

Nowcasting de pobreza en Costa Rica:
Metodología para estimar y pronosticar pobreza con
mayor periodicidad

Instituto de Investigaciones en Ciencias Económicas

Facultad de Ciencias Económicas

Universidad de Costa Rica

Autores:

Esteban Huang

Luis Oviedo

Saúl Paredes

Juan Robalino

Resumen

En Costa Rica, la publicación de datos de pobreza se realiza de forma anual, lo que hace que no sean tan actualizados ni periódicos como lo necesitan tomadores de decisiones e investigadores. Esta investigación utiliza el *nowcasting* para estimar la tasa de pobreza en forma mensual para Costa Rica durante junio 2020 y junio 2022. Esta metodología busca estimar variables en el presente y futuro cercano. El modelo propuesto muestra diferencias no significativas entre las estimaciones y los datos oficiales para la mayoría de estimaciones. Esta investigación hace tres contribuciones importantes. Primero, la variación significativa en los datos mensuales contribuye significativamente para la toma de decisiones, pues la naturaleza del dato anual oculta estas variaciones. Segundo, la publicación mensual disminuye el rezago en la información y, tercero, los resultados del modelo demuestran la importancia de la dinamicidad del mercado laboral para generar cambios en la pobreza.

Palabras clave

Ingresos, mercado laboral, pobreza, simulación, bienestar, pronóstico

Clasificación JEL

C23, C53, I3

1 Introducción

En Costa Rica, la publicación de datos de pobreza se hace a partir de la información recolectada en la Encuesta Nacional de Hogares (ENAH) por el Instituto Nacional de Estadística y Censos (INEC) que se implementa solamente una vez al año. Esto hace que las estadísticas sobre pobreza no sean tan actualizadas ni tan periódicas como lo necesitan tomadores de decisiones e investigadores. Aumentar la periodicidad y reducir el rezago entre un dato de pobreza y el otro hace posible conocer en forma más oportuna la situación económica del grupo con más desventaja de la sociedad y permite tomar decisiones informadas en el corto plazo.

Para resolver este problema, en esta investigación se plantea una metodología para estimar la tasa de pobreza en forma mensual para Costa Rica durante el periodo comprendido entre junio 2020 y junio 2022. Para esto se utiliza una metodología llamada *nowcasting* que tiene como propósito estimar el valor de variables estratégicas en el presente, el futuro cercano e incluso el pasado reciente. Esta metodología se ha utilizado cuando las estimaciones oficiales de alguna estadística están disponibles con un rezago considerable, y, a su vez, existe una mayor regularidad en la publicación de variables relacionadas.

Para este ejercicio, se utiliza la muestra anual de la ENAH, y se modela la evolución de los ingresos de los hogares con parámetros estimados que describen la situación del mercado laboral en forma mensual (como resultado de la muestra de tres meses móviles) a partir de las Encuestas Continuas de Empleo (ECE). Específicamente, se generan estimaciones de ingresos de los hogares de la ENAH tomando en cuenta los cambios en las contribuciones de las características sociodemográficas de la persona, en ingresos y en la probabilidad de estar ocupado, con datos de la ECE. Luego, se actualiza la línea de pobreza, que considera la variación en los precios, y se encuentra el porcentaje de hogares en pobreza y pobreza extrema.

Combinar información de dos encuestas presenta sus retos. Aunque las preguntas de la ENAH y la ECE son idénticas en muchos casos, debido a las diferencias en la implementación, es conocido que generan resultados diferentes. Afortunadamente, durante junio de cada año las dos encuestas se aplican simultáneamente, lo que permite compararlas y estimar las diferencias que tienen que ver con la forma en que se implementan. Luego, se utilizan estas estimaciones para corregir las

diferencias entre la ENAHO de junio y la ECE de cualquier otro mes para poder obtener solo los cambios que tienen que ver con la evolución del mercado laboral en el tiempo.

Al examinar las diferencias entre el pronóstico de este modelo y los datos observados en la ENAHO, se encuentra que éstas están dentro del intervalo de confianza para la mayoría de estimaciones. Para junio del 2021, la tasa de pobreza pronosticada con el modelo fue de 23,09% con intervalo de confianza al 95% que va de 22,05% a 24,12%, mientras que el dato observado en la ENAHO fue de 23,02%. La tasa de pobreza extrema pronosticada fue de 6,28% con intervalo de confianza que va desde 5,67% a 6,88%, mientras que el dato observado en la ENAHO fue de 6,27% para el mismo periodo. Por otro lado, para junio de 2022, la tasa de pobreza total y extrema pronosticada es de 23,49% y 6,96%, respectivamente, con intervalos de confianza que van desde 22,62% a 24,36% para el caso de pobreza total y 6,46% a 7,45% para el caso de pobreza extrema. Asimismo, las tasas oficiales de pobreza total y extrema para ese periodo corresponden a 23,04% y 6,40%, respectivamente. Lo anterior muestra que las diferencias puntuales son siempre menores al 10%.

Este documento hace tres contribuciones importantes. Primero, los resultados muestran que las tasas de pobreza pueden variar significativamente durante un año. Por ejemplo, la estimación de pobreza de enero del 2022 es estadísticamente más alta que la estimación de pobreza en marzo del mismo año. Esto demuestra que tener el modelo puede contribuir significativamente para la toma de decisiones, pues el dato anual oculta estas variaciones.

Segundo, los datos de pobreza de la ENAHO implementada en junio son publicados hasta octubre de cada año. Por lo tanto, no hay información de pobreza por un periodo prolongado. Por ejemplo, para septiembre del 2022 el último dato que se conoce sobre pobreza es el de junio del 2021, lo que representa un rezago de 15 meses. Con el modelo propuesto, el dato de pobreza estaría rezagado por un máximo de tres meses.

Tercero, el modelo permite demostrar que el mercado laboral es suficientemente dinámico para generar cambios en los niveles de pobreza en poco tiempo. A pesar de que los precios, a través del cambio en la línea de pobreza son relevantes, no se puede depender solamente de ellos para pronosticar los cambios en los niveles de pobreza.

Sin embargo, este modelo no toma en cuenta la variabilidad de ingresos ajenos a la actividad principal, tales como ingreso por empleo secundario, ingresos de capital, pensiones contributivas, transferencias privadas y estatales a través del año. Esto es relevante ya que una gran parte del ingreso depende de ayudas sociales, y sobre todo para los individuos con ingresos familiares cercanos a la línea de pobreza. Sería relevante para próximas investigaciones poder incluir la modelación de esta información.

Otro reto importante es que, aunque este estudio hace el esfuerzo por estimar las diferencias en las tomas de la muestra entre encuestas, es importante notar que implícitamente se asume que esta diferencia es estática en el tiempo, que el sesgo estimado para junio es el mismo para el resto de meses durante el año. Este supuesto podría no cumplirse en algunos meses, sin embargo, es un reto difícil de resolver hasta que haya datos de la ENAHO en otro mes.

