The French Labour Force Survey is overhauled in 2021: rationale behind the renovation and impact on the measurement of employment and unemployement

The Labour Force Survey is crucial for monitoring the labour market. It has undergone a major overhaul in 2021. In compliance with the new European regulation on social statistics, which came into force at the beginning of the year, INSEE renovated the survey questionnaire to improve European harmonization, but also to better meet the needs of its users and the new challenges of the labor market. At the same time, INSEE modernized the survey protocol: since the beginning of the year, after an initial face–to–face interview, INSEE offers the possibility of answering on the Internet for reinterviewing, in addition to the telephone. INSEE also revised the weighting method.

These changes were likely to modify the level of the indicators produced by the survey. In order to estimate the induced breaks, a vast methodological "Pilot survey" was conducted throughout the year 2020 and in the first quarter of 2021. The pilot survey allowed estimating backward–looking "break free" series of the main economic indicators, in order to continue to shed light on the labor market over the long term.

All in all, the employment rate for people aged 15 to 64 is raised by 0.8 points in 2020, while the unemployment rate is unchanged; the share of the so–called "halo around unemployment" among people aged 15 to 64 is raised by 0.8 points. The trends in all the break–adjusted indicators are virtually unchanged over the past, especially during the health and economic crisis of 2020.

Although its uses go far beyond this, the primary objective of the Labour Force Survey is to measure employment and unemployment according to the concepts defined by the International Labour Office (ILO). These concepts are based on factual definitions of employment and unemployment, independent of unemployment benefit schemes or labor market legislation. They thus provide a measure that is stable over time and comparable with other countries, insofar as most of the world's statistical institutes, particularly those of the European Union, use them.

As the French version of the European Labor Force Survey, this survey has been part of a European regulatory framework since 1973. This framework specifies the information that must be collected (including the questionnaire). It also specifies certain methodological

aspects, such as having a sufficiently large sample covering all the weeks of the year; or having a panel survey, i.e., the households are surveyed several times with reinterviews at regular intervals; or finally certain aspects of the protocol, for example, limiting the maximum collection time to five weeks. Other aspects of the survey are left to the discretion of the countries, which are free to use collection methods mediated by an interviewer (by telephone or face–to–face) or not (by internet or on paper).

Conducted since 1950, the French Labour Force Survey can capture long–term changes in the labour market. It has undergone many changes over the course of its history: conceptual changes, in order to conform to ILO or Eurostat guidelines or to better measure labor market transformations, but also changes in statistical engineering (sampling,

treatment of non–response, collection mode, etc.). It has also benefited from technical improvements with the increasing computerization of data collection and processing. The last overhaul of the Labour Force Survey was in 2013 [INSEE (2021)].

A change of questionnaire imposed by a new European regulation

The entry into force of a new European regulation on social statistics in January 2021 has necessitated a new redesign of the Labour Force Survey. The "IESS" regulation (Integrated European Social Statistics framework regulation) imposes a coherent and integrated framework for household surveys in order to improve their international comparability. This new regulation does not fundamentally change

the European requirements in terms of methodology or protocol, but harmonizes even further the information collected.

Thus, the new European regulation requires not only that new information be collected, such as the reasons for migration to France or work during studies, but above all that the central questions used to determine ILO activity status be asked in a precise order and manner. However, this did not result in too much change in the French questionnaire, which was already quite close to the new European framework.

Among the main changes in the questionnaire, a clarification by Eurostat of the interpretation of ILO criteria led to a change in the scope of employment. To improve comparability between countries, Eurostat proposes operational definitions of ILO concepts with even more precise criteria for classifying the population as employed, unemployed or inactive. These criteria concern, for example, the reasons for absence for people who have a job but have not worked during a given week (known as the "reference week") or a precise list of job–seeking activities for unemployed people.

With the entry into force of the new European regulation, people who declare having a job but being absent due to sick leave are now classified as employed, regardless of the duration of their absence. Previously, they were only considered to be in employment if they were expected to be absent for a year or less. Similarly, people reporting absence from work due to parental leave are now classified as employed if their expected absence is less than or equal to 3 months or if they receive a compensatory income, such as the French "shared child–rearing benefit" (Prepare), and in this case, without a duration condition. Previously, they were only considered to be in employment if they were absent for 3 months or less. In both cases, the changes lead to more people being classified as employed than before.

New information collected, a change in collection mode, a renovation of the weighting method

Beyond the European requirements, the redesign of the questionnaire was also an opportunity to respond to national expectations on subjects such as telework, new forms of employment (in particular situations of economic or organizational dependence of self–employed workers) and vocational training. The new Labour Force Survey is also the first household survey to implement the renovated French

socio–professional nomenclature [CNIS, 2019]. The exploitation of this new information requires time and preferably a full year of data. It will thus be carried out in a second phase, in 2022, after the release of the main cyclical labor market indicators, for which a quarterly exploitation is possible.

In order to introduce only a single break in measurement, the redesign of the survey, made necessary by the European regulation, was also an opportunity to embark on other developments desired at the national level. The first was to modernize the survey's collection protocol by offering the Internet as an additional response mode in reinterview ► **Box 1**. The second was to renovate the survey's weighting method ► **Box 2**.

An unprecedented large–scale operation to quantify measurement changes

Each change in the survey can affect the measurement of labor market indicators (employment, unemployment, halo around unemployment, etc.). In order to quantify the possible changes in measurement, the 2021 overhaul relied on a large–scale operation, a "Pilot" of the new survey conducted since the first quarter of 2020 ► **Box 3**. It has allowed INSEE to compare the old and new versions of the survey for five consecutive quarters.

