

Recomendaciones para la
publicación de estadísticas oficiales
a partir de encuestas de hogares en
medio de la pandemia COVID19

Statistics Division

ECLAC

June - 2020



Resumen

En el contexto de la pandemia mundial COVID-19, la CEPAL ha preparado una serie de recomendaciones sobre la recopilación de información sobre el mercado laboral a través de encuestas de hogares, prestando especial atención a la estrategia de muestreo (diseño + estimador) necesaria para este fin.



UNITED NATIONS

ECLAC

Basado en publicaciones COVID de la CEPAL

COVID-19 REPORTS

Recommendations for the publication of official statistics from household surveys in the context of the coronavirus disease (COVID-19) pandemic

April, 2020

Summary

In the context of the global COVID-19 pandemic, ECLAC is preparing a series of short publications with relevant policy recommendations. A number of recommendations are made in this note on gathering labour market information through household surveys, paying particular attention to the sample designs required for this purpose.



A. Introduction

Since the health emergency arising from the global COVID-19 pandemic, which has led most of the countries of the region to declare states of emergency, national statistical offices (NSOs) have had to interrupt abruptly the collection of primary information as part of many of their statistical operations, including household surveys. This note describes possible scenarios for the collection and analysis of household survey data and provides a set of recommendations regarding the capture, publication and dissemination of official statistics from those surveys, in the context of the international situation caused by the COVID-19 pandemic.

In those countries that have declared states of emergency, in order to stop the spread of the disease and protect those people most vulnerable to it, governments have taken difficult decisions that have affected society and the normal functioning of government institutions, educational entities, industry, transportation and commercial establishments, among others. In some cases, the general public has begun to collaborate with voluntary quarantines; but in light of the rapid spread of the disease, governments have had to impose curfews, movement restrictions and compulsory quarantines, among other health measures, and remote learning and working methodologies have been implemented or developed.

In view of this situation, the Statistics Division of ECLAC sent out a short questionnaire to the NSOs and central banks of Latin America and the Caribbean, to learn about the effects that the health emergency was having on the functioning of statistical institutions. Responses from 20 countries indicated that, as at 20 March, they had all declared some form of health emergency as a result of COVID-19. The statistical operations most affected by the situation were, surveys, followed by administrative records and censuses. With regard to household surveys, most respondents said that information collection would be postponed to later in the year in most cases. The main information collection difficulties are movement restrictions and the closure of establishments.

A series of videoconferences have been held since 24 March to assess the impacts of the COVID-19 pandemic on statistical operations.¹ These videoconferences have been convened by ECLAC, in conjunction with the International Labour Organization (ILO) and the National Institute of Statistics (INE) of Chile, in its capacity as coordinator of the virtual talks of the Knowledge Transmission Network (KTN) of the Statistical Conference of the Americas, chaired by the National Administrative Department of Statistics (DANE)

¹ See [online] <https://rtc.cea.cepal.org/en/conectados-rtc/planes-de-contingencia-adaptados-por-las-oficinas-nacionales-de-estadistica-ante-la>

COVID-19 REPORTS

Recommendations for eliminating selection bias in household surveys during the coronavirus disease (COVID-19) pandemic

May 2020

Summary

In the context of the global coronavirus disease (COVID-19) pandemic, the Economic Commission for Latin America and the Caribbean (ECLAC) is preparing a series of short publications with relevant policy recommendations. A number of recommendations are made in this note to address bias problems that may arise in household surveys carried out during this exceptional period, as a complement to the suggestions made in a previous note on the sample designs for this type of survey.

Introduction

In an attempt to slow the spread of COVID-19, countries have imposed movement restrictions on people, which have prevented the face-to-face collection of information for household surveys. In order to address this issue and to continue producing relevant and timely official statistics, some national statistical offices (NSOs) have resorted to conducting surveys by telephone or via the Internet. The document "Recommendations for the publication of official statistics from household surveys in the context of the coronavirus disease (COVID-19) pandemic" contains some possible lines of action for drawing the sample of households to participate in the surveys, in particular using a selected panel from a recent period for which the telephone contact information is available (ECLAC, 2020). This note complements the recommendations made in that document, by proposing two approaches to minimize the bias generated by non-response that will be encountered when carrying out surveys by telephone. In addition, a third approach is proposed for those instances in which it is difficult to obtain auxiliary information.

