



Tercera videoconferencia sobre impacto de COVID-19 en las encuestas de hogares: Avances y desafíos en las Encuestas de Fuerza de Trabajo

11 de junio de 2020, vía Webex

Minuta

En el contexto de las charlas virtuales “Conectados RTC”, el INE de Chile, la División de Estadísticas de CEPAL y la OIT convocaron a las Oficinas Nacionales de Estadísticas (ONE) de América Latina a una tercera videoconferencia para continuar con el seguimiento de temas metodológicos y el intercambio de experiencias en relación a las encuestas de hogares ante la contingencia del COVID-19, específicamente las encuestas de fuerza de trabajo.

En esta ocasión, el tema se centró en las recomendaciones y experiencias nacionales para la detección de sesgos y calibración de factores de expansión, así como los resultados obtenidos para los principales indicadores de mercado laboral. Se contó con presentaciones de la División de Estadísticas de la CEPAL, el DANE de Colombia, el INEGI de México, el IBGE de Brasil y el INE de Chile.

1. Bienvenida

Rolando Ocampo, Director de la División de Estadísticas, brindó unas palabras de bienvenida, agradeciendo la participación de las oficinas y el apoyo del INE de Chile y la OIT. Informó que ya estaba disponible el quinto documento metodológico de la serie de publicaciones realizadas por la División de Estadísticas, esta vez en el ámbito de precios y cuentas nacionales. Por su parte, Marcela Cabezas (OIT) destacó el trabajo realizado en conjunto con la CEPAL, señalando que si bien ha sido arduo, los resultados han sido gratificantes. También recordó que se encuentran disponibles las notas técnicas sobre encuestas rápidas y teletrabajo.

2. Presentación de la División de Estadísticas de CEPAL

Andrés Gutiérrez presentó las recomendaciones para la publicación de estadísticas oficiales de encuestas de hogares en medio de la pandemia COVID19, publicadas en dos documentos de la CEPAL. Se refieren a la recopilación de información sobre el mercado laboral a través de encuestas de hogares, prestando especial atención a la estrategia de muestreo (la combinación de diseño de muestreo y estimadores) requerida para este propósito.

La suspensión de la recopilación de datos cara a cara para encuestas de hogares ha llevado a cambiar la modalidad de recopilación a una modalidad telefónica. Para este fin, la recomendación es seleccionar una muestra de hogares de un período anterior para el cual los datos de contacto estén disponibles.

Sin embargo, este proceso puede estar sujeto a diferentes sesgos (de selección, cobertura y falta de respuesta). Se originan a partir de situaciones tales como: no todos los hogares de la muestra original

proporcionaron su información de contacto telefónico; algunos hogares proporcionaron su información de contacto, pero al 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; y no todos los hogares que proporcionaron su información de contacto están dispuestos a responder el cuestionario de la encuesta

En esta situación, el representante de la CEPAL mostró algunos ejercicios para estimar el efecto del sesgo y corregir el sesgo de no respuesta a través del “propensity score” y los métodos de calibración en dos etapas.

Comentarios

Desde Colombia consultaron qué se puede hacer en el caso que no se disponga de un panel; cuál podría ser el procedimiento para corregir. La respuesta por parte de CEPAL es que estas recomendaciones están basadas en el seguimiento a un panel. Pero en caso de no tener un panel, y sí un censo reciente, como es el caso de Colombia, pueden hacer simulaciones con el censo.

Desde el INEGI se comentó que, considerando que se trabaja con muchas variables, un modelo de calibración no puede responder a todo. La experiencia de México es que, aún realizando las calibraciones necesarias, hay desagregaciones que se ven afectadas y no es posible publicar de manera oficial.

Ecuador realizó una pregunta específica respecto de qué parámetros se deben tomar en cuenta para el seguimiento del panel que están efectuando para los meses de mayo y junio que proviene de levantamientos anteriores del año 2019. Específicamente, si se debe reemplazar o seguir con el mismo que se venía trabajando. Desde CEPAL se indicó que habría que analizar cuál es el desgaste que va teniendo el panel. La OIT ofreció la posibilidad de tener reuniones específicas con Ecuador para analizar su caso.

