

## **Implementación de la Muestra Probabilística**

Con la publicación de los datos preliminares de Enero 2014 en Marzo 2014, el programa de Estadísticas Sobre Empleo (CES, por sus siglas en inglés) completará la implementación de su muestra probabilística para todas las industrias del sector privado en Puerto Rico. La muestra probabilística es el modelo reconocido internacionalmente para muestras de encuestas. Anteriormente el programa CES utilizaba la técnica de muestreo por cuota, la cual estaba sujeta a sesgos significativos. La muestra probabilística garantiza la representación adecuada del universo de empresas no agrícolas a través de técnicas de selección aleatorias.

El rediseño de la muestra que Puerto Rico está adoptando ya se completó en los 50 estados y Washington, D.C. con la publicación de los estimados preliminares de Enero 2003 en Marzo del mismo año. Los objetivos principales de este rediseño fueron el desarrollar técnicas de muestreo y de estimación basadas en probabilidad, producir medidas de precisión uniformes e intervalos de confianza para la encuesta, y mejorar los métodos para estimar la apertura y cierre de establecimientos. La introducción de la muestra probabilística de CES está acompañada de técnicas mejoradas para el registro de nuevos patronos.

La muestra probabilística de CES está diseñada a nivel de los estados, lo cual minimiza la varianza en los estimados del empleo total en el sector privado no agrícola. Es una muestra aleatoria simple, estratificada por industria y tamaño, y agrupada por número patronal del programa de Seguro por Desempleo (UI). El marco de muestreo y la muestra en sí son actualizados dos veces al año, pero con un tiempo de rezago, a medida que los nuevos registros trimestrales de UI están disponibles.

## **Metodología de Estimación**

### *Estimador del Enlace Relativo Ponderado de CES*

Un estimador mejorado fue desarrollado para el programa CES para respaldar el diseño de la muestra probabilística. La diferencia principal del antiguo estimador de CES es la aplicación de un peso (valor) a cada unidad de la muestra en el proceso de estimación. Los pesos se obtienen a través de fracciones de muestreo de la población y son una característica estándar de los estimadores para muestras probabilísticas. El peso de una unidad es el inverso de la probabilidad de ser seleccionada en la muestra. El nuevo estimador para las series de empleo total es un estimador racional, conocido como el Enlace Relativo Ponderado.

### *Procedimiento de Estimación Robusta*

El programa CES utiliza el concepto de muestra pareada para producir los estimados. La muestra pareada está compuesta por todas las unidades de la muestra que hayan reportado datos para ambos meses, es decir, el mes de referencia y el anterior. En ocasiones la muestra pareada contiene un número reducido de observaciones que pueden tener un efecto grande y adverso en el estimado del cambio relativo. La influencia de esas observaciones puede atribuirse a unidades en la muestra con mucho peso, un cambio inusual en el nivel de empleo reportado, o al efecto combinado de estos factores. Si no se le aplica un ajuste, observaciones influyentes pueden causar cambios exagerados en los estimados mensuales, especialmente a niveles detallados para publicación. El procedimiento de Estimación Robusta está diseñado para reducir el efecto de observaciones influyentes en el estimado del cambio relativo mensual. A su vez, se reconoce que cambios inusuales e influyentes en la muestra pueden representar un comportamiento similar en la población de interés y la intervención desmedida en el proceso de estimación puede llevar a estimados parcializados. Esto se puede observar, especialmente, si la muestra es grande. Es por esto que el estimador está diseñado para reducir la volatilidad de los estimados causada por unidades (reportes) con cambios extremos, mientras controla la intervención para evitar que se produzcan estimados parcializados.

### *Modelo de Aperturas/Cierres Netos de CES*

Otro cambio en la metodología insertado o implantado en la muestra probabilística es el modelo de aperturas/cierres netos, utilizado para estimar la porción residual de nuevos empleos que no se contabilizan en la muestra. Existe un retraso inevitable entre la apertura de un nuevo establecimiento y su aparición en el marco de muestreo para que pueda ser seleccionado. Debido a que las nuevas firmas generan una porción de crecimiento en empleo mensualmente, se deben utilizar métodos no-muestrales para estimar este aumento de empleo. Investigaciones previas han indicado que aunque la porción del empleo total relacionada a la apertura y cierre de empresas es generalmente significativa, la contribución neta es relativamente pequeña y estable. Para contabilizar esta porción del empleo neto correspondiente a las aperturas y cierres, el BLS utiliza un procedimiento de estimación con dos componentes: el primer componente excluye del estimado la pérdida de empleos atribuidas a los cierres de negocios, para así compensar los aumentos en el empleo relacionados a la apertura de nuevos negocios. Esto se incorpora en el proceso de estimación al simplemente no reflejar las unidades que están cerrando, sino imputándolas al atribuirles la misma tendencia que las otras firmas en la muestra. Este paso representa la mayoría del empleo correspondiente a los nuevos negocios.