A. Detecting bias

Changing the household survey data collection modality from face-to-face interviews to a telephone- or web-based modality may have unintended consequences and, in particular, may generate biases (of selection, coverage and non-response) among survey respondents. In a scenario where a sample of households from a previous period is being used (hereinafter the "original sample") and where every effort is being made to contact those selected households, the process is inevitably exposed to the following difficulties:

- Not all the households in the original sample provided their telephone contact information.
- Some households provided their contact information, but at the time of the interview they do not live at the selected address.
- Some households provided their contact information, but they have since changed their contact telephone number.
- Not all households that provided their contact information are willing to answer the survey questionnaire.

Summary

Introduction

A. Detecting bias

B. Propensity score adjustment

C. Two-stage calibration method

D. Poststratification based on multilevel models

E. Conclusions

Bibliography



ECLAC



Restricciones de movimiento

- Desde la emergencia sanitaria derivada de la pandemia mundial COVID-19, las oficinas nacionales de estadística (ONE) han tenido que interrumpir abruptamente la recopilación de información primaria como parte de muchas de sus operaciones estadísticas, incluidas las encuestas de hogares.
- A la luz de la rápida propagación de la enfermedad, los gobiernos han tenido que imponer toques de queda, restricciones de circulación y cuarentenas obligatorias, entre otras medidas sanitarias
-



UNITED NATIONS

ECLAC

Encuestas de hogares

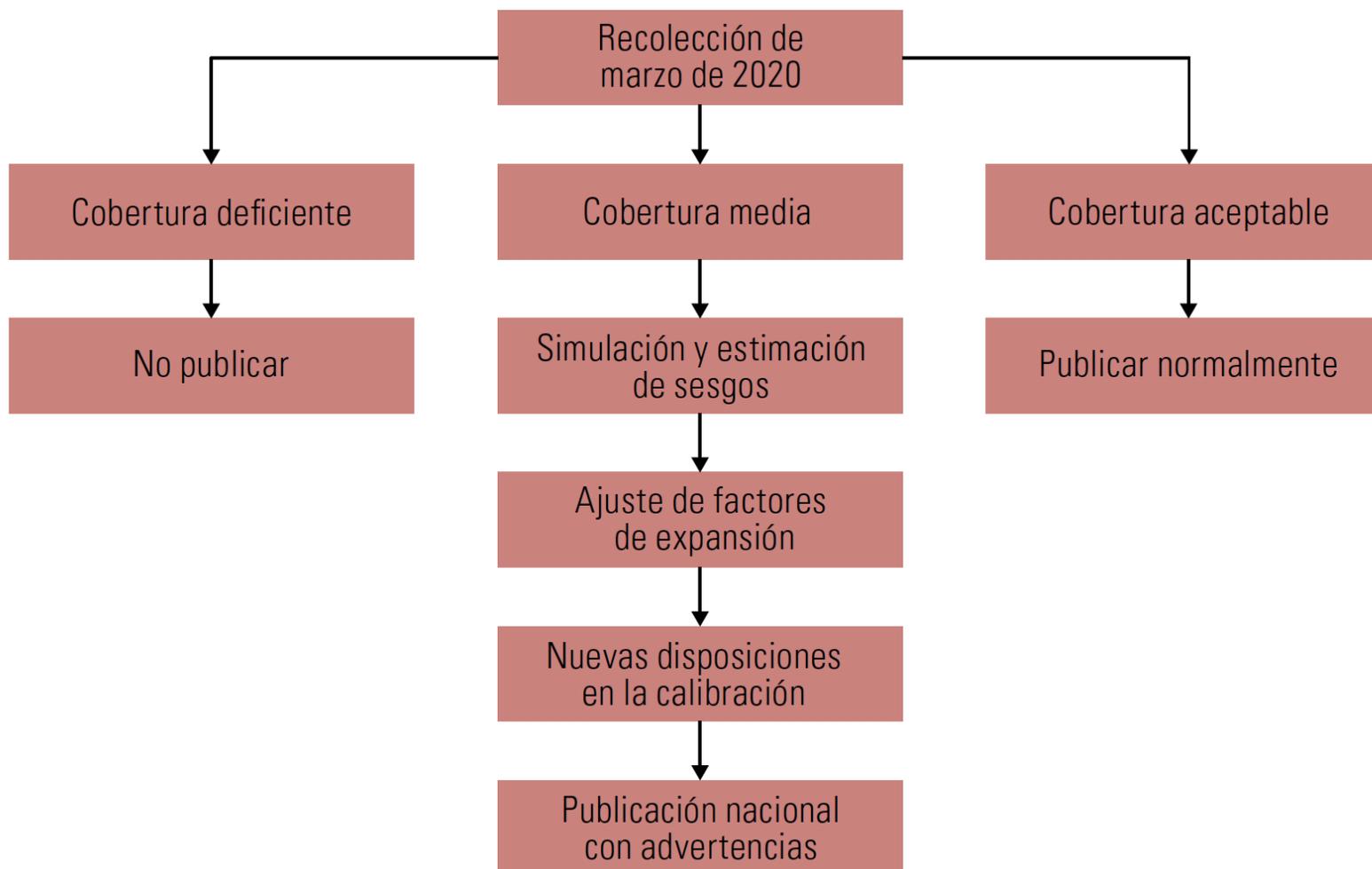
- Las restricciones de movimiento han llevado a la suspensión de la recopilación de datos presenciales para las encuestas de hogares hasta nuevo aviso.
- Las ONE están realizando encuestas por teléfono o Internet.
 - Este cambio en la metodología de la colección es necesario para que se sigan produciendo cifras de empleo y pobreza.
 - Estos indicadores son particularmente importantes en el contexto de la pandemia, dado el profundo impacto que las restricciones y cuarentenas de movimiento tienen en el empleo de las personas.