En la siguiente sección se desarrollan los antecedentes relacionados al uso del *nowcasting*. En la tercera sección se describen los datos utilizados para el análisis. En la cuarta sección se describe la metodología utilizada para las estimaciones. La quinta sección muestra los resultados. En la última sección, se concluye y discuten las limitaciones encontradas.

2 Literatura relacionada a la técnica del *nowcasting*

El uso del *nowcasting* fue impulsado principalmente, con el fin de medir el PIB per cápita y su crecimiento en tiempo real a partir de variables macroeconómicas disponibles con mayor periodicidad (Banbura & Runstler (2007), Angelini et al. (2008) y Barhoumi et al. (2008)).

Posteriormente, la técnica se utilizó combinando datos, tanto macroeconómicos como microeconómicos, para pronosticar otras variables. Por ejemplo, en India, se utilizaron tasas de traspaso estimadas del gasto final en consumo agregado de los hogares (componente de cuentas nacionales) al gasto en consumo en encuestas a nivel de hogar para calcular las nuevas tasas de pobreza según la línea de pobreza internacional (Newhouse & Vyas, 2018). Asimismo, en Bangladés, se combinaron proyecciones macroeconómicas, tales como el crecimiento poblacional, el cambio en la producción o el movimiento de las remesas con microdatos relacionados a la fuerza laboral para generar distintos escenarios que permiten evaluar, ex ante, los impactos sobre los ingresos laborales y no laborales de crisis económicas (Habib et. al., 2010).

El método también se ha utilizado en Uruguay para estimar el efecto redistributivo del COVID-19 a través de una microsimulación del cambio en el ingreso para el sector formal, el efecto de las transferencias monetarias impulsadas por el COVID-19, y la reducción del ingreso de los trabajadores informales e independientes. El efecto es simulado de acuerdo a características del hogar, así como de estadísticas macroeconómicas tales como contracción del PIB y de seguridad social; y haciendo uso de modelos probabilísticos que permiten asignar las caídas del ingreso agregado a los individuos (Braun & De Rosa, 2021).

Una gran ventaja de la técnica es que permite explotar la relación entre variables para predicción. En Kenia, por ejemplo, se usaron diferentes activos del hogar, medidos en encuestas demográficas y de salud, para estimar el consumo de los hogares cuando no está disponible de manera oficial. Para ello, se utiliza un modelo estimado en periodos en donde la información de activos y de consumo se encontraba disponible simultáneamente. Con la información de consumo, así como de la aplicación del modelo en diversas ocasiones, Stifel & Christiaensen (2006) son capaces de generar una distribución de bienestar (medido como pobreza o inequidad de los ingresos).

De forma similar, en Bolivia, se realizó una microsimulación dinámica para estimar la pobreza en diferentes periodos en el tiempo con representación nacional, usando una encuesta demográfica y de salud en conjunto con una encuesta de estándares de vida (Grosse et al., 2009). A diferencia de Stifel & Christiaensen (2006), las estimaciones realizadas para Bolivia permitieron hacer pronósticos de carácter regional de forma transversal en lugar de pronósticos en el tiempo, lo que muestra la versatilidad de la técnica en cuanto a objetivos.

Al simular las tasas de pobreza e ingreso, es común el uso de información del mercado laboral. Tal es el caso en Marruecos, donde se estima la pobreza para un periodo más amplio que el disponible de forma oficial, usando de métodos de imputación. Para esto, se estima un modelo del gasto per cápita en el hogar usando información disponible tanto en la Encuesta de Gastos de los Hogares como en la Encuesta de Fuerza laboral. Ambas encuestas contienen información sobre los regresores, sin embargo, solo la primera posee datos sobre el gasto para los años seleccionados por los autores (Doudich et. al., 2015).

En la misma línea, se estima el crecimiento del ingreso de los hogares en Argentina. Esto se hace con microdatos y asumiendo que el crecimiento del ingreso por trabajo es igual al crecimiento agregado del ingreso promedio del decil al que pertenece el hogar (Arakaki et al., 2020).

Esta investigación se asemeja en cuanto su metodología a las realizadas por Arakaki et al. (2020) y Doudich et. al. (2015), puesto que se hace uso de encuestas de fuerza laboral para realizar la actualización de las medidas de pobreza contenidas en encuestas de hogares. Mientras que, al tratar de conocer los resultados de políticas anticíclicas como las transferencias por la pandemia de COVID-19, se asemeja a la investigación realizada por Habib et. al. (2010) y Braum & De Rosa (2021).

3 Datos de pobreza en Costa Rica

En Costa Rica, los niveles de pobreza se miden con la ENAHO que es implementada por el INEC. Esto debido a que es la única encuesta que genera información, no solo sobre la situación laboral de cada uno de los miembros del hogar, sino también de los ingresos laborales y no laborales como las transferencias sociales que hace el gobierno. Además, es una encuesta representativa a nivel nacional, regional, por zona y por decil de ingreso. Estas encuestas se han implementado desde el 1976 y se han ido modificando, y cambiado su nombre.

Sin embargo, las ENAHO se implementan una vez al año en el mes de julio, momento en el cual se hacen preguntas sobre lo que ocurrió en junio. Normalmente, las estadísticas de pobreza son publicadas en el mes de octubre, luego de que las encuestas son procesadas por el INEC. Esto genera dos problemas sustanciales en términos de información.

Primero, la periodicidad del dato de pobreza es de un año. No hay información sobre cómo cambia la pobreza en los diferentes momentos del año. Las personas tomadoras de decisiones no tienen información sobre los niveles de fluctuación y si son necesarias o no medidas remediales más fuertes en ciertos meses del año. Así, se toman decisiones solamente con los datos sobre el mes de junio de cada año.