El segundo componente es un modelo de serie de tiempo ARIMA diseñado para estimar el empleo residual relacionado a la apertura/cierre de negocios que no ha sido tomado en cuenta en el proceso de imputación. La serie de tiempo histórica usada para crear y probar el modelo ARIMA provino de la base de datos universal, a nivel micro, del programa de seguro por desempleo (UI), y refleja el residuo actual de las aperturas/cierres de negocios por los pasados 5 años.

#### *Modelo Fay-Herriot de CES*

El programa CES utiliza un estimador basado en el modelo de Fay-Herriot (FH) para estimar el empleo por sectores a nivel Isla para aquellas series que no cumplen con el criterio de publicación basado en tamaño de muestra. El modelo de FH es un promedio ponderado del estimado basado en la muestra combinado con un pronóstico ajustado ARIMA. La estimación del nivel de empleo se obtiene aplicando el cambio relativo estimado al nivel de empleo del mes anterior.

#### *Modelo de Dominios Pequeños de CES*

El BLS ha establecido para publicación una cantidad mínima de industrias para cada estado, territorio y Áreas Estadísticas Metropolitanas (MSA). Además del número mínimo de series garantizadas, series de empleo adicionales son publicadas si contienen suficiente muestra. Las series de empleo garantizadas que no cumplan con el criterio de tamaño de muestra para su publicación se estiman con un modelo de regresión conocido como 'Modelo de Dominios Pequeños de CES (SDM)', desarrollado por el Negociado de Estadísticas del Trabajo Federal. El SMD es un modelo de mínimos-cuadrados ponderados compuesto por: (1) un estimado basado en la muestra disponible para la serie en cuestión; (2) una proyección ARIMA de la tendencia del universo basada en 10 años de datos históricos del Censo Trimestral de Empleo y Salarios y (3) un estimado con la tendencia a nivel Isla para la industria en cuestión para ser aplicado en las Áreas Estadísticas Metropolitanas. Los pesos (valores) para cada uno de estos componentes son calculados mensualmente. En adición a las series de empleo garantizadas para publicación, el SDM se usa para estimar las series de industrias a dos dígitos NAICS que no tengan suficiente muestra para estimados directos.

#### *Modelo de Fay-Herriot Modificado de CES*

Las series de empleo garantizadas para publicación para las Áreas Estadísticas Metropolitanas (MSA) que no cumplan con el criterio de tamaño de muestra, se estiman utilizando un modelo Fay-Herriot modificado (MFH). El modelo MFH estima el cambio relativo en el empleo con un promedio ponderado del estimado basado en la muestra y un valor promedio histórico ajustado. El nivel de empleo se obtiene aplicando el cambio relativo estimado por el modelo al nivel de empleo del mes anterior.

## **Ajustes Estacionales**

Los datos de empleo y salario de CES para Isla y Áreas son ajustados estacionalmente por el método de dos pasos. El BLS utiliza el programa ARIMA X-12 para remover el componente estacional a los cambios mensuales del empleo. Este proceso utiliza las tendencias estacionales del total de empleo en el censo para ajustar los datos de empleo históricos ya revisados, a la vez que incorpora las tendencias estacionales relacionadas a la muestra para ajustar los estimados de empleo de la misma. Al tomar en cuenta los diferentes patrones estacionales encontrados en los datos de empleo históricos revisados y los estimados de empleo basados en la muestra, esta técnica produce mejores series estacionalmente ajustadas con respecto al análisis del cambio mensual en el empleo.

## **Confiabilidad de los estimados**

Los estimados de empleo, horas y salarios de la muestra son basados en una encuesta, en datos administrativos y en modelos estadísticos, por lo cual están sujetos a errores de muestreo y de otros tipos. (El error de muestreo es una medida de variabilidad muestral – esto es, la variación asociada al hecho de encuestar una muestra en vez de la población completa). Los datos de la encuesta también están sujetos a errores no relacionados al proceso de muestreo, tales como los que se pueden introducir mientras se recopilan y procesan los datos. Los estimados no derivados directamente de encuestas muestrales están sujetos a errores adicionales relacionados con el proceso de estimación utilizado. Las medidas de error muestral (desviaciones estándar) para los datos de CES de Empleo Asalariado No Agrícola y Super Sectores a nivel Isla y de Empleo Asalariado No Agrícola a nivel de las Áreas están disponibles en [www.bls.gov/sae/790stderr.htm](http://www.bls.gov/sae/790stderr.htm). El BLS utiliza un intervalo de confiabilidad de 90% para determinar si los cambios en el nivel de empleo son estadísticamente significativos para cambios de 1, 3 y 12 meses.

Para controlar ambos errores, el de muestreo y el de no-muestreo, los estimados de CES son revisados anualmente utilizando como referencia al universo de empleo. Información de revisiones recientes para los estados está disponible en la página electrónica de BLS [www.bls.gov/sae/](http://www.bls.gov/sae/).