BE-FAST Deliverable 2.1 – Literature review and inventory of nowcasting approaches employed to perform distributional analyses in Belgium for the period 2020 – 2022[1]

Julie Janssens and Sarah Marchal

Abstract

During the COVID-19 crisis, researchers interested in gauging the socio-economic impact of the pandemic on household and individual incomes heavily turned to nowcasting methodologies to overcome the lack of timely observational data on incomes. Leveraging the most recent macroeconomic statistics, nowcasting techniques have facilitated the updating of information on the income distribution before the pandemic to the situation in 2020. To monitor the situation in Belgium, different nowcasting techniques with different degrees of detail, in line with the availability of external, aggregate data have been employed. In this deliverable, we present an inventory of the various nowcasting approaches employed for conducting distributional analyses in Belgium from both national and international papers. We describe and compare the different techniques on the basis of the following aspects: the policy measures incorporated in the simulation model; the publication date and period of analysis; the input data and data used for nowcasting (including a discussion on the level of detail and timeliness of the data); the nowcasting method and level of detail included in the modelling; and the way monetary variables are uprated. Additionally, we compare the main findings of the different papers regarding the impact of the pandemic on employment, incomes, poverty and inequality. Overall, we find that despite the variations in nowcasting method and data used, the conclusions drawn from different studies show a considerable degree of similarity. The final section of the paper outlines a plan for comparing the nowcasted results with the ex-post observed distributional impact of the crisis, and the challenges that are associated with this.

[1] We thank Maisarah Wizan, Johannes Derboven and Wim van Lancker for thoughtful comments and suggestions.

1 Introduction

Throughout the COVID-19 crisis, researchers all over the world were confronted with a similar problem: how can we monitor as swiftly as possible the ongoing socio-economic impact of the pandemic? In the face of the daunting and disruptive shock that COVID-19 quickly turned out to be, the usual one to three year time lag of administrative and survey data made common approaches and data sources seem all but useless. In response, researchers and governments rolled out ad-hoc surveys among civil society organizations, such as foodbanks, or surveys drawing on a non-probability sample to monitor the impact of the broader impact of the pandemic. While these surveys did provide important information on the most pressing needs in the early phases of the crisis, their lack of representativeness or focus on very specific issues limited their use for a distributional analysis of the socio-economic impact. Alternatively, and where possible, standing longitudinal panels studies were extended with ad hoc online and phone surveys (e.g. the UKHLS panel study, see Brewer and Tasseva (2020)). In countries where online probability sample surveys were already active at the time, these were used to gather repeated information on citizen's behavior and social change.

Most often however, distributional analyses of the socio-economic impact of the pandemic on household and individual incomes drew on 'nowcasted' data. Depending on the available microdata and aggregate statistics, researchers turned to various modelling approaches to proxy the current income and labour market situation as accurately as possible, and to assess the effectiveness of the social policies introduced to compensate for the loss of earnings due to the imposed lockdown measures. The results of these nowcasting exercises however varied substantially depending on the level of detail applied in the process (Capéau et al., 2022). This level of detail depended to some extent on the availability of external statistics. If the first information at hand is lacking sufficient detail to properly calibrate nowcasting models, the accuracy of the estimates will be affected. When more details become available, the accuracy of the estimates will likely improve. Policymakers need to have a sense of the margin of error between the two time periods to judge when and for what questions it is appropriate to use nowcasting models.

In this first deliverable of WP2 of the BE-FAST project, we present an inventory of the various nowcasting approaches employed for distributional analyses throughout 2020 – 2022 in Belgium. We include exercises from national research teams (including the COVIVAT consortium[2]) and federal administrations, as well as comparative exercises conducted by European institutions. We describe the methodology behind each of these exercises, their timeliness as well as their main results. In the next deliverable of the BE-FAST project, we will compare the nowcasted results with post hoc available observational data.

This paper is organized as follows. Section 2 lays out the main elements of a nowcasting exercise. Section 3 lists the nowcasting analyses focused on the COVID-19 period (2020-2022) in Belgium. These form the basis of our inventory. Section 4 then compares the main characteristics of the different Belgian nowcasting analyses. Finally, in section 5, we lay out our future work plan.

[2] An academic consortium of research teams from the University of Antwerp and the KU Leuven, funded by the Federal Public Service Social Security, with the objective to bring the socioeconomic impact of the pandemic and the accompanying sanitary measures to light. See www.covivat.be for papers and results.

2 Nowcasting

Nowcasting has been described as '*the estimation of current indicators using data on the past income distribution combined with other information including the latest available macroeconomic statistics*' (Navicke et al., 2014). Nowcasted indicators are used in different domains. Nowcasts and forecasts of national account statistics, such as GDP figures, are common (Piette, 2016). With regard to social outcomes, the Flash Estimates of Income and Poverty Indicators published by Eurostat are a well-known example (Leulescu et al., 2023). However, nowcasting techniques exist in different forms.

O'Donoghue and Loughrey (2014) distinguished three different modelling choices that may be necessary when nowcasting an entire income distribution: (1) the use of ageing techniques to account for changes to the population and labour market structure; (2) uprating the monetary amounts in the underlying (outdated) data, for instance indexation to adjust for wage growth; and (3) updating or adjusting the tax-benefit rules in a microsimulation model to account for policy changes. In what follows we briefly discuss the different options for each of these elements. For this description, we heavily build on Neelen et al. (2022).

2.1 Adjusting population characteristics

A major element of nowcasting social outcomes on untimely micro-level data, is the adjustment of population characteristics present in the underlying data to the relevant period. To this end, one can either use static or dynamic ageing techniques. The literature usually refers to reweighting as the prime "static ageing" technique (Immervoll et al., 2005). This implies altering the weights of the observations in the underlying data in order to reflect differences in the composition of the population that have occurred over time. This technique is frequently used when there is a relatively small time lag between the data year and the year one is interested in. The technique is considered less appropriate for long-term predictions (Dekkers, 2012; Dekkers & Liègeois, 2012).

Static ageing is relatively easy to implement but its success depends upon the availability of up-to-date information on the true size of population sub-groups in the year of interest (O'Donoghue & Loughrey, 2014). Often, this sub-group information is not available. In addition, information that was captured by the original weights can be lost, especially when the variables used for the reweighting process differ in some respects to the variables used to construct the original weights (O'Donoghue & Loughrey, 2014). Finally, static aging may distort the joint distribution of important variables, in particular alternative sources of income. In the context of rapidly changing labour markets, reweighting would assign the characteristics of those (un)employed in the baseline to the 'new' (un)employed. In reality, they may differ in important ways, for instance regarding the duration of (un)employment and the resulting eligibility for benefit receipt (Neelen et al., 2022).

Alternatively, one can turn to dynamic ageing techniques. One then allows individuals to change their characteristics due to endogenous factors within the model (O'Donoghue, 2001). More in particular, dynamic ageing techniques model the factors driving the changing characteristics of individuals over time; instead of exclusively relying on trends in statistical aggregates. As such, a dynamic model typically simulates inter-temporal transitions in the population, thus modelling the evolution of individual and household characteristics (e.g. age) as well as their behaviour and specific life events over time (e.g. birth, marriage, education and labour market participation). This changing structure can be combined in the model with evolutions in market mechanisms, tax-benefit policies and the macro-economic environment. If one wants for instance, to take account of changes in the labour market status between the year the available (input) data refer to, and the most recent available information, one might estimate a system of econometric equations to include these changes in the data (Li et al., 2014).

With regard to changing the labour market status in the context of dynamic ageing of the underlying data, a distinction can be made between two nowcasting methods: parametrically estimated labour market status changes and non-parametrically estimated labour market status changes. In the first method, the adjustment of the labour market status in the underlying available input data is based on the probability that observed individuals' labour market status will change, derived from *parametric models* that build on alternative, more current microdata. Probabilities will be estimated through various regressions with explanatory variables such as gender, age and income (Neelen et al., 2022). In the second method the nowcasting methodology does not include the estimation of individual-level probabilities. Instead, labour market status changes are allocated *non-parametrically*, based on information about labour market status change rates in specific strata from external aggregate statistics.

2.2 Uprating market income

Uprating or indexation means that monetary amounts are adjusted to changes over time to account for e.g. income growth or inflation (Neelen et al., 2022). To do so, two choices can be made: either income levels are calibrated to aggregate information from National Accounts, or indices are used to account for changes in external control totals (e.g. by using consumer prices indices, or different indices for different income components if they evolve in a different way in the real world) (O'Donoghue & Loughrey, 2014).

When income levels are calibrated to aggregate national accounts information, this can seriously affect the income distribution. Non-reporting or underreporting of particular sources of income, particularly amongst capital and self-employment income, can lead to under-estimation of incomes relative to the national accounts, (O'Donoghue et al., 2000; Sutherland, 2001).

In case the data allows for it, one could choose to make use of indices by income source or demographic category in the indexation of income. Immervoll et al. (2005) showed that a higher level of disaggregation in income indexation is preferred above the uniform uprating of market incomes.

2.3 Updating the tax-benefit rules in place

Nowcasting approaches that build on microsimulation models evidently adjust the taxes and benefits in line with the most recent legislation. The policies included in the underlying microsimulation models can however differ. Often, the level of detail that is included in microsimulation models depends on the level of detail in the underlying input data, i.e. the extent to which this input data actually allows to estimate often complex benefit rules. An obvious example is the eligibility to unemployment benefits, that usually depends on the length of the previous work career. According to common eligibility rules, this length usually surpasses the reference period included in survey data.

3 Nowcasting COVID 19 in Belgium

One of the big challenges for the distributive analyses of the impact of the COVID-19 social distancing and social policy measures for Belgium, was the unavailability of representative data at the time of monitoring. Both detailed micro-level administrative records as well as representative survey data only come available with a certain lag. At the start of the pandemic the most recent available, representative survey data was the EU-SILC for 2018, which contains incomes based on the situation in 2017.

Various efforts were taken to remedy this situation. Importantly, ad-hoc (non-representative) surveys were organized. Highly noteworthy in this regard is the Corona Study[3]. This study was carried out by researchers from the University of Antwerp, in collaboration with the ULB, the KUL and the UHasselt, as soon as the lock-down measures were implemented. The survey was repeatedly organized over a period of two years, first on a weekly basis but later with larger timespans in between, to collect information on people's health, their willingness to behave according to the lockdown measures, and the impact these measures had on their personal wellbeing and financial situation. As regards administrative data, the Working Group Social Impact COVID-19 (WG SIC) and the Taskforce Vulnerable Groups (TVG) managed to rapidly disclose administrative data from the relevant social security institutions, offices and departments in Belgium, to inform policymakers and the policy process through regular monitoring reports.