There are many possible sources of measurement disruption: changes in the questionnaire, in the collection method with the introduction of the Internet, in the weighting method, with possible cross effects. The Pilot survey makes it possible to estimate the overall change in each indicator by direct comparison between the old and new surveys, without systematically trying to quantify the contribution of each change ("global" estimate). Nevertheless, whenever possible, INSEE analyzes separately certain effects to describe the mechanisms at work.

Compared to the previous Labour Force Survey, the employment rate for people aged 15 to 64 is increased by 0.8 points

With the new version of the Labour Force Survey, the **employment rate** for people aged 15 to 64 is increased by 0.8 percentage points on average in 2020 (► **Chart 1**, or +270,000 persons in employment). Conceptual changes regarding persons in employment but absent due to sickness or parental leave explain most of the increase. Changing the criteria for these reasons for absence raises the employment rate for people aged 15 to 64 by 0.7 points: +0.5 points for sick leave and +0.2 points for parental leave.

Since many more women than men are on parental leave, their employment rate (15 to 64 years old) is revised further up (+1.4 points, of which +0.4 points are related to this effect), especially for middle–aged women (25–49): +1.5 points, of which +0.8 points are related to this effect ► **Chart 2**. In addition, the change in sick leave increases the employment rate of older workers more than the average: +1.1 points in total for people aged 50 to 64, of which +1.0 points is related to this change. This is particularly the case for women aged 50 to 64 (+1.7 points, of which +1.2 points is related to this change in concept).

In addition, the new sampling frame includes information that was not available in the previous one, notably on household income. By taking advantage of this information, the new weighting method makes it possible to better correct for nonresponse to the survey ► **Box 2**. This has the effect of increasing the employment rate for people aged 15 to 64 by 0.3 points in total, since this new method better accounts for populations with a high employment rate.

In addition, when the survey was redesigned, the implementation of the "usual residence concept" for

► 1. Measurement breaks by ILO status

In percentage points

	Employment rate	Unemployment share	Halo share	Inactives outside halo share
Global effect	**0.8**	**0.1**	**0.8**	**– 1.7**
Change in employment concept	0.7	0.0	– 0.1	– 0.6
- *of which parental leave*	*0.2*	*0.0*	*0.0*	*– 0.2*
- *of which sick leave*	*0.5*	*0.0*	*– 0.1*	*– 0.4*
Change in weighting method	0.3	0.0	– 0.1	– 0.1
"Proxy" effect(*)	0.2	0.1	0.1	– 0.4
Other effects (questionnaire. protocol...)	– 0.3	0.0	0.9	– 0.6

(*) including the change in usual residence rules ► **Box 2**.
Reading: the employment rate for people aged 15 to 64 is 0.8 points higher in 2020 in the new survey compared to the old survey.
Note: due to rounding, the sum of rows may slightly differ from the total.
Scope: France excluding Mayotte, people living in private household aged 15 to 64.
Source: INSEE, Labor Force Survey.

determining the scope of the survey was revised, essentially concerning students of legal age living both in their own accommodation for their studies and with their parents ► **Box 2**. In the old survey, these students were mainly attached to their parents' housing. Starting in 2021, INSEE implements a "mixed" solution: students with two residences are surveyed for their two dwellings, taking care that they are not overrepresented.

This change in the implementation of usual residence, but also the internet mode as a possible mode of reinterview, lead to a clear reduction in the response by "proxy", i.e. by a third party who answers in their place, of people aged 15 to 24 (53% of responses by proxy in the new survey, against 65% in the previous version). Since the respondents are in a better position than their relatives to say whether they have done at least one hour of paid work, these changes lead to an increase in the overall employment rate of 0.2 percentage points, particularly for young people, both male and female. Among young people who spontaneously declared themselves to be students, the employment rate is notably raised by 5 points. Thus, the new usual residence rule and the reduction of the proxy response lead to a greater weighting of students in their own housing and to a better identification of their employment situation.

Finally, the other changes (which include the redesign of the questionnaire and the possibility of responding by Internet) contribute to a 0.3 points decrease in the employment rate for people aged 15 to 64, with this difference varying according to sex and age categories. However, it is difficult at this stage to separate the effects of the change of questionnaire from those of the new protocol and the possibility of responding by internet ► **Box 3**. These effects are, by their very nature, closely linked. When answering on the internet, being able to read the new instructions specific to a question may lead respondents to choose different answers from those they would have chosen with the previous version of the question, or the ones they would have made if questioned directly by an interviewer. The split between protocol and questionnaire effects will require further study, but will not affect the "global" estimate made by comparing the old and new surveys.