3. Presentación del DANE de Colombia

Gabriel Osorio presentó tres tipos de sesgos detectados en la encuesta de mercado laboral, referidos a la temporalidad, selección de hogares y estructura de los hogares. Con respecto al primero, indicó que se trata del peso que tienen las semanas en la encuesta: las encuestas presenciales se suspendieron en la última semana de marzo, con lo cual se perdió una semana de recolección, lo que generó un sesgo estacional. Por su parte, el sesgo de la selección de hogares se refiere al sesgo ocasionado por los hogares que atendieron la encuesta en el recuento y ofrecieron su número de teléfono. En ese sentido, estos hogares que otorgaron su número no necesariamente tienen características similares a los que no atendieron la encuesta. Y los hogares que no dieron el teléfono posiblemente no tenían personas desempleadas o inactivas. En cuanto al sesgo en la estructura de los hogares, se refiere al sesgo generado por encuestas incompletas.

Con respecto a la estrategia que siguieron para controlar estos sesgos, el presentador señaló que en el caso del sesgo temporal, se hizo representar la segunda semana de referencia con la primera, y se ajustaron las

demás semanas para que tuvieran pesos iguales. En cuanto al sesgo de selección de hogares, se corrigieron a nivel de segmento el número de hogares completos que posiblemente se perdieron. A posteriori se calibraron por estructura demográfica; y se revisó el estrato socioeconómico y la estructura de personas por hogar. Si se evidencia un sesgo se ajusta. Por su parte, en el caso del sesgo de estructura de los hogares, se definió en la base un número de variables suficiente para asegurar que la encuesta se definiera completa y se asegurara que la base final tuviera completas las variables para todos los miembros del hogar.

Geovanni Portilla dio a conocer los resultados obtenidos para los meses de marzo y abril, en donde es posible observar un aumento de casi dos puntos porcentuales en la tasa de desempleo para el mes de marzo, respecto del año anterior, y casi 10 puntos más para el mes de abril.

4. Presentación del INEGI de México

Octavio Heredia hizo una presentación sobre la Encuesta Telefónica de Ocupación y Empleo (ETOE). Esta encuesta se diseñó para monitorear el mercado laboral durante el periodo de contingencia sanitaria, y apoyar la toma de decisiones, el diseño y seguimiento de las políticas laborales. Se trata de una submuestra de viviendas de la Encuesta Nacional de Ocupación y Empleo (ENOE) del primer semestre de 2020, donde se cuenta con un número de contacto telefónico.

Indicó que debido a que la submuestra de ETOE son las mismas de ENOE, conservan sus factores de expansión, los cuales son ajustados a través del colapsamiento de estratos, es decir, genera nuevos estratos y son construidos con el objetivo de corregir la subestimación de la población de interés y demás variables. Estos pseudo-estratos son construidos de modo que contengan al menos dos UPM, los que a su vez deberán contener suficientes viviendas en su interior para mantener los factores de expansión sin valores atípicos.

Los factores de expansión se ajustan por no respuesta atribuida al informante a nivel de estrato, en cada uno de los dominios. Posteriormente, se corrigen por técnicas de calibración, a fin de asegurar que en cada dominio de interés de la encuesta se mantengan los totales poblacionales determinados por ENOE del primer trimestre de la 2020.

En cuanto a las etapas de calibración de la ETOE, en una primera etapa de calibración, se ajustaron las estructuras demográficas de la submuestra por edad, sexo y nivel de instrucción a lo que estima el primer trimestre de 2020 de la ENOE. En una segunda etapa de calibración, se ajustó a las variables: población total, población ocupada y población subocupada por posición en la ocupación, población desocupada, así como población no económicamente activa. En la última etapa se calibró de forma implícita a la Población Económicamente Activa y a la Población de 15 años y más. Esta información se obtuvo de las viviendas de la ENOE del primer trimestre de 2020.

Finalmente, dio a conocer las cifras, en donde a partir de la encuesta ETOE de abril, se evidencia un aumento en la tasa de desocupación respecto de los resultados de la ENOE de marzo.

5. Presentación del IBGE de Brasil

Cimar Azeredo comenzó dando a conocer las cifras de algunos indicadores de mercado laboral, en donde se evidencia un aumento de la tasa de desocupación del trimestre feb-mar-abr, respecto del correspondiente a nov-dic-ene, a partir de la Encuesta Nacional de Muestra de Hogares Continua (PNAD en portugués).

