral. En su caso, entiende las diferencias en términos de su relación con la pertenencia a una clase social diferente, argumento que aplicaría a la situación menos acuciante de los países del continente analizado por el autor.⁶ Pereira et al. (2013) en su aplicación del modelo *inflated-truncated beta* (véase Sección 5 de este documento) al pago de las tarjetas de crédito argumentan que la independencia de los procesos es una cuestión empírica. Puede mencionarse también a Schwiebert y Wagner (2015), quienes estudian específicamente la cuestión de la dependencia entre procesos en un modelo de dos partes para la estimación de proporciones. Proponen una generalización del modelo de dos partes que permite la consideración de dependencia entre los procesos, comprendiendo al caso de la independencia como un caso especial.

En el apartado siguiente se presentan el diseño y la implementación de la medición de impacto del programa de vivienda de emergencia de TECHO como una manera de ilustrar el procedimiento y las implicancias de las decisiones metodológicas.

4. Uso de la medida de Alkire-Foster en evaluación de impacto: el caso de TECHO

Para que la discusión acerca de la conveniencia de la forma de modelaje de la medida resulte más comprensible, este documento procura dar respuesta a estos interrogantes mediante la aplicación de la medida de Alkire-Foster para la medición de impacto de un programa apuntado a reducir la pobreza multidimensional.⁷ Las medidas de impacto consolidadas tienen un uso cada vez mayor en la literatura de evaluación de impacto, como una estrategia para formular conclusiones sobre el efecto general de un programa y para resolver el incremento en la probabilidad de cometer error tipo I cuando se realizan múltiples pruebas de hipótesis (Anderson, 2008). Kling, Liebman y Katz (2007), por ejemplo, construyen indicadores resumen de las distintas dimensiones evaluadas, mediante el cálculo de puntuaciones-z (estandarización de las variables). Para cada indicador de impacto se resta la media del indicador del grupo de control y se divide por el desvío estándar del indicador del grupo de control. Si Y_k es el indicador k de K indicadores de impacto dentro de una dimensión, μ_k y σ_k son, respectivamente, la media y el desvío estándar del grupo de control, el resultado normalizado es $Y_k^* = (Y_k - \mu_k) / \sigma_k$ y el índice de resumen es $Y^* = \sum_k Y_k^* / K$.⁸ Sin embargo, estas operaciones matemáticas dan lugar a un indicador agregado que generalmente no tiene un sentido conceptual de fondo. Asimismo, los indicadores se incorporan al índice agregado de manera directa y no ponderada. Lo novedoso de esta propuesta consiste en la construcción de un indicador (la medida de Alkire-Foster) construido sobre la base de criterios normativos en lugar de estrictamente empíricos. A continuación, se describen el programa de vivienda de emergencia de TECHO, el diseño de la

⁶ Beduk (2018) utiliza un ZIP (zero-inflated poisson) para estimar su modelo ya que emplea un recuento de privaciones en lugar de utilizar una proporción de privaciones ponderadas. El uso de este modelo es menos generalizable, quedando inhabilitado si el umbral que utilizara no fuese $k=1$, o si las ponderaciones dieran lugar a valores no enteros de la variable de recuento.

⁷ Otro ejemplo de una posible aplicación, fuera de la evaluación de impacto, podría ser la verificación de brechas urbanorurales en la pobreza multidimensional.

⁸ Dado que muchas veces, los indicadores de impacto tomados en cuenta muestran direccionalidades distintas, es necesario invertir el signo para que valores mayores de cada indicador signifiquen resultados más beneficiosos en términos de la evaluación (Kling, Liebman y Katz, 2007), y de esta manera los efectos no se contrarresten en el agregado.

evaluación de impacto del programa y la medida de pobreza multidimensional de Alkire-Foster en dimensiones vinculadas con la vivienda.

4.1. Programa de vivienda de TECHO

El programa de vivienda de emergencia de TECHO provee una casa simple y pensada para ser transitoria a familias que habitan en villas y asentamientos en situación de profunda precariedad. La vivienda de 18m² está hecha de paneles de madera prefabricados y techo de chapa con aislante, y se apoya sobre 15 pilotes que la aíslan de la humedad del suelo y de las inundaciones que son frecuentes en estos barrios. No provee conexiones de agua, saneamiento o electricidad. El costo de la vivienda es de aproximadamente \$1.000 dólares.

El proceso de selección de beneficiarios de TECHO consiste, en primer lugar, en la realización de una encuesta de detección en cada barrio, con el objetivo de relevar información sobre la composición de los hogares y las características de los miembros, la situación económica del hogar y las características de la vivienda. Luego se utiliza esta información en conjunto con información cualitativa para clasificar los hogares según su nivel de necesidad. Los hogares que se clasifican como de riesgo alto y medio alto son elegibles para participar del programa.

4.2. Diseño de la evaluación de impacto

En 2013 se comenzó una investigación del Departamento de investigación de la Facultad de Ciencias Económicas de la UCA en conjunto con la organización TECHO para evaluar el impacto del programa de vivienda de emergencia en el bienestar de los hogares beneficiados. Como puede consultarse en Mitchell, Macció y Mariño Fages (2019) se aplicó un diseño cuasi-experimental de *pipeline* (Coleman, 1999; Bali Swain y Varghese, 2009 y Deininger y Lui, 2013) que consiste en la comparación ex post entre los hogares tratados y un grupo de control conformado por los hogares que estuvieron aguardando su turno (*“in the pipeline”*, en inglés) para recibir una vivienda. Los casos fueron relevados en 34 asentamientos informales del Gran Buenos Aires y Gran La Plata.

La evaluación utilizó información recolectada a través de dos cuestionarios: una encuesta línea de base aplicada a los hogares (tanto los casos tratados como los de control) después de haber sido seleccionados para el programa, pero antes de recibir la vivienda, y una encuesta de seguimiento aplicado a los hogares tratados aproximadamente un año después de recibir la vivienda. El grupo de tratamiento está conformado por los 334 hogares que recibieron la vivienda de TECHO durante la primera mitad de 2014⁹ y el grupo de control por los 352 hogares que la recibieron durante la primera mitad de 2015. Las pruebas de balanceo entre los hogares tratados y de control mostraron que no hubo una diferencia estadísticamente significativa entre grupos para ninguno de los indicadores considerados. Se utilizó la siguiente ecuación de regresión para estimar el efecto promedio del tratamiento en los hogares tratados (δ):

$$Y_i = \alpha + \beta X_i + \gamma Z_i + \delta T_i + \varepsilon_i \quad (5)$$

donde $Y_i = 1$ es el indicador de impacto; T_i es el indicador de tratamiento; X_i es un conjunto de características pre-tratamiento, Z_i es un conjunto de variables dummy de barrio; ε_i es el término de error. Se tienen en cuenta errores estándar robustos con *clustering* a nivel de barrio.

