

Continuidad del levantamiento de las encuestas de hogares tras la coyuntura de la enfermedad por coronavirus (COVID-19)

Diciembre de 2020

Resumen

En el contexto de la pandemia mundial de la enfermedad por coronavirus (COVID-19), la Comisión Económica para América Latina y el Caribe (CEPAL) está elaborando una serie de publicaciones cortas con recomendaciones de política relevantes para este período. En la presente nota, la tercera referente al levantamiento de las encuestas de hogares, se plantean diversas consideraciones para la continuidad de este tipo de instrumentos estadísticos, en un escenario en el que empiezan a levantarse en la región las restricciones a la movilidad impuestas para contener el avance de la pandemia.

Introducción¹

En el contexto de la pandemia de COVID-19, la mayoría de los países de América Latina y el Caribe adoptaron medidas para restringir la movilidad de los ciudadanos. Estas restricciones impidieron la aplicación presencial de las encuestas de hogares y llevaron, en algunos casos, a suspender los levantamientos y, en otros, a continuar la realización de encuestas por vía telefónica.

El cambio de modalidad de recolección de datos de un levantamiento presencial a uno telefónico, que fue necesario para dar continuidad a la producción de estadísticas sobre el empleo y los ingresos, particularmente importantes en el contexto de la pandemia, implicó nuevos desafíos para las oficinas nacionales de estadística y otros organismos públicos encargados de la realización de las encuestas. Estos desafíos fueron compartidos y analizados en diversos seminarios virtuales enmarcados en el trabajo de la Red de Transmisión del Conocimiento de la Conferencia Estadística de las Américas de la Comisión Económica para América Latina y el Caribe, así como en la XIX Reunión del Comité Ejecutivo de la Conferencia (véase CEPAL, 2020a y 2020b).

La División de Estadísticas de la CEPAL planteó además algunas recomendaciones para garantizar la calidad de las estadísticas ante el cambio en la metodología de la recolección y ha brindado asistencia técnica a varios países para su implementación (véase CEPAL, 2020c y 2020d). En estas se prestó especial atención a las propiedades inherentes a la estrategia de muestreo, definida como la dupla conformada por el diseño de muestreo de las encuestas y los procedimientos de estimación utilizados para realizar la inferencia a partir de los datos recolectados. En general, se recomendó:

- Basar la publicación de las estadísticas oficiales provenientes de las encuestas de hogares en la selección probabilística de la muestra y no en modelos predictivos.
- Conformar un panel de seguimiento probabilístico, basado en los meses de recolección presencial más recientes, para realizar encuestas de forma telefónica, manteniendo como población de referencia la contenida en el marco de muestreo maestro de la oficina nacional de estadística.

¹ La información empleada para la elaboración de este documento, salvo que se indique otra cosa, es la disponible hasta el 31 de octubre de 2020.

Resumen

Introducción

A. Sugerencias para el levantamiento mixto de las encuestas de hogares en el retorno a una nueva normalidad

B. Comparabilidad e impacto del COVID-19

C. Conclusiones

Bibliografía

- En los casos en los que las tasas de respuesta fueron bajas, publicar las estadísticas a nivel nacional, evitando las desagregaciones habituales y haciendo explícito el período de recolección de la información.
- Utilizar modelos de corrección del sesgo de selección y cobertura que permitieran ajustar los factores de expansión teniendo en cuenta la información auxiliar disponible en el panel de seguimiento.

En los meses transcurridos desde el inicio de la pandemia, la mayoría de los países de la región han hecho frente adecuadamente a los retos impuestos por este contexto excepcional, manteniendo la continuidad de la publicación de estadísticas oficiales sobre el mercado de trabajo, mediante la realización de encuestas por vía telefónica o mixta. El cese gradual de las restricciones a la movilidad de las personas plantea un nuevo escenario, en el que se empiezan a retomar paulatinamente los levantamientos presenciales de las encuestas de hogares. En este documento se plantean algunas consideraciones relevantes para este período. Se pone énfasis en que los procedimientos que se propusieron y ejecutaron en la coyuntura de las restricciones a la movilidad no deberían ser abandonados, sino que deberían adoptarse como procesos regulares para las encuestas de hogares. Asimismo, se propone mantener la flexibilidad ante la posibilidad de nuevos confinamientos y considerar un período pospandemia en el que no se dependa exclusivamente de los levantamientos presenciales.