In Belgium, the COVIVAT project set out to track the impact of the COVID-19 shock and the income support measures on the household income distribution right from the onset of the crisis in 2020. One of the tools used to assess the impact of the COVID-19 shock was the use of nowcasting methods on the most recently available survey data at that time (EU-SILC 2018) combined with aggregate administrative data on benefit receipt. In line with the increasing availability of (ever more detailed) external data, consecutive COVIVAT working papers and policy notes have applied different nowcasting methods, as well as different levels of detail, in order to update the EU-SILC 2018 to reflect as well as possible the situation in 2020.

In a first COVIVAT nowcasting paper, **Marchal et al. (2021)** assessed the impact of the lockdown in April 2020 in Belgium on individual and household incomes in that month. The paper used a parametric model and data from the corona study to estimate the probability of a person becoming temporarily unemployed (or receive a bridging rights) as a result of the pandemic, and calibrated the results using external aggregated data.

In a second COVIVAT nowcasting paper, **Capéau et al. (2021)** calculated the impact of the COVID-19 shock on the incomes of employees for the entire year of 2020. In contrast to the parametric model used in Marchal et al. (2021), a non-parametric model was used to randomly allocate labour market statutes on the basis of external aggregated administrative data.

In a third COVIVAT nowcasting paper, **Derboven et al. (2021)** built further on the method developed in Capéau et al. (2021), but focused on the change in monthly incomes due to COVID-19 shock and the (stabilizing) impact of COVID-19 policy measures instead of on the yearly impact of the COVID-19 shock.

In a next COVIVAT paper, **Capéau et al. (2022)** compared two different nowcasting approaches to analyse the impact of the COVID-19 shock and the compensation measures taken to preserve citizens' income, on the annual change in disposable incomes. Both methods built on random allocation, yet they differed in the level of detail and heterogeneity in the simulation of the impact of the shock on incomes and the modelling of the compensation policies.

[3] https://www.uantwerpen.be/nl/projecten/coronastudie/resultaten/

Finally, **Neelen et al. (2022)** present the last nowcasting exercise that was performed in the context of the COVIVAT project. This method builds further on the simulation model used in Derboven et al. (2021) and therefore looks at the monthly impact of the COVID-19 shock and compensation measures on monthly incomes. As in Derboven et al. (2021), they apply a non-parametric allocation, but use detailed external data (especially for the self-employed) as well as transition information for the two COVID-19 peaks in 2020.

In addition to the papers that have been written by the COVIVAT consortium, in this inventory we also include other nowcasting exercises that have been performed to gauge the socio-economic impact of the COVID-19 shock on citizens' incomes in Belgium. The Belgian Federal Planning Bureau studied the monthly impact of the Belgian government's compensation measures on household incomes in the period March to May 2020 (**Thuy et al., 2020**), using a non-parametric approach on a large administrative dataset.

Almeida et al. (2021) analysed the impact of the COVID-19-crisis on households' income in the European Union member states, including Belgium. They used economic forecasts for 2020 issued by the European Commission and a counterfactual scenario in which they simulate a COVID-19 year in which no policy measures were taken by EU Member states to cushion the effect of the crisis. Consequently, the information from the macroeconomic forecasts was used to reweight the most recent available EU-SILC data at the household level to mimic aggregate employment figures in each scenario and change the number of (un)employed in the EU-SILC dataset.

Also **Christl et al. (2021)** analysed the impact of the pandemic and the cushioning effect of fiscal policy for all EU member states (including Belgium). They used the Labour Market Adjustment (LMA) add-on in EUROMOD to adjust labour market conditions to reflect those of 2020 in a non-parametric way. Detailed aggregate labour market statistics were used to simulate transitions from work into unemployment and monetary compensation schemes.

Eurostat publishes Flash estimates of income inequalities and poverty indicators for all EU member states, and has evidently continued to do so during the pandemic (Eurostat, 2021, 2022, 2023). These indicators are calculated based on microsimulation and nowcasting techniques and have a release date appreciably earlier than the yearly EU-SILC results. The calculation method differs by country and has changed over the years. Nonetheless, the main methodology used to produce the flash estimates is the estimation of yearly changes on the basis of auxiliary information already available for the target year implemented on the EU SILC. For Belgium, before the COVID-19 crisis (2019 flash estimates), the method used to adjust population and labour market characteristics was a (static) reweighting method on the basis of EU-LFS data (Leulescu et al., 2020). When the COVID-19 pandemic started, several methodological changes were implemented and Eurostat changed to a dynamic method of nowcasting in which employment transitions were estimated on the basis of a parametric model for the 2020 FE and a non-parametric approach for the FE of 2021 (Leulescu et al., 2021; Leulescu et al., 2022). After the pandemic, the methodology of the FE 2022 came back to pre-pandemic standards, but some developments were consolidated in the estimation process (e.g., the use of a dynamic, parametric model to calculate probabilities of labour transitions) (Leulescu et al., 2023).

4 Inventory and comparison

In what follows, we compare the different characteristics of the nowcasting exercises that are included in our analysis. We discuss successively the following components of the nowcasting exercise:

- The policy measures modelled (section 4.1)
- The publication date and period of analysis (section 4.2)
- The adjustment for changes in population and labour market characteristics (section 4.3), including an in-depth discussion of the data (section 4.3.1) and methodology used for nowcasting (section 4.3.2)
- The uprating of monetary amounts (section 4.4)
- The main findings (section 4.5)

4.1 Policy measures

In the earliest weeks of the COVID-19 pandemic, the federal and regional governments in Belgium implemented important support measures to absorb the income shock for employees, self-employed, and – more general – households due to the crisis. To protect employees and self-employed against the loss of earnings, the federal government announced an extension of both the temporary unemployment scheme and the bridging right. Both benefits existed already before the pandemic and granted compensation in case an employee or self-employed was – due to "force majeure" – unable to perform her (self-)employment activity. From March 13th, 2020 onwards the federal government relaxed the eligibility criteria for these benefits and all persons who had to stop their work activity due to the Covid-19 crisis were entitled to the temporary unemployment benefit or bridging right. Also the access criteria to the ordinary unemployment benefit were (somewhat) relaxed, while the degressivity of the benefit was halted. Throughout 2020, the federal level took further measures, such as the supplement to means-tested benefits in the Summer of 2020, and the double bridging right in Autumn 2020.

Also the regional governments implemented support measures during the pandemic. The most sizeable ones were those granted to businesses that had to interrupt their activities because of the lockdown measures (for example the *hinderpremie, compensatiepremie, vlaams beschermingsmechanisme, 'indimnité compensatoire*, see Capéau et al., 2022 for details). Still, also individuals and households were targeted, by utility and rent premia and social supplements within the child benefit schemes.

Table 1 shows for the nowcasting exercises included in our inventory the policy measures that were studied in each of the papers. In every microsimulation model a standard range of social policies and taxes are modelled. However, potential extensions to the standard microsimulation model are possible if researchers invest in extending the standard microsimulation model with more policies. The two microsimulation models that are used in the papers included in our inventory are the EUROMOD and EXPEDITION microsimulation model. Only one paper (Thuy et al., 2020) makes use of the EXPEDITION microsimulation model. This model was developed by the Federal Planning Bureau in order to simulate the effects of the 2019 election programs (Federaal Plan Bureau, 2018). It uses pseudonymized administrative data from the Crossroads Bank for Social Security (CBSS) for a large representative sample of the population. Consequently, the model allows to estimate the expected direct effects of certain social policy reforms on public spending and/or revenues and on the income distribution. The model replicates policies in the domain of sickness and disability benefits, unemployment benefits, social assistance benefits, child benefits, social security contributions, and personal income taxes. As such, the temporary unemployment benefit and bridging right were already included in the model before the start of the pandemic. Thuy et al. (2020) updated the policies in line with the new COVID-19 eligibility criteria. Further pandemic measures were not added to the model (also as the paper focused on the first months of the pandemic).

Table 1. Microsimulation model used and policy measures modelled in the different papers

Paper	Model	Reported extensions
Thuy et al. (2020)	EXPEDITION (TU, BR)	
Almeida et al. (2021)	EUROMOD (TU, BR)	
Marchal et al. (2021)	EUROMOD (TU, BR)	Monthly April 2020
Christl et al. (2021)	EUROMOD (TU, BR)	
Capéau et al., (2021)	EUROMOD (TU)	Supplement long-term unemployment
Derboven et al. (2021)	EUROMOD (TU, BR)	Supplement social assistance beneficiaries, supplement long-term unemployment, Flemish energy premium, Brussels rent premium, Walloon water premium, Flemish and Brussels child benefit supplement
Eurostat (2021) – 2020 FE	EUROMOD (TU, BR)	
Capéau et al., 2022 (less detail)	EUROMOD (TU, BR)	Supplement long-term unemployment
Capéau et al., 2022 (more detail)	EUROMOD (TU, BR)	Supplement long-term unemployment *Hinderpremie, compensatiepremie, vlaams beschermingsmechanisme, indimnité compensatoire*
Neelen et al., (2022)	TU, BR,	Supplement social assistance beneficaries, supplement long-term unemployment , Flemish energy premium, Brussels rent premium, Walloon water premium, Flemish and Brussels child benefit supplement
Eurostat (2022) – 2021 FE	EUROMOD (TU, BR)	
Eurostat (2023) – 2022 FE	EUROMOD (TU, BR)	

TU = temporary unemployment benefit, BR = bridging right

All other papers included in our inventory use the EUROMOD model. EUROMOD is a tax-benefit microsimulation model for the European Union that enables researchers and policy analysts to calculate the effects of taxes and benefits on household incomes and work incentives for the population of each country. The Belgian version uses the EU-SILC as underlying database. The EUROMOD country reports for Belgium (Assal et al., 2021, 2022; Assal et al., 2020) describe the inclusion of the bridging right and the temporary unemployment system to the EUROMOD system in the pandemic years. Other federal or regional compensation measures that were implemented during the pandemic were, according to the country reports, not included in the updated version of the model due to the lack of information in the underlying input data set to identify beneficiaries for these measures.

As such, all papers in our inventory analyse the impact of the pandemic taking account of the temporary unemployment and bridging right benefits. Capéau et al. (2021) zoom in on the employed population, and therefore do not study the bridging right.

A number of papers report the inclusion of additional policies (see Table 1). For instance, Derboven et al. (2021) and Neelen et al. (2022) include the social assistance premium, as well as the regional measures intended to support individual households, such as the various energy premia and supplements in the child benefit systems. Capéau et al. (2022) is the only paper that models specific regional support measures targeted at businesses, based on estimated turnover losses for self-employed.