⁹ La encuesta de seguimiento fue aplicada a un total de 293 hogares. Mitchell, Macció y Mariño Fages (2019) evalúan la sensibilidad de los resultados a la pérdida de 41 casos y concluyen que la atrición de casos no tuvo un efecto relevante sobre los resultados.

Mitchell, Macció y Mariño Fages (2019) muestran que el programa de vivienda produjo un efecto de gran magnitud y estadísticamente significativo en indicadores de privacidad, seguridad, relaciones interpersonales, salud psicológico y percepción de bienestar. En la dimensión de salud física, se encuentra que el programa produjo una reducción en la prevalencia de tos y congestión. Los indicadores de sueño también mejoraron, pero estos efectos dejaron de ser estadísticamente significativos luego de ajustar por aplicar pruebas de hipótesis múltiples. Asimismo, se muestra que los resultados son robustos a la aplicación de pruebas por posibles fuentes de sesgo (atracción de casos, diferencias de barrio y de estación) y un método de evaluación alternativa (apareamiento por puntaje de propensión – *propensity score matching*), además de una prueba de falsificación.

Mitchell y Macció (2018) muestra los resultados preliminares de la medición del efecto del programa de vivienda de TECHO en una medida agregada de privación asociada a la vivienda usando el método de Alkire-Foster. A diferencia de los mencionados antecedentes, este documento propone una discusión metodológica de los métodos econométricos empleados para la demostración del efecto del programa.

4.3. Medida de pobreza multidimensional asociada a la vivienda

Sobre la base de los indicadores de impacto utilizados en Mitchell, Macció y Mariño Fages (2019) se elaboró una medida agregada, equiponderada en dimensiones e indicadores dentro de cada dimensión (Tabla 1). Mitchell y Macció (2018) presentan todos los argumentos normativos y metodológicos que fueron empleados para la construcción de esta medida.

El siguiente gráfico presenta la distribución de la variable de recuento dividida por el máximo número posible de privaciones totales. Se trata de una proporción ponderada de indicadores en situación de privación. Si bien todavía no se ha aplicado el segundo umbral para el cálculo de la medida AF, puede notarse que la proporción de hogares con un tercio de las privaciones o más es de 69,1%.

Gráfico 1a: Distribución de la variable c_i

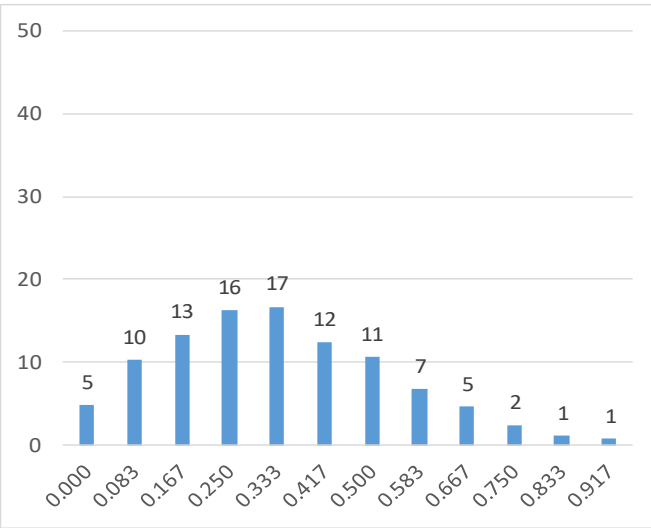
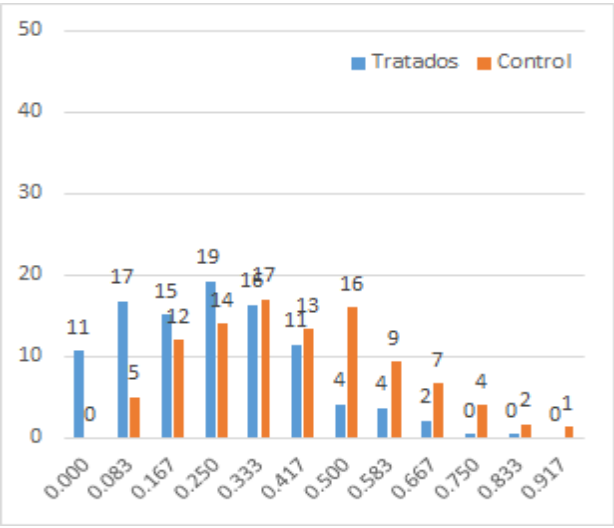


Gráfico 1b: Distribución de la variable c_i según situación del hogar



Fuente: Elaboración propia sobre la base de UCA-TECHO (2015).

Tabla 1: Dimensiones, indicadores y ponderación empleados para construir la medida AF empleada en la evaluación de impacto multidimensional del programa de TECHO

Dimensión	Indicador	Definición del indicador	Ponderación
Salud física	Salud respiratoria	Al menos un miembro sufrió tos o congestión frecuentemente durante el último año	0.0833
	Dolor de articulaciones	Al menos un miembro del hogar tuvo dolor en las articulaciones durante el último mes	0.0833
Salud psicológica	Salud emocional	El respondiente tuvo sentimientos negativos (ansiedad, depresión, etc.) durante el último mes	0.0833
	Stress debido a conflictos en el hogar	El respondiente se sintió estresado por los conflictos en el hogar durante el último mes	0.0833
Sueño	Calidad del sueño	Al menos un miembro se levanta frecuentemente durante la noche	0.0833
	Cantidad de sueño	El respondiente duerme menos de 6 horas por noche	0.0833
Privacidad	Privacidad dentro de la casa	Los miembros del hogar no tienen lugar para vestirse en privado	0.0833
	Privacidad en el sueño	Al menos una cama está hacinada (más de una persona por plaza)	0.0833
Relaciones interpersonales	Relaciones sociales con los demás miembros del hogar	Conflicto frecuente en el hogar debido a la falta de espacio	0.0833
	Relaciones sociales con personas fuera del hogar	El hogar no recibe amigos o familiares en la casa porque se siente incómodo	0.0833
Seguridad	Seguridad física	El respondiente teme que la vivienda se derrumbe	0.0833
	Seguridad respecto del robo	La vivienda sufrió un robo durante el último año	0.0833

Fuente: Mitchell y Macció (2018).

Sobre la base de esta distribución de las privaciones en dimensiones del bienestar vinculado con la vivienda, se presentan a continuación las estimaciones obtenidas de las medidas de Alkire-Foster y las diferencias observadas entre sus valores para los grupos de tratamiento y control. Debe tenerse en cuenta que se definió un umbral de pobreza multidimensional igual a $k=33\%$, equivalente a un tercio de los indicadores ponderados.

Tabla 2: Medidas Alkire-Foster para los grupos de control y tratamiento, con $k=33\%$

	H	A	M0
Grupo de control	0.691	0.501	0.346
Grupo de tratamiento	0.384	0.427	0.164
Diferencia observada	-0.307	-0.074	-0.182

Fuente: Elaboración propia sobre la base de UCA-TECHO (2015).

