The revision of the Labour Force Survey data: the context of the COVID-19 pandemic

Luís Correia [luis.correia@ine.pt], Statistics Portugal
Célio Oliveira [celio.oliveira@ine.pt], Statistics Portugal

1   Abstract

Following the measures adopted by the competent authorities to contain the COVID-19 pandemic, Statistics Portugal decided, between the 1st half of March 2020 and the end of the 2nd quarter of 2022, to suspend the face-to-face interviewing, replacing it exclusively with telephone interviewing. This suspension had an impact on response rates and on the profile of respondents. In order to minimize the impact, the quarterly sample size was increased, and the sampling frame was restricted to dwellings with telephone.

Although Statistics Portugal monitored the impact of suspending the face-to-face interviewing and provided supplementary indicators to monitor the evolution of the labour market, it was decided, at that time and following Eurostat's guidelines, not to introduce any significant additional changes that could lead to incomparability in the estimates at European level and even create a break in series.

The resumption of the mixed-mode data collection from the 3rd quarter of 2022 onwards contributed to a gradual reversal of the response rates to the levels observed in the pre-pandemic quarters. In addition, some year-on-year changes were different from what would be expected. Although other factors related with the dynamics of the labour market could explain some of these changes, it was believed that these were mainly due to the reintroduction of the face-to-face interviewing.

In this context, Statistics Portugal, after investigation of different solutions, decided to revise the estimates from the 2nd quarter of 2020 to the 2nd quarter of 2023 by incorporating the variable "educational attainment" in the calibration weighting, in addition to the information usually used (monthly estimates of the resident population by sex, age group and region). Also, estimates for the 3rd quarter of 2023, the last quarter affected by the suspension of face-to-face interviewing, were obtained under the same methodology.

This paper aims to focus on the methodological solution of using the educational attainment in the calibration weighting, and on the impact of the revision regarding the results on the population aged 16 to 89 by labour status (employment, unemployment, and inactivity). Dissemination and communication strategies to the users are also presented.

2   The context of the COVID-19 pandemic and its impact on the Labour Force Survey

As a result of the COVID-19 pandemic, Statistics Portugal made a crucial change in the Labour Force Survey (LFS) data collection mode. In response to health concerns and the public health policies adopted by the competent authorities, face-to-face interviewing was suspended, and telephone

interviewing became the exclusive mode of data collection.[1] This change followed the same measure adopted by other European Union countries and took place in the first half of March 2020 and lasted until the end of the 2nd quarter of 2022.

The immediate consequence was the decline in the response rate, despite all measures taken to minimize this effect (Figure 1). This shows the importance of face-to-face interviewing in the 1st wave in updating contacts, among other advantages[2].

Figure 1. Response rates in the Labour Force Survey
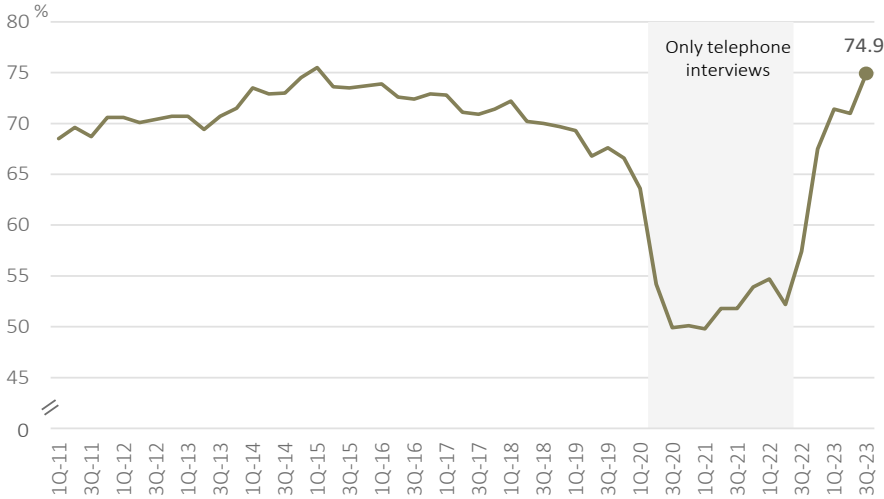
Source: Statistics Portugal, Labour Force Survey.

