Ficha técnica Encuesta Anual de Actividad Económica Año 2019

La presenta nota aborda únicamente los aspectos metodológicos que tienen que ver con el diseño muestral de la encuesta del año 2019. El resto del marco conceptual utilizado se puede encontrar detallado en el documento "Metodología Encuesta Anual de Actividad Económica 2013 y 2014", publicado en http://www.ine.gub.uy

1. Introducción

Fueron elegibles para participar en la Encuesta Anual de Actividad Económica del año 2019 (EAAE 2019) todas las empresas formales del país[1] con diez o más personas ocupadas[2] o con un promedio mensual de ventas mayor a $U 160.670.500[3] y su actividad económica principal, según la CIIU REV 4, se encuentra comprendida en las secciones de actividad B, C, D, E, G, H, I, J, K, L, M, N, P, Q, R y S.

Para los cálculos finales se agregaron aquellas empresas que siendo usuarias de alguna zona franca del país, tuvieran un VBP anual mayor a $U 160.670.500 o que realizaran alguna clase de actividad perteneciente a la industria manufacturera, según la CIIU REV 4.

2. Diseño muestral

La EAAE es una encuesta de tipo de panel. La muestra para el año 2019 se encuentra compuesta por todas aquellas empresas elegibles en la EAAE 2018 que continúan cumpliendo los criterios de elegibilidad en 2019, más un subconjunto de empresas nuevas (*refresh*) que fueron seleccionadas en base al directorio de empresas del año 2019. El *refresh* tiene como objetivo poder seguir brindando estimaciones confiables para el año 2019, mientras reduce el impacto producto de la atrición que sufren todos los paneles. Las empresas incluidas en el refresh fueron seleccionadas utilizando un diseño muestral similar al utilizado para la selección las empresas incluidas en los años anteriores (panel). Dicho diseño es aleatorio, directo y estratificado.

Los estratos están definidos a partir de dos niveles de información. En un primer nivel, las empresas son clasificadas (agrupadas) de acuerdo a su actividad económica principal a

[1] A excepción de las empresas que operan en zonas francas.
[2] Según datos del Banco de Previsión Social.
[3] Según datos de la Dirección General Impositiva.

nivel de CAE[4]. En un segundo nivel, dentro de cada CAE, las empresas son segmentadas en tres estratos de tamaño:

- Estrato 1: comprende a las empresas con personal ocupado entre 10 y 19.
- Estrato 2: comprende a las empresas con personal ocupado entre 20 y 49.
- Estrato 3: comprende a las empresas con Personal ocupado mayor o igual a 50 y/o promedio mensual de ventas mayor a $U 160.670.500.

Las empresas clasificadas en el estrato de tamaño 3 son de inclusión forzosa (censo a priori).

3. Tamaño de muestra teórico y tasa de respuesta

El tamaño de muestra teórico para el año 2019 se situó en 4.300 empresas, de las cuales 568 son empresas nuevas (refresh). Por otra parte, la tasa de efectividad se situó en 79% aproximadamente, lo que se traduce en una muestra efectiva (elegibles respondentes) de 3.417 empresas.

4. Ponderación

Los datos de la muestra son ponderados para de obtener estimaciones tanto a nivel de la población total como para otros dominios de estudio.

El ponderador final para la empresa i viene dado por:

$$w_i = d_i \times \mathrm{nr}_i \times \mathrm{cal}_i \times \mathrm{trim}_i$$

donde d_i es el ponderador original o basado en el diseño para la empresa i, nr_i es el ajuste por no respuesta cal_i es el ajuste proveniente de la calibración y trim_i es el recorte o trimming de los pesos finales por la presencia de empresas atípicas.

[4] Clasificación de Actividades Económicas preliminar: clasificación creada por el Banco Central del Uruguay en el proceso de desarrollo del Plan de Adecuación a Estándares Internacionales (PLAE).

Ponderador original

El ponderador original se define como el inverso de la probabilidad de selección de la empresa en la muestra, es decir, viene determinado por el diseño de la muestra. El ponderador de la empresa i que pertenece al estrato de diseño h viene dado por

$$d_{i,h} = \frac{1}{(i \in \text{muestra})} = \frac{1}{\pi_i} = \frac{1}{n_h/N_h} = \frac{N_h}{n_h}$$

donde N_h y n_h es la cantidad de empresas en el marco y en la muestra respectivamente perteneciente al estrato de diseño h .

Ajuste por no respuesta

La existencia de no respuesta implica que los ponderadores originales deban ser ajustados para corregir la subestimación provocada por la misma. El ajuste por no respuesta fue realizado a nivel de estrato y viene dado por el inverso de la tasa de respuesta del estrato. Por lo tanto, el ponderador de la empresa i ajustado por no respuesta perteneciente al estrato h viene dado por:

$$w_{i,h}^{\text{nr}} = d_{i,h} \times \text{nr}_{i,h} = \frac{N_h}{n_h} \times \frac{n_h}{m_h} = \frac{N_h}{m_h}$$

donde m_h es el tamaño de muestra efectivo en el estrato h.

Ajuste por calibración.

En este paso los ponderadores ajustados por no respuesta son calibrados a información conocida a partir del marco muestral, de forma que las estimaciones provenientes de la encuesta coincidan de forma exacta con ella. La información utilizada para la calibración es:

- Total de empresas por sección de actividad
- Total de empresas por estrato de diseño
- Total de empleados (personal ocupado) según sección de actividad.

Una vez definido el set de variables auxiliares para la calibración y los totales poblacionales, se utilizó el *raking truncado* para calcular los ajustes cal_i, en donde se puso como restricción que los ajustes se encontraran comprendidos entre 0.5 y 2.

Recorte o trimming de los pesos finales

Una vez computados los pesos ajustados por calibración, pueden existir empresas muy grandes en términos de VAB o VBP, que tengan ponderadores distintos de 1. En estos casos, a posteriori, los ponderadores calibrados fueron redefinidos igual a 1, debido a que las empresas en cuestión se consideran únicas y por ende deberían representarse solamente a sí mismas. Esto busca evitar estimaciones inestables de los indicadores de interés.

Precisiones

La estimación puntual de los coeficientes de variación[5] para las estimaciones del total de VAB, VBP y PO se encuentra por debajo del 1%.

[5] El coeficiente de variación (CV) es una medida utilizada para describir la precisión de una estimación. Mientras más pequeño sea el CV, más precisa es la estimación. El CV de una estimación, es el ratio entre el Error Estándar de la estimación (SE) respecto a la propia estimación.