La incidencia de la pobreza multidimensional es del 69% para el grupo de control y 38% para el grupo de tratamiento. La diferencia observada entre ambas incidencias es de 30 puntos porcentuales. Por su parte, la intensidad de la pobreza en el grupo de control muestra que estos hogares tienen en promedio un 50% de los indicadores en situación de privación (en promedio, acumulan 6 de las 12 privaciones evaluadas). Entre los hogares tratados, este valor se reduce a 42%. Finalmente, cuando se ajusta la tasa de recuento mediante la intensidad de la pobreza, se observa un M_0 de 0.346 para los hogares de control y de 0.164 para los hogares tratados, una diferencia de 0.182 puntos.

5. Propuestas para el modelaje de las medidas Alkire-Foster

En esta sección se avanza de manera progresiva sobre las alternativas de modelaje, demostrando su aplicación al ejemplo de la evaluación de impacto del programa de vivienda de emergencia de TECHO.

5.1. Estimación mediante mínimos cuadrados ordinarios

La alternativa más simple para la estimación del impacto en las tres medidas es el método de mínimos cuadrados ordinarios (MCO). Para medir el impacto en H (tasa de recuento de la pobreza multidimensional) se puede estimar:

$$h_i = \alpha + \beta X_i + \gamma Z_i + \delta_1 T_i + \varepsilon_i \quad (6)$$

Que es la ecuación (5) donde la variable dependiente se reemplaza por $h_i = 1$ si el hogar es multidimensionalmente pobre, o 0 en caso contrario.

La utilización de MCO para la estimación de impacto en variables dicotómicas está ampliamente difundida, particularmente cuando se busca evaluar el efecto promedio de una intervención en los tratados (Angrist y Pischke, 2009). El coeficiente del indicador de tratamiento puede interpretarse como el efecto en la media de la variable dependiente (Wooldridge 2002, p. 454-457). Sin embargo, el modelo lineal de probabilidades (MLP) -la aplicación de MCO a variables dependientes dicotómicas- tiene un conjunto de dificultades: la distribución no normal de las perturbaciones, su varianza heteroscedástica, y el hecho de que el valor pronosticado no se encuentra acotado entre 0 y 1, e incluso cuestiones operativas como el valor cuestionable del coeficiente de determinación R^2 como una medida adecuada de la bondad de ajuste (Wooldridge, 2002).

Para medir el impacto en A (promedio de privaciones de los pobres), se puede estimar:

$$c_i = \alpha + \beta X_i + \gamma Z_i + \delta_2 T_i + \varepsilon_i, \quad \text{para } c_i \geq k \quad (7)$$

La variable dependiente en este caso es el indicador de recuento o proporción de privaciones y la regresión se estima para la submuestra determinada por los hogares multidimensionalmente pobres (cuyo c_i supera el umbral). La dificultad principal de este análisis consiste en el truncamiento que se realiza al restringir la muestra de esta manera. De esta manera, hay implícito un sesgo de selección no aleatorio.

Finalmente, para medir el impacto en M_0 (tasa de recuento ajustado), se puede estimar:

$$M_0: c_i(k) = \alpha + \beta X_i + \gamma Z_i + \delta_3 T_i + \varepsilon_i \quad (8)$$

En este caso la variable dependiente está censurada por debajo del umbral. Las dificultades se observan en la acumulación de un exceso de ceros y la censura simultánea que ocurre para valores de c_i inferiores al umbral.

5.2. Regresión logística para modelar H

Un camino seguro para el análisis de la pobreza multidimensional es la estimación de una regresión logística que permita modelar la probabilidad de que un hogar sea considerado multidimensionalmente pobre. En el contexto de la evaluación de impacto, por su parte, la regresión logística habilitaría el análisis del efecto del tratamiento sobre la probabilidad de que un hogar sea multidimensionalmente pobre.

En un modelo logístico, la probabilidad de que un hogar sea considerado multidimensionalmente pobre puede modelarse mediante:

$$P_i = \frac{1}{1 + e^{-z_i}} = \frac{e^{z_i}}{1 + e^{z_i}} \quad (9)$$

donde $z_i = \beta_0 + \beta_1 x_i$. De esta manera, los valores de probabilidad están acotados entre 0 y 1. La estimación mediante la regresión logística de este valor permite una alternativa más adecuada para reemplazar a la estimación mediante MCO presentada en (6).

El procedimiento para realizar esta estimación implica la recodificación de la variable presentada en el Gráfico 1, asignando valor 0 para los hogares por debajo del umbral (no pobres) y valor 1 para los hogares que acumulan privaciones en o por encima del umbral de pobreza multidimensional. Según los datos que se presentan en los Gráficos 2, observamos el 55% de los hogares multidimensionalmente pobres que también surgen del Gráfico 1. La incidencia de la pobreza multidimensional entre los hogares tratados es de 38%, en tanto que para los hogares del grupo de control alcanza un 69%, resultados que replican los observados en la Tabla 2.

Gráfico 2a: Situación de pobreza multidimensional

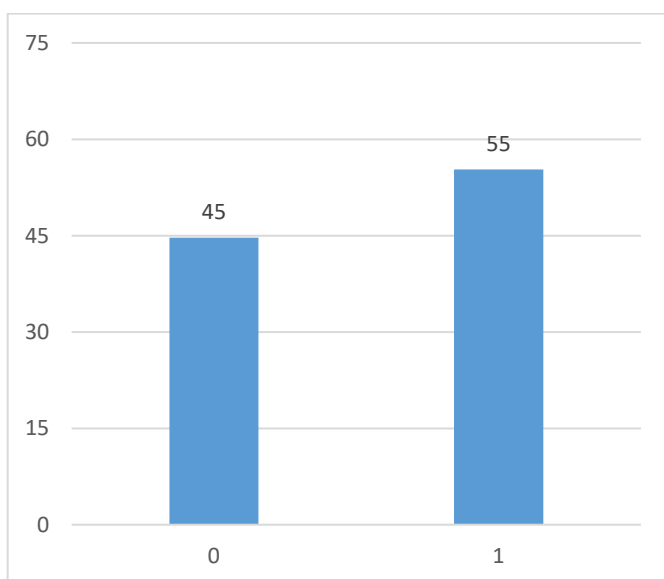
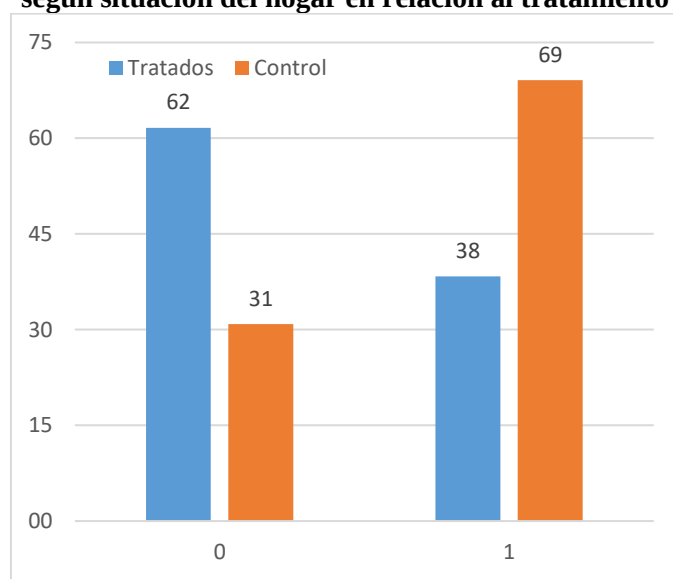


