

Box 2: Methodological Updates and Primary Changes to the Comprehensive Household Survey (GEIH)

Didier Hermida
Karen L. Pulido-Mahecha*

The Integrated Household Survey (GEIH for its initials in Spanish) serves as the primary source for official statistical labor market information in Colombia. The survey, based on the geostatistical framework incorporated in Colombia's 2005 national census and referred to in this box as the GEIH-M05, has been updated in line with a new national census conducted in 2018. This update takes into account changes in demographics, territorial organization and population distribution in recent years. The new survey, referred to in this box as GEIH-M18, also incorporates recommendations from international bodies, such as the International Labor Organization (ILO) and the Organization for Economic Cooperation and Development (OECD), related to characterization of working-age populations and the statistical visibility of other groups. This update keeps Colombian labor market statistics in line with international standards.

While the GEIH-M18 includes several important updates, particularly regarding the definition of informality, the measurement of labor income, and tax payment data, this box is focused on those changes with the potential to modify the levels and/or behavior of the main labor market aggregates.

1. Primary changes to the GEIH

Data collection under the auspices of the GEIH-M18 began in parallel with the GEIH-M05 in 2021. The GEIH-M18 replaced the previous survey entirely in 2022. Among the main changes was a significant methodological redesign, which most important aspects are described below.

a. Update to the geostatistical framework

The GEIH-M05 was based on a geostatistical framework² from 2005. Demographic and population shifts in the country since then (e.g., new neighborhoods, population density) have precipitated the need for adjustments to the survey in order for it to continue to be representative of the population as a whole. The GEIH sample design has thus been updated based on the 2018 census.

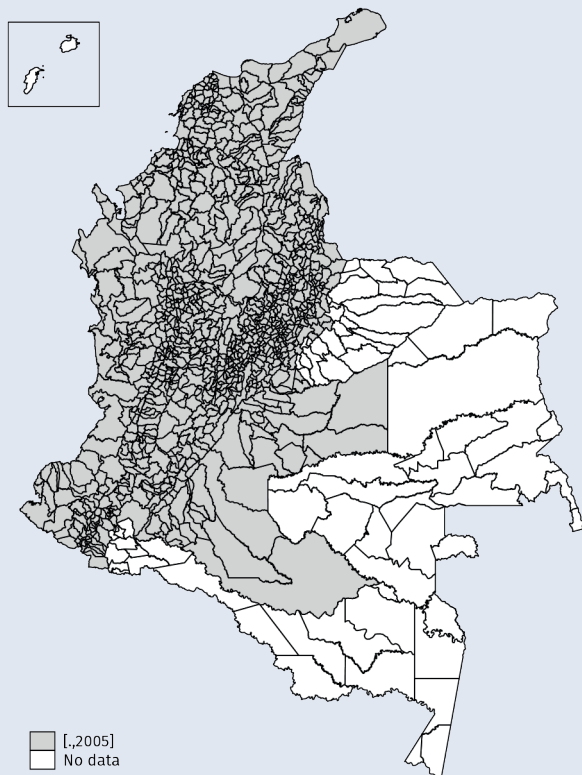
* The authors are members of Banco de la República's Labor Market Analysis Group (GAMLA) and Programming and Inflation Department (DPI). The views and opinions expressed herein do not necessarily reflect those of the bank or its board of directors.

1 This box is based on a special section in *Banco de la República's Labor Market Report*, No. 22. For a more detailed description of the changes and series effects associated with the updated GEIH-M18 framework, as well as for a provisional linkage and seasonal adjustment proposal for the main aggregates of the labor market, see the full report at: <https://www.banrep.gov.co/es/reporte-mercado-laboral>

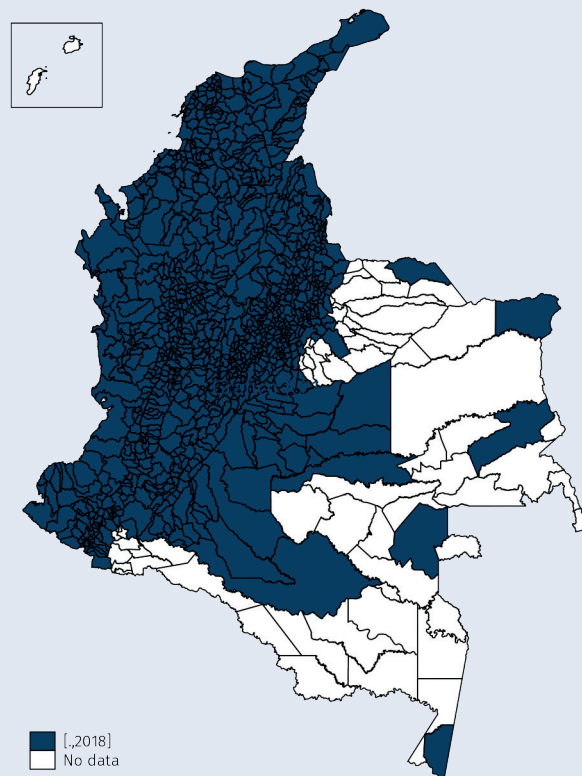
2 It is the sample used in the household surveys which consists of a combination of political-administrative categories and geostatistical areas and is used across all stages of the statistical operation (DANE, 2020).