UNITED NATIONS

ECLAC

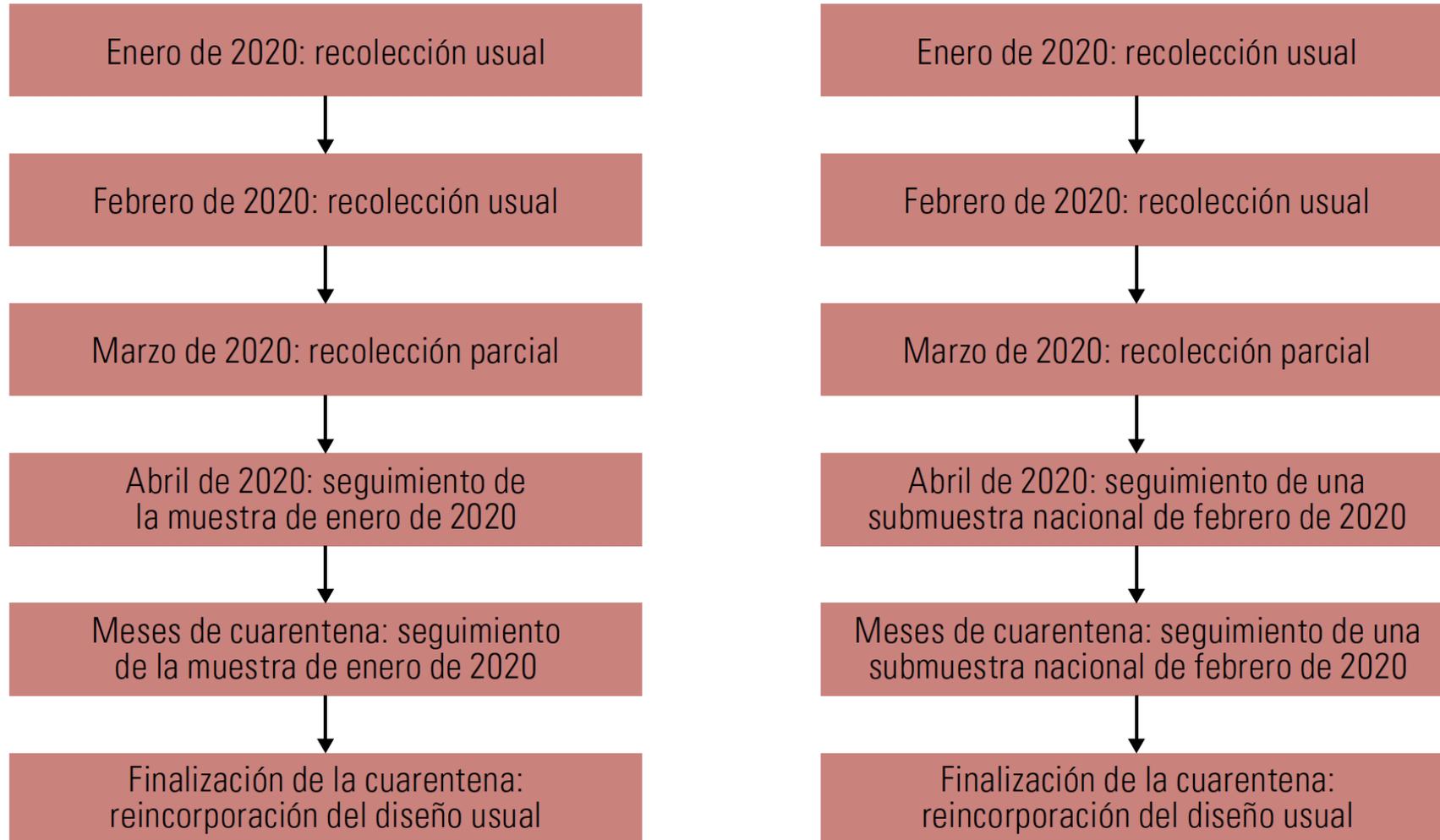
Marzo 2020 - Posibles escenarios y recomendaciones



UNITED NATIONS

ECLAC

Meses de cuarentena: posibles escenarios y recomendaciones



UNITED NATIONS

ECLAC

Detección del sesgo

- Cambiar la modalidad de recopilación de datos de la encuesta de hogares de entrevistas presenciales a una modalidad telefónica puede generar sesgos (de selección, cobertura y no respuesta) entre los encuestados.
 - La recomendación es definir una muestra de hogares de un período anterior y ponerse en contacto con los hogares seleccionados para que proporcionen información actualizada.



UNITED NATIONS

ECLAC

Sin embargo...

- No todos los hogares de la muestra original proporcionaron su información de contacto telefónico.
- Algunos hogares proporcionaron su información de contacto, pero en el momento de la entrevista no viven en la dirección seleccionada.
- Algunos hogares proporcionaron su información de contacto, pero han cambiado su número de teléfono de contacto.
- No todos los hogares que proporcionaron su información de contacto están dispuestos a responder el cuestionario de la encuesta



UNITED NATIONS

ECLAC

Pérdida de respondientes efectivos

Supongamos que el 85% de la muestra proporcionó datos de contacto y que la probabilidad de que un hogar contactado responda a toda la encuesta es del 80%, entonces:

- Las respuestas se obtendrían de sólo el 68% de la muestra original.
- Tendríamos un efecto de desgaste en el panel (pérdida de participantes cuanto más tiempo se utilice el panel).
 - Habrá hogares que dejarán de responder a la encuesta porque se les contacta repetidamente.



UNITED NATIONS

ECLAC

Pérdida de respondientes efectivos

- En este escenario, es muy probable que los hogares encuestados no tengan características similares a los hogares no respondientes o no cubiertos.
- Las razones de la falta de respuesta del hogar a la encuesta pueden estar asociadas con el fenómeno que se está midiendo:
 - Hay más desempleados en los hogares no demandados.
 - Las tasas de no respuesta son más altas entre los hogares que viven en la pobreza).