Gráfico 2b: Situación de pobreza multidimensional según situación del hogar en relación al tratamiento



Fuente: Elaboración propia sobre la base de UCA-TECHO (2015).

5.3. Regresión GLM para c_i

Alkire et al. (2015) proponen los modelos lineales generalizados (GLM, por sus siglas en inglés) como un marco analítico suficientemente amplio para abarcar el análisis de regresión con las medidas Alkire-Foster. Los autores argumentan que estos modelos dan cuenta de la naturaleza acotada y discreta de la variable de interés, permitiendo la estimación de un conjunto de modelos con variable dependiente cuantitativa o cualitativa, como los modelos de regresión lineal, los modelos de elección discreta logit o probit, e incluso modelos para datos de proporciones. Se trata de un marco general que permite el estudio de determinantes de la pobreza multidimensional tanto a nivel micro como macro.

El modelo GLM fue originalmente propuesto por Papke y Wooldridge (1996) para la estimación de proporciones. Lo que sigue procede de Alkire et al. (2015). La familia de modelos GLM supone la predicción de una función de la media condicional de la variable dependiente como una combinación lineal de un conjunto de variables explicativas. La regresión lineal se estima mediante un caso especial del GLM en el cual la esperanza condicional de la variable dependiente se modela mediante la función identidad. Se puede extender el dominio de aplicación a casos en los que la variable dependiente no es continua o no está distribuida normalmente, así como para casos en los que se intenta modelar distribuciones asimétricas. En los modelos GLM, el término de error no necesariamente es aditivo, sino que se atribuye a la variable dependiente en sí misma. Esto implica que es necesario especificar la distribución condicional de la variable dependiente, dados los valores de las variables explicativas, denominado $f_Y(\mathbf{y})$. Estas distribuciones habitualmente corresponden a la familia lineal exponencial, como es el caso de la gaussiana, binomial, poisson, gamma, entre otras (si bien también existen extensiones a las familias no exponenciales). El GLM toma la forma:

$$g(\mu_{Y_i/x_i}) = \eta_i = \beta_0 + \sum_{j=1}^K \beta_j x_{ij} \quad (10)$$

donde la parte sistemática de la predicción -el predictor lineal- (η_i) es ahora una función (g) de la esperanza condicional de la variable dependiente μ_{Y_i/x_i} . La función $g(\cdot)$ es una función diferenciable, a la que se denomina función *link*. Esta función transforma la esperanza condicional de la variable dependiente en el predictor lineal, que es una función lineal de las variables explicativas que podría tomar cualquier forma. De esta manera, es posible incluir variables continuas o categóricas, interacciones, transformaciones de variables continuas, etc. En el caso de la estimación de las medidas AF, el interés principal radica en estimar la media condicional μ_{Y_i/x_i} , la cual se puede obtener invirtiendo la función link:

$$\mu_{Y_i/x_i} = g^{-1}(\eta_i) = g^{-1}(x_i \boldsymbol{\beta}) \quad (11)$$

La elección de la función link está relacionada con la distribución de la variable dependiente, en particular del rango en que se extiende esta variable. Para las medidas AF, la variable dicotómica que supone la condición de pobre/no pobre puede describirse mediante una distribución Bernoulli y modelarse mediante un link logit. En el caso de que se emplee la proporción ponderada de privaciones, cuyos valores están restringidos al intervalo de la unidad, la distribución binomial puede resultar apropiada. La función link tendría el objetivo de mapear el intervalo $\{0,1\}$ del primer caso, o el $[0,1]$ del segundo caso, en el intervalo $(-\infty, +\infty)$.

5.4. El modelo de regresión “*truncated inflated beta*”¹⁰

Si bien acordamos con la propuesta del GLM como marco general para la estimación de los determinantes de pobreza, entendemos que estos modelos no permiten modelar la censura de la variable dependiente en el umbral que identifica la pobreza multidimensional. Por ese motivo, sobre la base de Pereira, Botter y Sandoval (2012, 2013) proponemos la utilización del modelo de regresión beta inflado y truncado (*truncated inflated beta*). Los autores presentan el ejemplo del pago mensual de las tarjetas de crédito, en donde las personas pueden decidir no pagar nada (variable con valor cero), pagar la proporción mínima establecida por la empresa (τ), lo cual supone la inexistencia de valores en el intervalo entre cero y este valor, o pagar cualquier proporción en el intervalo $(\tau, 1)$. De hecho, el documento permite incluso la existencia de un umbral variable (el límite mínimo de pago puede variar entre las instituciones bancarias). La variable resultante en este caso guarda mucha similitud con la variable c_i censurada para valores inferiores al umbral. A continuación, se brindan las principales características del modelo.

Pereira, Botter y Sandoval (2012, 2013) presentan el modelo de regresión truncado e inflado, que está definido para el valor cero y para el intervalo $[\tau, 1]$. Se trata de un mixture model (modelo de mezclas): por un lado, se presenta una distribución multinomial con masas de probabilidad en 0, τ , y 1, y por el otro, una distribución beta definida en el intervalo abierto $(\tau, 1)$. Por su parte, Smithson y Verkuilen (2012) presentan una mezcla de C componentes formados por distribuciones beta para dar cuenta de la asimetría y la heteroscedasticidad en las distribuciones, mediante su propuesta de modelar la media (posición) y la varianza (dispersión) de manera separada con sus propios conjuntos de predictores. Gray y Hernández Alava (2018) combinan ambas generalizando la versión inicial del modelo beta inflado y truncado. Esta generalización permite modelar variables dependientes que están delimitadas dentro de un intervalo y pueden sufrir un truncamiento ya sea al inicio o al final de la distribución. Según los autores, las ventajas de este método son varias:

1. La mezcla de varias distribuciones beta permite modelar la existencia de múltiples modas.
2. Permite modelar variables de respuesta que sufren una discontinuidad entre uno de los valores límites (ya sea el inferior o el superior) y la parte continua de la distribución.
3. A diferencia del modelo beta tradicional, se pueden modelar las observaciones que se encuentran sobre los valores límites (0 o 1) o bien en el punto de truncamiento.