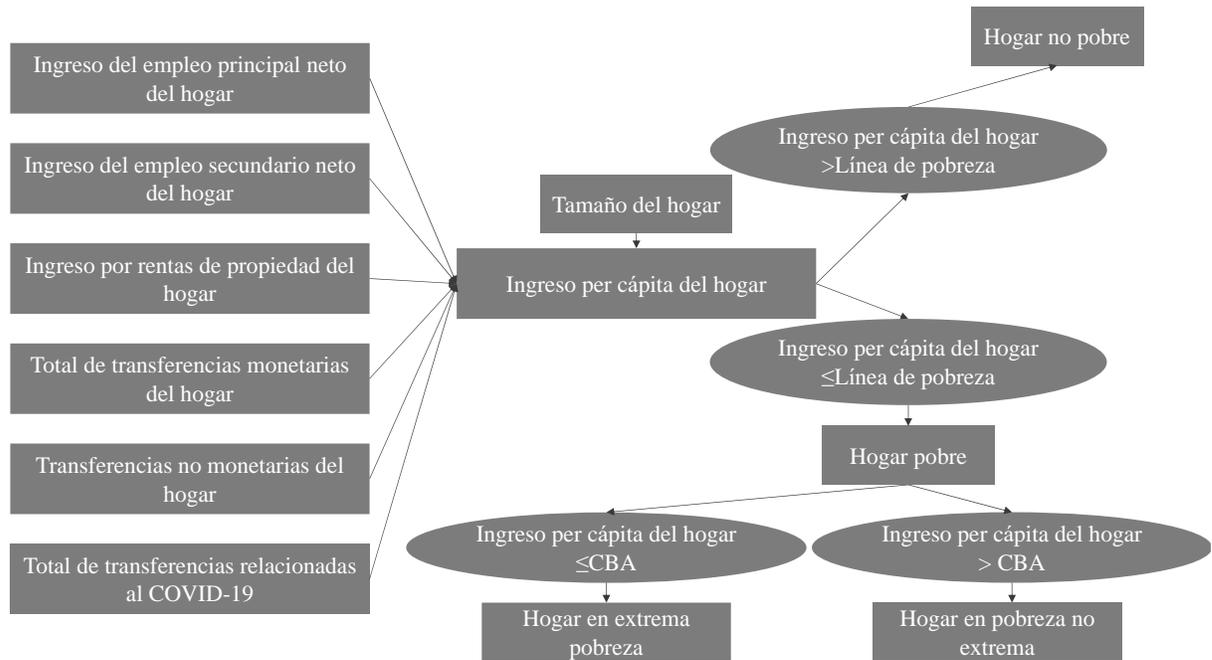
Segundo, el nivel de rezago de la información de pobreza puede llegar a ser importante. Dado que el reporte de pobreza del mes de junio se hace público en octubre, el menor rezago entre el dato y el reporte es de cuatro meses. Luego, el rezago empieza a crecer por 11 meses más. Esto hace que

para septiembre del siguiente año el nivel de rezago sea de 15 meses. Esto implica que, por ejemplo, algunas decisiones presupuestarias, o bien, decisiones de políticas relacionadas a la pobreza tengan que tomarse con información desactualizada.

Para definir la tasa de pobreza se usa el ingreso per cápita del hogar¹ (*ipcn*). Este se compone de la suma de todos los ingresos de los miembros del hogar: ingreso del empleo principal neto del hogar (*ipnh*), ingreso del empleo secundario neto del hogar (*isnh*), ingreso por rentas de propiedad del hogar (*irph*), total de transferencias monetarias del hogar (*ttmh*), transferencias no monetarias del hogar (*tnmh*) y para el caso único del 2020, transferencias relacionadas al COVID-19 (*COVIDtth*). La suma resultante se divide entre la cantidad de miembros del hogar (*TamHog*). Posteriormente, el ingreso per cápita del hogar se compara con una línea de pobreza (LP), calculada a partir del valor de una cesta compuesta por bienes y servicios, para el caso de la línea de pobreza total y, únicamente por alimentos (CBA), para el caso de la línea de pobreza extrema o línea de indigencia. El hogar se cataloga como pobre si el ingreso per cápita del hogar es inferior o igual a la línea de pobreza total y como hogar en extrema pobreza si el ingreso per cápita del hogar es inferior o igual a la línea de pobreza extrema. En caso contrario, se categoriza como no pobre. Este proceso se describe en la Figura 1.

¹ Los componentes del ingreso per cápita del hogar, tales como ingresos de empleo secundario, renta por propiedad, transferencias monetarias y transferencia no monetarias se expresan en la nomenclatura utilizada por el INEC.

Figura 1. Diagrama de las etapas seguidas para la estimación de las tasas de pobreza



Fuente: Elaboración propia con base en la ENAHO del INEC.

El peso de los componentes que determinan el ingreso per cápita del hogar se detallan en la Tabla 1. El componente más importante, tanto en hogares pobres como no pobres, corresponde al ingreso del empleo principal neto, mientras que el menos relevante corresponde al ingreso del empleo secundario neto del hogar.

Tabla 1. Participación promedio de los componentes en el ingreso total del hogar neto (ENAHO 2020)

Componente	Hogares pobres	Hogares no pobres
Ingreso del empleo principal neto del hogar	38,36%	67,48%
Ingreso del empleo secundario neto del hogar	0,24%	0,57%
Ingreso por rentas de propiedad del hogar	0,64%	3,23%
Total de transferencias monetarias del hogar	41,44%	21,40%
Transferencias no monetarias del hogar	4,70%	1,77%
Total de transferencias relacionadas al COVID-19	14,61%	5,54%
Cantidad de hogares	406.265 (26,16%)	1.184.922 (73,84%)

Fuente: Elaboración propia con base en la ENAHO del INEC.

4 Metodología

Para llevar a cabo este análisis se utiliza las ENAHO implementadas en el 2020, 2021 y 2022. Esta encuesta contiene información sobre características a nivel de hogar e individual, así como ingresos laborales y no laborales para los meses de junio de cada año. La encuesta pregunta sobre todas las fuentes de ingreso del hogar, lo que permite hacer la estimación de pobreza y pobreza extrema para junio de cada año.

Los datos que permiten ampliar la periodicidad de las estimaciones de pobreza se encuentran en la ECE. Esta encuesta se aplica de forma continua en todos los meses, pero se publica de forma trimestral. Brinda información sobre el estado del mercado laboral en el país, así como características del individuo y del hogar que están relacionadas a la condición de ocupación y a los ingresos por trabajo. Sin embargo, carece de información suficiente para la estimación de la pobreza, puesto que no contiene información sobre ingresos no laborales del hogar. A pesar de que su publicación se hace de manera trimestral, esta encuesta es un agregado de datos mensuales, lo cual permite utilizar datos mensuales, resultantes de trimestres móviles, de junio de 2020 hasta junio de 2022.

Combinando las dos encuestas se puede aumentar la periodicidad de las estimaciones de pobreza. Se modela el ingreso del hogar por ocupación principal proveniente del mercado laboral con la ECE para algún mes en particular utilizando el agregado trimestral de los dos meses anteriores y el mes en cuestión², y, con las otras fuentes de ingreso de la ENAHO (que se asumen constantes), se hace una estimación del ingreso total del hogar. Una vez que se obtienen las estimaciones de los ingresos mensuales del hogar para la muestra de la ENAHO, se utilizan las líneas de pobreza mensuales para definir los porcentajes de hogares en pobreza y pobreza extrema.

Para poder implementar este proceso de estimación se requiere que (1) las dos encuestas tengan las mismas variables que determinan los ingresos del hogar en el mercado laboral y (2) que haya un momento donde las dos encuestas sean implementadas simultáneamente. Lo primero asegura que se pueda predecir adecuadamente los cambios en los ingresos por trabajo en un hogar entre un

² Este ejercicio supone que el trimestre móvil es representativo del último mes que lo compone.

mes y otro. Lo segundo permite eliminar las diferencias entre las dos encuestas se deben a factores como las diferencias en la estrategia de muestreo y de implementación de la encuesta.