Finally, a number of papers (Marchal et al. (2021), Derboven et al. (2021) and Neelen et al. (2022)) takes steps to proxy monthly incomes as well as possible, rather than annual averages.

4.2 Publication date and period of analysis

Table 2 shows the papers included in our inventory in order of their earliest publication date. The Belgian Federal Planning Bureau was the first to publish results on the impact of the COVID-19 shock on the monthly disposable household incomes of the persons eligible for a temporary unemployment benefit or bridging right. They studied the period March-May 2020 and were able to publish their results within the same year. The JRC of the European Commission followed suit with their first analysis of the impact of the pandemic in Belgium (and other EU member states) as they submitted their first results in September 2020 which resulted in a publication by June 2021. The first COVIVAT paper that studied the impact of COVID-19 on the incomes nowcasted for April 2020 was published in January 2021. Consecutive COVIVAT papers were published in line with external administrative data that became gradually available and which were used in the nowcasting methodologies (see section 4.4). The Flash Estimates of Eurostat are published with a six month to one-year delay. Usually the first estimates of the previous income year are published around July-August of the next year. The methodological note and final estimates follow in December of that year. Note that this is appreciably faster than the release of the final EU-SILC data which is usually available with a two year delay. The Flash Estimates for 2021 were not published for Belgium. In the remainder of our paper, we will therefore remove this publication from our review.

Table 2. Publication data and period of analysis of the different papers

Paper	Period of Analysis	Publication date
Thuy et al. (2020)	March - April - May 2020 (monthly impact)	September, 2020
Almeida et al. (2021)	2020 (annual impact)	June, 2021 (received sept. 2020)
Marchal et al., (2021)	April 2020	January, 2021
Capéau et al. (2021)	March - December 2020 (annual impact)	April, 2021
Christl et al. (2021)	2020 (annual impact)	July, 2021
Derboven et al. (2021)	March - December 2020 (monthly impact)	December, 2021
Eurostat (2021) – 2020 FE	2020	September, 2021
Capéau et al. 2022 (- detail)	March - December 2020 (annual impact)	April, 2022
Capéau et al. 2022 (+ detail)	March - December 2020 (annual impact)	April, 2022
Neelen et al. (2022)	March - December 2020 (monthly impact)	October, 2022
Eurostat (2022) – 2021 FE	2021	Not published for BE
Eurostat (2023) – 2022 FE	2022	June, 2023

4.3 Adjusting population characteristics

As discussed in section 2, different methods can be used for the adjustment of changes in population characteristics (such as labour market status) from the base year to the target year. The method that is used is inextricably linked to the (type of) data that is available. Both the level of detail of the external data and the timeliness of the data are important in this respect. In section 4.3.1 we discuss the data that are used in the different nowcasting exercises by presenting successively: (1) the input data used for the base year; (2) the external (aggregate) data used for nowcasting; (3) the level of detail of the external data; and (4) the timeliness of the external data.

Next, section 4.3.2 provides an in-depth discussion of the nowcasting methodology used in the different papers. We compare the different papers in terms of: (1) the ageing technique employed; (2) the nowcasting methodology used; (3) the level of detail included in the modelling; (4) the transitions modelled; and (5) the adjusted population groups and way income is adjusted.

4.3.1 Data

Table 3 gives an overview of the data that are used in the different nowcasting exercises.

Input data

Most papers used the most recent version of the EU-SILC that was available at the time of their nowcasting analysis. The EU-SILC is a yearly survey carried out in all EU member states on the income and living conditions of private households. It contains a representative sample of private households in each country. For Belgium the sample size is generally around 14 000 individuals. Only one paper in our inventory uses another source of input data. Thuy et al. (2020) make use of pseudonymised administrative data underlying the EXPEDITION microsimulation model. This dataset contains data from different administrative sources for a large representative sample (N= 601683) of the Belgian population on the 1st of January 2012 (Federaal Planbureau, 2018).

Aggregate statistics

Except for the paper of Almeida et al. (2021) which uses Forecasts of the European Commission to reweigh the SILC microdata, all papers in our inventory use of aggregate administrative data. In papers that use a stratified sampling, non-parametric approach (cf. Table 4), aggregate administrative data are used as the only source for nowcasting (e.g. Thuy et al., 2020; Derboven et al., 2021; Capéau et al., 2021; Capéau et al., 2022; Neelen et al., 2022), while in others administrative data are used in combination with survey data (respectively from the Corona Study or Labour Force Survey) (Marchal et al., 2020; Christl et al., 2020; Eurostat, 2021; 2022; 2023).

The National Employment Office (NEO), the National Social Security Office (NSSO) and the NISSE (National Institute for the Social Security of the Self-employed) provided information on the number of persons that received the temporary unemployment benefit or bridging right during the pandemic to the Belgian researchers.

Table 3. Data used for nowcasting

Paper	Source input data	Source external statistics	Description external statistics	Timeliness external statistics
Thuy et al. (2020)	Administrative data (EXPEDITION, 2012)	NEO, NSSO, NISSE, USS	E: linked RVA-RSZ data, detailing TU by occupational status (2), gender (2), daily wage level (cont.) and parity committee SE: BR, general recipient numbers	Available Jul 2020
Almeida et al. (2021)	EU-SILC 2017	Forecasts European Commission		Available Nov 2019, Spring 2020
Marchal et al. (2021)	EU-SILC 2018	Corona study, NEO, NSSO, NISSE	E: TU, by sector (21), age group (5 years), gender, province, region, occupational status (2) SE: BR, by occupation code (>60), age group (5 years), region, type of SE and gender	Available Nov 2020
Christl et al. (Feb, 2021)	EU-SILC 2018	LFS, administrative data (available in EUROMOD)	E: TU by sector (up to August) SE: BR by sector (Q1-Q2)	Available Sept 2020
Capéau et al. (April, 2021)	EU-SILC 2018	NEO, NISSE	E: TU by gender, age group (4), sector (22), daily wage (5) and numbers of days in TU (5) + monthly info on transitions	Available Jan 2021
Derboven et al. (December, 2021)	EU-SILC 2018	NEO, NISSE	E: TU by gender, age group (4), sector (22), daily wage (5) and numbers of days in TU (5) + monthly info on transitions SE: BR by occupation code (>60), age group (5 years), region, type of SE and gender (M/F)	Available Jan 2021 (TU), Feb 2021 (BR)
Eurostat (2021) – 2020 FE	EU-SILC 2018	LFS 2020, administrative data	E: number of TU (admin.) [tbc] SE: number of BR (admin.) [tbc]	tbc
Capéau et al. (April 2022) (LD)	EU-SILC 2018	NEO	E: TU by sector SE: BR by sector	Available Jan 2021
Capéau et al. (April 2022) (MD)	EU-SILC 2018	NEO, NSSO, NISSE	E: TU by gender, age group (4), sector (22), daily wage (5), number of days in TU (5), LM status previous month SE: BR by sector (and aggregate transitions)	Available Jan 2021
Neelen et al., (2022)	EU-SILC 2018	NEO, NISSE, CBSS, agg. peak-to-peak transition rates KUL.	E: TU by gender, age group (4), sector (22), daily wage (5), number of days in TU (5), LM status (7) + monthly transitions SE: BR by gender, age, sector, income (2019), LM status (2) + monthly transitions	Available May 2022 (CBSS)
Eurostat (2023) – 2022 FE	EU-SILC 2020	LFS 2022, administrative data, aggregate LFS data	E: number of TU (admin.) [tbc] SE: number of BR (admin.) [tbc]	Version 1, published June 2023

Note: E = employees, SE = self-employed, TU = temporary unemployment benefit, BR = bridging right; LD: less detail; MD: more detail; FE: flash estimates; LM: labour market. Tbc: to be confirmed – the notes on the Eurostat Flash estimates describe the methodology in broad terms, not zooming in on specific countries. Where necessary for the final FAST deliverable, we will contact Eurostat with more detailed follow-up questions throughout this project.

Thuy et al. (2020) were the first to receive statistics from the NEO and NSSO on the number of persons in temporary unemployment, in high detail, including information on labour status, gender, parity committee and daily wage level for the months March, April and May 2020 (see Table 3). These numbers were made available in June 2020. In addition, information was received on the number of days the temporary unemployment benefit was paid to employees. For the self-employed individuals, the authors received aggregate statistics on the number of bridging right payments carried out in the months March to May; these aggregate numbers were not differentiated by socio-demographic characteristic.

In November 2020, the COVIVAT consortium received the first aggregate statistics from both the NEO and NISSE. These data included the numbers of persons that received a temporary unemployment or bridging right during the first months of the pandemic (from April to August 2020) (see Table 2). The NEO data showed the share of temporary unemployed individuals in each sector that was less than 6 days, 6 to 12 days, 13 to 19 days, 20 to 25 days and 26 days or more temporary unemployed.

By the time of writing the second (Capéau et al., 2021) and third COVIVAT nowcasting paper (Derboven et al., 2021) (January, 2021), the researchers had received more detailed external data on the number of persons in temporary unemployment. An important improvement relative to the external data previously used in Marchal et al. (2021), was that the external data referred to more detailed subpopulations (specifically: also defined by daily wage category) and were available on a monthly basis conditional upon the status in the previous month. These data referred to the entire year 2020. Data for the very last months of 2020 was provisional or unavailable, as the data were made available already early January 2021. In Capéau et al. (2021) the authors also used statistics from the National Social Security Office (NSSO) on the number of flex-workers per sector for the first two quarters of 2020.

With regard to data on self-employment, both in Marchal et al. (2021), Derboven et al. (2021) and Capéau et al. (2022), external aggregated data provided by the NISSE were used to assign the bridging right for the self-employed. Percentages of those receiving a bridging right were available by sector, gender and age group. General monthly transitions, over the entire populace of the self-employed, were also provided. The data were received early February 2021.

For the final COVIVAT paper, Neelen et al. (2022) received in May 2022 updated aggregate external administrative data from the Crossroads Bank for Social Security (CBSS). This CBSS data comprised monthly statistics for both employees and self-employed for the period March - December 2020. The main improvements of the updated data were the final numbers of (temporary) unemployed by subpopulation (especially for the last quarter of 2020), the more fine-grained wage categories for employees (i.e. higher up in the wage distribution) and the availability of monthly information on self-employment and transition rates by income category level. Finally, Neelen et al. (2022) received aggregate peak-to-peak transition rates from the KU Leuven, which allowed to re-anchor the monthly transitions in order to better proxy those whose labour market status and earnings were affected in both COVID-19 waves in 2020.