Siguiendo a Gray y Hernandez Alava (2018), podemos suponer una variable de respuesta y_i que está definida en el punto a y en el intervalo $[\tau, b]$, donde $a < \tau < b$. Si suponemos tres diferentes vectores columna de covariables x_{i1} , x_{i2} y x_{i3} , la densidad condicional de y_i puede escribirse como:

$$g(y_i/x_{i1}, x_{i2}, x_{i3}) = \begin{cases} P(y_i = a/x_{i3}) & \text{if } y_i = a \\ P(y_i = \tau/x_{i3}) & \text{if } y_i = \tau \\ P(y_i = b/x_{i3}) & \text{if } y_i = b \\ [1 - \sum_{s=a,\tau,b} P(y_i = s/x_{i3})]h(y_i/x_{i1} x_{i2}) & \text{if } y_i \in (\tau, b) \end{cases} \quad (12)$$

Las probabilidades $P(y_i/x_{i3})$ se obtienen de un modelo logit multinomial:

$$P(y_i = r/x_{i3}) = \frac{\exp(x'_{i3}\gamma_r)}{1 + \sum_{s=a,\tau,b} \exp(x'_{i3}\gamma_s)} \quad \text{para } r = a, \tau, b \quad (13)$$

¹⁰ Una opción de estimación muy próxima a este modelo es el ZOIB (zero-one inflated beta).

Las x_{i3} son variables que afectan la probabilidad de ocurrencia de uno de los valores límites de la variable dependiente (ya sea 0, 1 o el umbral de truncamiento), y γ_r son los coeficientes correspondientes.

Luego, la función de densidad de probabilidades $h(\cdot)$ es una mezcla de C componentes de distribuciones beta con medias μ_{ci} y parámetro de precisión ϕ_c

$$h(y_i/x_{i1} x_{i2}) = \sum_{c=1}^c [p(c/x_{i2})f\{y_i; \mu_{ci}(x_{i1}; \beta_c), \phi_c, \tau, b\}] \quad (14)$$

donde $f(\cdot)$ es la densidad beta definida en (12). Un logit multinomial se define para establecer la probabilidad de membresía a cada clase latente, es decir, a cada componente:

$$P(c = x_{i2}) = \frac{\exp(x'_{i2}\delta_c)}{\sum_{j=1}^C \exp(x'_{i2}\delta_j)} \quad (15)$$

donde x_{i2} es un vector de variables que afectan la probabilidad de membresía a cada componente, δ_c son los coeficientes correspondientes, y C es la cantidad de clases que se usan en el análisis.

Cuando se incluye un solo componente, la estimación de este modelo es equivalente a un modelo en dos partes, que se ajusta de manera conjunta bajo el supuesto de independencia condicional entre las dos partes del modelo (Grey y Hernandez Alava, 2018).

Gráfico 3a: Distribución de la variable $c_{i(k)}$

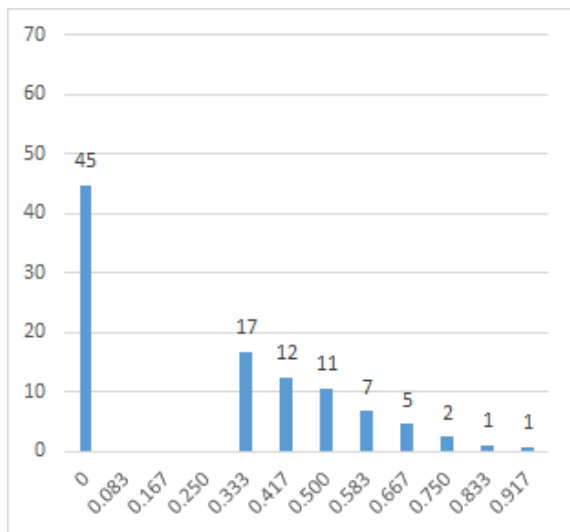
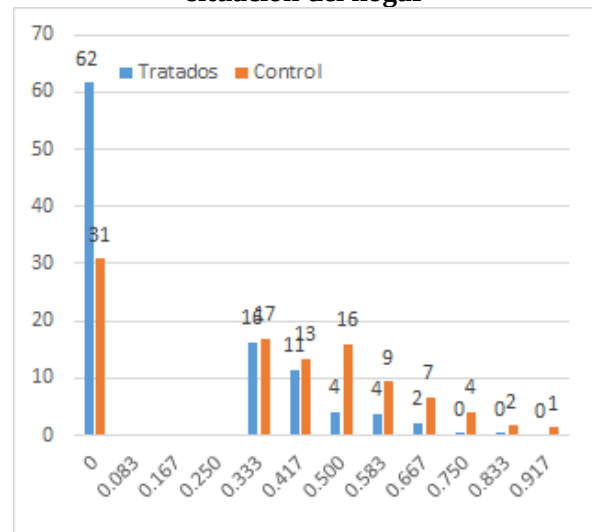


Gráfico 1b: Distribución de la variable $c_{i(k)}$ según situación del hogar



Fuente: Elaboración propia sobre la base de UCA-TECHO (2015).

6. Resultados

Tomando en cuenta las alternativas presentadas en el apartado anterior, se presentan aquí los resultados de la estimación de los diferentes modelos descriptos. En términos generales, los resultados parecen confirmar un efecto significativo del programa de vivienda de emergencia de TECHO en la disminución de la incidencia y la intensidad de la pobreza multidimensional.

Las primeras 3 columnas del cuadro presentan los resultados de las estimaciones mediante MCO. La estimación del coeficiente que acompaña a la variable dicotómica de tratamiento da cuenta de la diferencia entre la medida para los hogares tratados y los del grupo de control. Se observa que todas las diferencias son negativas y significativamente distintas de cero, siendo su magnitud muy similar a la diferencia observada que se consignó en la Tabla 2.

La columna correspondiente a los resultados de la regresión logística permite ver únicamente el efecto sobre la incidencia de la pobreza, sin tomar en cuenta su intensidad. En este caso se informan las razones de probabilidades (OR: odds ratio); un OR por debajo de la unidad implica menores chances de ser considerado multidimensionalmente pobre entre los tratados que en el grupo de control. Específicamente, la magnitud estimada indica que la probabilidad de que un hogar sea clasificado como multidimensionalmente pobre es de un cuarto (1/4) entre los tratados respecto de los hogares del grupo de control. Nuevamente se observa que la diferencia es estadísticamente significativa.

Los resultados de la estimación GLM muestran un signo coherente, pero su interpretación debe estar asociada al hecho de que la variable dependiente corresponde a la distribución completa de privaciones ponderadas. En ese sentido, no se ejerce el efecto del umbral, y el coeficiente representa entonces la disminución en la proporción de privaciones condicional al tratamiento.