4.1 Modelo

El objetivo del *nowcasting* es generar la estimación de alguna variable no disponible en una dimensión (temporal o espacial) explotando la relación previamente establecida con otras medidas o variables en un contexto en donde ambos grupos de variables se encuentren disponibles simultáneamente. La simultaneidad permite determinar modelos que posteriormente son extrapolados a una dimensión en donde solo uno de los grupos de variables se encuentra disponible, para estimar el valor de la variable faltante. En este caso, el objetivo es estimar el parámetro θ_{t+1}^g en el periodo t de la fuente g cuando hay disponibilidad de información en el periodo anterior θ_t^g y con información de otra fuente de g' sobre el parámetro $\theta_{t+1}^{g'}$ y $\theta_t^{g'}$. Dado lo anterior, es posible encontrar el valor en $t + 1$ de la encuesta g , mediante el siguiente proceso interno $\Delta\theta_{t+1}^g = \theta_{t+1}^g - \theta_t^g$.

Entonces, definiendo $g = ENAHO$ y $g' = ECE$, para cualquier t , la relación entre encuestas puede escribirse como $\theta_t^{ENAHO} = \theta_t^{ECE} + d$, donde d es la diferencia entre las encuestas, que se asume invariante en el tiempo. Con este supuesto y partiendo de que $\theta_{t+1}^{ENAHO} = \theta_{t+1}^{ECE} + d$ y $\theta_t^{ENAHO} = \theta_t^{ECE} + d$, se obtiene que $\Delta\theta_{t+1}^{ENAHO} = \Delta\theta_{t+1}^{ECE}$. Esto implica que así se puede estimar θ_{t+1}^{ENAHO} a partir de:

$$\theta_{t+1}^{ENAHO} = \theta_t^{ENAHO} + \Delta\theta_{t+1}^{ECE} \quad (1)$$

Mediante este método se estima el ingreso de las personas que se encuentran laborando, las probabilidades de estar ocupados y los niveles de ocupación en la economía. Con esos tres factores se actualizan los ingresos por hogares de la muestra y se calcula los niveles de pobreza.

4.2 Ingresos de las personas trabajadoras

Para estimar los ingresos por el concepto de mercado laboral se utiliza la siguiente ecuación:

$$y_{it}^{ENAHO} = \beta_t^{ENAHO} X_{it}^{ENAHO} + \mu_{it}^{ENAHO} \quad (2)$$

Donde y_{it}^{ENAHO} corresponde al logaritmo del ingreso del empleo principal neto de la persona ocupada i , los cuales se pueden descomponer en factores observables en la encuesta $\beta_t^{ENAHO} X_{it}^{ENAHO}$ y en factores no observables μ_{it}^{ENAHO} . El vector de factores observables está compuesto por las características de los individuos de la ENAHO, X_{it}^{ENAHO} , y el vector de parámetros β_t^{ENAHO} . Las características de las personas incluidas en el análisis son: sexo, condición de jefatura de hogar, edad, región, país de nacimiento, nivel de instrucción³, condición de empleado público, rama de actividad en la que se desempeña y formalidad en el empleo principal.

De tal forma que siguiendo el proceso explicado en la sección 4.1 se puede estimar β_{t+1}^{ENAHO} de la siguiente manera:

$$\beta_{t+1}^{ENAHO} = \beta_t^{ENAHO} + \Delta\beta_{t+1}^{ECE}$$

donde, para estimar β_t^{ENAHO} , se utiliza la ecuación (2) y para estimar $\Delta\beta_{t+1}^{ECE}$ se utiliza la siguiente especificación:

$$y_{it,t+n}^{ECE} = \beta_t^{ECE} X_{it}^{ECE} + \Delta\beta_{t+n}^{ECE} X_{it+n}^{ECE} I_{it+n} + \mu_{it,t+n}^{ECE} \quad (3)$$

donde I_{it+n} corresponde a una variable dicotómica con valor de 1 si la observación i corresponde al periodo $t + n$ o 0 si corresponde al periodo t , entendiendo t como el agregado móvil de tres meses (t junto a los dos meses inmediatamente anteriores a este). Asimismo, el vector $\Delta\beta_{t+n}^{ECE}$ contiene la diferencia temporal entre β_t^{ECE} y β_{t+n}^{ECE} , vinculada a las características individuales.

Posteriormente, y asumiendo que la diferencia entre la recolección y procesamiento de los datos entre ambas encuestas es constante como se mencionó en la sección 4.1, es posible proyectar el ingreso de las personas para cada observación i usando la ecuación (4):

$$y_{it+n}^{ENAHO} = (\beta_t^{ENAHO} + \Delta\beta_{t+n}^{ECE}) X_{it}^{ENAHO} + \mu_{it}^{ENAHO}. \quad (4)$$

³ Agrupado en menos de primaria, primaria completa, secundaria completa o universidad.

Las ecuaciones (2), (3) y (4) se implementan por separado para personas de la zona rural y de la zona urbana, y para aquellos en cuenta propia, por un lado, y, por el otro lado, aquellos que son asalariados y empleadores.

4.3 Número de personas ocupadas

Para definir el cambio en el número de ocupados, se sigue la estrategia explicada en la sección 4.1. Se calcula el número de ocupados por zona y por nivel de educación⁴. Luego, se calcula los cambios en la cantidad de ocupados con las ECE y se determina cuántas personas más deberían de entrar o salir del mercado laboral en la ENAHO.

Una vez obtenidos los cambios absolutos en cada uno de estos estratos, es necesario escoger cuáles individuos se excluyen o incluyen de la masa de ocupados. Para ello, se asigna una probabilidad de ocupación para individuo en la muestra, con el fin de parear los cambios en la masa de ocupados de forma absoluta.

Para lo anterior, se utiliza una relación inicial de variables independientes y ocupación, y bajo el supuesto de invariabilidad de la diferencia muestral, se estiman los cambios temporales en las relaciones de variables y la ocupación con un modelo similar al de la sección 4.2. Una vez estimados estos dos componentes, se asigna una probabilidad a cada individuo considerando los impactos temporales. En este caso, la matriz X está conformada por el sexo, condición de jefatura de hogar, edad, región y país de nacimiento.

Una vez obtenidas las probabilidades para el periodo $t + n$, se procede a incluir o excluir el número de individuos equivalente al cambio en el número de ocupados en cada estrato escogiéndolos según su probabilidad de estar ocupado. Es decir, si el cambio en el número de ocupados para un estrato en particular fue positivo, se seleccionan aquellos individuos que no estaban ocupados y que tengan la probabilidad más alta de estar ocupados de acuerdo con la ecuación generada para el periodo $t + n$. Si, por el contrario, el cambio en el número de ocupados en un estrato en particular

⁴ Los niveles que se utilizan son los siguientes: menos de primaria, con primaria y con secundaria o más.

es negativo, se procede a elegir aquellos individuos cuya condición es ocupado y posean la menor probabilidad de estar ocupados.