Por otro lado, se refirió a la encuesta PNAD COVID-19, la que tiene por objetivo estimar el número de personas con síntomas asociados a la enfermedad y cuantificar algunos de los impactos de la pandemia en el mercado laboral brasileño. Esta encuesta se basó en la muestra de hogares de la PNAD Continua del primer trimestre de 2019.

La muestra total de la encuesta se dividió en 4 semanas, para obtener resultados semanales. La planificación de la encuesta es del tipo panel, y la muestra de la encuesta se mantiene fija durante los meses que se repite la encuesta. Señaló que en un principio, la intención es repetir la investigación durante cuatro meses, que puede extenderse dependiendo de la duración de la pandemia.

Especificó que para la asignación de hogares de muestra por semana de recolección, se utilizó un método de muestreo de hogares estratificado simple, con asignación equitativa, y describió los procesos de ponderación de la muestra y calibración de peso para totales de la población.

Para la corrección de falta de respuesta mejorada, señaló que la estrategia es ajustar los modelos de regresión logística para la propensión a responder. Considerando las propensiones estimadas por tales modelos, entonces será posible calcular los factores de corrección de la falta de respuesta con un mejor rendimiento para reducir los posibles sesgos debido a la falta de respuesta diferencial potencialmente observada en la investigación.

6. Presentación del INE de Chile

Iván Touron indicó que realizaron ejercicios para conocer los posibles impactos que tendría la crisis sanitaria sobre la Encuesta Nacional del Empleo. Estos ejercicios fueron en dos vías: por un lado, se hizo un análisis de aproximación del sesgo, y por otro, un análisis de caracterización de la muestra para la que se tenía información de contacto.

Con respecto a la primera línea, se realizaron ejercicios sobre el último trimestre de recolección en condiciones normales (dic-ene-feb). Para ene-feb-mar, se aproximó el sesgo en los indicadores principales, simulando la pérdida muestral evidenciada en marzo, sobre el mes de febrero que pertenece al trimestre de dic-ene-feb. Como resultado se obtiene una muestra DEF' afectada de manera similar al trimestre ene-feb-mar, cuyas estimaciones pueden compararse con el resultado oficial del trimestre dic-ene-feb. Para este proceso se simuló 100 posibles muestras sobre febrero. Para feb-mar-abr se continuó el ejercicio anterior sobre DEF', pero considerando en enero las mismas unidades respondientes de abril. Como resultado se obtuvo el denominado DEF''.

En cuanto al análisis de caracterización de la muestra con información de contacto, se comparó la distribución de la muestra según las subpoblaciones vinculadas con el mercado laboral, para la muestra

total y para la que se dispone de información de contacto. Se calcularon para las variables y categorías de calibración: sexo, edad, región, población extranjera, área urbana o rural. Una comparación de las estimaciones principales entre la muestra total y la muestra con información de contacto. Y la relación entre disponibilidad de información de contacto y la probabilidad de estar ocupado.

David Niculcar dio a conocer algunos indicadores de mercado laboral. Señaló que también hubo un aumento de la desocupación. Entre los desafíos futuros, destacó la necesidad de profundizar en el análisis de la presencia de las subpoblaciones de interés dentro de la muestra para la que se dispone información de contacto; la construcción de la metodología para el tratamiento de la pérdida de cobertura y la no respuesta telefónica; el análisis y publicación de nuevas preguntas sobre impactos en los ingresos del trabajo; y el análisis de matrices de transición.

Comentarios

A modo de cierre, Andrés Gutiérrez (CEPAL) enfatizó que con las desagregaciones, aun cuando se utilicen calibraciones, se pierde precisión y el intervalo de confianza puede ser mucho más amplio, por lo que es necesario tener presente las limitaciones de las encuestas. La recomendación de CEPAL es que si no se tiene una cobertura suficiente, no se publiquen desagregaciones y se entreguen cifras solamente a nivel nacional.

Por su parte, Marcela Cabezas (OIT) agregó que es importante comunicar al público usuario respecto de que las desagregaciones se podrían haberse visto afectadas.

Rolando Ocampo (CEPAL) remarcó que el trabajo de las ONE en estos momentos es fundamental para medir la catástrofe que esta pandemia está generando en el empleo, y felicitó a las oficinas por el nivel de información que están entregando para la toma de decisiones.