Finalmente, el inflated-truncated beta se estimó para un único componente, suponiendo una varianza constante (es decir, sin modelar los determinantes de la varianza de la parte beta de la mezcla). Se estableció un límite inferior de 0, y un límite superior de 1. Esto fue realizado en atención a los límites teóricos de la variable y no a los observados, dado que no se presentan hogares que sufran privación en todos los indicadores de manera simultánea. Se estableció el umbral de censura en 0,333 (correspondientemente con el umbral de pobreza multidimensional seleccionado). No se estimó inflación en el umbral ni en el límite superior, pero sí se estableció la existencia de inflación en el límite inferior de cero¹¹.

Los resultados se presentan en las últimas columnas de la Tabla 3. En la última columna se observa la estimación de la inflación en ceros, que implica la estimación de una regresión logística. A diferencia de la presentada anteriormente, en este caso se modela la probabilidad de *no ser* multidimensionalmente pobre. El coeficiente que se asocia al indicador T debe ser antilogaritmado para poder interpretarse de manera comparable al odds ratio anterior. De esta manera, $e^{1.333} = 3.79$, es decir, un hogar tiene casi cuatro veces menos chances de ser no pobre en el grupo de tratados que en el grupo de control¹². Cuando se calcula la probabilidad de una observación de pertenecer a la parte beta de la mezcla, se obtiene que la probabilidad promedio de un caso no tratado (control) de tener una proporción de privaciones igual o mayor que el umbral es de 69%, en tanto que la probabilidad promedio de un caso tratado de tener una proporción de privaciones igual o mayor que el umbral es del 38%. Estas estimaciones coinciden con los valores observados para H de la Tabla 2.

¹¹ Los cálculos fueron realizados con STATA, empleando el comando *betamix* (Gray y Hernandez Alava, 2018).

¹² Este es un resultado muy similar al presentado arriba para la regresión logística.

Tabla 3: Resultados de la estimación de los modelos alternativos para la medida AF

	MCO			Regresión logística	Fractional regression	ITBeta		
	H Coef. err. std	A Coef. err. std	M0 Coef. err. std	H OR err. std	GLM Coef. err. std	μ Coef. err. std	ϕ Coef. err. std	PM_lb Coef. err. std
Tratamiento	-0.286 *** (0.065)	-0.088 *** (0.021)	-0.180 *** (0.034)	0.264 *** (0.079)	-0.735 *** (0.121)	-1.085 *** (0.239)		1.333 *** (0.3)
Cantidad de núcleos familiares	0.030 (0.036)	0.004 (0.022)	0.018 (0.022)	1.128 (0.193)	0.083 (0.078)	-0.126 (0.222)		-0.121 (0.171)
Núcleo completo	0.169 *** (0.055)	0.047 (0.029)	0.102 *** (0.033)	2.396 *** (0.671)	0.402 *** (0.117)	0.313 (0.237)		-0.874 *** (0.279)
Edad del referente	-0.001 (0.002)	0.001 (0.001)	0.000 (0.001)	0.992 (0.01)	0.001 (0.004)	-0.004 (0.011)		0.008 (0.01)
Referente varón	-0.057 (0.061)	-0.073 ** (0.027)	-0.058 * (0.031)	0.723 (0.213)	-0.203 * (0.12)	-0.695 ** (0.284)		0.325 (0.294)
Referente extranjero	0.138 (0.104)	0.054 (0.034)	0.087 (0.054)	1.895 (0.929)	0.367 ** (0.182)	0.569 * (0.327)		-0.639 (0.49)
Referente del interior	0.035 (0.067)	0.052 ** (0.022)	0.041 (0.032)	1.177 (0.399)	0.122 (0.1)	0.413 ** (0.189)		-0.163 (0.339)
Referente con educación primaria	0.054 (0.054)	-0.009 (0.021)	0.021 (0.028)	1.296 (0.345)	0.083 (0.096)	-0.205 (0.263)		-0.260 (0.266)
Referente con secundaria incompleta	0.134 ** (0.05)	0.029 (0.024)	0.082 *** (0.028)	1.938 *** (0.483)	0.304 *** (0.096)	0.156 (0.185)		-0.662 *** (0.249)
Referente con secundaria completa	0.099 (0.06)	0.001 (0.03)	0.046 (0.033)	1.623 * (0.477)	0.216 * (0.121)	-0.299 (0.245)		-0.484 * (0.293)
En el hogar hay una mujer embarazada	-0.030 (0.073)	-0.001 (0.029)	-0.009 (0.037)	0.858 (0.308)	0.021 (0.112)	0.048 (0.211)		0.153 (0.359)
El referente trabaja	0.035 (0.043)	0.002 (0.018)	0.012 (0.022)	1.204 (0.259)	-0.028 (0.08)	-0.022 (0.166)		-0.186 (0.215)
El cónyuge trabaja	-0.133 ** (0.05)	-0.026 (0.027)	-0.083 ** (0.032)	0.500 ** (0.137)	-0.314 *** (0.12)	0.003 (0.304)		0.694 ** (0.275)
Otro miembro del hogar trabaja	0.009 (0.035)	0.006 (0.012)	0.002 (0.021)	1.053 (0.183)	0.004 (0.067)	0.127 (0.143)		-0.052 (0.173)
Proporción de personas de 0 a 5 años	0.037 (0.127)	-0.031 (0.062)	0.007 (0.073)	1.111 (0.675)	0.081 (0.247)	-0.534 (0.81)		-0.105 (0.606)
Proporción de personas de 6 a 17 años	0.088 (0.118)	-0.007 (0.059)	0.044 (0.077)	1.532 (0.873)	0.292 (0.264)	0.102 (0.539)		-0.427 (0.569)
Alguien el el hogar tiene asma	0.091 (0.06)	-0.003 (0.02)	0.042 (0.028)	1.579 (0.489)	0.119 (0.082)	-0.097 (0.243)		-0.456 (0.309)
Ingresos per cápita del hogar (4to trim 2014)	0.000 ** (0)	0.000 * (0)	0.000 ** (0)	1.000 ** (0)	0.000 *** (0)	0.000 * (0)		0.000 ** (0)
_cons	0.532 *** (0.147)	0.476 *** (0.093)	0.257 *** (0.093)	1.264 (0.917)	-0.775 ** (0.355)	-0.182 (0.819)	0.543 *** (0.116)	-0.234 (0.725)
Dummies de barrio	Se incluyen	Se incluyen	Se incluyen	Se incluyen	Se incluyen	Se incluyen		Se incluyen
R2/pseudo R2	0.207	0.260	0.248	0.162				
n	546	302	546	544 ξ	546		546	
AIC	695.96	-382.50	-42.40	662.85	0.99		-910.42	
BIC	773.41	-315.71	35.05	740.23	-3149.98		-764.13	

