Nowcasting of the poverty rate using microsimulation: estimations based on French data

Kevin Schmitt, Insee, kevin.schmitt@insee.fr

**Abstract**

*Every year in September* ***N+2****, INSEE publishes the poverty rate and the main indicators of inequalities in standard of living for year* ***N****. This delay is unsatisfactory for meeting the social requirements of users of these indicators. Of the 21 months between the end of the year under consideration and the publication of the poverty rate, about three-quarters of this time is taken up collecting tax and social data, and about one quarter with statistically matching Labour Force Survey (LFS) data, from which the Tax and Social Incomes Survey (ERFS) is produced. Nowcasting consists of producing an earlier indicator (in autumn* ***N+1****) of the poverty rate for the target year* ***N*** *based on the ERFS* ***N-1****. The method to be used here is microsimulation, which simulates individuals’ standard of living by imputing benefits and contributions on scales, and allows to take account of any legal changes in the social and tax legislation. The exercise is based on the INES model, which simulates the majority of French social security and tax legislation, based on any year of the ERFS. To implement nowcasting, one important step is ageing population by uprating incomes (using surveys about wages, aggregated tax data, inflation...) and calibration weighting (using margins from LFS and census). Reverse ageing is also used so that evaluations for year* ***N*** *and* ***N-1*** *(and thus annual evolutions) are only based on the ERFS* ***N-1*** *(that is minimising the sample bias). In this paper, we present the methodology and assess the quality of the early indicators thus produced. Indeed, we compare the results that would have been produced by microsimulation with those that were in fact disseminated from the ERFS. When applied to the target years 2010 to 2015, this method produced estimations similar to the actual figures published the following year for the poverty rate and the main indicators of inequalities.*

**Keywords:** Flash Estimates, Nowcasting, Microsimulation, Poverty and Inequalities indicators

**1. Introduction**

In September N+2 of each year, the French Institute of Statistics (Insee) releases the final estimates of equivalised disposable income, poverty rate and the main inequalities indicators for year N, based on the Tax and Social Incomes survey (ERFS). It is therefore 21 months after the end of the period under consideration that the first results on standards of living are disseminated. This delay, which may seem long given the needs of the users, is mainly due to the specific features of the information system. In fact, the ERFS is the result of statistical matching between the Labour Force Survey, tax records from the fiscal administration (income tax and housing tax) and the administrative records of social benefits. The calendar for producing ERFS is as follows: tax information and social records are transferred 12 and 15 months respectively after the end of the period under consideration (December N+1 and April N+2); once these data have been received, Insee carries out statistical matching which is completed by the end of June N+2. The two next months are devoted to statistical processing to produce key indicators of poverty and inequality.

It is not possible to reduce the delay in producing standard of living by very much. As a result, in order to meet the demand for early indicators and thus monitor policies targeting poverty, Insee developed a method of producing rate of change indicators by microsimulation. This exercise is called nowcasting by analogy with the term forecasting, but focusing here on a period that is already past. The method used is based on the INES microsimulation model. In this paper, we present the methodology and assess the quality of the early indicators thus produced.

**2. Microsimulation method**

*2.1. Review of INES in normal use*

The INES microsimulation model is used to simulate the vast majority of scaled contributions and social benefits used in calculating standard of living. In normal use, the model calculates individual-level disposable income in N+2 (the most recent complete year at the time of the update) based on data of year N in the ERFS. The last update of the model was completed in late summer 2017, which allows to simulate the social and tax legislation of year 2016 based on ERFS 2014.

The microsimulation performed in INES consists, first of all, in carrying out static ageing on the database (changing the socio-demographic structure and adjusting incomes not dependent on scales) and, second, in simulating scaled benefits and contributions that affect households’ standard of livings (without using information obtained from matching with social and tax files).

A specific feature of French legislation which is taken into account in the INES microsimulation model is that income from previous years is taken into account when calculating certain social benefits and contributions. Income tax paid in a given year N is calculated from the income and the situation of the previous year N-1. Housing allowances are calculated on the basis of income from two years previously (N-2).