Unemployment rate is unchanged between the two surveys

Overall, **unemployment** is almost unchanged in the new survey compared to the previous one. The **share of unemployed persons** among people aged 15 to 64 is revised upwards

► 2. Measurement breaks on the employment rate, by sex and age

In percentage points

	Women				Men				Total
	15-64	15-24	25-49	50-64	15-64	15-24	25-49	50-64	**15-64**
Global effect	**1.4**	**– 0.3**	**1.5**	**1.7**	**0.2**	**1.2**	**– 0.6**	**0.5**	**0.8**
Change in employment concept	1.0	0.1	1.2	1.2	0.4	0.0	0.3	0.7	0.7
- of which parental leave	*0.4*	*0.1*	*0.8*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.2*
- of which sick leave	*0.6*	*0.0*	*0.4*	*1.2*	*0.4*	*0.0*	*0.3*	*0.7*	*0.5*
Change in weighting method	0.3	– 0.3	0.5	0.4	0.2	0.2	0.4	– 0.1	0.3
"Proxy" effect(*)	0.2	1.1	0.1	– 0.1	0.1	0.7	0.1	– 0.2	0.2
Other effects (questionnaire. protocol...)	– 0.1	– 1.2	– 0.2	0.3	– 0.5	0.3	– 1.4	0.0	– 0.3

(*) including the change in usual residence rules ► **Box 2**.
Reading: the employment rate for women aged 15 to 64 is 1.4 points higher in 2020 in the new survey compared to the old survey.
Note: due to rounding, the sum of rows may slightly differ from the total.
Scope: France excluding Mayotte, people living in private household aged 15 to 64.
Source: INSEE, Labor Force Survey.

► 3. Measurement breaks on the share of unemployed, by sex and age

In percentage points

	Women				Men				Total
	15-64	15-24	25-49	50-64	15-64	15-24	25-49	50-64	**15-64**
Global effect	**0.1**	**1.5**	**-0.3**	**0.1**	**0.0**	**0.0**	**0.1**	**0.0**	**0.1**
Change in employment concept	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
- of which parental leave	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
- of which sick leave	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
Change in weighting method	– 0.1	0.1	– 0.1	– 0.1	0.0	– 0.1	– 0.1	0.1	0.0
"Proxy" effect(*)	0.1	0.2	0.1	0.1	0.1	0.2	0.1	0.1	0.1
Other effects (questionnaire. protocol...)	0.1	1.2	– 0.3	0.1	– 0.1	0.0	0.0	– 0.2	0.0

(*) including the change in usual residence rules ► **Box 2**.
Reading: the share of people aged 15 to 64 unemployed is 0.1 points higher in 2020 in the new survey compared to the old survey.
Note: due to rounding, the sum of rows may slightly differ from the total.
Scope: France excluding Mayotte, people living in private household aged 15 to 64.
Source: INSEE, Labor Force Survey.

very slightly (+0.1 points in 2020, i.e. +30,000 unemployed persons), and the **unemployment rate** is unchanged, as the number of persons in the denominator is also raised with employment. This quasi–stability of the unemployment share and rate between the two versions of the survey is true for all gender and age categories, except for women aged 15 to 24 (+1.5 points for the unemployment share; ► **Chart 3**.

The share of people in the halo around unemployment is significantly increased

The new survey increases the share of persons in the **halo around unemployment** by 0.8 points or +310,000 persons in 2020. This situation refers to persons who are jobless but want a job, who are not available within 15 days, or who are not actively looking for work in the four weeks ending with the reference week. The upward revision of this situation is mainly due to the combined effects of the change in questionnaire and protocol. More respondents report having conducted job searches in the new survey, but not necessarily being available. In particular, when they responded *via* the

Internet, the examples of steps that were added as instructions before the dedicated question ("submit your CV on the Internet, take a competitive examination, take steps to find premises or funding, etc.") may have encouraged them to report searches. When they had not looked for work, slightly more of them also said that they nevertheless wanted to work.

This surplus of people in the halo around unemployment concerns all gender and age groups, but particularly young people. For the same reasons that led to an increase in their employment rate, the effect of the decrease in the proxy leads to an increase in the share of jobless young people who declare that they want to work. It contributes 0.4 points to the 1.8 points increase in the share of youth in the halo ► **Chart 4**.

The share of inactive persons outside the halo has clearly declined in the new survey

The counterpart of all these changes in employment, unemployment and the **halo around unemployment** is that the share of **inactive persons outside the halo** in the new Labour Force Survey has dropped

significantly (– 1.7 points in 2020). First of all, this reflects the fact that persons now classified as employed in the ILO sense for a long absence due to sickness or parental leave were previously classified as inactive outside the halo. Indeed, they had no reason to declare that they were looking for or even wishing for a job other than the one to which they were entitled upon their return from leave. These conceptual changes thus contribute to a 0.6 percentage point decrease in the share of people aged 15 to 64 in non–halo inactivity, especially middle–aged women because of parental leave and older men and women because of sick leave ► **Chart 5**.

The new weighting method also results in a slight decrease in non–halo inactivity (– 0.1 points for people aged 15 to 64). In addition, the overall effect of the decrease in proxies (including the usual residence rule) is a decrease in non-halo inactivity: respondents are in a better position than their relatives to indicate whether they want a job or whether they have taken steps in this direction. The share of non-halo inactivity has therefore fallen by 0.4 points, especially among young people. Finally, the other effects of the change of questionnaire and protocol also lead to a decrease in the share of non-halo inactive persons, because with the new survey more unemployed persons declare that they would like to have a job.