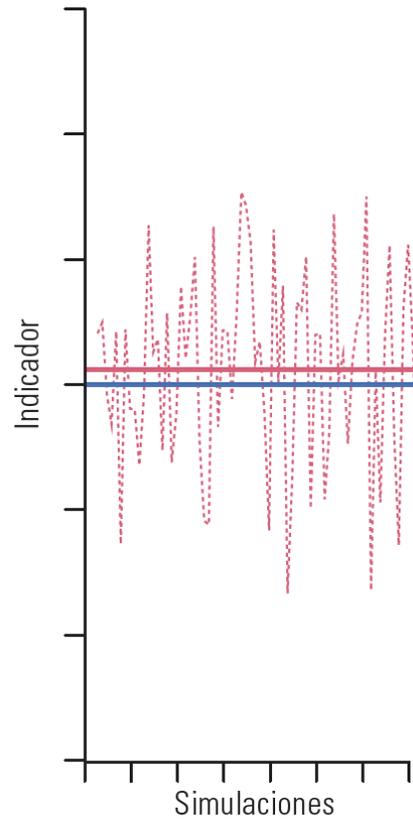


UNITED NATIONS

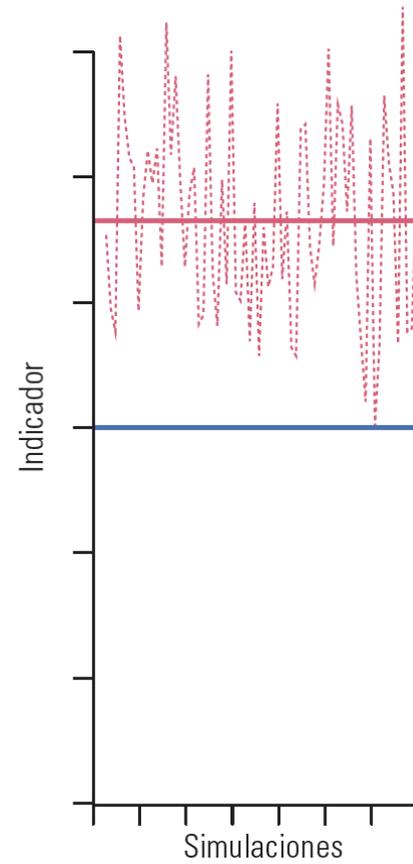
ECLAC

Estimación del efecto del sesgo

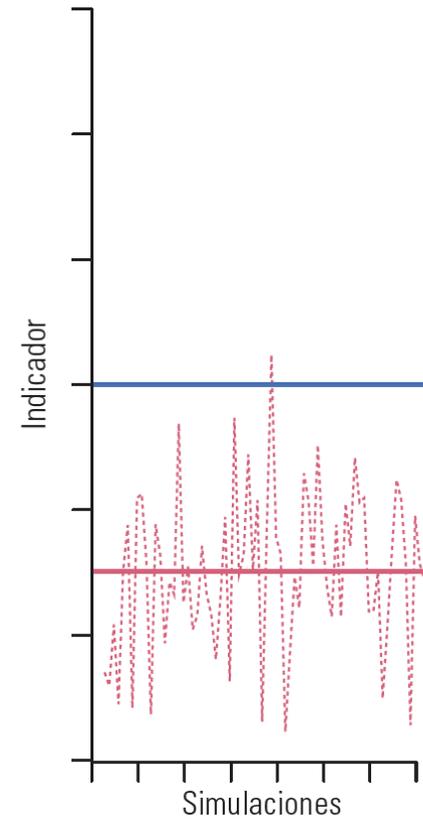
A. Sin sesgo



B. Sesgo positivo



C. Sesgo negativo



— Promedio de las simulaciones — Estimación publicada



UNITED NATIONS

ECLAC

Corrección del sesgo por ausencia de respuesta

Propensity score

- Este enfoque propuesto por Rosenbaum y Rubin (1983) es útil para esclarecer la estructura de la no respuesta y, por lo tanto, para corregir el sesgo de cobertura y el sesgo de no respuesta.