Only one paper written by the COVIVAT consortium makes use of survey data (i.e. the online Corona study) in their nowcasting exercise. Additionally, both the JRC and Eurostat use survey data from the Labour Force Survey (in combination with administrative data) in their nowcasting methodologies. Marchal et al. (2021) use in their paper survey data from the Corona Study. The data used here are those of the seventh wave, which was fielded on April 28th. The variables used from the Corona study were: gender, age, education, parttime work, occupation status and sector, recoded in dummy variables or categorical variables.

For producing the Eurostat flash estimates, a combination of Labour Force Survey (LFS) data and administrative data are used. From the LFS, longitudinal quarterly data on employment status are used, aggregated by gender, age group, (economic) sector and type of contract (temporary or permanent) (only for employees). In addition, Eurostat used administrative data, provided by the different EU Member states, on the number of beneficiaries of different wage compensation schemes. Finally aggregated EU-LFS statics were used for calibration for the FE 2022 (see below). The level of detail of these statistics depends on the country and the sample size, but most often employed were: sex (male/female), age group (16-24/25- 64), status in employment (employees/self-employees), sector (7 aggregations of sectors) and type of contract (permanent/temporary) (Leulescu et al., 2021; Leulescu et al., 2023).

In the technical report of the JRC (Christl et al., 2021) similar data to those used by Eurostat, are used for producing the Eurostat flash estimates, consisting of a combination of LFS data provided by Eurostat and administrative data from national sources provided by EUROMOD National Teams. The LFS data were used on a quarterly basis for the year 2020 and disaggregated by gender and occupation (employees vs self-employed), i.e. with a lower level of detail compared to the LFS statistics used to produce the flash estimates. The administrative data on the number of employees with monetary compensation schemes were available to the researchers up to August 2020. These statistics were available at the sectorial level and are further disaggregated by duration of the monetary compensation spell and the reduction in the hours worked during the spell. For the self-employed, the number of persons with compensation schemes were available by economic sector, but were only available for the first two quarters of 2020.

Level of detail and timeliness of the data

Summarizing the information above, we can make a few observations regarding the level of disaggregation and timeliness of the external data used for nowcasting in the different papers. First, there is a marked difference between data on employees and self-employed individuals. For employees, data are generally rapidly available at the level of detailed subpopulations. In fact, the first information on the number of employees that received the temporary unemployment benefit were already available after 3 to 6 months following the pandemic outbreak in Belgium (Thuy et al., 2020; Marchal et al., 2021). After 9 months, quasi-final numbers on the persons in temporary unemployment were available for different subpopulations, also defined by previous daily wage (Capéau et al., 2021; Derboven et al., 2021). Month-to-month transitions between the different statuses (employed, temporary unemployed by days of temporary unemployment, unemployed) were available at the subpopulation level. For the self-employed, it was generally harder for the researchers to obtain information by (previous) income level, although data by other (often very detailed) categories were delivered swiftly (cf. above). Still, it was only two years after the start of the pandemic that complete information on monthly transitions by different subpopulations (i.e. by gender, age, sector and – importantly - yearly income of 2019) became available to the researchers of the COVIVAT consortium (Neelen et al., 2022).

4.3.2 Nowcasting methodology

When comparing the nowcasting methodologies employed in the different papers in our inventory, table 4 summarizes for all papers: (1) the ageing technique applied; (2) the nowcasting method used; (3) the level of detail included in the modelling; (4) the transitions modelled; (5) the adjusted population groups and the way income is adjusted.

Table 4. Nowcasting methodology

Paper	Ageing technique	Nowcasting method	Level of detail in modelling	Transitions modelled	Adjusted groups and adjustment of income
Thuy et al. (2020)	Dynamic	Non-parametrically, random allocated labour market status	Stratified sampling at gender (2), statute (2), daily wage (5) and committee level	E -> TU SE -> BR	E + SE Income adjusted in line with period TU/BR
Almeida et al. (2021)	Static	Reweighting method based on economic forecasts on GDP and employment			
Marchal et al. (2021)	Dynamic	Parametrically-estimated labour market status (estimation probability of change in status) + calibration on administrative data	Logit model using gender (2), age (2), educational attainment (2), occupation (4), work regime (2) and sector (NACE1)	E -> TU SE -> BR *(April SILC to 'April 2020')*	E + SE Monthly incomes Income adjusted in line with TU/BR
Christl et al. (2021)	Dynamic	Non-parametrically, random allocated labour market status	Stratified sampling at sectorial level ; gender and occupation (only UN)	E -> TU SE -> BR E/SE -> U	E + SE
Capéau et al. (2021)	Dynamic	Non-parametrically, random allocated labour market status	Stratified sampling at the sectorial level, gender (2), age (4), wage-category (3), labor market status previous month	E -> TU* TU* -> E E -> U U -> E TU* -> U U -> TU* *Month to month*	E + FLEXI Income adjusted in line with period TU and working hour reduction
Derboven et al. (2021)	Dynamic	Non-parametrically, random allocated labour market status	Stratified sampling at the sectorial level (22), gender (2), age (4), wage-category (3), labor market status previous month (E); Stratified sampling at sectorial level (SE)	E -> TU* TU* -> E E -> U U -> E TU* -> U U -> TU* SE -> BR *Month to month*	E + SE Monthly incomes
Eurostat (2021) –	Dynamic	Parametrically-estimated labour market status (estimation probability of	Logit model using gender (2), age group, sector (10), occupation (4) and type of contract (2)	U -> E E/SE - > STU STU -> LTU	Active population

Paper	Ageing technique	Nowcasting method	Level of detail in modelling	Transitions modelled	Adjusted groups and adjustment of income
2020 flash estimates		change in status) + calibration on administrative data		E/SE -> TU/BR *Quarterly, net transitions*	
Capéau et al., 2022 (less detail)	Dynamic	Non-parametrically, random allocated labour market status	Stratified sampling at the sectorial level	E -> TU SE -> BR	E + SE Monthly incomes
Capéau et al., 2022 (more detail)	Dynamic	Non-parametrically, random allocated labour market status	Stratified sampling at the sectorial level (22), gender (2), age (4), wage-category (3), labor market status in previous month (E); Stratified sampling at sectorial level (SE)	E -> TU* TU* -> E E -> U U -> E TU* -> U U -> TU* SE -> BR BR -> SE	E + SE Monthly incomes + self-employment income
Neelen et al., (2022)	Dynamic	Non-parametrically, random allocated labour market status	Stratified sampling at the sectorial level, gender (2), age group (4), wage-category (3), labor market status previous month (E) stratified sampling at sectorial level, gender (2), yearly income category (SE), peak-to-peak transition)	E -> TU* TU* -> E E -> U U -> E TU* -> U U -> TU* TU* -> TU* SE -> BR BR -> SE *Month to month*	E + SE Monthly incomes
Eurostat (2023) – 2022 flash estimates	Dynamic	Parametrically-estimated labour market status (estimation probability of change in status) + calibration on administrative data	Logit model using gender, age group, sector, occupation and type of contract + calibration based on aggregate data (gender(2), age (2), occupation (2), sector(7), contract type (2))	UN -> E E/SE - > STU E/SE - > LTU STU -> LTU *Quarterly, net transitions*	Active population

E = employees, SE = self-employed, TU = temporary unemployment benefit, BR = bridging right, *Different transitions from and to temporary unemployment were modelled taking into account the days of temporary unemployment each month. Following five categories were used: 1-5, 6-12, 13-19, 20-25, 26+ days of temporary unemployment in a particular month.

Ageing technique, nowcasting method and level of detail in modelling

Table 4 shows the different ageing techniques and nowcasting methods that are used in the different papers. A few observations can be made. First, only one paper, Almeida et al. (2021), uses a static ageing technique. The authors reweigh the underlying EU-SILC microdata at the household level by using macroeconomic forecasts of the European Commission and applying the predicted changes in total (un)employment, wages received, population structure and other macroeconomic indicators in order to replicate different scenarios of interest (e.g. COVID-19 and no-COVID19 scenarios). Only the characteristics of the active population were adjusted for changes as no reweighting was applied for pensioners, students, inactive persons, and sick or disabled persons.

All other papers included in our review use a dynamic ageing technique to adjust labour market statutes of the active population during the pandemic. One can choose between parametrically or non-parametrically approaches to simulate transitions between labour market statuses. Marchal et al. (2021) is the only COVIVAT paper that uses a parametrically estimated model, based on the online Corona Study[4] (see above). The information from the survey allowed to estimate the odds of becoming temporary unemployed or receiving a bridging right in April 2020, on the basis of gender, age, educational attainment, occupation, work regime and sector. This model was then applied to EU-SILC 2018 data in order to identify the likelihood of individual observations becoming temporary unemployed or receiving the bridging right in April 2020. In a second step, aggregated administrative data delivered by Belgian administrations were used to calibrate the results.

Second, Table 4 shows that the consecutive COVIVAT papers used a dynamic ageing method and a non-parametric approach to adjust changes in labour market statuses. This implies that different labour market statuses (e.g. temporary unemployment, ordinary unemployment, bridging right) were randomly allocated in the EU-SILC data in line with external, administrative data on the occurrence of theses statuses in the population. The precise subpopulations used for the stratified sampling (see column 4 in Table 4) depended on the level of detail and the discerning population characteristics present in the external data (as reported in Table 3), but also on the level of detail possible to allocate on the SILC sample. Most papers were able to take sector, gender, age, wage-category and status in the previous month into account. In order to be able to allocate labour market statutes in line with the more detailed administrative data, the researchers inflated the EU-SILC with a factor 10.

For the self-employed, the stratified sampling was initially applied by sector, gender and age group. However, as there was no disaggregated information available about monthly transitions, the chances for a self-employed person to receive a bridging right remained stable throughout the entire period studied (Derboven et al., 2021). Transitions into and out of bridging right status as well as the overall number of affected changes, would therefore always affect the same observations. In Capéau et al. (2022) this was remedied by applying aggregate transition probabilities. As for the employees, allocated statuses could change from one month to another in line with transition rates in the external data. This paper also added far more variation and empirical substance to the modelling of income losses for the self-employed (see below). Neelen et al. (2022) finally used monthly transition probabilities at the disaggregated level to allow movements into and out of bridging right. Additionally, the new administrative data available to Neelen et al. (2022) allowed to recalibrate the nowcasting

[4] https://www.uantwerpen.be/nl/projecten/coronastudie/resultaten/

through labour market status transition information, for both employees and self-employed, from April to November 2020, the two peaks of the COVIVAT social distancing measures in 2020.