► **4. Measurement breaks on the share of "halo around unemployment", by sex and age**

In percentage points

	Women				Men				Total
	15-64	15-24	25-49	50-64	15-64	15-24	25-49	50-64	**15-64**
Global effect	**0.7**	**1.9**	**0.6**	**0.2**	**0.8**	**1.7**	**0.8**	**0.3**	**0.8**
Change in employment concept	– 0.2	0.0	– 0.2	– 0.1	– 0.1	0.0	– 0.1	– 0.1	– 0.1
- of which parental leave	*– 0.1*	*0.0*	*– 0.1*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*
- of which sick leave	*– 0.1*	*0.0*	*– 0.1*	*– 0.1*	*– 0.1*	*0.0*	*– 0.1*	*– 0.1*	*– 0.1*
Change in weighting method	– 0.1	0.0	– 0.1	– 0.1	– 0.1	– 0.3	– 0.2	0.0	– 0.1
"Proxy" effect(*)	0.2	0.5	0.1	0.1	0.0	0.2	0.0	0.0	0.1
Other effects (questionnaire. protocol...)	0.8	1.4	0.8	0.4	1.0	1.9	1.1	0.4	0.9

(*) including the change in usual residence rules ► **Box 2**.
Reading: the share of people aged 15 to 64 in the halo around unemployment is 0.8 points higher in 2020 in the new survey compared to the old survey
Note: due to rounding, the sum of rows may slightly differ from the total.
Scope: France excluding Mayotte, people living in private household aged 15 to 64.
Source: INSEE, Labor Force Survey.

► **5. Measurement breaks on the share of " non-halo inactives", by sex and age**

In percentage points

	Women				Men				Total
	15-64	15-24	25-49	50-64	15-64	15-24	25-49	50-64	**15-64**
Global effect	**– 2.3**	**– 3.0**	**– 1.9**	**– 2.1**	**– 1.1**	**– 2.9**	**– 0.3**	**– 0.8**	**– 1.7**
Change in employment concept	– 0.8	– 0.1	– 1.0	– 1.0	– 0.3	0.0	– 0.2	– 0.7	– 0.6
- of which parental leave	*– 0.4*	*– 0.1*	*– 0.7*	*0.0*	*0.0*	*0.0*	*0.0*	*0.0*	*– 0.2*
- of which sick leave	*– 0.5*	*0.0*	*– 0.3*	*– 1.0*	*– 0.3*	*0.0*	*– 0.2*	*– 0.7*	*– 0.4*
Change in weighting method	– 0.2	0.2	– 0.3	– 0.2	0.0	0.2	– 0.2	0.0	– 0.1
"Proxy" effect(*)	– 0.5	– 1.8	– 0.3	– 0.1	– 0.3	– 1.0	– 0.3	0.1	– 0.4
Other effects (questionnaire. protocol...)	– 0.8	– 1.4	– 0.3	– 0.8	– 0.4	– 2.1	0.3	– 0.2	– 0.6

(*) including the change in usual residence rules ► **Box 2**.
Reading: the share of " non-halo inactives" among people aged 15 to 64 is 1.7 points lower in 2020 in the new survey compared to the old survey.
Note: due to rounding, the sum of rows may slightly differ from the total.
Scope: France excluding Mayotte, people living in private household aged 15 to 64.
Source: INSEE, Labor Force Survey.

Long–term unemployment declines in the new survey with a change in definition

The new Labour Force Survey revises downwards by 0.7 percentage points the share of long–term unemployment in the population. On average, over the five quarters of overlap between the two versions of the survey, it falls from 2.2 percent of all people aged 15 to 64 to 1.5 percent. To be considered long–term unemployed, one must be considered unemployed, have been looking for work for at least one year, and have not held a job for at least one year. However, in the old survey, the last job held as a reference for calculating this jobless time was limited to "regular" jobs. This concept was not consistent with the ILO concept of employment, for which the regular or occasional nature of the job is not relevant. In accordance with the clarifications made at the European level, in the new survey, the last job held and described must henceforth also include occasional jobs. This change in scope mechanically reduces the duration since the last job and the number of long– term unemployed is lowered.

In addition, the impact on the employment rate depends on the employment status. The permanent employment rate (including civil servants) thus revised upward by 0.5 points, which is less than the increase in the overall employment rate. Two main factors play a role. The changes in the concept of employment contribute upwards (people absent from work for a long duration sick leave or parental leave, now classified as employed, are more often permanent employees). On the contrary, in accordance with the CNIS recommendations [Gazier *et al*, 2016], the category "permanent employment" no longer is the default category for people in employment without a contract, or whose contract information is unknown. The temporary employment rate (fixed–term contracts, temporary work) revised downward by 0.8 points, notably because internship and apprenticeship contracts are now accounted for in a separate category rather than in the "temporary employment" category.

Finally, the share of people underemployed due to short–time work or temporary lay–off, which was anecdotal before the economic crisis and jumped in 2020, is revised downwards by 2.0 points on average in 2020. This is mainly due to changes in the questionnaire. Whereas the previous survey used to identify these situations through questions about the causes of changes in the reference week's schedules, the new survey now proceeds through questions on the reasons for absence, with a modified list of reasons. At an average of 4.3 percent in 2020 in the new survey, the level of underemployment due to short–time working or temporary lay–off remains at an unprecedented level in 2020.

Over the long term, the trends in most of the indicators adjusted for the break are virtually unchanged