$$\text{logit}(\hat{\phi}_k) = \mathbf{x}_k \hat{\beta}$$

- El factor de expansión ajustado tomaría la siguiente forma:

$$w_k = \frac{d_k}{\hat{\phi}_k}$$



UNITED NATIONS

ECLAC

Corregir del sesgo por ausencia de respuesta

Calibración en dos etapas

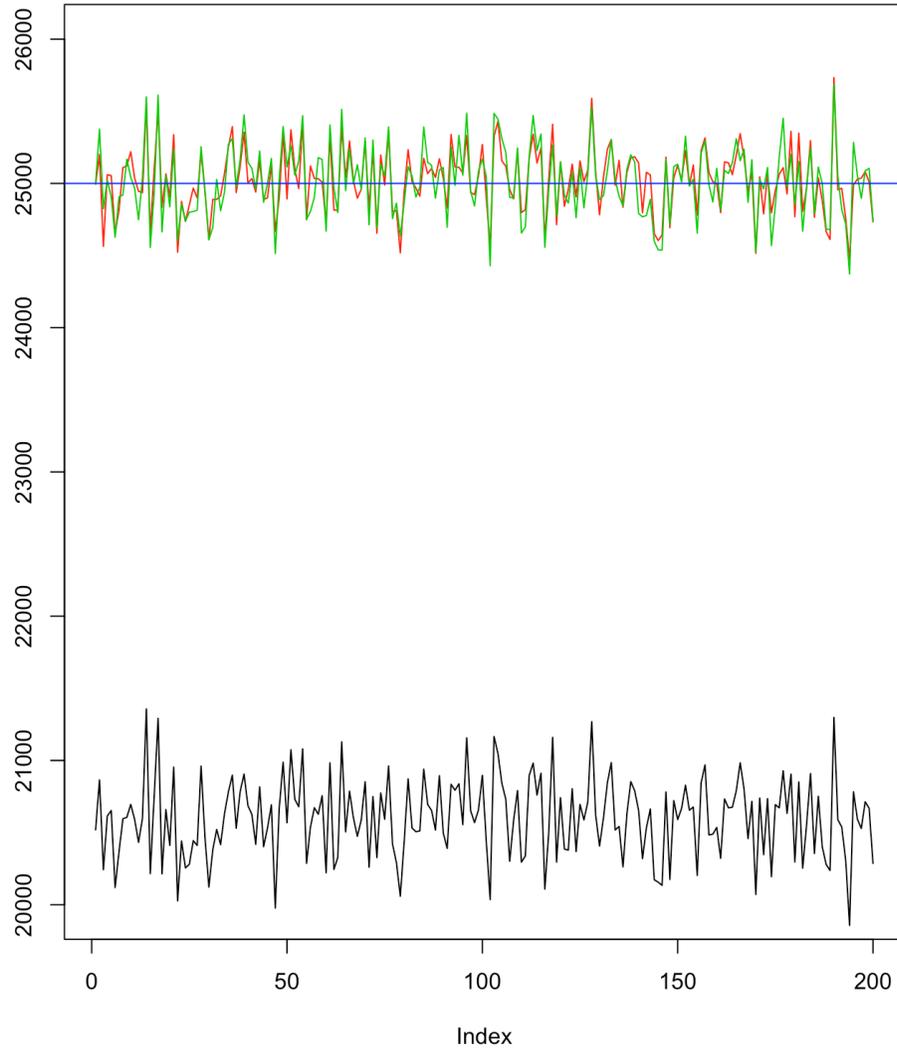
Sarndal (2007) afirma que cuando las encuestas por muestreo se ven afectadas por la falta de respuesta, es posible utilizar el enfoque de calibración.