Map B2.1

A. Coverage – GEIH 2005 framework



B. Coverage – GEIH 2018 framework



Note: monthly data
Source: GANE (GEIH); calculations by Banco de la República.

b. Geographical expansion

The GEIH-05 covered 23 departments plus Bogotá on a continual and monthly basis. This implied that national labor market indicators reflected the aggregate results of the territories covered but not the total number of departments in Colombia (Map B2.1, Panel A). The GEIH-M18 expands the geographical scope of the survey, adding nine department capitals in the Amazonas and Orinoco regions and on the island of San Andrés (Map B2.1, Panel B)³. For its part, coverage of small municipalities and dispersed rural populations remains unchanged from the GEIH-M05.

c. Increase in the working-age population threshold

The GEIH-M05 classified the working age as 10 years or older in small municipalities and rural areas, and as 12 years or older in urban areas. Although this definition was based on characteristics of the Colombian labor market, it contrasted with those stipulated in international agreements related to the fight against child labor. To address this disparity, the GEIH-M18 adopts ILO Convention 138, establishing the working age for young people who have met education requirements. For Colombia, this threshold was set at 15 years, with no discrepancy between urban and rural areas. Data collection on the labor conditions of young people between 5 and 17 years old will continue to be collected in the annual module for child labor.

d. Characterization of the working-age population

The new survey made changes to the characterization of working age population. These were primarily associated with a battery of questions used to classify a person as employed. It also adds questions to provide indicators of the potential labor supply in populations that were previously characterized as inactive, which is now called the population outside the labor force.

First, regarding the employed population, the GEIH-M05 included two types of non-remunerated workers: employees of a family business⁴ and non-remunerated workers who were not related to their employers. Under the GEIH-M18 the latter group is no longer considered part of the employed population, and it is now a requirement to confirm whether these workers receive payment in money or in kind. Employees in family businesses will continue to be classified as employed.

Second, the GEIH-M18 defines the employed population absent from work, in terms of the reasons for their absence and their expectations for returning to work, factors that were not considered in the GEIH-M05. The new survey asks people who are absent from their place of employment to provide the reason for that absence. If this is due to illness, employer permission, temporary leave, holidays, vacations or training, that person is automatically classified as employed. However, if the absence is

3 The nine capitals are: Mocoa, Leticia, Yopal, Arauca, San José del Guaviare, Mitú, Puerto Inírida, Puerto Carreño and San Andrés, increasing overall population coverage from 96.9% to 97.8%.

4 People who in the survey week worked at least one hour in a business belonging to a family member but did not receive payment.

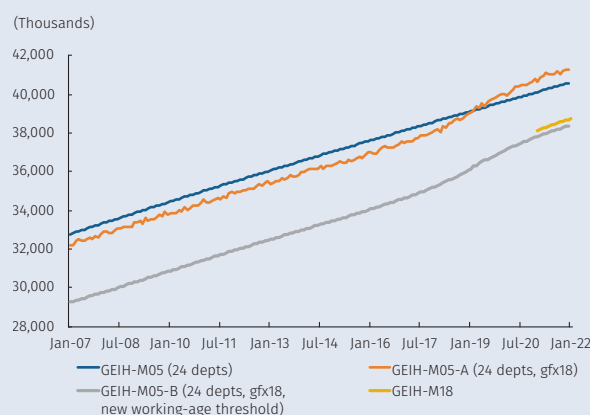
due to other causes, such as a temporary suspension, reduction in economic activity or disruptions caused by a natural disaster, that person's employment status will depend on the expectations for their return to work. If the lapse is expected to be four months or less⁵, the person is considered to be employed. For longer periods, that person could only be classified as employed if they are working for a family business.