Siguiendo la misma línea, si hay un cambio positivo en alguno de los estratos, es necesario proyectar un ingreso por trabajo principal a aquellos que engrosan el número de ocupados. Por lo que se realiza el procedimiento de la sección 4.2 excluyendo las variables que no están disponibles para individuos no ocupados, es decir, se estima la relación inicial de variables, la corrección temporal y la asignación de ingreso proyectado incluyendo, únicamente, sexo, condición de jefe de hogar, edad, región, país de nacimiento y título educativo.

4.4 Ayudas especiales por la pandemia

Para el año 2020 los ingresos adicionales que generó el gobierno fueron de suma importancia, sobre todo para aquellos con ingresos bajos. Por eso se modela la asignación del Bono Proteger, el cual se otorgó en dos opciones, un bono de 65 mil colones y otro de 125 mil colones, dependiendo de las condiciones del beneficiario. A partir de julio de 2020, los informes oficiales del Ministerio de Trabajo y Seguridad Social muestran una disminución consistente en la entrega de ambos bonos. Entonces, con el fin de ajustar por estos cambios en los datos oficiales, se generó una regresión para obtener la probabilidad de recibir el bono y se excluyen del beneficio aquellos con la mayor probabilidad de recibirlo, pues son estos los que, potencialmente, verían acabado el periodo beneficio con mayor prontitud. Este ejercicio se realiza hasta que el número relativo de excluidos en las bases de datos sea equivalente al número relativo de excluidos en los datos oficiales.

4.5 Determinación de pobreza

Una vez concluidas las estimaciones descritas, se agregan todos los ingresos por empleo principal para obtener el ingreso del empleo principal neto del hogar proyectado ($ipnh_{t+n}$) y, descartando transferencias privadas por COVID-19 e incluyendo los Bonos Proteger para aquellos que continúan recibéndolo, se obtiene el total de transferencias relacionadas al COVID-19 proyectado ($COVIDtth_{t+n}$). Adicionalmente, para calcular el nuevo ingreso per cápita del hogar ($ipcn_{t+n}$) se hace uso de la ecuación (5), bajo el supuesto de que las demás fuentes de ingreso, ingreso del empleo secundario neto del hogar ($isnh$), ingreso por rentas de propiedad del hogar ($irph$), total de transferencias monetarias del hogar (tth) y transferencias no monetarias del hogar ($tnmh$),

así como el tamaño del hogar ($TamHog$) permanecen constantes al menos durante un año. Esto es posible puesto que las fuentes de ingreso modeladas (en conjunto) representan la mayor proporción del ingreso total del hogar neto (ver Tabla 1).

$$\widehat{ipcn}_{t+n} = \frac{\widehat{ipnh}_{t+n} + isnh_t + irph_t + ttmh_t + tnmh_t + \widehat{COVIDtth}_{t+n}}{TamHog} \quad (5)$$

Adicionalmente, se considera los cálculos mensuales del INEC sobre el cambio en los precios de las cestas de bienes y servicios que determinan las líneas de pobreza para las zonas rural y urbano. Con las cestas actualizadas que definen las líneas de pobreza y el ingreso per cápita del hogar se calcula la proporción de hogares por debajo de estas.

Como producto de las etapas anteriores, se obtienen estimaciones puntuales de las tasas de pobreza para cada uno de los meses considerados en la investigación. No obstante, resulta importante proveer un intervalo de confianza para cada una de esas estimaciones puntuales, puesto que permite brindar un rango de la estimación puntual. Para lo anterior, se utilizan simulaciones de Montecarlo basadas en la técnica de *bootstrapping*. Se toman 1.000 muestras aleatorias con reemplazo de los hogares en la ENAHO, cada una de estas muestras contiene el mismo número de hogares de la ENAHO, con la diferencia de que algunas observaciones de la muestra se encuentran más de una sola vez y otras son excluidas totalmente. Una vez obtenidas, se procede a calcular las tasas de pobreza para cada una de las 1.000 muestras. Posteriormente, se estima un intervalo de confianza al 95% tomando en cuenta la mitad de la diferencia entre la estimación puntual de los percentiles 97,5 y 2,5. Este proceso se realiza para cada tasa de pobreza mensual estimada.

5 Resultados

En esta sección se presentan los resultados de la metodología planteada en el apartado anterior. En el apartado 5.1, se muestra el promedio de personas ocupadas por hogar en percentiles seleccionados basándose en la metodología desarrollada en la sección 4.3, manteniendo el ingreso constante al momento de determinar los deciles. Mientras que en la sección 5.2, se discuten los cambios en los niveles de ingreso de los hogares, siguiendo la metodología descrita en la sección 4.2 e incorporando los cambios en la ocupación. Por otro lado, en la sección 5.3, se describe la evolución a través del tiempo de la línea de pobreza, el valor de la canasta que determina la condición de pobreza de un hogar, según se describe en la Figura 1. Por último, en el apartado 5.4,

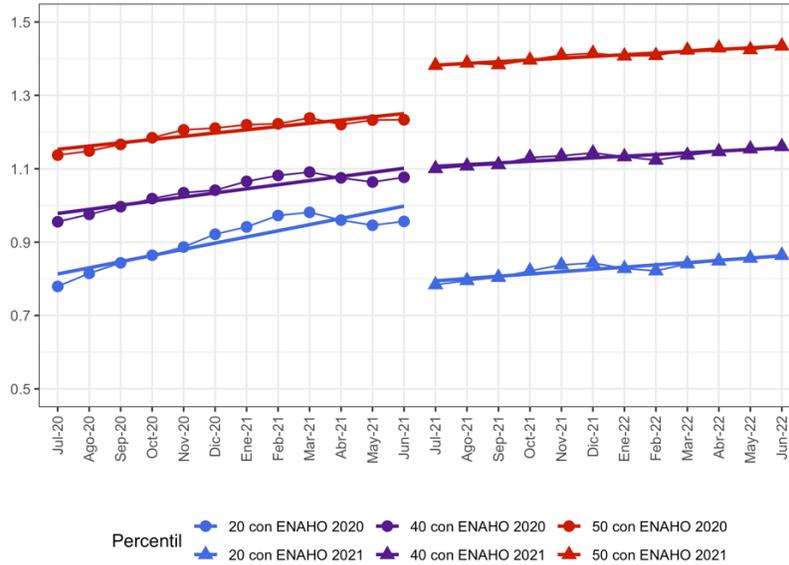
se incluyen todos estos aspectos para describir las tasas de pobreza simuladas y sus intervalos de confianza; considerando cambios en ingreso principal por trabajo, condición de ocupación y transferencias relacionadas al COVID-19.