It should also be noted that, due to the entry into force of the Integrated European Social Statistics Framework in the European Union, in January 2021, a gradual increase in the sample size was introduced between the 1st quarter of 2021 and the 2nd quarter of 2022. This increase corresponded to an oversample of one-third of the initial gross sample size, which made it possible to meet the new precision requirements, as well as to compensate the response rate decline during this period. Currently, the LFS sample includes 30,096 dwellings, and information is obtained on all persons residing in the dwelling.

In this context, Statistics Portugal monitored the impact of suspending face-to-face interviewing on the main labour market indicators, as well on the profile of the respondents. To mitigate the impact, it was decided to restrict the sampling frame to dwellings for which telephone numbers were available. At the same time, studies were carried out on introducing additional information in the calibration weighting, namely the "educational attainment", in addition to the variables usually used (monthly estimates of the resident population by sex, age group and region).

---

[1] Under standard conditions, the information is obtained directly, through a computer-assisted interview conducted by an interviewer, using a mixed collection mode: the initial interview is done face-to-face by an interviewer visiting the dwelling and the other subsequent five are done by telephone if certain requirements are met.

[2] The LFS sample (consisting of private dwellings of usual residence) is divided into six subsamples (waves or rotations), and in each quarter a subsample is replaced by a new one (selected under the same design) after being observed for six consecutive quarters.

The choice of this variable relied on the potential bias introduced by a respondent selectivity effect associated with telephone interviewing, by a decrease in the share of persons (aged 16 to 89 years old) with no education or with primary or lower secondary education (from 56.9% in 2019 to 49.8% in 2022), and an increase in the share of persons with upper secondary (from 23.1% in 2019 to 25.8% in 2022) or tertiary education (from 20.0% in 2019 to 24.5% in 2022).

However, since countries were in a period of great instability, Eurostat issued several methodological guidelines, including not to make any changes in the calibration methodology. This recommendation was intended to avoid introducing additional sources of incomparability in the estimates at the European level and potentially creating a break in the data series.

In these circumstances, during the pandemic period, Statistics Portugal decided to disseminate supplementary indicators in the Press Releases to monitor the evolution of the labour market, always warning when the impacts of the measures to contain the COVID-19 pandemic could interfere with the normal evolution.

3 Estimates revision and underlying methodological approach

3.1 Return to mixed-mode data collection (CAPI and CATI)

Following the positive evolution of COVID-19 health policies, the return to normality was restored. With the resumption of the mixed-mode data collection (face-to-face in the 1st rotation/wave) from the 3rd quarter of 2022 onwards, a gradual reversion on the response rates to the levels observed in the pre-pandemic period took place (see Figure 1). At the same time, unexpected year-on-year changes in some indicators were observed. In May 2023, Statistics Portugal addressed these concerns in the Press Release "Methodological note on the labour force survey data: the context of the COVID-19 pandemic"[3].

Starting in the 3rd quarter of 2022, the year-on-year change of the total and employed population aged 16 to 89 with tertiary education became negative (Figure 2). In the same way, the population outside the labour force with tertiary education showed a trend towards the pre-pandemic levels from the 4th quarter of 2022 onwards. However, the unemployed population aged 16 to 74 with tertiary education showed recurrent fluctuations throughout the data series.

The indicator regarding tertiary educational attainment rate (which corresponds to the share of people aged 30 to 34 with tertiary education) also decreased from the 3rd quarter of 2022. These results were close to the figures observed before the COVID-19 pandemic. Additionally, when comparing the results of this indicator for the year 2021 with those obtained in the Census of that same year, safeguarding the differences between the two statistical operations, there was an overestimation from the LFS (which was not the case in 2011). In 2011, the tertiary educational attainment rate was 28.6% from Census and 26.7% from LFS, 1.9 percentage points (pp) lower; in 2021, the rate was 36.9% based on the Census and 43.0% based on the LFS, 6.1 pp higher.