A. Sugerencias para el levantamiento mixto de las encuestas de hogares en el retorno a una nueva normalidad

Los operativos telefónicos demostraron ser una valiosa opción cuando existen restricciones de movilidad y pueden conciliarse con los operativos presenciales. Esta experiencia puede ser aprovechada en el futuro, para que la logística de las encuestas contemple varias formas de acercarse al individuo seleccionado, ya sea de manera presencial, telefónica o digital.

Los operativos mixtos permiten una mayor flexibilidad en el contexto actual. Ante el riesgo de que vuelvan a aumentar los casos de COVID-19 y sea necesario restablecer restricciones a la movilidad de las personas, este tipo de operativos permiten dar continuidad al levantamiento de información. Además, mantener la modalidad de levantamiento telefónico ofrece nuevas posibilidades para aumentar la eficiencia de los operativos y reducir sus costos.

En el mediano plazo y a medida que los nuevos modos de recolección de información vayan siendo aceptados por los respondientes, es esperable que las encuestas de hogares se realicen combinando distintos modos de recolección, incluso cuando la coyuntura de la pandemia se considere superada.

Por otra parte, se debe tener en consideración que existen temas que no se prestan de forma adecuada para una recolección telefónica, así como grupos de la población que no serán fácilmente accesibles por vías distintas a la presencial, por ejemplo, quienes residen en asentamientos marginales y zonas rurales. En estos casos, las entrevistas presenciales parecen ser la forma más adecuada para obtener información².

A continuación se presentan algunas consideraciones para el levantamiento de las encuestas de hogares, que se plantean teniendo en cuenta la continuidad de los modos mixtos para la recolección de la información.

1. Selección de la muestra y códigos de disposición

En general, los operativos mixtos seguirán supeditados a la selección usual de unidades primarias de muestreo con un correspondiente submuestreo hasta llegar a las viviendas, hogares e individuos. Esta selección está basada en la medida de probabilidad inducida

² Naturalmente, en una entrevista presencial se deben tomar todas las medidas necesarias para minimizar el riesgo de transmisión de la enfermedad por coronavirus. Estas incluyen: monitoreo continuo y sistemático del estado de salud del equipo que realiza labores en campo; disponibilidad de una indumentaria básica que incluya todos los protocolos de bioseguridad para la protección del entrevistado y del entrevistador, y disponibilidad de los elementos necesarios de sanitización y desinfección, entre otras.

por el diseño de muestreo complejo (la mayoría de las veces estratificado y multietápico) sobre el marco de áreas cartográficas en el país. Esta selección es exhaustiva en la medida en que puede ser seleccionada cualquier área de la malla cartográfica del país y, por consiguiente, cualquier hogar o persona.

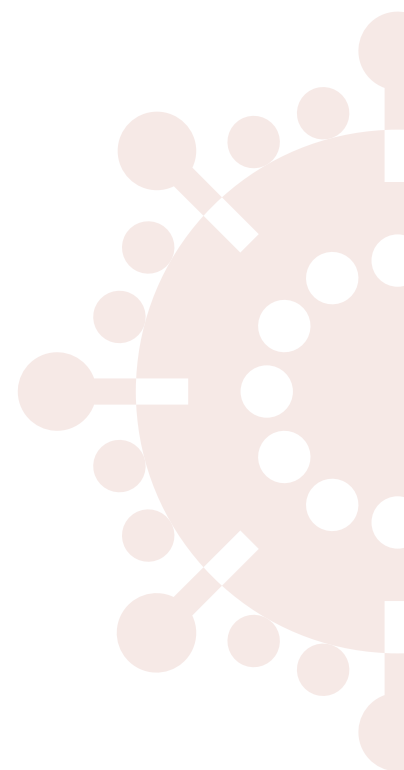
En principio las actualizaciones cartográficas en que se empadronan y enumeran las viviendas con hogares particulares deben ser reforzadas con operativos en los que se levante de forma exhaustiva la información de contacto de las viviendas preseleccionadas: números de teléfonos móviles y de líneas fijas y correos electrónicos. Además, en estos operativos es deseable que se levante el número de ocupantes del hogar, su información sociodemográfica básica y, en la medida de lo posible, información que permita vincular a los residentes del hogar con registros administrativos externos.