5.1 Cambios en ocupación

En la Figura 2 se encuentra el promedio de personas ocupadas en los hogares que pertenecen al percentil 20, 40 y 50, desde julio de 2020 a junio de 2022, las series se presentan en dos periodos distintos, antes y después de junio de 2021, ya que posterior a dicha fecha la ENAHO presenta información actualizada. El promedio de ocupados por hogar fue calculado utilizando la metodología desarrollada en la sección 4.3, no obstante, la determinación de los deciles obedece al dato oficial del ingreso total del hogar neto de la ENAHO para junio de cada año que corresponda. Es decir, se muestra únicamente, el resultado de los movimientos de la masa de ocupados de forma desagregada manteniendo constante el decil de ingreso que se tenía al principio del periodo. Las series para la fecha posterior a junio de 2021 muestran un cambio de nivel, esto obedece a que el cambio en ocupación es simulado a partir de un modelo aplicado a submuestras resultantes de la intersección de la zona y nivel de educación de las observaciones. Por otro lado, si la estimación se realizara a nivel agregado, estas diferencias se reducirían a la vez que se pierde precisión en la simulación.

En los dos periodos de forma separada, se observan tendencias consistentemente crecientes para los tres diferentes percentiles elegidos, lo que quiere decir que el número de ocupados por hogar ha ido en aumento, al menos para los percentiles seleccionados. Por otro lado, el crecimiento promedio es mayor para el percentil de ingreso más bajo, en comparación con los percentiles 40 y 50, indicando una mayor incursión, en promedio, en la masa de ocupados por parte del grupo con menos ingresos en los últimos dos años, esto sucede al comparar los percentiles en el mismo periodo.

Figura 2. Evolución mensual del promedio de ocupados por hogar según percentiles de ingreso seleccionados

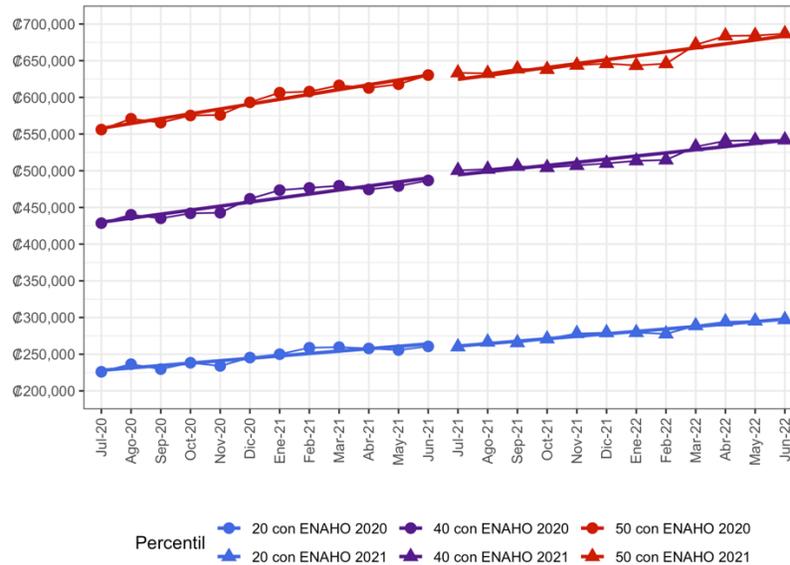


Fuente: Elaboración propia con base en la ENAHO y ECE del INEC.

5.2 Cambios en ingreso

En la Figura 3 se muestra el ingreso total del hogar neto para el hogar en el percentil 20, 40 y 50 desde julio de 2020 a junio de 2022, las series se presentan en dos periodos distintos, antes y después de junio de 2021, ya que posterior a dicha fecha la ENAHO presenta información actualizada. Estos datos fueron calculados utilizando la metodología desarrollada en la sección 4.2. Se incluyen los cambios estimados en el número de ocupados por hogar.

Figura 3. Evolución mensual del ingreso total del hogar neto para el hogar en el percentil de ingreso seleccionado



Fuente: Elaboración propia con base en la ENAHO y ECE del INEC.

Como es de esperar, los montos del ingreso del hogar para los hogares ubicados en los percentiles seleccionados, son consistentemente mayores conforme aumenta el percentil de ingreso analizado. Asimismo, la tasa de crecimiento es mayor para los percentiles mayores, lo anterior es consistente a lo que sucede con el número de ocupados por hogar, al menos para el percentil 40 y 50. No obstante, a pesar de mostrar tendencia decreciente en ocupación, el percentil 20 muestra una tendencia creciente en el ingreso del hogar.

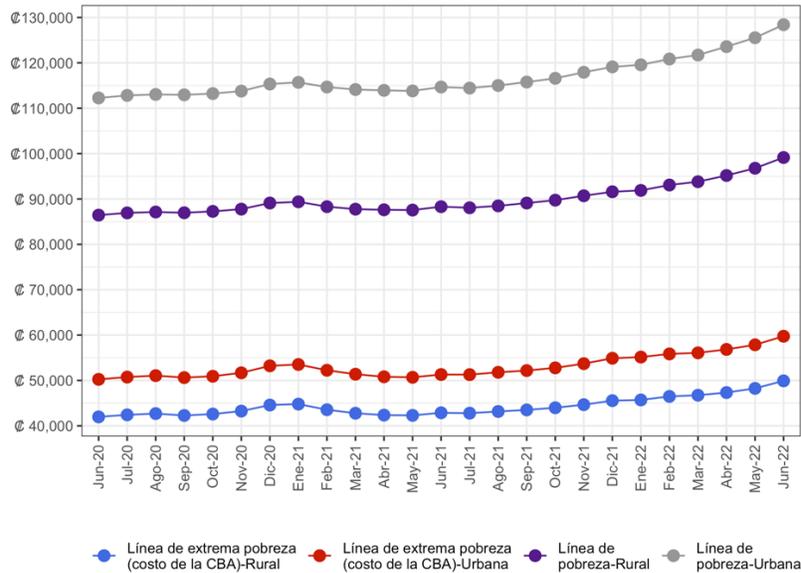
5.3 Línea de pobreza

En la Figura 4 se muestra la línea de pobreza y línea de extrema pobreza desde junio de 2020 a junio de 2022 tanto para la zona urbana como rural. Estos datos son calculados por el INEC. Estas líneas permiten calcular la proporción de hogares en pobreza y su variación obedece a la variación del índice de precios al consumidor (IPC).

Previo a mayo de 2021, se observa que las ambas líneas muestran un comportamiento estable. A partir de esa misma fecha, se aprecia un crecimiento sostenido de todas las líneas. Lo anterior es consistente con la evolución de la inflación. Puesto que, la evolución de precios fue estable hasta

febrero de 2021, y fue a partir de marzo de 2021, en donde los precios han sufrido aumentos sostenidos.