Finally, both the JRC and Eurostat used dynamic nowcasting methodologies to analyse the impact of the COVID-19 pandemic in EU Member states. In the JRC Technical Report of February 2021, Christl et al. (2021) used the Labour Market Adjustment (LMA) add-on in EUROMOD to simulate transitions from (self)-employment into unemployment or monetary compensation schemes in order to nowcast labour market conditions of 2020 in the underlying EU-SILC 2018 data. The LMA add-on in EUROMOD was developed for the specific purpose of nowcasting and adjusts labour market characteristics and incomes of those prior identified to change their labour market status. The LMA add-on allows for the modelling of transitions using random allocation and aggregate statistics (see Table 4 for more information on the specific categories used).

The approach used by the JRC is similar to the approach used by Eurostat in the production of the Flash Estimates (FE). Where initially a static nowcasting methodology (i.e. reweighting) was used for producing the 2019 flash estimates (at least for Belgium), Eurostat has changed their approach to a dynamic nowcasting methodology during the pandemic. In the dynamic approach, changes in employment were modelled by simulating individual labour market transitions using a parametric (FE 2020) or non-parametric (FE 2021) model. For both years, transitions between employment, short-term unemployment, long-term unemployment and into monetary compensation schemes were modelled[5]. In all years, transitions were modelled on a quarterly basis, in order to capture infra-annual movements between statuses. For the production of the 2020 FE, the individual probability of being employed, unemployed or inactive was modelled with a logistic regression on EU-LFS longitudinal data. The logistic model was then applied to the SILC data to identify observations for transitions. The main covariates used to identify profiles of workers entering transitions were age, sex, education, economic sector, occupation and type of contract (temporary vs permanent). For producing the 2021 FE, Eurostat switched to a non-parametric allocation of labour market transitions based on the LFS target rates by stratum. The reason for this switch was related to changes in the LFS longitudinal data that prevented the individual probabilities to be calculated. Finally, after the pandemic, Eurostat switched back to a dynamic, parametric model in which the probability to lose/find employment was estimated via a logit model and EU-LFS longitudinal data. The information used from the EU-LFS was the same as for the production of the 2020 FE (i.e. age, sex, education, economic sector, occupation and type of contract). The allocated transitions were calibrated by different strata on the basis of aggregate statistics. The level of detail of these statistics depends on the country and the sample size. The most often employed were: sex and age group (16-24/25- 64), status in employment (employees/self-employees), sector (7 aggregations of sectors) and type of contract (permanent/temporary), but the methodological notes make no specific mention of the detail applied to the Belgian sample.

Transitions modelled

When individual transitions between labour market statuses are modelled, the type and number of transitions may differ. The one transition that is modelled in all (dynamic) papers is the transition from employment to temporary unemployment for employees. Except for Capéau et al., 2021, who explicitly focus on the employed population, all papers also include the transition from self-

[5] In 2020, four transitions are modelled, in 2021 also a fifth transition, from employment to long-term unemployment was added.

employment towards receipt of the bridging right. Aside from the temporary unemployment benefit and bridging right, most papers also looked at ordinary unemployment, and modelled the transition from (self)-employment to ordinary unemployment (Christl et al., 2021, Capéau et al., 2021; Derboven et al., 2021, Capéau et al., 2022, Neelen 2022, Eurostat, 2021; 2022). Eurostat (2021; 2022; 2023) further distinguishes between the transition to short-term unemployment (STU) vs. long-term unemployment (LTU) and models the transition between both statuses (i.e. from STU to LTU). Additionally, Eurostat also includes the transition to partial unemployment in terms of reduced working hours. While for employees, both the transition to compensation measures (e.g. temporary unemployment) and reduced working time is modelled, for self-employed the effect of the pandemic is only modelled in terms of reduced working hours. An important difference between Eurostat (and by extension other papers using the LFS data) and the COVIVAT papers is that both Eurostat and Christl et al. (2021) only model transitions on quarterly basis, whereas COVIVAT papers (from Capéau et al., 2021 onwards) model month-to-month transitions between different statuses [6]. This evidently also leads to variation in the overall duration of (temporary) unemployment.

In Capéau et al. (2021; 2022[7]), Derboven et al. (2021) and Neelen et al. (2022) transitions from and to temporary unemployment were modelled, distinguished by the numbers of days in temporary unemployed in a particular month (i.e. 1-6 days, 7-12 days, 13-19 days, 20-26 days, + 26 days). Marchal et al. (2021) did not simulate transitions by number of days, however the length of the unemployment spell was taking into account by assigning the number of days on the basis of aggregate unemployment figures, once those individuals that make the transition to temporary unemployment were identified. A similar approach was used in Thuy et al. (2020) to assign the numbers of days of temporary unemployment after the random allocation by subpopulation using aggregate information on employment time. The administrative and LFS data used by Eurostat and Christl et al. (2021) are disaggregated by duration of the monetary compensation spell and reduction in the hours worked during the spell which allowed to take into account the duration of those transitions as well.

Adjusted population groups and adjustment of incomes

Two final aspects that are relevant when comparing nowcasting methodologies, are 1) the population groups for whom changes in labour market status are adjusted; and 2) the way incomes of the affected groups are adjusted in accordance with their changes in status.

From the previous section, we can infer that the papers included in our inventory focused on adjusting the characteristics of the active population in Belgium. Given their focus on the large influx into temporary unemployment and the bridging right, most papers only adjusted the labour market status of employees and self-employed individuals. This exclusive focus on both compensation measures, however implies that for individuals who are not eligible for these benefits, the impact of the crisis is not assessed. Examples are persons who saw their weekly working hours heavily reduced, without becoming (temporary) unemployed, such as flexijobbers or persons on temporary contracts. In the Eurostat nowcasting exercises, individuals who are still employed but temporarily absent from work

[6] The allocation of month-by-month transitions, on a disaggregated level that includes a distinction based on prior wage, necessitated the identification of monthly wages based on the EU-SILC. In Marchal et al. (2021) and Derboven et al. (2021) this is done on the basis of the labour market status reported for every month in the EU-SILC and retracing the annual income sources reported in the EU-SILC to the months in which they were likely to be received.

[7] In the detailed model.

or working reduced hours due to the COVID-19 crisis were included in their analysis. Next, Capéau et al. (2021) was the only paper that devoted explicit attention to measuring the impact of the crisis on the situation of non-standard workers, in particular of those working under the flexi-job system. Publicly available statistics from the NSSO for flex-workers per sector were used to assign flex-worker status to observations in the EU-SILC who had an employment of at least 4/5th in the first quarter. Of this group, observations were randomly selected in line with external data in order to assign who would lose their flexi-job throughout 2020, giving a likely more realistic image of the reduction in working hours faced by this group.

Next, once transitions in labour market status are allocated, incomes have to be adjusted accordingly. We distinguish here between the modelling of earning losses for employees and losses in self-employment incomes.

Overall, the earnings of employees were generally reduced in line with the number of months and days that they were allocated to become temporarily unemployed (or the number of months that they became fulltime unemployed). Overall, their earnings were set to zero for the affected days[8]. Capéau et al. (2021) included a further, non-temporary-unemployment related reduction of working hours, that also led to a proportionate decrease in earnings. Also, Eurostat reduced earnings proportionally to the number of month lost either due to spells of unemployment or spells of partial unemployment (i.e. a reduction of working hours).

For self-employed individuals assumptions on the losses in earnings were perhaps less straightforward. In all COVIVAT papers except for Capéau et al., 2022, personal incomes of affected self-employed individuals were set to zero in months when the bridging right was received, assuming that self-employed individuals did not gain any income in months that they received a bridging right. Self-employed without a bridging right were assumed to be unaffected. It was only in the paper of Capéau et al. (2022) that this assumption was relaxed and the impact of the COVID-19 crisis on self-employment earnings was modelled in far more detail, taking account of variation in fixed costs and heterogeneity in turnover losses at the sectorial level. Capéau et al. (2022) did so by adding this extra information to the underlying EU-SILC data, based on aggregate statistics at the sectorial level on cost-income and income-turnover ratios. They further divided costs into fixed and variable costs based on estimates of the share of fixed costs in turnover. Capéau et al. (2022) further assumed fixed costs to be constant during 2020, while variable costs changed in proportion to their turnover for every month. They used information on the impact of the COVID-19 crisis on the turnovers of self-employed provided by the Economic Risk and Management Group (ERMG). The losses in turnover were then used to generate a percentage change in turnover, which was further used to adjust the gross self-employment income. Furthermore, fixed costs were deducted, as well as the variable costs in proportion to the turnover for every month, in order to come to their final gross self-employment income. Capéau et al. (2022), also made a second income concept, where the regional compensation measures were added to the income explained above.

[8] Upon the return to employment, formerly affected employees were generally assumed to return to their prior income level, in the same sector. While this is a realistic assumption for those becoming temporary unemployed, it may be less evident for those transitioning out of unemployment. Most papers concentrated the earnings loss among this group.

In general, the reduced earnings were used as input for the tax benefit system, leading to the inclusion of relevant tax reductions, social supplements and benefits in the final net disposable incomes.

4.4 Modelling of income and uprating of monetary amounts

Finally, a last essential aspect of nowcasting is the uprating of monetary amounts. To update market incomes and non-simulated benefits and taxes of the base year to price levels of the simulation year (i.e. 2020 for most papers, except for Eurostat 2022, 2023), all papers use a similar approach and generally make use of EUROMOD uprating factors. These factors include well-known indices such as the consumer price index and health index, as well as specific indices for market incomes. For Belgium these indices are published by Belgostat (http://www.nbb.be/belgostat) or by the relevant public services themselves. While Eurostat initially used the same approach in the production of the flash estimates, they introduced in 2017 a more differentiated uprating of wages and salaries using uprating factors disaggregated by economic activity and/or sector. Also, Eurostat included a correction in the uprating of employees' wages below the national minimum wage. Finally, for self-employment income the 'Gross mixed income' from the National Accounts (ESA, 2010) is used as uprating factor.

4.5. Main findings

As a final part of our comparison between the different nowcasting exercises, in this section we focus on the various findings that emerged from the different papers. Before we do so, it is important to highlight that a real comparison of individual results across the different nowcasting exercises is difficult to make. As is evident from our inventory so far, the different nowcasting exercises use different data and methodologies. Overall, microsimulation outcomes interpret their results relative to a baseline scenario, but as the underlying data and methodologies differed, so did the underlying baselines against which the (simulated) results of the COVID-19 shock were compared. Therefore, in what follows we will only compare the main findings of the different papers, instead of discussing all results in detail.