Although there might be other factors behind these results, it was considered that they were due, in a great extent, to the return to face-to-face interviewing.

---

[3] Available at http://www.ine.pt/xurl/dest/612944288.

Figure 2. Resident and employed (16 to 89) and unemployed (16 to 74) population with tertiary education.
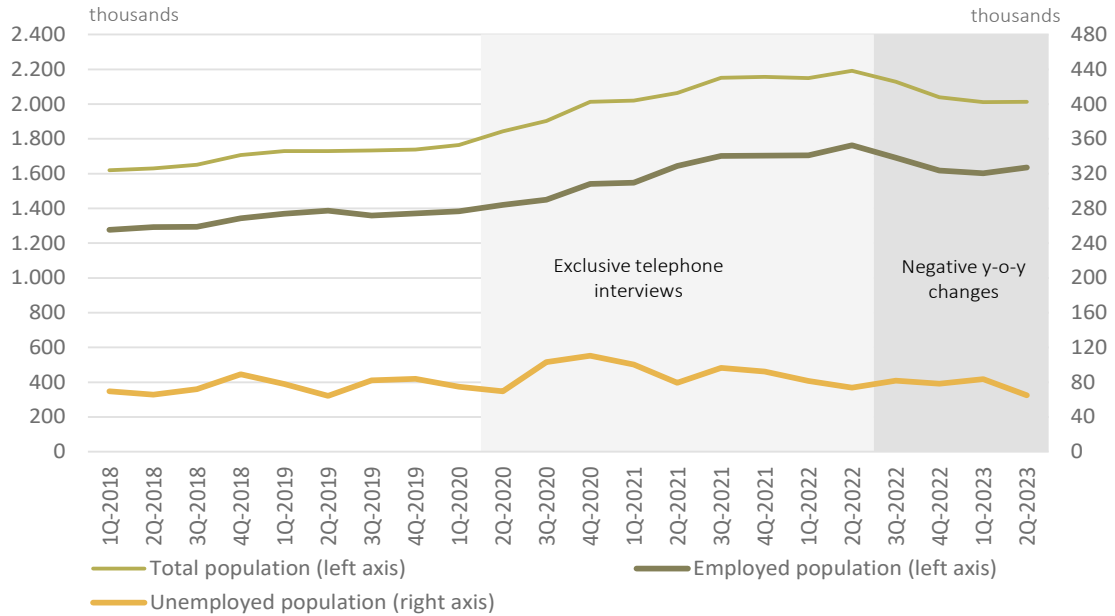
Source: Statistics Portugal, Labour Force Survey.

3.2 Methodological approach

In this context, Statistics Portugal resumed the work of including additional information in the calibration weighting methodology in addition to the information usually used (monthly estimates of the resident population by sex, age group and region). After some simulations it was decided to add only the variable "educational attainment" broken down by NUTS 2, and six educational levels: no level of education, 1st cycle primary education, 2nd cycle primary education, lower secondary education, upper secondary and post-secondary non-tertiary education, and tertiary education.

Given the absence of external data sources on "educational attainment" (only Census 2021 data), the control totals were obtained from the LFS itself by means of a linear regression model using the explanatory variables time and quarter, where the response variable was the estimate of each education level from the 1st quarter of 2011 to the 1st quarter of 2020 (period not affected by the pandemic).

$$Y = \beta_0 + \beta_1 T + \beta_2 Q_1 + \beta_3 Q_2 + \beta_4 Q_3 + e,$$

The Q_1, Q_2 and Q_3 are dummy variables representing the first, second and third quarter and T the sequentially time periods from the 1st quarter of 2011 to the 1st quarter of 2020 (T=1, …, 37).