In order to continue to publish results with invariant methods and concepts over time – in other words, so that no change in the published indicators comes from a change in the statistical measure – it is necessary to backcast the series over the long term ► **Box 4**. Once the breaks have been estimated over the overlap period of the two versions of the survey, the backcasting exercise is based on the assumption, in the general case, that the cumulative
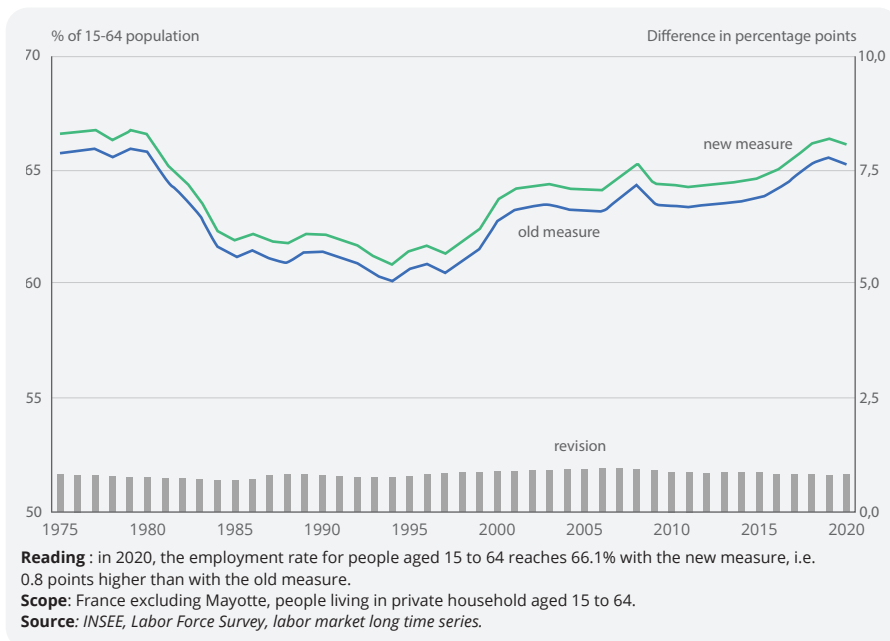
effects of the changes would apply with a constant multiplicative factor, if the survey conditions had been identical since the beginning of the estimation periods. The series thus appear most often translated, as in the case of the employment rate for people aged 15 to 64 ► **Chart 6**. For parental leave, a specific backcasting exercise had to be carried out, leading to a less constant effect of the backcast over time on the revision of the female employment rate. Moreover, since INSEE conducted the backcasting exercise at a fine level of detail (five–year age groups) and separately for women and men, changes in the composition of the population are likely to modify the aggregate indicators on age or gender groupings. Overall, since the measured break in the unemployment rate is almost zero, the new unemployment rate curve for people aged 15 or more is almost identical to the old series ► **Chart 7**.
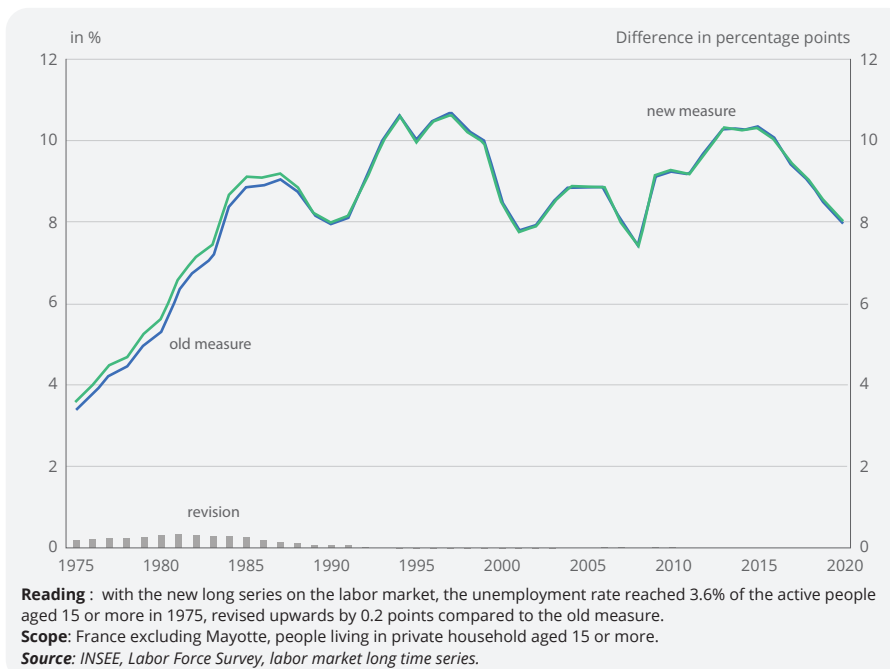
Given the principles of backcasting, the past annual or quarterly variations of the main economic aggregates are for the most part almost unchanged. This is notably the case for the diagnosis of the consequences of the health and economic crisis of 2020 on the labour market. ●

**Département de l'emploi
et des revenus d'activité (Insee)**

► **6. Employment rate for people aged 15 to 64, old and new measure, revisions**

Reading : in 2020, the employment rate for people aged 15 to 64 reaches 66.1% with the new measure, i.e. 0.8 points higher than with the old measure.
Scope: France excluding Mayotte, people living in private household aged 15 to 64.
Source: *INSEE, Labor Force Survey, labor market long time series.*

► **7. Unemployment rate, old and new measure, revisions**

Reading : with the new long series on the labor market, the unemployment rate reached 3.6% of the active people aged 15 or more in 1975, revised upwards by 0.2 points compared to the old measure.
Scope: France excluding Mayotte, people living in private household aged 15 or more.
Source: *INSEE, Labor Force Survey, labor market long time series.*

► **Box 1. Changes in protocol and the proposal of the Internet as an additional mode of response in reinterviewing**

The Labour Force Survey is a continuous survey: the sample is spread over all the weeks of the year. Each weekly sample is associated with a so-called reference week, in relation to which the respondents describe their employment situation. To avoid recall bias and to ensure quality data, the collection period is very short: until 2020, respondents had to respond within two weeks and two days after the reference week. With the redesign, this period is extended to three weeks, which is still below the European ceiling.