When comparing the results sections of the different papers included in our inventory, overall, we find that most papers report on the following aspects of the impact of the COVID-19 shock: the impact of the shock on changes in employment; the impact of the shock on earnings/market income; the distribution of the shock; the (cushioning) impact of the policy measures implemented, and the impact of the shock on poverty and/or inequality. However, the specific focus of the papers is often different. For example, some papers only report the results for a very short period (e.g. Marchal et al., 2021; Thuy et al. 2020), while other papers look at the impact of the COVID-19 crisis for the whole year 2020. Among the papers that measure the impact of COVID-19 for the entire year 2020, some look at the impact on a month by month basis (e.g. Derboven et al., 2021) while others assess the combined annual impact (e.g. Capéau et al., 2021; Eurostat, 2021, 2022, 2023). Also, various income concepts are used in the different papers (i.e. individual vs. household income; gross vs. disposable income, etc.).

In what follows, we compare the main findings regarding the impact of the COVID-19 shock on employment status, (disposable) incomes and poverty and/or inequality.

Table 5 shows the impact of the COVID-19 shock on the labour market. In particular, it shows the estimated percentages of persons becoming temporarily unemployed or receiving a bridging right as a consequence of the COVID-19 outbreak and lockdown measures. For employees, the percentage temporarily unemployed varies between 21,7% and 34,4% of all employees. For the population in self-employment, percentages are on average higher, ranging between 32,3% and 52,2%. Differences are likely due to differences in denominators and period under focus, as well as the level of detail applied to, and the size of, the underlying data. In the first months of COVID-19 crisis (March-May), the first papers that studied the change in employment among self-employed individuals, reported that about half of all self-employed individuals ceased their activities due to COVID-19 and received a bridging right (Thuy et al., 2020; Marchal et al., 2021). In later estimates, this percentage was significantly reduced (Capéau et al., 2022; Neelen et al., 2022). Both the JRC and Eurostat did not publish any numbers on the persons that changed employment status due to COVID-19, however they did report the estimated impact on overall employment income.

Table. 5. The impact of the COVD-19 shock on changes in employment statutes: percentages of persons who became temporary unemployment or received a bridging right

Paper	% Temporary unemployed	% Bridging right
Thuy et al. (2020)	21.7% of E (Mar-May)	51.2% of SE (Mar-May)
Almeida et al. (2021)	-	-
Marchal et al., (2021)	27.9% of E (April)	52.0% of SE (April)
Capéau et al., (2021)	34.4% of E	-
Christl et al. (2021)	23,1% of E (2020)	32,3% of SE (2020)
Derboven et al. (2021)	25% (Mar-May)[9] of E + SE	
Eurostat (2021) – 2020 FE	-	
Capéau et al., 2022 (less detail)	26% of E + SE (2020)	
Capéau et al., 2022 (more detail)	38% of E + SE (2020)	
Neelen et al., (2022)	22,5% (Mar-May)[10] of E	33,2% (Mar-May)[11] of SE
Eurostat (2023) – 2022 FE	-	

Table 6 shows the impact of the COVID-19 shock on incomes. The second column of the table shows the percentage change (decrease) in individual/household income as a consequence of COVID-19 outbreak and associated lockdown measures in the hypothetical scenario that no compensating measures of the government were implemented as reaction on the COVID-19 crisis. The third column shows the part of the income distribution that bore the largest losses in absence of the compensation measures (i.e. progressive or regressive pattern). The fourth and fifth column of Table 6 show the same numbers, but after taking into account the compensation measures implemented by the government. In this way, the results in Table 6 also provide insight in the cushioning effect of the compensation measures to reduce the income losses of the (affected) population. Due to the different income concepts and target populations used in the different papers, it is not possible to compare the numbers reported in Table 6 in absolute terms. However, and first, Table 6 clearly shows that the compensation measures introduced by the government were to a large extent effective in compensating part of the income losses for the affected population. Second, all papers that report on a hypothetical scenario of income losses without compensation measures (left columns), show a regressive pattern in individual/ household income losses, with higher relative income losses in the first half of the income distribution. Further, the majority of the papers in our inventory show that the policy measures were able to reduce

[9] In Derboven et al. (2021) monthly percentages of persons in TU or with a BR are calculated for the period March-December 2020. In order to ease the comparison with the other papers, we report here the average percentage for the first three COVID-19 months (March, April, May).

[10] In Neelen et al. (2020) monthly percentages of persons in TU or with a BR are reported, but for ease of comparison we show here an average percentage for the period March-May.

[11] In Neelen et al. (2020) monthly percentages of persons in TU or with a BR are reported, but for ease of comparison we show here an average percentage for the period March-May.

or even offset the regressive effect, resulting in larger income losses for the higher income quintiles after the compensation measures were implemented. Only Derboven et al. (2021) and Neelen et al. (2022) did not observe a clear income gradient in the cumulative losses in disposable income for those who were affected in March 2020 as they found the highest losses to be situated in both income quintile 1 and 5. The two simulation methods employed in Capéau et al. (2022) showed different results regarding the level of detail used in the modelling. In the less detailed model, the impact on incomes (after the compensation measures) was highest for the highest income quintiles, whereas in the less detailed model, a slightly higher average impact for the lowest income quintiles was observed. Overall, we can conclude that most papers show that the policy measures taken by the governments were likely to be effective at reducing both the size and the regressivity of the COVID-19 pandemic.

Table. 6. The impact of the COVD-19 shock on changes in income and the distribution of the shock among the income distribution, before and after the compensation measures

Paper	Before policy		After policy	
	Change in income	*Pattern*	*Change in income*	*Pattern*
Thuy et al. (2020)			-0,7% in monthly, disposable income (pop.)	Progressive
Almeida et al. (2021)	-5.5% in annual equivalised disposable hh income (pop.)	Regressive	-2% in annual equivalised disposable hh income (pop.)	Highest losses in D10
Marchal et al., (2021)			-4% in monthly, disposable hh income (april) (pop.)	Progressive
Capéau et al., (2021)	-4,9% in annual, gross income (all E) -15,1% in annual, gross income (affected E)	Regressive	-1,0% in annual, disposable income (all E) -3,1% in annual, disposable income (affected E)	Progressive
Christl et al. (2021)	-7.1% annual, disposable household income (pop.)	Regressive	-1.3% in annual, disposable household income (pop.)	Progressive
Derboven et al. (2021)			-5% (Mar-Dec) in mean cumulative disposable individual income (affected E)	Highest losses in Q1 and Q5
Eurostat (2021) – 2020 FE			-3,5% in average, gross income (working pop.)	Slightly regressive
Capéau et al., 2022 (- detail)			-2,2% in annual equivalised disposable household income (pop.)	Progressive
Capéau et al., 2022 (+ detail)			-3,6% in annual equivalised disposable household income (pop.)	Slightly regressive
Neelen et al., (2022)			-3% (Mar-Dec) in mean cumulative disposable individual income (affected E)	Highest losses in Q1 and Q5
Eurostat (2023) – 2022 FE			-5% < -2% in median equivalized disposable income compared to 2021 (pop.)	

Finally, Table 7 provides an overview of the results of the papers that measured the impact of the COVID-19 shock on income inequality and poverty. Almost half of the papers reported the Gini index in case of a hypothetical COVID-19 scenario without policy measures implemented, and in the scenario with policy measures implemented. Instead of comparing exact numbers, we focus on comparing general trends. From the results in Table 7 we can observe that in absence of policy responses, the

COVID-19 pandemic would have triggered an increase in inequality. Policy measures, however, were able to counteract the inequality increasing effect of the COVID-19 pandemic, as inequality in the scenario including policy measures decreased (Almeida et al., 2021; Christl et al., 2021; Capéau et al., 2021) or stayed more or less the same (Marchal et al., 2021). Capéau et al. (2022) is the only paper in which the Gini index increased even after policy changes, which the authors explain by the higher degree of detail that was included in the nowcasting model.

With regard to poverty, the JRC, European Commission (Almeida et al. 2021), Eurostat and some of the COVIVAT papers reported on the change in the At-Risk-of-Poverty (AROP) rate as a result of the COVID-19 crisis. Not surprisingly, most papers found that the AROP rate would increase significantly due to the COVID-19 pandemic compared to their baseline scenario's. When accounting for policy measures, however, this increase was less pronounced (Almeida et al., 2021; Christl et al., 2021). Also here, the nowcasting approach has a large influence on the results. Capéau et al (2022) show that using a less detailed simulation model leads to quasi-unchanged poverty figures, while once more detail is inserted, an increase in the nowcasted AROP rate was observed. Finally, Eurostat published for every year Flash Estimates of the AROP using a rounded uncertainty interval. For both 2020 FE and 2022 FE, this interval ranged between -1,2% and 1,2% compared to the previous income year.

Table. 7. The impact of the COVD-19 shock on changes in poverty and/or inequality before and after the compensation measures

Paper	Impact on poverty/inequality	
	Before policy	After policy
Thuy et al. (2020)	-	-
Almeida et al. (2021)	+0.003 Gini +3.8 pp in AROPE (compared to no-covid baseline)	-0.001 Gini +1.9 pp in AROPE (compared to no-covid baseline)
Marchal et al., (2021)	+0.050 Gini	Stable 8,6% below poverty line (April)
Capéau et al., (2021)		-3,1% Gini
Christl et al. (2021)	+0.007 Gini + 2.0 pp in AROPE	-0.004 Gini +0.2 pp in AROPE fixed line (-0.2 pp floating line)
Derboven et al. (2021)		13% below poverty line (April) 9,7% below poverty line (Mar-May)
Eurostat (2021) – 2020 FE		-1,2% - 1,2% AROPE
Capéau et al., 2022 (less detail)		-0.006 Gini +0.14 pp. AROP
Capéau et al., 2022 (more detail)		+0.009 Gini +2.21 pp. AROP
Neelen et al., (2022)	-	-
Eurostat (2023) – 2022 FE		-1,2% - 1,2% AROPE

5 Way forward

5.1. Aim

The inventory above clearly shows the similarities in and differences between the various nowcasting efforts that aimed to track changes in the income distribution for Belgium in the period 2020 – 2022.

These efforts all stemmed from highly motivated and conscientious researchers making the most of the data that was available at the time, to make estimates of the income distribution as swiftly as possible. It is therefore with some reservation that we set out to conduct a "hindsight" exercise. The aim of this exercise is not to criticize the work done, but to take stock of the different methods used and to distill valuable lessons for the future. When governments are in dire need of up-to-date information, what are viable options? Which data are needed, and what are realistic time frames? Which investments should be made upfront in order to have reliable short term estimates at hand during an upheaval? And the focus of WP2's forthcoming exercise: what turned out to be the margin of error, in light of those determinants? In combination with the results from the other BE-FAST work packages, this information should contribute to a timeline and a proposed approach for the future.