Although not perfect, this solution was considered the best one at that moment to correct the potential bias introduced by the self-selection effect when moving from a face-to-face interviewing to a telephone based one. It allowed to: i) ensure the trend and some seasonality observed in the quarters prior to the pandemic as illustrated in Figure 3.; ii) obtain predicted values for the 1st and 2nd quarters of 2021 close to those obtained in the Census 2021 and iii) obtain predicted values for the 1st and 2nd quarter of 2023 close to those observed in the 3rd quarter of 2023.

It should be noted that this change in the LFS calibration weighting was an exception and limited in time and covered only the period affected by the suspension of face-to-face interviewing (from the 2nd quarter of 2020 to the 2nd quarter of 2023). In addition, the estimates of the 3rd quarter of 2023 released in November 2023, the last quarter with one wave affected by the suspension of face-to-face interviewing, were also obtained following the same methodology.

Figure 3. Upper secondary and post-secondary non-tertiary education
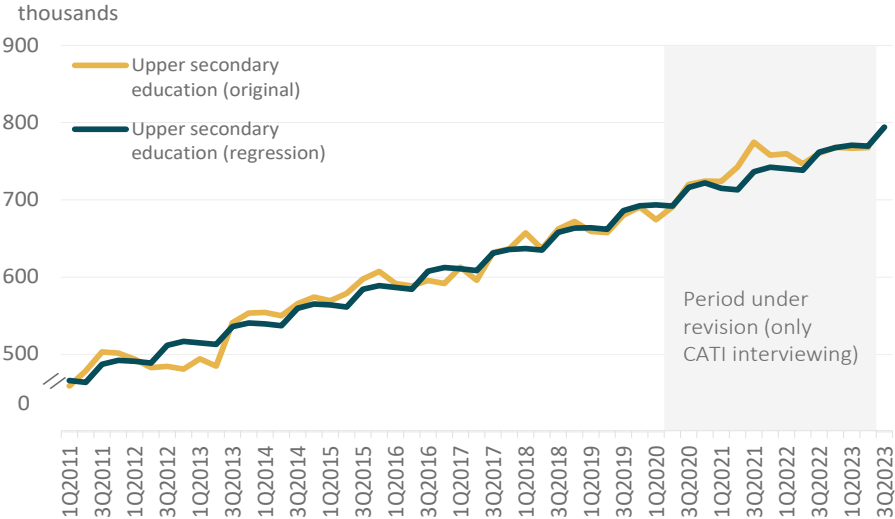
Source: Statistics Portugal, Labour Force Survey.

4   Impact of the revision

The figures and tables disseminated by Statistics Portugal showed the impact of the revision on the population aged 16 to 89 by labour status (in labour force: employment and unemployment; and outside labour force), also restricting the analysis to those with tertiary education.

These figures underlined the methodological note mentioned before (published in May 2023) and show the effect of returning to the mixed-mode data collection. The return to face-to-face interviewing in the 1st wave allowed to increase the response rates to the pre-pandemic levels at the beginning of 2023 (Figure 1), which in turn led the estimates also to the pre-pandemic levels to the point that the absolute differences between the original and revised figures were very small in the quarters of 2023.

Focusing on the population aged 16 to 89 with tertiary education, the revision reversed the signal of the year-on-year changes since 2021 from negative to positive, either in the total population or in the employed population. This result is in line with what was observed in the pre-pandemic period (Figure 2). However, between the 2nd quarter of 2021 and the 1st quarter of 2022, there were negative year-on-year variations in the population outside the labour force aged 16 to 89 with tertiary education, as well as in the unemployed population from the 3rd quarter of 2021, which, as mentioned before, already showed fluctuations throughout the series.

As a consequence of the revision, the indicator for the tertiary educational attainment rate was revised downwards in all periods covered, between 0.1 pp in the 2nd quarter of 2023 and 2.6 pp in the

2nd quarter of 2022, close to the value observed in the 1st quarter of 2020 (37.4%). In the 3rd quarter of 2023, the tertiary educational attainment rate was 39.4%, 1.3 pp more than in the previous quarter and 0.6 pp less than the Europe 2020 target of 40%. Looking at the annual figures, in 2020 the tertiary educational attainment rate was revised downwards (0.6 pp less than the originally published estimate), standing at 39.0%, when the Europe 2020 Strategy target was of 40.0% for that year. Similarly, the estimates for 2021 and 2022 were also revised downwards and now stand at 41.6% and 41.2%, respectively.