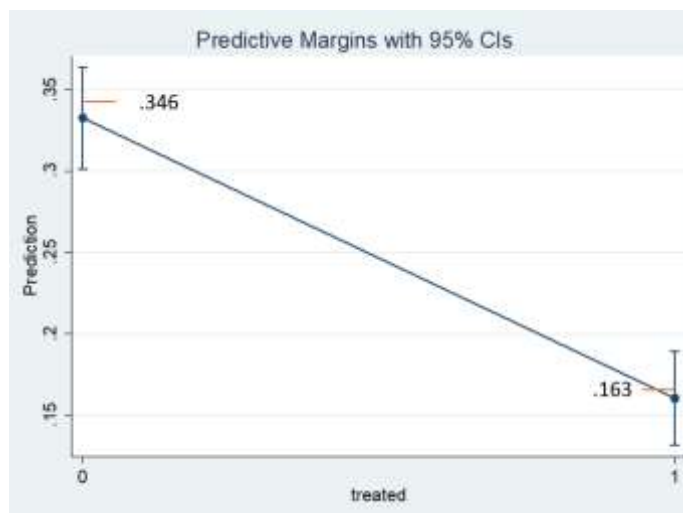
Nota: los criterios AIC y BIC, así como el coeficiente de determinación y pseudo R cuadrados, no son todos comparables entre sí.

ξ Dos observaciones permitían una predicción perfecta por barrio y fueron eliminadas del análisis.

Como se indicó, la modelación de la parte beta de la mezcla se realiza con un solo componente. El coeficiente de la variable de tratamiento para μ (la media de distribución beta de este componente) igual a -1.085 nos indica que recibir la vivienda de TECHO reduce de manera significativa la proporción ponderada de privaciones por arriba del umbral de pobreza. La varianza del componente, estimada de manera no condicional a las variables explicativas, está fija en 1.72 (si se exponencia el valor informado para encontrar el valor del parámetro sin transformar). El valor pronosticado de la variable dependiente para la muestra mediante el modelo informa una media de aproximadamente 0.25, muy similar a la media observada de 0.264. Estos parámetros describen una distribución beta con pronunciada asimetría por derecha, lo cual resulta descriptivo de la distribución informada en el Gráfico 3.

En términos de la posibilidad del modelo de pronosticar la medida M_0 de Alkire-Foster, se presentan en el Gráfico 4 los efectos marginales pronosticados condicionales al tratamiento, consignando a efectos de su comparación el valor observado M_0 de la Tabla 2.

Gráfico 4: Efectos marginales pronosticados del tratamiento



Es relevante también realizar una comparación del efecto del resto de las variables explicativas. Una visión general de la Tabla 3 muestra similitudes entre los modelos que buscan estimar la incidencia (MCO_H, regresión logística y PM_lb de la mezcla del modelo beta inflado-truncado), entre los que buscan estimar la acumulación de privaciones por encima del umbral (MCO_A y μ de la mezcla), y entre los que toman ambos en cuenta (MCO_M₀ y GLM), si bien con algunas diferencias en niveles de significatividad y magnitud del efecto. Por ejemplo, el hecho de que el hogar tenga núcleo incompleto (no esté presente el cónyuge) tiene una influencia significativa sobre la probabilidad del hogar de estar privado en dimensiones vinculadas a la vivienda, pero no incrementa de manera significativa la cantidad de privaciones en hogares pobres. El hecho de que el referente sea varón no afecta sobre la probabilidad de que el hogar sea pobre pero sí influye disminuyendo la acumulación de privaciones por encima del umbral, y esto es evidente tanto en los modelos para A como los que apuntan a M₀. El ingreso monetario, en cambio, es relevante en todos los modelos. Esta evidencia de la existencia de diferentes factores asociados a la incidencia y a la intensidad de la pobreza abona la hipótesis de procesos independientes y apoya la necesidad de estimar ambos aspectos de manera conjunta pero diferenciada.

Por otro lado, los resultados sobre el efecto en la cantidad de privaciones sobre el umbral también varían de acuerdo al modelo elegido. Por ejemplo, el coeficiente correspondiente al hecho de que el referente sea extranjero es estadísticamente significativo en la parte μ de la mezcla del modelo beta inflado-truncado, pero no en la regresión MCO. También se observan diferencias para el nivel educativo secundario completo del referente en cuanto a su influencia en la probabilidad de que el hogar sea pobre: es significativo en la regresión logística y en su correspondiente parte del modelo beta inflado-truncado, pero no es significativo cuando se analiza H mediante mínimos cuadrados ordinarios.

7. Consideraciones finales

Este documento busca poner en consideración un conjunto de opciones metodológicas para la estimación microeconómica de los determinantes de la pobreza multidimensional cuando la variable dependiente es una medida multidimensional de pobreza construida mediante el método de Alkire-Foster. A manera de ilustración se emplea la evaluación de impacto multidimensional del programa de vivienda de emergencia de la organización TECHO.

Se presentan un conjunto de métodos alternativos que logran una estimación parcial de algún miembro de la familia de medidas. La estimación mediante mínimos cuadrados sufre de algunas restricciones que pueden superarse tomando los debidos recaudos, aunque el truncamiento que ocurre al estimar A sugiere la existencia de sesgos en la estimación. La regresión logística, por su parte, replica los métodos empleados para la tradicional estimación de la pobreza monetaria, pero permite solamente la estimación de la tasa de incidencia de pobreza multidimensional, dejando sin explicar los factores que influyen sobre la acumulación de privaciones por encima del umbral. Por su parte, el uso del modelo lineal general (GLM) es flexible a la consideración de la intensidad de privaciones pero no permite tomar en cuenta el umbral de pobreza, un factor fundamental constitutivo de la medida Alkire-Foster. La sugerencia, entonces, es la utilización de una *inflated-truncated beta regression*, un modelo de mezclas que permite tomar en cuenta ambos factores.

Este modelo resulta adecuado en razón de la posibilidad de tomar en cuenta tres aspectos de la medida Alkire-Foster. En primer lugar, la posibilidad de estimar de manera conjunta H , A y M_0 , al mismo tiempo en que es posible determinar los efectos diferenciados sobre cada uno. En segundo lugar, la incorporación explícita del umbral de pobreza multidimensional, y de esta manera la modelación de la censura existente para valores inferiores a este umbral y la consecuente inflación en los ceros. Finalmente, la consideración tanto teórica como empírica de la independencia entre los procesos generadores de la pobreza multidimensional, diferenciando la probabilidad de ser considerado multidimensionalmente pobre del hecho de acumular privaciones por encima del umbral.