Another particularity of the survey is that it is a housing panel. INSEE interviews the inhabitants of the sampled dwellings six quarters in a row, in order to estimate changes more robustly and to produce longitudinal analyses. Until 2020, INSEE interviewed households face-to-face in the first and last interviews and by telephone in-between. With the new survey, the first interview remains face-to-face, but for the reinterviews, the Internet is proposed as an additional response mode to the telephone. In the general case, the first three days of the collection period are reserved for the response by Internet. From Thursday of the first week of collection, the INSEE interviewer can call the household to conduct the survey by telephone. The collection website remains accessible to the household as long as it has not responded and at the latest until the end of the third week of collection. Note also that the sixth interview, which was face-to-face until then, becomes a reinterview like the others.

This multi-mode protocol is intended to encourage responses from households, which are increasingly using the Internet in their daily lives, and to lighten the load on interviewers. However, even with reinterviewing, the role of the interviewers remains crucial to obtain the response of households less familiar with the Internet or more sensitive to telephone follow-up.

During the first quarter of 2021, among the respondents whose households were reinterviewed, 41% responded by Internet, the others having responded by telephone. This result is close to those observed in the 2020 Pilot survey. The overall collection rate for the first quarter of 2021 is close to that of 2019, although slightly lower due to the health crisis conditions for the first face-to-face interview. In 2020 the collection rate for the Pilot survey and the old version of the survey were broadly comparable, both of which were affected by the difficulties associated with the health crisis.

► **Box 2. The redesign of the weighting method**

The purpose of weighting a survey is to reconstitute the total population (not just the population that responds to the survey) by correcting for possible non-response bias (if the non-respondents have particular characteristics) and by controlling for sampling fluctuations (the sample may not be strictly representative of the reference population according to a particular criterion). To calculate the weight assigned to each respondent, auxiliary information from the sampling frame or from external data is used. As the sampling frame for the Labour Force Survey is different for metropolitan France and the French overseas departments, the weighting method differs accordingly.

Starting in 2021, in metropolitan France, INSEE weights the Labour Force Survey data according to a new method. This change of method aims first at taking advantage of the data of a new sampling frame, richer in characteristics of dwellings and their occupants. Since the fourth quarter of 2020, the Labour Force Survey sample for metropolitan France has been drawn from a new sampling frame (Fideli), following the same principles as before. This new sampling frame is very similar to the one used previously, but has new information that allows for a better understanding of nonresponse [Courrier des statistiques, July 2021]. The change in method is also intended to apply the requirements of the new European regulation, particularly on population margins, and more generally to improve on certain aspects of the old weighting method.

New information in the sampling frame to better model survey nonresponse

For metropolitan France, INSEE changed the weighting method from a one-step method (single calibration on variables in the sampling frame and on external data) to a two-step method, as was already done in the DOM: first a correction for sampling bias and nonresponse, then calibration on external data.

In addition, the calculation of weights uses new information from the sampling frame (place of birth, household composition, income, etc.). These characteristics are linked to both the response behavior and the variables of interest (being employed). The dwellings occupied by persons receiving high pensions or high incomes from work (without being in the last decile), or couples with children have, all other things being equal, a higher probability of responding to the survey. This new information complements the information already used (geographical location, occupation status, type and size of dwelling, etc.) and makes it possible to better characterize respondents and non-respondents in order to better correct for nonresponse bias.

In the DOM, the weighting method has been marginally adjusted (notably calibration by interview rank), but without being renovated in its general principle, as the sampling frame (Census of Population) has not been modified.

The rules of "usual residence" have evolved to better take into account the situation of young people

The implementation of the "usual residence concept" of individuals, which aim to manage cases of multi-residence in order to respect the scope of household surveys and to deal with double counting, have also been revised (for metropolitan France and the DOM). These rules define the population of interest, which must be weighted.

At the beginning of the questionnaire, all persons living in the dwelling, even occasionally, are identified. However, not all of them fill in an individual questionnaire and therefore not all of them participate in the calculation of the indicators. Either because they do not fall within the scope of the survey (e.g., if they reside in a community: retirement home, hostel, etc.), or to avoid double counting (e.g., a child living alternately with separated parents).

When INSEE redesigned the survey, the questionnaire was changed to better identify dual residence situations, and some usual residence rules were revised, mainly the one concerning adult students living both in housing for their studies and with their parents. In the old survey, these students were essentially attached to their parents' accommodation; this improved their response rate (parents were quicker to respond to the survey), but generated a high proportion of proxy responses (as students would frequently be absent when the interviewer visited their parents, their parents responded in their place). However, the response by a third party is of lesser quality, especially for the questions aimed at determining the activity status, which are very precise. For example, at a given age and level of education, the employment rate of dual-residence students is higher when they are surveyed in their student housing than in their parents' housing. Starting in 2021, a "mixed" solution is implemented: students in dual residence are surveyed in their two dwellings, and then their weight is divided by two.

► Box 3. The Pilot survey, a large-scale operation to estimate measurement breaks

Experience with previous redesigns of the Labour Force Survey has shown that even simple changes in the order of questions can cause breaks in the measurement of key labour market indicators. With the 2021 overhaul, which involves the questionnaire, the protocol, and the weighting method, such breaks were also likely. In order to provide sufficiently precise information on measurement breaks, INSEE took the necessary steps to carry out an unprecedented preparatory operation: a complete "pilot survey". It was designed to align perfectly with the new survey in terms of questionnaire, protocol and processing method, and would be in the field at the same time as the traditional Labour Force Survey, one year before the required changeover, i.e. as early as the first quarter of 2020. In order to reconcile the need for precision on the one hand (to be able to detect significant impacts), and the need for reasonable cost on the other hand, a quarter of the sample of the survey in production was taken to carry out the Pilot survey in 2020.