Figure 4. Tertiary educational attainment rate
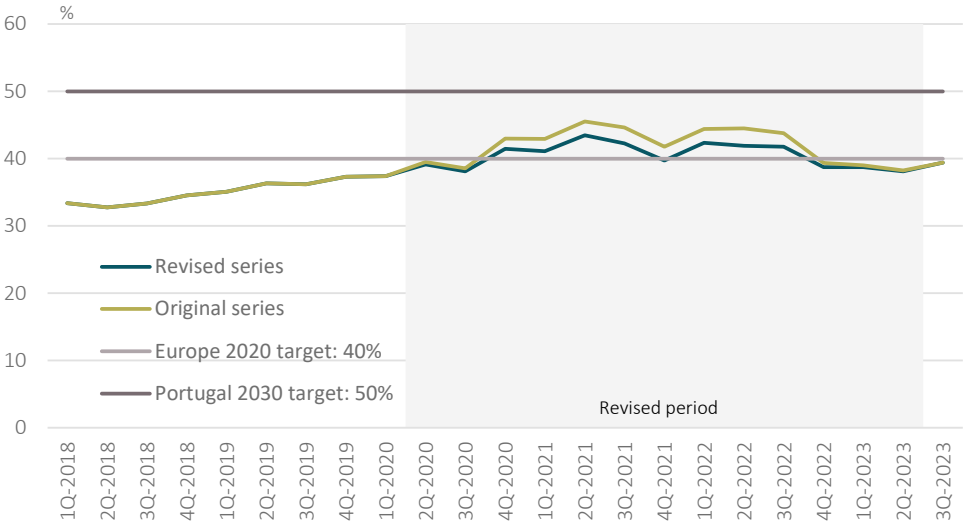
Source: Statistics Portugal, Labour Force Survey.

Analysing the totals for the population in labour force and for the employed population, one can conclude that the first figures were overestimated between the 2nd quarter of 2020 and the 4th quarter of 2022 and underestimated in the first two quarters of 2023. It should be noted, however, that the maximum overestimation was 0.8% (39.9 thousand people) and 1.0% (47.1 thousand), respectively, in the 2nd quarter of 2022, and that the highest underestimation was 0.2% in both populations (9.5 thousand and 9.3 thousand, respectively) in the 2nd quarter of 2023. On the other hand, the unemployed population was revised upwards in the period under analysis, with a maximum difference of 2.1% (7.3 thousand people) in the 2nd quarter of 2021. These revisions were reflected in the unemployment rate, which was underestimated between the 1st quarter of 2021 and the 2nd quarter of 2022, by 0.1 pp and 0.2 pp.
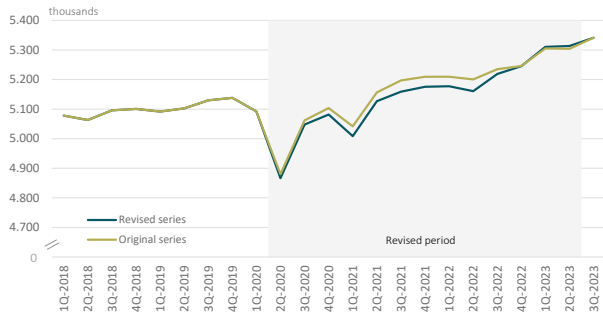
Figure 5. Labour force