5.2. First checks done in COVIVAT

We do not start this exercise out of the blue. In fact, the previous nowcasting exercises done during the pandemic already took as much stock as possible from one another.

A particularly noteworthy approach is the paper by Capéau et al. (2022), that was also included in the inventory above. This highly insightful analysis brought to light the differences in results obtained by a more versus a less detailed nowcasting approach. The authors bring together a number of assumptions that were present in the earliest COVID-19 nowcasting works, including a random allocation of temporary unemployment and bridging right by sector and month only (cf. Christl et al. 2021), an across-the-board assumption on the impact of COVID-19 on the incomes of self-employed (i.e. concentrated among those receiving bridging right, who are assumed to see their income fall to zero for the months receiving the BR, used in *inter alia* Marchal et al. (2021)), a 100% overlap of the affected population from month to month, and no variation in the duration of temporary unemployed within one month. Their detailed scenario on the other hand, takes account of the number of days of temporary unemployment within one month, and differentiates the probability of becoming temporary unemployed not solely by sector, but also by labour market status in the previous month, age, gender and previous daily wage level. They add transitions to and from ordinary unemployment and also include variation in the transition to and from bridging right for the self-employed. Importantly, they model income losses for the self-employed based on external data on income-costs and income-turnover ratios by sector, and aggregate turnover loss derived from both VAT returns and an ad hoc survey from the National Bank.

They show that more detailed assumptions[12], that allow for both more heterogeneity in the allocation process itself, but as a consequence also in the impact of the COVID-19 shock on employees and self-employed, lead to a higher incidence of those affected by temporary unemployment and bridging right at the annual level (be it for a shorter duration, and hence also confronted with more limited earnings losses). At the same time, among those that are affected, the income gradient is different. When more detail was incorporated in the allocation process, a pronounced negative income gradient appeared, with those affected overrepresented among the lower three (gross earnings) quintiles[13]. In addition,

[12] In combination with an "inflated" SILC, so as to accommodate this heterogeneity.

[13] A similar observation was made by Derboven et al. (2021), when comparing their findings for April 2020, based on a random allocation by wage level, sector, gender and age (following the approach by Capéau et al., 2021) to the findings in Marchal et al. (2021), based on a parametric model. The income information in the external statistics used for the stratified sampling of labour market transitions, resulted in more lower-wage employees being affected, with evident consequences for the overall income gradient of the incidence of temporary unemployment.

in the more detailed scenario, the affected employees and self-employed individuals in the lower quintiles also experience larger relative earnings losses.

As main conclusion, Capéau et al (2022, p 1) state “*that policy-relevant conclusions differ dependent on the level of detail and heterogeneity introduced in the nowcasting techniques. However, even our more detailed nowcasting technique is far from capturing all heterogeneity in income losses and coverage of compensation, which is why we plea for details and heterogeneity in the monitoring of the impact of the COVID-19 crisis, and the evaluation of compensatory policies, and more broadly, for a continued effort to utilize administrative data in microsimulation-based policy-oriented research*.”

Other COVIVAT studies were less ambitious in their comparison of different nowcasting scenarios. Still, as the underlying allocation process changed, authors took care to assess the impact of these changes on comparable output.

Derboven et al. (2021) report in their annex the causes behind the different findings between Derboven et al. (2021) and Marchal et al. (2021). As is clear from this contribution, the main difference between both papers in terms of methods was the switch from a parametric modeling of employment to temporary unemployment labour market transitions for April 2020 in Marchal et al. (2021), to a non-parametric (month-to-month) allocation of labour market status change, by wage, sector, gender and age in Derboven et al. (2021). As Derboven et al. (2021) therefore estimated the monthly income distribution for March to December 2020, it was possible to compare the results for April 2020 with the previous report. The overall conclusions, specifically with regard to the impact of the welfare state and other incomes in the households, remained the same. Still, there were some differences apparent regarding the size of specific income decreases by income quintile. The inclusion of explicit income information in the non-parametric model led to more affected persons in (household income) quintiles two and three, and less in quintiles one and five. The increased detail regarding the duration of temporary unemployment in turn led to an increase in income losses in quintiles one and three. New aggregate administrative data of the duration of temporary unemployment by wage category showed (and made it possible to apply in the allocation) that low wage individuals, when affected, experienced a higher number of temporary unemployment days in a given month, leading to higher income losses.

Neelen et al. (2022) finally included information obtained from the Datawarehouse Labour Market and Social Security, with both month-to-month transitions alongside a slightly finer range of wage categories, as well as information on repeated temporary unemployment, ordinary unemployment or bridging right incidence beyond the month-to-month focus in the form of “peak-to-peak” transition rates from April to November 2020. Also the results from this further finetuning were compared to previous results reported in Derboven et al. (2021). Overall, temporary unemployment incidence turned out to be higher according to the peak-to-peak allocation in April and November, but was very similar in the other months (or somewhat lower in December and the summer months). The share of temporary unemployed that was affected both in the fourth and in the second quarter of 2020 was, in line with the purpose of the peak-to-peak transition, higher in Neelen et al. (2022) than in Derboven et al. (2021), but overall the difference was fairly limited. Not surprising in light of this result, the further finetuning of the allocation method did not lead to largely differing conclusions with regard to the income gradient and cumulative income losses over the year 2020.

5.3. Work plan

In the remainder of the BE-FAST project, and in addition to the checks already done in the framework of the COVIVAT project, this work package sets out to assess to what extent the nowcasted findings are in line with the actually observed trends, both in the EU-SILC data that have since become available, and in the work status trajectories based on individual-level administrative data that were reported in Vinck et al. (2023).

A specific element that we will take on board is to what extent estimates improved as more elaborate nowcasting techniques became possible with the availability of ever more detailed data. As these detailed data were at the time only available with a certain lag, we hope to learn from what level of detail onwards additional finetuning may not have brought additional insights, given the inherent insecurity of the nowcasting techniques that were possible and used. At the same time, we use this exercise as an opportunity for a number of further robustness checks.

Specifically, and at this stage, we foresee the following research tasks. Please note that not all these tasks will be possible (or in fact, relevant) for each study included in the inventory above.

1. Comparison of main findings to the trends observed in the EU-SILC

We will compare the findings summarized in section 4.5 above to the actual trends in incomes and income distribution observed and measured in the EU-SILC files that have since become available. The nowcasted data above either uses nowcasted administrative data (Thuy et al.) or the EUROMOD input file based on SILC 2018 (with incomes for 2017) or SILC 2020 (the more recent Eurostat flash estimates) to estimate the situation in 2020, 2021 or 2022. By now, it has become possible to compare the findings for 2020 and 2021 with those observed in SILC 2021 (incomes 2020) and SILC 2022 (incomes 2021)[14].

There are a number of challenges when comparing the levels and trends observed in the nowcasted data with the post hoc available EU-SILC data. We list these challenges below, along with a number of options to address those challenges.

Monthly and annual incomes

A sizable number of COVIVAT publications focus on changes to monthly incomes throughout 2020 (cf. supra). This focus on monthly incomes does hinder a straight comparison of the reported findings (and summarized in section 4.5) with the information included in the SILC, that is designed to capture as well as possible an annual income concept.

As both the severity of the lockdown measures, the shock on the labour market and the social protection measures were highly changeable from one month to the other, it made sense to try to assess the situation in a specific month. This was even more so as the external data used to nowcast the situation, usually included the share of affected employees and self-employed by month, with additional information regarding their previous month's labour market status. The focus on the situation in each month furthermore meant that no assumptions needed to be made on the labour market status and incomes in the remainder of the year[15]. At the same time, we should note that this approach is not in line with the intended use of the underlying SILC data set, that does not aim to have

[14] https://ec.europa.eu/eurostat/documents/203647/771732/Datasets-availability-table.pdf

[15] This argument was especially relevant for the earliest analyses (e.g. Marchal et al., 2021), with work focused in 2020. Later contributions that looked back on the situation in 2020 did have more information available on the entire year.

a representative sample at the monthly level. This indeed led to some discrepancies in the nowcasting exercises, as certain sectors were underrepresented in the underlying SILC data, especially in particular months (e.g. Neelen et al., 2022).

The monthly nowcasting used in nearly all COVIVAT exercises exploited the information on monthly labour market status available in the SILC. This information was used to derive monthly (and even daily) wages from the annual income information in the SILC, which were necessary to allocate the share of temporary unemployed and ordinary unemployed in line with the available aggregate external administrative statistics. The monthly labour market status in April (in the SILC reference period) was used as starting point for the month-to-month transitions imposed on the SILC 2018 data in line with the external statistics. In the contributions focusing on monthly income changes, the impact of the various federal and regional support measures was furthermore implemented in the months in which they were relevant. Also, net monthly incomes took account of the withholding tax instead of the personal income tax, in order to detect immediate cash flow problems for population groups at risk.

For the post hoc comparison, this means that a large number of the COVIVAT papers reported findings that cannot one on one be compared to the results that can be directly obtained from the SILC. Evidently, there are two solutions. Either we convert the SILC 2021 and 2022 annual incomes to monthly incomes in a way similar to the approach adopted for the policy notes and working papers. However, as the monthly labour market statuses in the SILC do not distinguish between temporary unemployment in addition to the standard labour market statuses, this would require making additional assumptions (and again, allocating temporary unemployment statuses). We therefore opt to derive an annual version for the monthly (nowcasted) microdata included in Neelen et al. (2022) and Derboven et al. (2021).

Comparison to a baseline

Nowcasted results are essentially products of microsimulation exercises. This means that it is not straightforward to directly compare micro simulated outcome statistics to summary measures derived from observational microdata.

Capéau et al. (2022) already refer to these problems in the annex to their paper, when they compare the obtained (simulated) change in poverty rate between their baseline and nowcasted scenario, to the figures that were at the time just published by Statbel, prior to the release of the microdata to the research community. Simulation scenarios assess the impact of (socio-demographic and social policy) changes relative to a baseline. In the exercises included in this overview, this was generally a "2020 without COVID-19", although it was constructed in different ways[16]. Evidently, a 2020 without COVID-19 is not available in the actually observed data.

An obvious challenge when assessing the robustness of the findings summarized in section 4.5 therefore relates to the identification of a useful "reference baseline" to assess the simulated and observed changes against.