Finally, the GEIH-M18 introduced survey questions that allow for the identification of potential participants in the labor market⁶, in line with resolutions from the 19th International Conference on Labor Statistics (ILO, 2013). These questions classify persons out of the labor force either as unavailable job seekers, available potential job seekers, or potential job seekers who wish to work. The primary factor in characterizing this population is its interest in joining the labor force.

2. Main effects of these changes on labor market indicators

Four series were created in order to analyze the effects of the primary changes to the GEIH-M18 on labor market indicators and relevant populations: 1) the official series from the GEIH-M05, with data through December 2021; 2) the GEIH-M05 correcting for expansion factors (GEIH-M05-A), with populations estimated in line with projections from the 2018 census to allow for measurement of the effect of the discrepancy in population projections from the GEIH-M05; 3) the GEIH-M05 corrected for expansion factors and incorporating the new working-age threshold (GEIH-M05-B), which allows for the measurement of the effect of excluding young people between 10 and 14 from the working age population; and 4) the new official series (GEIH-M18) based on labor market data beginning in January 2021. It is important to note that the GEIH-M5 only includes information from 23 departments plus Bogotá, while the GEIH-M18 series also includes the nine new department capitals which have been added to the geographical scope of the survey.

Graph B2.1
Working-age population
National aggregate (Jan 2007 to Jan 2022)



Note: monthly data

Source: DANE (GEIH); calculations by Banco de la República.

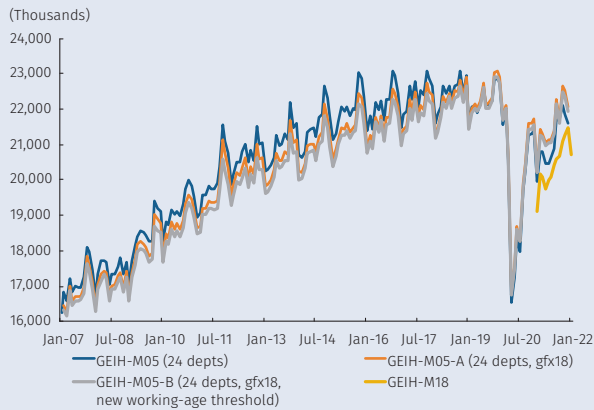
Graph B2.1 shows these series for the working-age population. The first years of the series corrected for expansion factors suggest lower population levels (GEIH-M05-A) due to the fact that results from the 2018 census showed slower demographic growth. However, beginning in 2021, amid a surge in migration to Colombia from Venezuela (Tribin-Urbe et al., 2020; Bonilla-Mejía et al., 2020), growth in the working-age population in this series becomes more significant, even surpassing the original series of the GEIH-M05.

Those series that exclude nearly 2.9 million people between 10 and 14 years old from the working-age population, based on the revised age threshold, naturally show lower levels for this population. This can be seen clearly in the difference between the series that is only corrected for expansion factors (GEIH-M05-A) and the series that is both corrected for expansion factors and incorporates the new working-age population threshold (GEIH-M05-B). The GEIH-M18 incorporates the population from

5 The four-month time horizon established by DANE is based on Article 51 of the Colombian Labor Code, which establishes that labor contracts can be suspended for no more than 120 days due to the temporary closure of the business or other reasons outside employers' control.

6 As a result, the inactive population is now denominated "persons outside the labor force."

Graph B2.2
Employed population
National aggregate (Jan 2007 to Jan 2022)



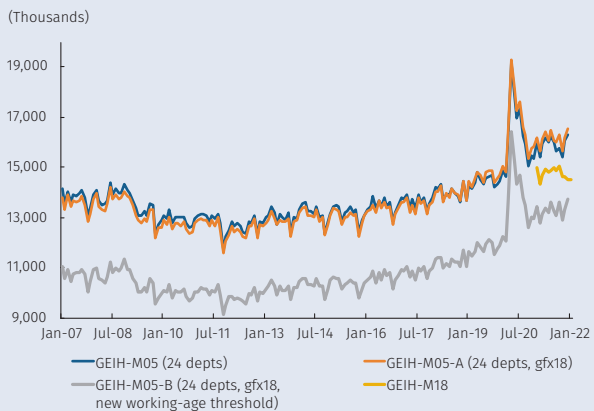
Note: monthly data
 Source: DANE (GEIH); calculations by Banco de la República.

the nine new department capitals and as a result suggests slightly higher levels for the working-age population than the GEIH-M05-B series.