In the end, the Pilot survey has two components:

• **an early switch to the new survey**: a quarter of the sample, across all interviewing ranks, switched to the new survey in the first quarter of 2020 - this switch foreshadowed the switch in the first quarter of 2021 for the rest of the sample - ;

• **an oversample in the old survey**: introduced progressively from the fourth quarter of 2019, an oversample of households was reinterviewed in the old survey up to and including the first quarter of 2021. A complete control sample in the old survey was thus available in the first quarter of 2021, when the new survey went into production.

In the first quarter of 2021, the three quarters that had remained in the old survey switched to the new one and joined the part that had switched one year earlier. The survey sample thus returned to its "full size".

The main objective of the Pilot survey was to provide a set of information that would allow INSEE to estimate as accurately as possible the measurement breaks induced by the new survey. With the Pilot survey, by comparing the indicators measured *via* the old survey and *via* the new survey, five quarterly measures of breaks in series were available, from the first quarter of 2020 to the first quarter of 2021. Having several quarters available makes it possible to measure the break with greater precision, but also to detect possible seasonality in the impacts.

INSEE measures the overall measurement break directly by comparing the two versions of the survey. Separate analyses of some components are also possible:

• the impact of the **new weighting method** can be studied in isolation, by applying the old method to the new survey data ► **Box 2**;

• the effect of the **usual residence rules** can be distinguished by simulating the effect of the new rule in the new survey and comparing it to the old one; the "**proxy**" **effect**, which is closely related to this change of rule, can be measured by comparing the results of the new survey with those of the new survey in which the weights of the respondents are recalculated so that the proxy rate is set to that of the old survey;

• the impact of **conceptual changes** can be isolated by counting the number of people for whom the change in criteria changes their employment status (people on sick leave or people on parental leave);

• Among the other changes, it is possible to distinguish effects related to the **change of questionnaire**, by comparing the new and old version in the first interview, for which the protocol changes little. However, the estimate is based on smaller samples, and is therefore more fragile. Similarly, specific estimates are also possible to isolate mode effects linked to responding by internet, by separating a "**measurement**" **effect** (the fact that the same person responds differently on the internet or to the interviewer) and a "**selection**" **effect** (responding rather than not responding because the internet mode is now available). However, the estimation of these **mode effects** is also more fragile because it is based on assumptions that allow simulating what the response would have been in a counterfactual scenario without change. The results are sensitive to the choice of the determining variables. For this reason, these effects are not distinguished here but will be the subject of further expert work.

► Box 4. Backcasting methodology

A "backcasting" exercise for macroeconomic indicators (from the national accounts, key labor market indicators, etc.) is a statistical operation which consists in adjusting past observations to match the "new measure" of an economic phenomenon. For the overhaul of the Labour Force Survey, this means adjusting past observations to make it "as if" the questionnaire had always been administered in the same way as in the first quarter of 2021, with the same concepts, the new weighting method and the same protocol. The objective is to create a series whose changes reflect only changes in the economic phenomenon of interest and not changes in the measurement of this phenomenon.

It should be noted that modifying a quantity in 1975 because of the possibility of responding now *via* the Internet may seem paradoxical, since in 1975 the Internet did not exist. However, the appearance of the Internet and its expansion should not have an impact on the backcast series. Although the hypothesis may seem very theoretical, it is the only way to guarantee the comparability of results over a long period.

A "pivot" coefficient to correct for most breaks

In practice, in the general case for a given indicator, INSEE estimates an average "pivot" coefficient over the five quarters of overlap between the

► Employment rate for women aged 35 to 39 years old, old measure and new backcast measure
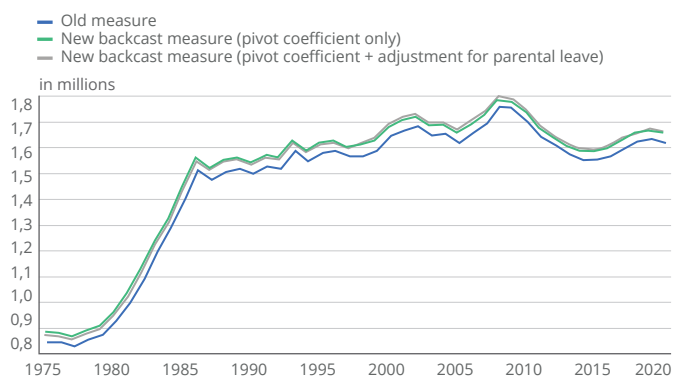
Reading: in 2020, there were 1,620,000 women aged 35-39 in employment with the old version of the Labor Force survey; this estimate has been increased to 1,660,000 with the new version of the survey.
Scope: France excluding Mayotte, women living in private household aged 35 to 39.
Source: INSEE, Labor Force Survey, labor market long time series.

old and new surveys, allowing for the transition from the old to the new measure. This coefficient measures the magnitude of the break over the overlap period between the two surveys, in a multiplicative form. It is then applied to the entire past series in order to simulate a series in the new measure. This is equivalent to assuming that the magnitude of the break is proportional over each year. For example, for women aged 35-39, employment is higher in 2020 by 2.6 percent with the new Labour Force Survey, so the coefficient 102.6 is applied each year. The same exercise is carried out for each of the activity situations, then the levels are proportionally recalibrated each year on the total population of women aged 35-39 (► **Chart**, new backcast measure - pivot coefficient only-).