Las encuestas de hogares regularmente implementadas en la región pueden clasificarse en dos grandes grupos: aquellas que tienen levantamientos transversales independientes y aquellas que contemplan algún tipo de rotación en su muestra. En el caso de las encuestas con estructuras de rotación, en que las viviendas son parte de la muestra por un determinado número de períodos, la modalidad de levantamiento mixto podría servir para una apertura inicial presencial de las entrevistas, seguida de un levantamiento no presencial en los períodos subsiguientes. De esta forma, aunque sea factible hacer los levantamientos de forma presencial, se podrá recurrir a la información de contacto telefónico para realizar la entrevista por otros medios y, aun cuando no se pueda efectuar la entrevista, siempre será posible ajustar el peso del muestreo de la encuesta para eliminar los sesgos de selección debidos a la falta de respuesta y a la falta de cobertura, pues se tendrá acceso a información más detallada recogida en períodos anteriores.

Las encuestas que no tengan incorporados esquemas rotativos en sus diseños también pueden beneficiarse de un levantamiento mixto, siempre y cuando antes de iniciar la encuesta se logre el levantamiento de la información de contacto y de algunas características básicas de los hogares, para incorporarla en la identificación y eliminación del sesgo de cobertura y selección. De esta forma, las actualizaciones cartográficas serán esenciales para el ajuste de los factores de expansión. Sin embargo, las bondades de los esquemas rotativos resaltan a la vista en este tipo de coyunturas, porque la búsqueda activa de nuevos respondientes se aliviana en comparación con la que tiene lugar en el caso de los levantamientos transversales.

Es importante tener en cuenta la definición de los códigos de disposición (individuos elegibles respondientes, individuos elegibles no respondientes, individuos con elegibilidad desconocida e individuos no elegibles) para cada uno de los tipos de levantamiento y su comparabilidad. Si bien los códigos de disposición cuentan con clasificaciones (AAPOR, 2016). estándares usadas en la mayoría de las encuestas en la región (presenciales y telefónicas), se recomienda enfáticamente hacer un empalme exhaustivo de cada posible evento de elegibilidad, su significado según el modo de recolección empleado y su correspondiente clasificación en las encuestas. Debe evitarse la adopción incorrecta de los códigos de disposición telefónicos tradicionales en los levantamientos mixtos.

Nótese que los códigos de disposición telefónicos están definidos sobre el supuesto de la marcación aleatoria de números y, por ende, un número seleccionado podría corresponder a un hogar (que es parte de la población objetivo) o a otro tipo de entidad que no sea parte de la población objetivo. A manera de ejemplo, en el caso de los códigos de disposición de las encuestas telefónicas se consideraría que un número de teléfono de una línea que estuviera ocupada correspondería a una situación de “elegibilidad desconocida” (no es posible saber si forma parte de la población objetivo o no). En el caso de las encuestas a hogares que llevan a cabo las oficinas nacionales de estadística, en las que los números de contacto telefónico están vinculados directamente con las unidades de muestreo en un marco de áreas, este evento no puede categorizarse como de “elegibilidad desconocida”, sino como un “elegible no respondiente”. Por ello, es pertinente revisar de forma exhaustiva los eventos asociados a los códigos de disposición tanto de las encuestas presenciales como de las encuestas telefónicas y realizar una clasificación apropiada para un modo de levantamiento mixto.



2. Pertinencia del modo mixto de levantamiento para todas las encuestas

En la mayoría de casos, las encuestas telefónicas no permiten captar todos los constructos que se indagaban en las encuestas de hogares presenciales. El tiempo disponible para una entrevista telefónica es menor que aquel disponible para una entrevista presencial, por lo que se requiere acortar los cuestionarios. De otra forma, es posible que las encuestas no puedan ser concluidas apropiadamente por cansancio del informante o que las tasas de respuesta efectiva decrezcan. En los meses recientes, en varios países de la región se utilizaron cuestionarios reducidos en levantamientos telefónicos de las encuestas de fuerza laboral, en los que se privilegió por lo general la información sobre el mercado de trabajo y los ingresos, y se excluyeron otras variables que se recogen habitualmente.