Figura 4. Evolución mensual de las líneas de pobreza

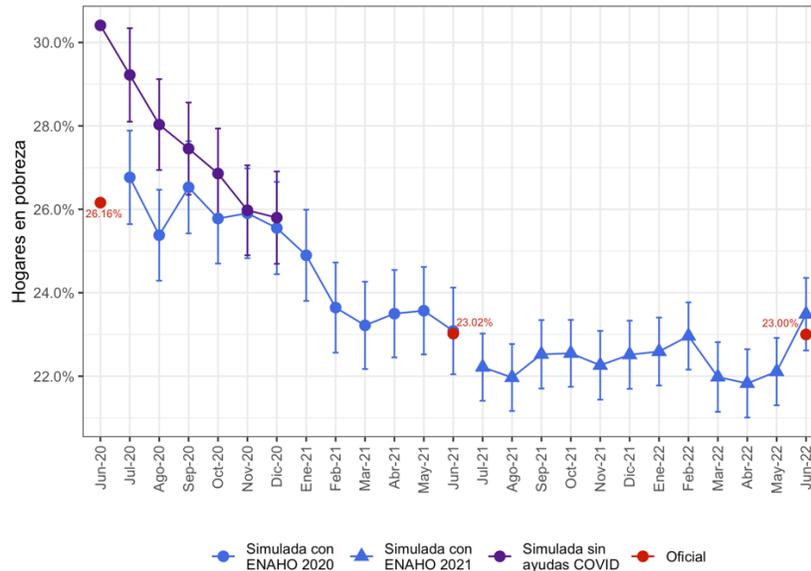


Fuente: Elaboración propia con base en las líneas de pobreza del INEC.

5.4 Tasas de pobreza e intervalos de confianza

Las Figuras 5 y 6 muestran el comportamiento de la tasa de pobreza total y pobreza extrema, respectivamente. La línea en azul corresponde a la serie de pobreza simulada, utilizando la ENAHO más cercana en una ventana de 6 meses previos al mes en cuestión. La línea en morado representa la simulación de la tasa de pobreza o pobreza extrema asumiendo ausencia de transferencias relacionadas al COVID-19. Por otro lado, las barras que acompañan los puntos simulados corresponden a los intervalos de confianza para cada mes simulado al 95%. Por último, los puntos denotados en rojo representan los datos oficiales para los meses competentes. Las simulaciones obedecen a la inclusión de los cambios en el ingreso principal por trabajo, masa de ocupados y transferencias estatales por COVID-19, cuando corresponde (línea azul desde julio 2020 a diciembre 2020). Las series simuladas se presentan en dos periodos distintos, antes (puntos) y después (triángulos) de junio de 2021, ya que posterior a dicha fecha la ENAHO presenta información actualizada.

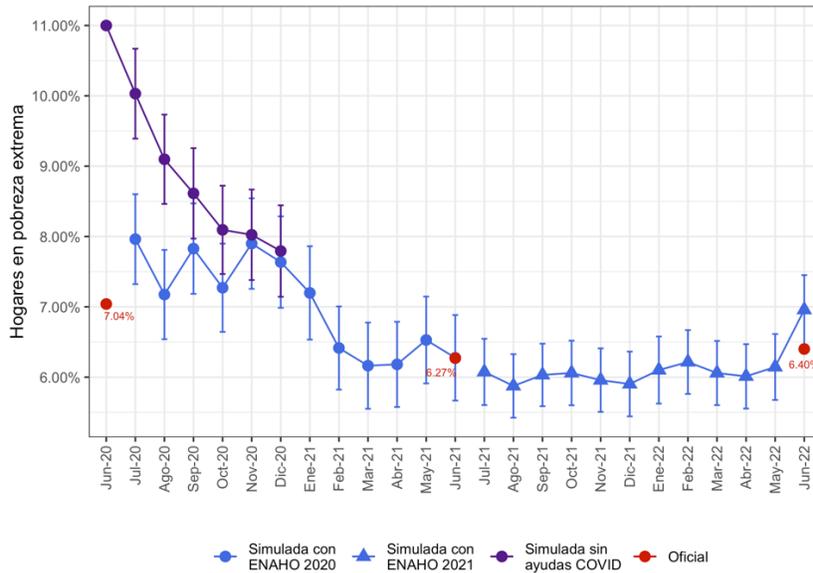
Figura 5. Evolución de la pobreza total oficial y simulada



Fuente: Elaboración propia con base en la ENAHO y ECE del INEC.

La tasa de pobreza total simulada más alta se observa en julio de 2020, mientras que la más baja en abril de 2022. A partir del primer mes simulado, se aprecia una tendencia decreciente de la tasa de pobreza hasta agosto de 2021, consistentes con la evolución de la pandemia de COVID-19. Posterior a esta fecha se observa una estabilización y a partir de abril 2022 se visualiza un crecimiento de la tasa de pobreza total. El intervalo de confianza de 95% fue calculado a partir de 1.000 muestras aleatorias con repetición, en cada mes simulado. De la Figura 5 se aprecia que las tasas de pobreza oficial, tanto para junio de 2021 como para junio de 2022, se encuentran dentro del intervalo de confianza de 95%.

Figura 6. Evolución de la pobreza extrema oficial y simulada



Fuente: Elaboración propia con base en la ENAHO y ECE del INEC.

Al igual que la tasa de pobreza total, la tasa de pobreza extrema simulada más alta se observa en julio de 2020, no obstante, la más baja en agosto de 2021. Desde el primer mes simulado hasta diciembre de 2020, la tasa de pobreza extrema muestra una tendencia estable alrededor de 7,63%. A partir de esta fecha muestra una tendencia decreciente hasta marzo de 2021. Desde esa fecha hasta abril de 2022, la tasa de pobreza extrema parece estabilizarse en torno a 6,10%, para mostrar un ligero aumento hasta junio de 2022.

Al igual que con la pobreza total, el intervalo de confianza de 95% fue calculado a partir de 1.000 muestras aleatorias con repetición para cada mes simulado. Similarmente, la tasa oficial de pobreza extrema para junio de 2021, se encuentra dentro del intervalo de confianza, sin embargo, el dato oficial para junio de 2022, se encuentra excluido del intervalo de confianza.

Como es de esperar, las magnitudes de ambas tasas considerando las alteraciones de las transferencias públicas relacionadas al COVID-19, son sistemáticamente menores en comparación con las estimaciones obtenidas en ausencia de transferencias, lo que realza la importancia de las transferencias coyunturales para aliviar la incidencia de la pobreza en los hogares. Esta diferencia se reduce con el tiempo debido a que los supuestos considerados en la simulación final contemplan

la disminución de los Bonos Proteger y la exclusión de transferencias privadas relacionadas al COVID-19, lo anterior se satisface para ambas tasas de pobreza.

6 Conclusiones y limitaciones

Aumentar la periodicidad y reducir el rezago entre un dato de pobreza y el otro hace posible conocer en forma más oportuna la situación económica del grupo con más desventaja de la sociedad y permite tomar decisiones informadas en el corto plazo. Para resolver este problema, en esta investigación se plantea una metodología para estimar la tasa de pobreza en forma mensual para Costa Rica durante el periodo comprendido entre junio de 2020 y junio de 2022.