The break between the old and new surveys often differs according to the characteristics of the populations. This is why it is important to measure and understand the determinants of the breaks (primarily gender and age) and then backcast at the level of employment subsets by gender and age, rather

►►►

than at the level of the entire given aggregate. This necessity leads to backcasting by gender and five-year age group. However, at this level of detail, the measurement break is based on small sample sizes and is therefore more fragile. Therefore, it is not the coefficient specific to each age group that has been applied, but rather a coefficient smoothed over three consecutive age groups. By re-weighting each year by the weight of each age and gender group, the backcasting reflects the changes in the composition of the population on the aggregate indicators.

Absences due to parental leave: a conceptual break that requires an ad hoc backcast

However, the assumption underlying the general method is not always satisfactory. In particular, in the new Labour Force Survey, people who report being employed but absent from work because of parental leave can be counted as employed regardless of the duration of the leave, provided they receive an allowance such as the French "shared child-rearing benefit" (Prepare). This conceptual change amounts to increasing the employment rate of women aged 25 to 49 by 0.8 points in 2020. However, it is not relevant to apply this increase in employment since 1975, because at the beginning of the period there were no financial incentives for parental leave. Their appearance in 1984 and their rise to prominence at the end of the 1980s led to an *ad hoc* retropolation, in order to have series in new measure, which better transcribe the changes in parental leave since 1975. For this purpose, external sources (other surveys, administrative sources) were mobilized. In the above example of women aged 35 to 39, this amounts to proposing a new backward-looking measure "adjusted" for the effect of parental leave. At the beginning of the period when the measures did not exist, the final series then lands between the old series and the backcast series without adjustment.

Quarterly series backcast in a second step, after checking that seasonality has not been modified by the new survey

By compiling the individual observations from the five quarters of coverage, the estimate of the break is more robust than from a single quarter. This is why the annual series are backcast first, using this "global" break. New quarterly series are then constructed in a consistent manner, i.e., in such a way that the average of the four quarters coincides perfectly with the annual series.

However, a change in questionnaire, collection mode or concept can also affect the seasonality of a quarterly series. A very different break in a given quarter from the collection year may indicate that the seasonality has changed. However, the thorough analysis of the quarterly series led to the finding that there was no discernable change in seasonality for any series. This conclusion will have to be re-examined with the addition of new quarters. One consequence of this finding is that the pivot coefficient can be identical each quarter, and thus, for the year 2020, the quarterly estimates backcast are not identical to those directly derived from the Pilot survey. Another consequence is that the past seasonality of the series has been preserved, so that the usual seasonal adjustment (SA) models have also been retained.

► Definitions

The **employed** population includes persons aged 15 years or older who worked (even for one hour) during a given "reference" week, as well as those who are employed but did not work during that week for a number of listed reasons (leave, sick leave, temporary lay-off, etc., sometimes with duration criteria). People in employment are employees, self-employed, employers, or helpers in the family business.

The **employment rate** is the ratio of the number of employed persons to the corresponding total population. It can be calculated for the entire population of a country, but is most often limited to the "working age" population (the convention is most often the 15-64 age group), or to a portion of the population (e.g., women aged 25-29).

An **unemployed person in the ILO sense** is a person aged 15 or older who simultaneously meets three conditions: being unemployed during a given week; being available to take a job within two weeks; and having actively looked for a job in the last four weeks or having found one starting in less than three months. The active steps considered are varied: studying job advertisements, going to a trade show, mobilizing one's social network or taking advice from the Pôle emploi, etc.

The **unemployment rate** is the ratio of the number of unemployed people to the number of active people (employed or unemployed). The **unemployment share** is the ratio of the number of unemployed to the corresponding total population.

The **halo around unemployment** is composed of people who are inactive in the ILO sense, but close to the labor market. These are jobless persons who are looking for a job but are not available for work within two weeks and jobless persons who want to work but have not taken any active steps to look for a job in the previous month, whether or not they are available. The "halo around unemployement" is thus larger than the « potential additionnal labour force » often used in Eurostat publications. It covers the two European categories "Jobless persons seeking a job but not immediately available for work" (so-called "halo 1") and "Jobless persons available for work but not seeking it" (halo 2) but also inactive people who want a job and are neither seeking nor available (halo 3).

ILO inactive persons are persons aged 15 years or older who are neither employed nor unemployed as defined by the ILO. They may be persons in the unemployment halo or persons classified as "non-halo inactive".

► Learn more

- **Lamarche P.**, **Lollivier S.**, « Fidéli, l'intégration des sources fiscales dans les données sociales », *Courrier des statistiques* n°6, INSEE, juillet 2021.
- **Guillaumat-Taillet F.**, **Tavan C.**, « Une nouvelle enquête Emploi en 2021. Entre impératif européen et volonté de modernisation », *Courrier des statistique*s n°6, INSEE, juillet 2021.
- **INSEE**, Methodological note on the Labor force survey, juin 2021.
- **Cnis**, « La rénovation de la nomenclature socioprofessionnelle (2018 2019) », *Rapport du groupe de travail du Cnis* n° 156, décembre 2019.
- **Gazier B.**, **Picart C.**, **Minni C.**, « La diversité des formes d'emploi », *Rapport au Cnis* n° 142, juillet 2016.

Depuis **75 ans**
l'Insee vous parle de **vous**
Insee