Por ello, en un modo mixto de levantamiento es recomendable tener dos versiones del cuestionario, que permitan calcular conjuntamente los principales indicadores de la encuesta. Así, en el momento en que algún área entre en confinamiento y no se permita realizar encuestas presenciales, se utilizará el cuestionario adaptado para la recopilación telefónica de información. La conciliación no significa simplemente acortar el cuestionario para la entrevista telefónica, sino también revisar si hay preguntas que puedan considerarse prescindibles en el cuestionario para la entrevista presencial.

A su vez, hay encuestas de hogares cuyo diligenciamiento telefónico durante la pandemia no parece viable. Tal es el caso de las encuestas de presupuestos familiares (ingresos y gastos), que requieren de un contacto extenso y por varios días entre el entrevistador y el respondiente. Obtener esta información de forma telefónica impone una carga muy alta al encuestado. A su vez, la misma pandemia puede acentuar problemáticas complejas que forman parte de los constructos que se investigan por medio de encuestas de victimización o encuestas de uso del tiempo; en el primer caso, los protocolos estrictos para favorecer la seguridad del respondiente no pueden aplicarse en un levantamiento telefónico; en el segundo caso, la misma pandemia puede redundar en que, en períodos de confinamiento, la estructura habitual de las actividades regulares de la población se vea trastocada.

3. Eliminación del sesgo

Uno de los desafíos más importantes ante el cambio en el modo de recolección de las encuestas es hacer frente al sesgo de cobertura y de selección. Se asume como un principio fundamental que los hogares que respondieron a la encuesta telefónica no tienen características similares a los hogares no respondientes o no contactados. Más aún, la falta de respuesta puede estar asociada con el fenómeno que se está midiendo. Por ejemplo, podría suceder que en los hogares que no proporcionaron información de contacto telefónico haya más personas desocupadas, o que las tasas de ausencia de respuesta sean más altas entre los hogares que cayeron en situación de pobreza debido a los efectos de la pandemia en sus ingresos. El documento “Recomendaciones para eliminar el sesgo de selección en las encuestas de hogares en la coyuntura de la enfermedad por coronavirus (COVID-19)” (CEPAL, 2020d) provee una referencia técnica sobre los procedimientos para detección y corrección de los posibles sesgos. Es recomendable que las oficinas nacionales de estadística adopten estos procesos de manera general para las encuestas de hogares, más allá de los períodos de cuarentena.

Por otro lado, las oficinas nacionales de estadística que tengan acceso a diferentes registros administrativos podrán usarlos para saber si los individuos seleccionados (respondientes y no respondientes) se encuentran registrados en sus bases de datos. De esta forma, los modelos estadísticos para controlar el sesgo por ausencia de respuesta se verán complementados de una forma más robusta. Por ejemplo, nótese que la información de si un individuo ha recibido ayuda del Estado o está afiliado a cierto tipo de protección social o no ha sido identificado en la base de aportes de pensiones en los últimos meses se convierte en una potente herramienta para detectar y eliminar el sesgo de selección que posiblemente pueda estar presente en este tipo de levantamientos mixtos.

En el futuro cercano, será posible estudiar las características de los respondientes presenciales en comparación con los respondientes telefónicos. Por ejemplo, antes de hacer las correcciones por ausencia de respuesta y la calibración de los factores de expansión, los países con levantamientos rotativos podrán comparar las estructuras poblacionales

expandidas de sus respondientes y detectar posibles sesgos. En algunos casos será posible incluir la información de registros administrativos como variables auxiliares en el ajuste de los factores de expansión. El solo hecho de que una persona o un hogar forme parte de un registro lo diferenciará del resto de los respondientes y se podrán establecer patrones sociodemográficos que recojan las diferencias entre respondientes y no respondientes. Una buena parte de la ausencia de respuesta podría venir de personas que sí solían responder en los levantamientos presenciales y la correcta identificación de estos patrones coadyuvará en la conservación de las buenas propiedades de los estimadores de las encuestas de hogares, de las cuales la más importante es el insesgamiento.