In normal use, income from two years previously is the income used in the ERFS. Income from the present year and the previous year is obtained by static ageing. This two-step ageing is used both to produce the years of income from which benefits are calculated and to project in the future (or the present). Specifically, from the ERFS of year N we construct a pseudo-ERFS for year N+1. For this year, a household *m* with socio-demographic characteristics  and income  in the population is represented in our sample by a household surveyed in year N, with characteristics  and income  (second stage of the ageing process) where the weighting has been modified in order to be representative of the population in N+1 (by calibration on margins which is the first stage of the ageing process) and where  is the mean change in income between years N and N+1 (changes depend on the nature of income). The operation is then repeated to construct a pseudo-ERFS for year N+2.

It is worth noting that due to static ageing of the ERFS over two years, there are no sudden variations in an individual’s situation: if a person is unemployed the whole year N, it is assumed that he was also unemployed for the whole of year N+2 with unemployment benefits that simply followed the average trend. This lack of any real individual dynamics suggests a better targeting of means-tested benefits and income tax than in the real life, since the order in the income scale is the same in N+2 as in N and the most severely deprived in N are still the most severely deprived in N+2, and are ultimately the best targeted. In the real life, a person who has lost his job, will pay a high level of income tax the following year compared with his earnings at that time. Conversely, a person whose income has risen sharply over the last two years will nevertheless be able to receive housing benefits, despite his current high income. These discrepancies are due to the French social and fiscal legislation and do not exist in the INES model because of static ageing.

In times of economic downturn, these individual transitions are more common, and thus there is a risk that the targeting error made by INES model may be intensified. This problem has to be taken into account in a nowcasting exercise where we are interested in changes.

*2.2. Diverting INES into nowcasting*

Nearly all the parameters for ageing (except for income derived from movable assets, extra earnings, property income and property deficits) and for producing the scales needed to calculate standard of living in N+1 are usually put into INES at the start of the summer of N+2. It is therefore possible to produce changes in poverty rates between N and N+1 as soon as INES is rebased according to ERFS N. ERFS N is available towards September N+2; if we assume that 1 month’s work is required to change the base, it is possible to gain 11 months over the current delays in publishing this indicator.

**Figure 1. Sequence of the production process for ERFS (top) and nowcasting the poverty rate by N+1 (bottom)**

The nowcasting exercise consists in producing a projection based on the latest information. In Autumn N+1, one single ERFS (that of year N-1) is used in order to produce both evaluations of the poverty rates in year N-1 and year N. The change in poverty rate for the target year N is then derived from the differences between these two evaluations.

Using one single ERFS has the immediate advantage of minimising the sample bias, given the imprecision associated with using two different ERFS. Moreover, nowcasting the poverty rate is an evolving process since the level of the poverty rate simulated with the INES model is weaker than the official one. This difference is due to the use of microsimulation: INES simulates the legislation on scaled benefits and contributions, by imputing any non-take-up randomly, while ERFS uses matched information that has actually been observed, correcting non-responses by imputation. It is also due to the ageing process: INES has to project socio-demographic structure and income over one or two years on the basis of a certain number of assumptions, whereas ERFS uses only the information available.

Specifically, the simulation for the year N-1 based on the ERFS N-1 is hereafter called “contemporary evaluation”, as the simulated legislation is identical to that used for the ERFS. In order to simulate the benefits and contributions of year N-1, “pseudo-ERFS” N-2 and N-3 are produced from the ERFS of year N-1. Thus the operation is no longer a static ageing but rather a static rejuvenation over two years. We use the term “reverse ageing” to refer to this process, as opposed to “classic ageing”.

On the other hand, the simulation for the year N based on the ERFS N-1 is called the “forward evaluation”, as the simulated legislation year is forwarded by one year compared with that of the ERFS. Population and income will be projected over only a single year instead of the usual two years with INES. This implies to construct pseudo-ERFS for years N-2 (reverse ageing) and N (classic ageing).

To carry out these two evaluations based on INES, the classic pattern of static ageing has therefore been diverted (Figure 2). More specifically, the method for projecting from N to N-1 and N-1 to N-2 is identical to that used to project from N to N+1: the calibration margins are the same and the coefficients for revising income are recalculated from the changes used to project from N to N+1 () and used in ways that respect the same rules for the total income trends.