Se utiliza la muestra anual de la ENAHO, y se modela la evolución de los ingresos de los hogares con parámetros estimados que describen la situación del mercado laboral en forma mensual a partir de un agrupamiento de las ECE de tres meses anteriores al mes a simular (incluido el mes en cuestión). Así, se generan estimaciones de ingresos de los hogares de la ENAHO tomando en cuenta los cambios en las contribuciones de las características sociodemográficas de la persona, en ingresos y en la probabilidad de estar ocupado, con datos trimestrales agrupados de la ECE. Luego, se actualiza la línea de pobreza, que considera la variación en los precios, y se encuentra el porcentaje de hogares en pobreza y pobreza extrema.

Esta metodología resulta ser efectiva para la estimación de los niveles de pobreza. Por ejemplo, para junio de 2021, la pobreza pronosticada con el modelo fue de 23,09% con intervalo de confianza que va desde 22,05% a 24,12%, y el dato observado en la ENAHO fue de 23,02%. En ese mismo periodo, la pobreza extrema pronosticada con el modelo fue de 6,28% con intervalo de confianza que va desde 5,67% a 6,88%, y el dato observado en la ENAHO fue de 6,27%. Mientras que para junio de 2022 la pobreza pronosticada con el modelo fue de 23,49% con intervalo de confianza que va desde 22,62% a 24,36%, y el dato observado en la ENAHO fue de 23,04%, asimismo, la pobreza extrema pronosticada con el modelo fue de 6,96% con intervalo de confianza que va desde 6,46% a 7,45%, y el dato observado en la ENAHO fue de 6,40%.

Este documento hace tres contribuciones importantes. Primero, los resultados muestran que las tasas de pobreza pueden variar significativamente en los meses de un año. Por lo tanto, es importante tener información de los niveles de pobreza en forma más periódica. Segundo, los datos de pobreza de la ENAHO implementada en julio (que preguntan sobre los ingresos de junio) se

publican hasta octubre de cada año. Por lo tanto, no hay información de pobreza por un periodo prolongado. Por ejemplo, para septiembre del 2022 el último dato que se conoce sobre pobreza fue el de junio del 2021, lo que representa 15 meses de rezago. Con el modelo propuesto, el dato de pobreza estaría rezagado por un máximo de tres meses. Tercero, el modelo permite demostrar que el mercado laboral es suficientemente dinámico para generar cambios en los niveles de pobreza en poco tiempo. A pesar de que los precios, a través del cambio en la línea de pobreza son relevantes, no se puede depender solamente en ellos para pronosticar los cambios en los niveles de pobreza.

Sin embargo, este modelo presenta también limitaciones y oportunidades de mejora. Primero, no toma en cuenta la variabilidad de ingresos ajenos a la actividad principal, tales como ingreso por empleo secundario, ingresos de capital, pensiones contributivas, transferencias privadas y estatales a través del año. Esto es relevante ya que una gran parte del ingreso depende de ayudas sociales, y sobre todo para los individuos cercanos a la línea de pobreza. Segundo, se asume que las diferencias entre la ENAHO y la ECE son estáticas en el tiempo. Este supuesto podría no cumplirse en algunos meses, sin embargo, es un reto difícil de resolver hasta que haya datos de la ENAHO en otro mes. Tercero, la simulación de la ocupación puede mejorar, aplicando los modelos a datos mucho más agregados, sin embargo, esto conlleva la pérdida de precisión aunado a la poca variabilidad de la ocupación relacionada a sus valores bajos.

Referencias

- Angelini, E., Rünstler, G. & Bańbura, M. (2008). “Estimating and forecasting the euro area monthly national accounts from a dynamic factor model”, Working Paper Series 953, European Central Bank. <https://ideas.repec.org/p/ecb/ecbwps/2008953.html>
- Arakaki, G., Chamussy, L. & Vezza, E. (2020). “Nowcasting Poverty in Argentina : A Methodological Note (English)”, Washington, D.C. World Bank Group. <http://documents.worldbank.org/curated/en/283681593599582508/Nowcasting-Poverty-in-Arentina-A-Methodological-Note>.
- Brum, M. & De Rosa, M. (2021). “Too little but not too late: nowcasting poverty and cashtransfers’ incidence during covid-19’s crisis”, *World Development*, 140:105227.
- Caruso,G., Lucchetti, L., Malasquez, E., Scot,T. & Castaneda, R. (2017). "But ? what is the poverty rate today? testing poverty nowcasting methods in Latin America and the Caribbean", Policy Research Working Paper Series 8104, The World Bank. <https://ideas.repec.org/p/wbk/wbrwps/8104.html>
- Castaneda, R., Gerszon, D., & Newhouse, D. (2019). “Nowcasting Global Poverty”, Special IARIW-World Bank Conference “New Approaches to Defining and Measuring Poverty in a Growing World”, Washington, DC, November 7-8, 2019. <https://unctad.org/system/files/official-document/NowcastingGlobalPoverty.pdf>
- Doudich, M., Ezzrari, A., Van der Weide, R. & Verme, P. (2013). “Estimating Quarterly Poverty Rates Using Labor Force Surveys : A Primer”, Policy Research Working Paper; No. 6466. World Bank, Washington, DC. <https://openknowledge.worldbank.org/handle/10986/15599> License: CC BY 3.0 IGO
- Grosse, M., Klasen, S. & Spatz, J. (2009). “Matching Household Surveys with DHS Data to Create Nationally Representative Time Series of Poverty: An Application to Bolivia”, Discussion Papers, No. 21, Georg-August-Universität Göttingen, Courant Research Centre - Poverty, Equity and Growth (CRC-PEG), Göttingen

- Habib, B., Narayan, A., Olivieri, S. & Sanchez-Paramo, C. (2010). “Assessing Ex Ante the Poverty and Distributional Impact of the Global Crisis in a Developing Country : A Micro-simulation Approach with Application to Bangladesh”, Policy Research working paper; no. WPS 5238. World Bank. <https://openknowledge.worldbank.org/handle/10986/3725> License: CC BY 3.0 IGO.
- Newhouse, D., & Vyas, P.. (2018). “Nowcasting poverty in India for 2014-15: A Survey to Survey Imputation Approach”, Global Poverty Monitoring Technical Note Series 6, The World Bank. <https://ideas.repec.org/p/wbk/wbgpmt/6.html>
- Rastrigina, O., Leventi, C., and Sutherland, H. (2015). “Nowcasting: Estimating developments in the risk of poverty and income distribution in 2013 and 2014”, EUROMOD Working Paper EM12/15, Colchester
- Rünstler, G & Banbura, M. (2007). “A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP”, No 751, Working Paper Series, European Central Bank. <https://EconPapers.repec.org/RePEc:ecb:ecbwps:2007751>
- Stifel, D. & Christiaensen, L. (2006). “Tracking Poverty over Time in the Absence of Comparable Consumption Data”, Policy Research Working Paper; No. 3810. World Bank, Washington, DC. <https://openknowledge.worldbank.org/handle/10986/8790> License: CC BY 3.0 IGO