B. Comparabilidad e impacto del COVID-19

En términos generales, la pandemia obligó a que los países efectuaran cambios en varios aspectos de la metodología del levantamiento y análisis de la información, que se resumen a continuación:

- Cambió el modo de levantamiento de presencial a telefónico (o mixto, en algunos casos), así como las definiciones de la estructura de elegibilidad de las viviendas seleccionadas y sus correspondientes códigos de disposición.
- Cambió el esquema de supervisión de los encuestadores y, en algunos casos, se suprimieron las actualizaciones cartográficas del número de hogares particulares en las unidades primarias de muestreo seleccionadas.
- Se introdujo un nuevo esquema de ajuste de factores de expansión, buscando eliminar el sesgo de cobertura (no se contaba con números telefónicos de contacto para todos los hogares de los levantamientos anteriores) y de ausencia de respuesta (algunos hogares contactados telefónicamente no contestaron el cuestionario).
- Se revisaron los esquemas de calibración de los factores de expansión y, en aras de la flexibilidad de la metodología de estimación, se restringió el número de restricciones de calibración.

Estas diferencias plantean un desafío al que deben responder las oficinas nacionales de estadística, sobre la comparabilidad de las series generadas durante el período excepcional de la pandemia. Teniendo en cuenta que ambos hechos (recolección telefónica y pandemia) ocurrieron al mismo tiempo, uno como consecuencia del otro, no es una tarea sencilla separar los efectos de cada modalidad (en persona y telefónica) y el posible impacto que se introdujeron en las cifras oficiales informadas por las oficinas nacionales de estadística a partir de las encuestas de hogares.

No obstante, no se debe caer en el falso dilema de escoger entre comparabilidad e insesgamiento. Por un lado, las diferencias que se listaron anteriormente muestran con claridad que los esquemas de planeación, levantamiento, ejecución y análisis de las encuestas antes y después de la pandemia no son comparables. Por otro lado, la pregunta de interés no es si existe o no comparabilidad en los métodos, sino si estos cambios tuvieron un impacto significativo en la estimación de las estadísticas oficiales. La respuesta a esta pregunta redundará en la adecuación de las series temporales producidas y alimentadas con cada levantamiento que las oficinas nacionales de estadística ejecutan. De esta forma, podría encontrarse que, gracias a todos los cambios y ajustes introducidos en el levantamiento y análisis de las encuestas de hogares para la eliminación de los sesgos, no existe un impacto significativo en las cifras producidas por la entidad, con respecto a los levantamientos anteriores.

Por supuesto que, como consecuencia de la pandemia, la magnitud de la estimación de los indicadores sociales cambió. Son claros los ejemplos en el total de personas activas en el mercado laboral, personas desocupadas, personas cuyos ingresos se han visto afectados y nuevas personas en condición de pobreza. La comparabilidad de las cifras no está supeditada a estos cambios de magnitud, sino a la posible introducción de una sobreestimación o una subestimación que el cambio de modo de recolección, y sus posteriores ajustes, podrían haber causado.

Una forma certera de separar el cambio atribuible al modo de recolección del efecto de la pandemia es mantener por un tiempo prudencial los levantamientos telefónicos y mixtos en el período posterior a la pandemia. Esta medida puede aplicarse de varias formas. Las oficinas nacionales



de estadística pueden realizar levantamientos independientes paralelos con tamaños de muestra adecuados. También es posible mantener una submuestra telefónica con representatividad nacional, mientras que los levantamientos presenciales incorporarían mayores tamaños al requerir desagregaciones más profundas. Otra opción puede ser incorporar paulatinamente el levantamiento presencial y reducir al mismo tiempo el levantamiento telefónico.

En las encuestas que tienen grupos de rotación se podrían considerar diferentes modos de recolección en estos grupos. Asumiendo que cada uno de los grupos de rotación constituye una muestra representativa de la realidad del país que está dispersa en todo el territorio nacional, y considerando que, con muestras grandes, como las que tienen las encuestas de hogares a nivel nacional, los estimadores de muestreo tienen una distribución normal, es posible realizar contrastes de hipótesis sobre los indicadores más importantes en los grupos de rotación. De esta manera se podría estimar el impacto del cambio de modo de recolección en las cifras y definir las subsecuentes acciones que deberían tomarse para que las cifras producidas durante el período de pandemia se adecuara a la serie temporal del indicador. Estas acciones permitirán evaluar mediante modelamiento estadístico si el cambio de modo de recolección tuvo un efecto (y, si lo hubo, en qué sentido) en la estimación de los indicadores sociales.