Graph B2.2 suggests that correcting for expansion factors produces a slightly lower level in the employed population (GEIH-M05-A) than in the original series (GEIH-M05). In recent years this difference has been offset by the migration effects mentioned above. The effects of incorporating the change in the working-age threshold are insignificant (GEIH-M05-B), as this excludes less than 1.0% of the population that was previously considered employed, corresponding to young people between 10 and 14 years old.

As seen in Graph B2.2, the new official series for the employed population (GEIH-M18) continues to be lower than the series based on the GEIH-M05. This is despite incorporating more than 179,000 employed people based on the new urban coverage. Although the new survey considers neither absent workers who do not expect to return to work in under four months or non-remunerated workers outside of family businesses as employed, microdata calculations suggest that the effects of these exclusions is low and would not significantly explain the gap between the statistical operations. Controlling for the factors mentioned, the difference in employment levels remains around 1.2 million jobs and could be a reflection of changes in the geostatistical framework. That is to say, the redesign of the sample itself and, in general, the entire statistical operation.

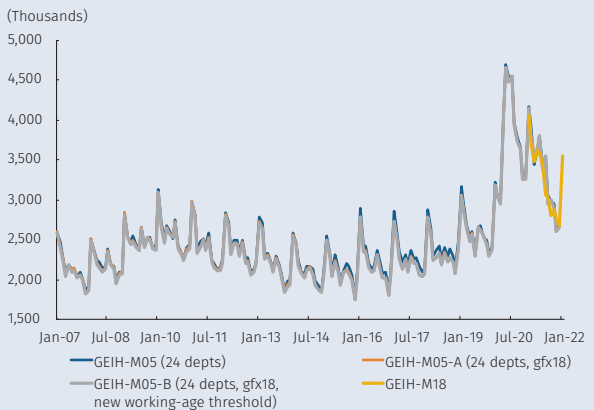
Graph B2.3
Population outside the labor force
National aggregate (Jan 2007 to Jan 2022)



Note: monthly data
 Source: DANE (GEIH); calculations by Banco de la República.

Graph B2.3 reflects changes in the number of persons outside the labor force. In this case, the correction for expansion factors has little impact. The revised working-age threshold, meanwhile, generates a significant decline. On average, the difference is 2.7 million between the series correcting for expansion factors and incorporating the new working-age threshold (GEIH-M05-B) and the series that only corrects for expansion factors (GEIH-M05-A). This is due to the fact that the majority of young people between 10 and 14 are in school but in GEIH-M5 were classified as being outside the workforce (inactive).

Graph B2.4
Unemployed population
National aggregate (Jan 2007 to Jan 2022)



Note: monthly data
 Source: DANE (GEIH); calculations by Banco de la República.

The number of persons outside the workforce in GEIH-M18 is higher and on average suggests an increase of 1.3 million people, of which only 119,000 can be explained by the expanded geographical scope of the survey. The remaining difference would be the result of the updated geostatistical framework, as persons outside the labor force are calculated as those who cannot be classified as either employed or unemployed, and as a result this population under the GEIH-M18 behaves inverse to the employed population.

The unemployed population changes when comparing the different scenarios of the GEIH-M05 (Graph B2.4) and its level estimated under GEIH-M18 is similar for 2021 even after adding 37,000 unemployed people as part of the expanded geographical scope of the survey.

References

- Bonilla-Mejía, L., Morales, M., Hermida, D., Florez, L. (2020). "The Labor Market of Immigrants and Non-Immigrants Evidence from the Venezuelan Refugee Crisis", *Borradores de Economía* No. 1119, *Banco de la República*.
- DANE (2020). *Manual de uso del Marco Geoestadístico Nacional en el Proceso Estadístico*, available at: https://geoportal.dane.gov.co/descargas/descarga_mgn/Manual_MGN.pdf
- DANE. (2022). *Enfoque conceptual y metodológico de la Gran encuesta integrada de hogares (GEIH), Marco 2018*, available at: <https://www.dane.gov.co/index.php/estadisticas-por-tema/mercado-laboral/empleo-y-desempleo>
- ILO (2013). *Resolución I: resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo*, International Labor Organization.
- Tribín-Urbe, A. M.; Adhvaryu, A.; Anzola-Bravo, C.; Ávila-Montealegre, O.; Bonilla-Mejía, L., Castro-Fernández, J. C.; ... Velásquez, S. (2020). "Migración desde Venezuela en Colombia: caracterización del fenómeno y análisis de los efectos macroeconómicos", *Ensayos Sobre Política Económica*, No. 97, No. 176, available at: <https://doi.org/10.32468/espe.97>