Las estimaciones realizadas para la aplicación muestran resultados coherentes con la distribución de la variable dependiente y además generan pronósticos aceptables de M_0 . Asimismo, permiten distinguir los efectos separados del tratamiento sobre la dicotomía pobre-no pobre y sobre la cantidad de privaciones por encima del umbral. Los resultados indican que el programa de vivienda de emergencia de TECHO causó una reducción estadísticamente significativa en las tres medidas de Alkire-Foster. La incidencia de la pobreza multidimensional (el índice de recuento) disminuyó cerca de 30 puntos porcentuales: mientras que en el grupo de control un 69% de los hogares están privados en las dimensiones relacionadas a la vivienda, el porcentaje disminuye al 39% para los hogares tratados. Se demostró que un hogar tiene casi cuatro veces menos chances de ser no pobre en el grupo de tratados que en el grupo de control. También se comprobó que recibir la vivienda de TECHO reduce de manera significativa la proporción ponderada de privaciones por arriba del umbral de pobreza. La disminución de la incidencia y de la intensidad de la pobreza en conjunto implican una disminución significativa en el recuento ajustado M_0 .

En esta aplicación, los resultados son robustos ante la utilización de diferentes métodos de estimación, pero en otras posibles aplicaciones en las que el programa que esté siendo evaluado no tenga un efecto tan fuerte, la elección del método de estimación puede alterar los resultados de manera sustantiva. En ese caso, las características e implicancias de cada modelo aquí discutidas son aún más pertinentes.

8. Referencias

- Alkire, S. y Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, 95(7-8), 476-487
- Alkire, S., Foster, J. E., Seth, S., Santos, M. E., Roche, J. M., y Ballon, P. (2015). *Multidimensional Poverty Measurement and Analysis*. Oxford: Oxford University Press.
- Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association* 103(484), 1481-1495.

- Angrist, J. y Pische, J. (2009). *Mostly harmless econometrics. An empiricist's companion*. Princeton: Princeton University Press.
- Bali Swain, R., and Varghese, A. (2009) Does Self Help Group Participation Lead to Asset Creation? *World Development*, 37(10), 1674-1682.
- Beduk, S. (2018). Understanding Material Deprivation for 25 EU Countries: Risk and Level Perspectives, and Distinctiveness of Zeros. *European Sociological Review*, 1-17. doi: 10.1093/esr/jcx089
- Bourguignon, F. y Chakravarty, S. (2003). The measurement of multidimensional poverty, *Journal of Economic Inequality*, 1 (1), 25-49.
- CEPAL (2014). Capítulo 1.b Análisis de la pobreza multidimensional, en Panorama Social de América Latina 2013. URL: https://repositorio.cepal.org/bitstream/handle/11362/35904/1/S2013868_es.pdf
- Coleman, B. (1999). The Impact of Lending in Northeastern Thailand. *Journal of Development Economics*, 60, 105-141.
- Deininger, K., and Liu, Y. (2013). Economic and Social Impacts of an Innovative Self-Help Group Model in India. *World Development*, 43: 149-163.
- Duclos, J.Y. y Araar, A. (2006). Poverty and Equity: Measurement, Policy and Estimation with DAD. Springer, NY. Disponible en URL: <http://www.idrc.ca/openebooks/229-5/>.
- Ferrari, S. y Cribari-Neto, F. (2004). Beta Regression for Modelling Rates and Proportions, *Journal of Applied Statistics*, 31:7, 799-815, DOI: 10.1080/0266476042000214501
- Foster, J., Greer, J. y Thorbecke, E. (1984). A Class of Decomposable Poverty Measures. *Econometrica*, 52(3): 761-766.
- Glewwe, Paul. (1990). Investigating the Determinants of Household Welfare in Côte d'Ivoire in 1985." Living Standards Measurement Study Working Paper No. 29, World Bank, Washington, DC.
- Gray, L. y Hernandez Alava, M. (2018). A command for fitting mixture regression models for bounded dependent variables using the beta distribution. *The Stata journal*, 18 (1), 51-75. ISSN 1536-867X
- Haughton, D., Haughton, J., Bales, S., Kim Chuyen, T. y Nga, N. (1999). Health and Wealth in Vietnam: An Analysis of Household Living Standards. Singapore: Institute of Southeast Asian Studies.
- Haughton, J. y Khandker, S. (2009). *Handbook on Poverty and Inequality*. Washington, DC: World Bank.
- INEC (2015). Índice de pobreza multidimensional: metodología / Instituto Nacional de Estadística y Censos, San José de Costa Rica, C.R.: INEC.
- Kling, J., Liebman, J. y Katz, L. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, 75(1), 83-119
- Ministerio de Desarrollo Social (2015). Informe de Desarrollo Social del Ministerio de Desarrollo Social, Chile. Disponible en URL: http://www.ministeriodesarrollosocial.gob.cl/pdf/upload/IDS_INAL_FCM_3.pdf
- MEF, MDF e INEC (2017). Índice de Pobreza Multidimensional de Panamá: Año 2017.
- Mitchell, A., Macció, J. y Mariño Fages, D. (2019) The Effects of Emergency Housing on Wellbeing: Evidence from Argentina's Informal Settlements, *European Journal of Development Research*, 31(3), 504–529. <https://doi.org/10.1057/s41287-018-0166-z>

- Mitchell, A. y Macció, J. (2018). ‘Evaluating the effects of housing interventions on multidimensional poverty: the case of TECHO-Argentina’, OPHI Working Paper 120, University of Oxford. Disponible en URL: <https://ophi.org.uk/evaluating-the-effects-of-housing-interventions-on-multidimensional-poverty-the-case-of-techo-argentina/>
- Pereira, G. H., Botter, D. A., y Sandoval, M. C. (2013). A regression model for special proportions. *Statistical Modelling*, 13(2), 125–151. <https://doi.org/10.1177/1471082X13478274>
- PNUD y OPHI (2019). Global Multidimensional Poverty Index 2019. illuminating inequalities. Disponible en URL: https://ophi.org.uk/wp-content/uploads/G-MPI_Report_2019_PDF.pdf
- Ramalho, J. y Vidiyal Da Silva, J. (2009). A two-part fractional regression model for the financial leverage decisions of micro, small, medium and large firms. *Quantitative Finance* 9(5), 621-636.
- Santos, M.E. (2010). Material de trabajo del OPHI’s Summer School on Capability and Multidimensional Poverty, Amman, Jordania. Disponible en URL: http://www.ophi.org.uk/ophi_events/2010-summer-school/
- Schwiebert, J. y Wagner, J. (2015). A Generalized Two-Part Model for Fractional Response Variables with Excess Zeros, Leuphana University Luneburg. URL: <https://www.econstor.eu/handle/10419/113059>
- Smithson, M. y Verkuilen, J. (2006). A Better Lemon Squeezer? Maximum-Likelihood Regression With Beta-Distributed Dependent Variables. *Psychological Methods* 11(1), 54-71.
- STPP y MINEC-DIGESTYC (2015). *Medición multidimensional de la pobreza*. El Salvador. San Salvador: Secretaría Técnica y de Planificación de la Presidencia y Ministerio de Economía, a través de la Dirección General de Estadística y Censos.
- Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge MA: MIT Press.