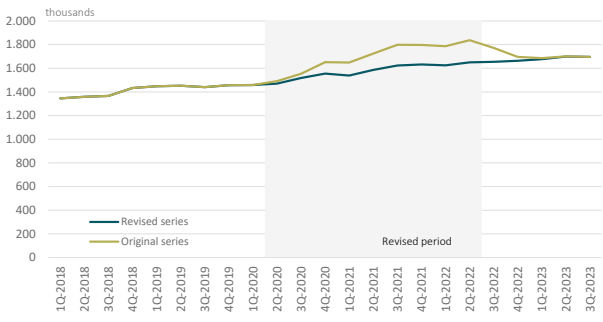## Figure 6. Labour force with tertiary education

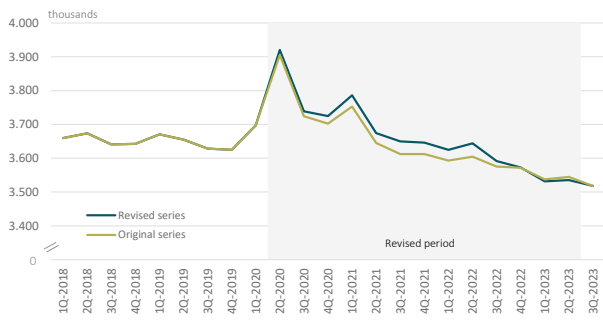## Figure 6. Outside labour force (16 and over)

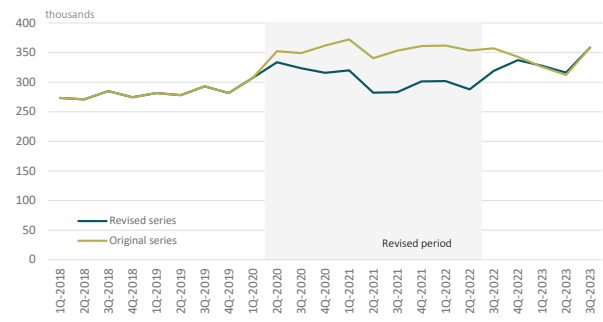## Figure 7. Outside labour force (16 to 89) with tertiary education
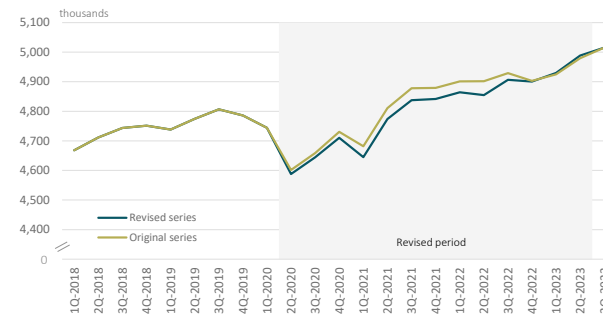

Figure 7. Employed population

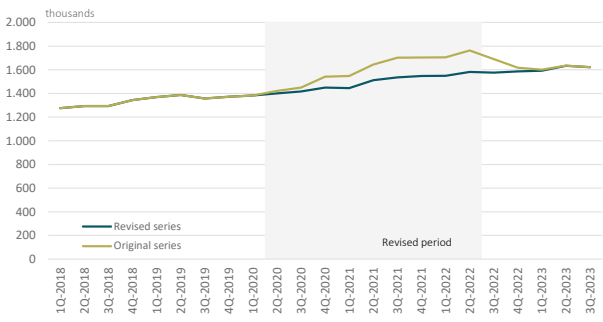## Figure 8. Employed population with tertiary education
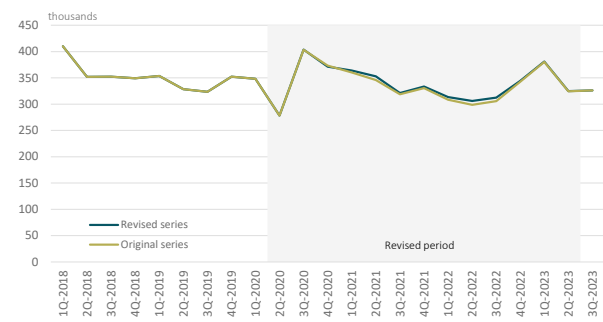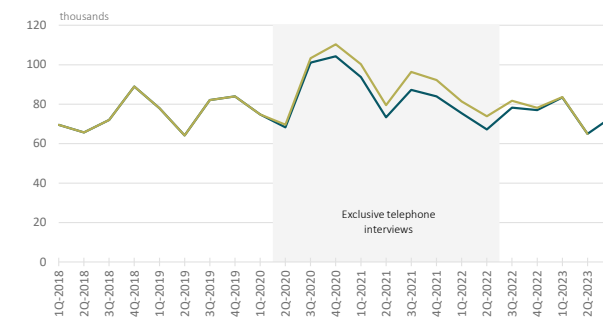