Al final, es probable que, gracias a los ajustes de eliminación del sesgo, las series temporales no sufran ningún quiebre. Si, como resultado de los levantamientos paralelos en la transición hacia la normalidad, se observa que el cambio de modo de recolección sí tuvo un efecto importante en la estimación de las cifras oficiales, se recomienda realizar un empalme de las series temporales de forma retrospectiva para poder hacer comparables las mediciones. Un argumento más a favor de los levantamientos paralelos es que, en todo caso, este empalme no podrá realizarse inmediatamente después de la superación de la pandemia, sino que deberá esperarse un período de tiempo prudencial para identificar correctamente los componentes que tuvieron mayor efecto en la estimación.

C. Conclusiones

Después de varios meses en que las encuestas de hogares se levantaron de manera telefónica, para hacer frente a las restricciones a la movilidad a consecuencia de la pandemia, se está ante un nuevo escenario, en el que se empiezan a retomar paulatinamente los levantamientos presenciales. En este documento se plantean algunas consideraciones relevantes para este período, que apuntan a la continuidad de algunos de los procedimientos adoptados, en algunos casos de forma temporal, y en otros como parte de los procesos regulares para las encuestas de hogares.

En particular, se plantea:

- Considerar un período de transición en el que se sigan realizando operativos telefónicos, manteniendo la flexibilidad ante la posibilidad de nuevos confinamientos en algunas zonas y disminuyendo gradualmente su realización en el tiempo.
- Evaluar la posibilidad de continuar efectuando levantamientos mixtos en este período de pandemia y estudiar exhaustivamente la viabilidad de incluir estos esquemas de levantamiento en el período postpandemia, mediante un rediseño de las encuestas.
- Realizar una conciliación de los cuestionarios para entrevistas presenciales y telefónicas a fin de mantener la comparabilidad en la medición del constructo en ambos modos de recolección.
- Incorporar procesos de recolección de información de contacto y datos básicos de los hogares durante las actualizaciones cartográficas, antes del diligenciamiento del cuestionario.
- Adaptar los códigos de disposición de las viviendas y hogares seleccionados, manteniendo la comparabilidad entre los levantamientos presenciales y telefónicos, rehaciendo la clasificación de acuerdo con la realidad de la selección sobre un marco de áreas.
- Adoptar los procedimientos técnicos de eliminación de sesgos de cobertura y selección que se ejecutaron en la coyuntura de restricciones a la movilidad como procesos regulares para las encuestas de hogares.
- Evaluar, mediante procedimientos estadísticos, el efecto del cambio de modo de recolección de las encuestas, sobre la base de la realización de operativos paralelos y, en caso de ser necesario, efectuar procedimientos de reajuste y empalme de las series temporales.

Bibliografía

AAPOR (American Association for Public Opinion Research) (2016), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* [en línea] https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.

CEPAL (Comisión Económica para América Latina y el Caribe) (2020a), “Planes de contingencia y recomendaciones para la producción estadística ante la emergencia sanitaria generada por el COVID-19”, Red de Transmisión del Conocimiento (RTC) [en línea] <https://rtc-cea.cepal.org/es/conectados-rtc/covid-19-planes-de-contingencia-de-las-oficinas-nacionales-estadisticas>.

_____(2020b), “XIX Reunión del Comité Ejecutivo de la Conferencia Estadísticas de las Américas de la CEPAL”, agosto [en línea] <https://www.cepal.org/es/eventos/xix-reunion-comite-ejecutivo-la-conferencia-estadisticas-americas-la-cepal>.

_____(2020c), “Recomendaciones para la publicación de estadísticas oficiales a partir de encuestas de hogares frente a la coyuntura de la enfermedad por coronavirus (COVID-19)”, *Informes COVID-19*, Santiago, abril.

_____(2020d), “Recomendaciones para eliminar el sesgo de selección en las encuestas de hogares en la coyuntura de la enfermedad por coronavirus (COVID-19)”, *Informes COVID-19*, Santiago, mayo.



Este documento es parte de un conjunto de informes elaborados por la Comisión Económica para América Latina y el Caribe (CEPAL) sobre la evolución y los efectos de la pandemia de COVID-19 en América Latina y el Caribe. Fue preparado por la División de Estadísticas, dirigida por Rolando Ocampo, bajo la coordinación general de Alicia Bárcena, Secretaria Ejecutiva de la CEPAL.

Copyright © Naciones Unidas, 2020



Comisión Económica para América Latina y el Caribe (CEPAL)
Economic Commission for Latin America and the Caribbean (ECLAC)
www.cepal.org