- En la primera etapa, la calibración de los pesos de muestra originales podría basarse en:
 - totales de edad, región subnacional, área y sexo, disponibles a partir de sólidas proyecciones de población.
- En la segunda etapa, la calibración de los pesos de la muestra telefónica podría basarse en:
 - totales de renta per cápita, estado de ocupación, rama de actividad y nivel de educación, obtenidos de la publicación de la encuesta original.

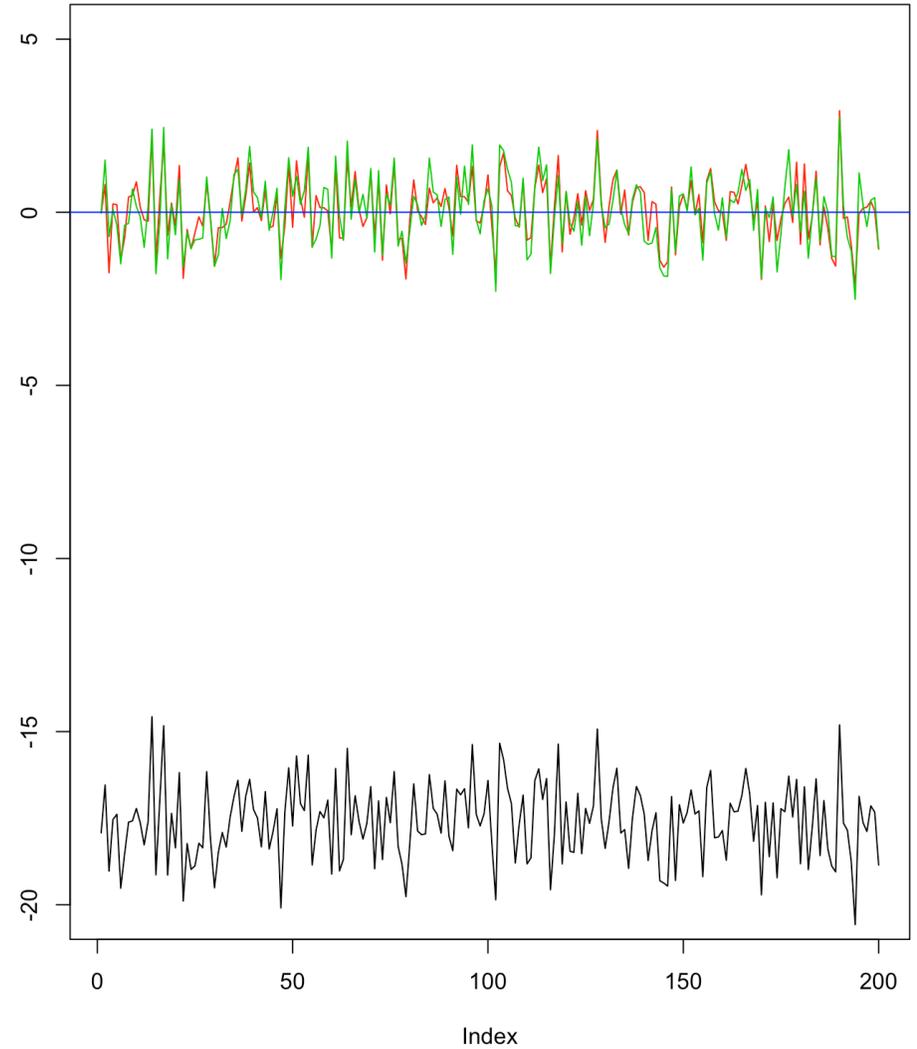


No se limite a ponderar

Expectation



Relative Bias (in percentage)



UNITED NATIONS

ECLAC

Muchas Gracias

Por favor, no dude en ponerse en contacto con la División de Estadística de la CEPAL si necesita asistencia en el diseño y análisis de las encuestas, detección de sesgo y/o calibración.

- Rolando Ocampo (rolando.ocampo@un.org)
- Xavier Mancero (xavier.mancero@un.org)
- Andrés Gutiérrez (andrés.gutiérrez@un.org)



UNITED NATIONS

ECLAC