Figure 9. Unemployed population

Figure 10. Unemployed population with tertiary education

Source: Statistics Portugal, Labour Force Survey.
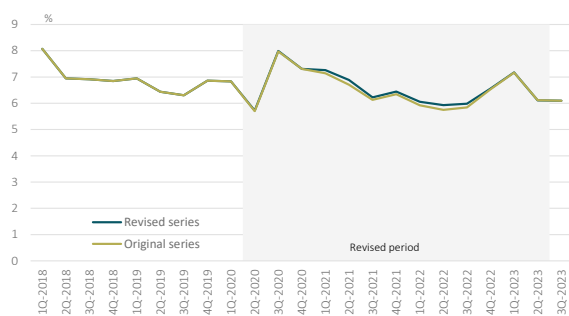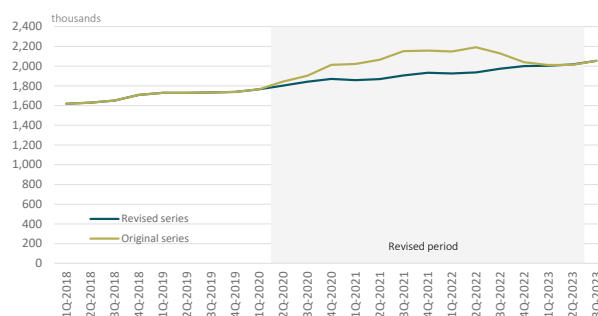
Figure 9. Unemployment rate

Figure 10. Resident population aged 16 to 89 with tertiary education

Source: Statistics Portugal, Labour Force Survey.

5 Brief overview and final remarks

Since the break of the COVID-19 pandemic, Statistics Portugal was aware of the challenges regarding the LFS. To provide a clear picture of the labour market, several changes and studies were made to minimize and assess the impact on estimates. Following Eurostat's guidelines, an effort was done to maximize the response rate by restricting the sampling frame only to dwellings with telephone numbers while maintaining the calibration weighting unchanged. Nevertheless, throughout the pandemic, monitoring and communication were in place to explain changes in the labour market, as well the challenges posed with the suspension of face-to-face interviewing.

The return to face-to-face interviewing in the 2nd quarter of 2022 resulted in an increase of the response rate to pre-pandemic levels. However, unexpected year-on-year changes in some key indicators regarding educational attainment were observed. Namely, the decrease of the employed population with tertiary education (-5.1% in the 4th quarter of 2022 and -7.3% in the 2nd quarter of 2023).

To correct the potential bias introduced by telephone interviewing, the estimates from the 2nd quarter of 2020 to the 2nd quarter of 2023 were revised, by adding the educational attainment to the calibration weighting. Population counts on this variable were obtained by means of linear regression model using data from the LFS from the 1st quarter of 2011 to the 1st quarter of 2020. In addition to the quarterly estimates, the annual and monthly estimates were also revised for the same period.

The revised estimates were released along with the results of the 3rd quarter of 2023 on 8 November 2023 (Press Release - 3rd quarter of 2023). [4] Together with this Press Release, a methodological note explaining the context, methodology and impacts of this revision was also made available. At national level, the LFS statistical indicators published in the Statistics Portugal Database have also been updated, as well as the anonymised microdata for research purposes.

---

[4] Available at http://www.ine.pt/xurl/dest/593942640.