**Figure 2. Diverting the traditional use of INES for the nowcasting exercise**

**3. Main results**

In the exercise carried out in this study, the goal is to publish the variation in the poverty rate between year N-1 and year N as soon as the ERFS for year N-1 becomes available, in other words in the autumn N+1. In October 2017, Insee published for the third time an advanced estimation for year N, using the experimental method based on microsimulation as described above.

In the following, we assess the quality of the early indicators thus produced by comparing the results that would have been produced by microsimulation with those that were in fact disseminated from the ERFS. When applied to the target years 2010 to 2015, this method produced estimations of the at-risk-of-poverty rate similar to the actual figures published the following year (see Table 1).

**Table 1. Variations of the 60% poverty rate estimated with ERFS and INES**

|  |  |  |
| --- | --- | --- |
| **Target year** ***(year N)*** | **Observed variation*****(in September N+2)*** | **Simulated variation*****(in Autumn N+1)\**** |
| 2010 | 0.6 | 0.4 |
| 2011 | 0.3 | 0.1 |
| 2012 | -0.4 | -0.3 |
| 2013 | -0.3 | -0.4 |
| 2014 | 0.1 | 0.2 |
| 2015 | 0.2 | 0.2 |
| 2016 | *in September 2018* | -0.3 |

\*: *simulated variations for the period 2010-2013 were performed in 2015*

*Sample: Ordinary households in Metropolitan France*

The direction of the annual change in poverty rate appears to have been correctly forecast each time for the 60% rate, and the predicted changes are always fairly close (in absolute terms) to the reference values. The ex post test for the last estimates (decrease by 0.3 points in the target year 2016) could be assessed in October 2018.

Similarly, between 2010 and 2015, the results of the nowcasting exercise are also conclusive with regards to some other inequalities indicators. Forecasts for changes in the Gini index were moving in the right direction and the predicted decile ratios P90/P10 were along the same lines as those observed one year later with ERFS (see Table 2). On the other hand, the nowcast of changes in other indicators was less satisfactory (50% poverty rate and poverty gap at 60%). This drop in the quality of estimates could be due to greater difference between INES and ERFS at the lower end of the standard of living distribution.

**Table 2. Variations of the Gini index and the decile ratio estimated with ERFS and INES**

|  |  |  |
| --- | --- | --- |
| **Target year N** | **Variation of the Gini index** | **Variation of the decile ratio P90/P10** |
| **Observed** ***(in September N+2)*** | **Simulated** ***(in Autumn N+1)\**** | **Observed** ***(in September N+2)*** | **Simulated** ***(in Autumn N+1)\**** |
| 2010 | 0.009 | 0.002 | 0.0 | 0.0 |
| 2011 | 0.003 | 0.002 | 0.1 | 0.0 |
| 2012 | -0.003 | -0.002 | 0.0 | 0.0 |
| 2013 | -0.014 | -0.007 | -0.1 | -0.1 |
| 2014 | 0.004 | 0.002 | 0.0 | 0.0 |
| 2015 | 0.003 | 0.003 | 0.0 | 0.0 |
| 2016 | *in September 2018* | -0.002 | *in September 2018* | -0.1 |

\*: *simulated variations for the period 2010-2013 were performed in 2015*

*Sample: Ordinary households in Metropolitan France*

**4. Conclusion**

Nowcasting using INES provides an estimate of the poverty rate 11 months earlier than the official estimate provided by ERFS. Between 2010 and 2015, nowcasting seems to have always predicted changes in the 60% poverty rate along the same lines as those observed one year later with ERFS. The results of the exercise are also conclusive with regards to Gini index and the interdecile ratio P90/P10. Otherwise, results for 50% rate and poverty gap at 60% are less satisfactory, and are not disseminated as flash estimates. Estimates would especially have to be considered with caution in times of crisis or recovery, as during these times the ageing applied to an ERFS dating back one year would not necessarily be relevant in representing reality.

The next flash estimates including the key indicators for year 2017 could be released as experimental data in October 2018 and the ex-post test for the previous year could be assessed at the same time.

**5. References**

Fontaine M. and Fourcot J. (2015), Nowcasting du taux de pauvreté par microsimulation, Document de travail n°1506, Insee.

Schmitt K. and Sicsic M. (2017), Estimation avancée du taux de pauvreté et des indicateurs d’inégalités – Résultats expérimentaux pour 2016, Insee Focus n°96.