In the margin, we mention an additional caveat related to comparing simulated and observed empirical data. Capéau et al. (2022) highlight that the disposable income concept from the simulation exercise

[16] The fact that people could be temporary unemployed even when they were in the original EU SILC 2018 data fully unemployed in a specific month led to readjusted baselines in later working papers and policy notes (e.g. Capéau et al., 2022; Wizan et al., 2023).

is different to the concept available in SILC. For pragmatic reasons, the annual version of EUROMOD calculates the impact of the applicable tax benefit rules, concentrated in one year. That means that tax returns that in reality will only be paid out in the following year, are included in the currently simulated one. The same goes for other income components that are in reality only relevant to citizens with a certain lag. The SILC includes the tax rebates that stem from the previous year, whereas EUROMOD already includes those. Capéau et al. (2022) hypothesize that the nature of specific COVID-19 measures, such as deferral of payments, may further contribute to this discrepancy between simulated and observed outcomes.

The issue regarding the disposable income concept reaches even wider. In fact, EUROMOD applies the tax benefit system as it is supposed to work. That means that all benefits and taxes are taken up, in a timely way. Even though there are some modifications to proxy the non-take-up of means-tested benefits, it is unlikely that the Belgian tax benefit system works in reality as well as the simulated version in EUROMOD. This is often seen as a reason why simulated poverty rates are lower than actually observed outcomes (e.g. Vinck & Verbist, 2022).

Also the Eurostat flash estimates recognize the discrepancy between simulated and observed social outcomes, and even explicitly account for it in the publication of their nowcasted social statistics. In fact, the nowcasted statistics are derived by first calculating the year-on-year change between the model based flash estimates from t-1 to t (with t the year the flash estimate refers to). Those models, and the resulting year-on-year change is based on the SILC income data referring to t-3 or t-2. Next, this year-on-year change on the aggregate outcome statistic is applied to the most recently available SILC based outcome statistic, to get a nowcasted estimate. Importantly, this nowcasted point estimate is not published as such. Instead, only the rounded uncertainty interval around the point estimate is shown, to stress the uncertainty of the nowcasted results.

To accommodate these challenges, we envision to carry out most of the sensitivity and robustness checks focusing on the observed changes between SILC2020 and SILC2021 (or alternatively, 2021 and 2022 for more recent estimates), and to compare those against the reported findings. To account for the discrepancy between simulated and observed data, we use the observed SILC data as input data to EUROMOD. This allows to focus on the changes due to the allocated changes in labour market status and related assumed changes in income, rather than on discrepancies that stem from an idealistic versus realistic working of the tax benefit system.

It is clear that this proposed approach hinges on changes observed between the SILC 2020 and the SILC2021. The COVID-19 pandemic evidently also impacted the SILC2020 rollout. We therefore also count on a comparison with the results from administrative data that are available to us (see section 3 below).

2. Robustness checks to the nowcasted data based on the EU SILC

We furthermore include a number of robustness checks to the nowcasted data, specifically to those that were part of the COVIVAT project[17].

A first, and obvious check that has by now become possible, is to assess the sensitivity of the nowcasting results to the underlying version of the EU-SILC that is being used. To this end, we compare

[17] This is evidently for pragmatic reasons, as we have the microdata and scripts available to perform these robustness checks.

selected nowcasting results with an alternative nowcasting on a more recent “origin” file, the SILC 2019 (with incomes referring to 2018) or the SILC 2020 (with incomes referring to 2019). This is closer to the period under focus, so it requires less assumptions on the uprating of incomes (or the changes in labour market status that took place from the reference period to the onset of the COVID-19 pandemic). At the same time, it allows to assess whether the change the SILC itself underwent from 2018 to 2019, with more reliance on register data, might affect the nowcasting results. Lohmann (2011) noted that register-based information in the EU-SILC has a tendency to include for a larger share of respondents income sources that do not one on one relate to the reported labour market statuses, which may be relevant in light of the method to derive monthly and daily wages (cf. footnote 5 above). If anything, and given our method, we would expect this change to lead to more smoothened incomes throughout 2020.

Second, the later COVIVAT papers that were built around a month-to-month transition by fairly detailed strata, applied an inflation factor of 10 to the underlying SILC data. This means that every SILC observation was duplicated ten times (while its weight was at the same time divided by 10), in order to allow for i) more variation in the assigned labour market transitions, and at the same time ii) prevent as much as possible an underestimation of the impact of the COVID-19 crisis through the allocation process on a fairly limited sample.

Finally, and following the example from the Eurostat flash estimates, we aim to more explicitly show the uncertainty around the nowcasted estimates (e.g. in line with the method proposed by Goedemé et al. (2013)).

3. Comparison to findings from administrative data

Evidently, the COVIVAT papers all took care in assessing the extent to which the nowcasted data aligned to the external aggregate statistics that were used as reference for each nowcasting exercise. Overall, these results were satisfactory, although some specific subgroups (such as those employed in the agricultural sector, or specific combinations of gender, sector and age) were in some cases less well proxied, due to an absence of relevant observations in the underlying SILC data. The administrative data that was available to the researchers was fully exploited to assess the likely accuracy of the nowcasted labour market statuses.

Since, linked micro-level administrative data have become available. Vinck et al. (2022) obtained administrative microdata from the Datawarehouse that allowed to zoom in on the trajectories of individuals throughout the entire year 2020, beyond the aggregate month-to-month transitions by stratum that were available to researchers for the nowcasting exercises. In addition, the microdata available to Vinck et al. (2022) includes transitions to statuses other than unemployment, temporary unemployment, (self-) employment and bridging right, but also allows to assess whether people became unemployed or inactive without access to an income support scheme, of fell back on social assistance. Their information also allows to assess the cumulation of transitions at the household level, whereas the aggregate statistics were only available at the individual level. Unfortunately, their information does not include income data.

An assessment of the accuracy of the nowcasting using these data therefore needs to focus on the cumulation of different labour market statuses (and specifically being temporary unemployed and receiving bridging right) over the year, and in each month[18]. In Neelen et al. (2022), we already performed a first rough comparison of the nowcasting performed in Neelen et al. (2022) and Derboven et al. (2021) with the results reported in Vinck et al. (2022). This first exploration showed that the most recent nowcasting exercise, that included a peak-to-peak transition from April to November to better proxy long-term affectedness, did come closer to the actually observed patterns in temporary unemployment than the nowcasting that relied solely on month-to-month transitions. At the same time, the share of temporary unemployed that was temporary unemployed in both the second and the fourth quarter of 2020, was still substantially lower than the numbers reported in Vinck et al. (2022). Also in terms of the cumulation of temporary unemployment at the household level, the administrative data show worse outcomes than were nowcasted according to both approaches.

As income data were not included in the data obtained by Vinck et al. (2022) it is not possible to further zoom in on these differences (and hence the margin of error included in the nowcasted data) by quintile. Still, it might be worth exploring the different outcomes in light of other socio-demographic criteria. Also, aggregate administrative data by alternative breakdowns can be used to further validate the allocation.

5.4. Further considerations

The insights obtained from this exercise need to inform us, together with the work that will be done in the other BE-FAST work packages, which changes in different domains, including to external data availability, modeling choices and processes, would lead to noticeable improvements to monitor future crises. The focus of this work package is foremost to take stock of the efforts that were done during the COVID-19 pandemic in terms of nowcasting the Belgian situation.

At the same time, this focus should not blind us to alternative approaches that have since become available, and that may be worthwhile to pursue further, in this project or others.

A first, and very important, consideration in this regard is the potential of the BELMOD model and underlying data. With the exception of Thuy et al. (2020), all nowcasting exercises focusing on Belgium in COVID-times were based on the EU-SILC data. In an effort to achieve sufficient representation of affected groups, the COVIVAT consortium “inflated” the underlying SILC data to allow for more variation. Still, certain groups are insufficiently represented. A standing large sample of the population, based on administrative data, may very well solve some of the issues encountered by previous analyses.

A second development to keep in mind is the introduction of The Social Study, a representative online panel that asks, in consecutive waves, after socio-demographic characteristics, migrant status, education level and income[19]. In Marchal et al. (2021), the probability of one’s labour market status being affected by COVID-19 was imputed on the SILC with a model that was developed on the online Corona study, that especially at the onset of the pandemic attracted sizable numbers of respondents. As response rates decreased in later months, while at the same time relatively detailed aggregate administrative data became available that allowed for rather fine-grained stratified sampling as an alternative method, subsequent studies no longer used a parametric model to identify those likely affected by the pandemic in the underlying SILC data. Still, a standing and representative panel could

[18] Note that the latter check is not fully possible with the SILC2021 and SILC2022. As such, the comparisons mentioned under paragraphs 1 and 3 are complementary.

[19] https://thesocialstudy.be/nl/website-voor-onderzoekers/

perhaps overcome some of the issues related to a sole reliance on administrative recipient data, not in the least the fact that such data do miss what happens to those that are uncovered by social security provisions.

Finally, it may be worth asking what kind of detail is expected (or even needed) from a nowcasting exercise. Given the uncertainty that surrounds survey-based results, one could consider whether certain indicators are more pressing than others to obtain from a nowcasting exercise. The exercise by Parolin et al. (2022) is worth mentioning here. While their exercise focuses on the United States (and in consequence builds on the rich and timely information in the Current Population Survey), their approach is interesting, as they succeeded in the swift production of monthly poverty rates. Rather than nowcasting the entire income distribution, they addressed the monthly poverty status as a missing variable problem, imputing the poverty status based on the changing labour market characteristics across the population. While income information is only asked in detail in the CPS in the May supplement, the monthly, more timely files ask labour market status and socio-demographic updates. Such a more creative use of monthly files was under the impetus of COVID-19 also started on the Belgian Labour Force Survey files[20], although not in relation to estimates of financial vulnerability[21].

While developing full-fledged nowcasting approaches on each of these alternative methods falls outside the scope of this work package, we do aim to construct the envisioned comparisons in such a way that the results also teach us more substantively and broadly on the gains to be expected from possible alternatives.

[20] https://statbel.fgov.be/nl/themas/datalab/maandelijkse-cijfers-over-de-arbeidsmarkt#documents

[21] Which evidently would be far less straightforward in the case of the LFS than it is in the case of the CPS, given the single focus on the labour market situation in the LFS, in contrast to the rich income information included in the CPS.

6 References

Almeida, V., Barrios, S., Christl, M., De Poli, S., Tumino, A., & van der Wielen, W. (2021). The impact of COVID-19 on households´ income in the EU. *The Journal of Economic Inequality, 19*(3), 413-431. doi:10.1007/s10888-021-09485-8

Assal, E.-M., derboven, J., Panagamuwa Gamage, S., Vanderkelen, J., & Van Houtven, S. (2021). *EUROMOD Country Report. Belgium (BE). 2018 - 2021. EUROMOD Version I4.0*. Seville: JRC.

Assal, E.-M., Derboven, J., Panagamuwa Gamage, S., Vanderkelen, J., & Van Houtven, S. (2022). *Belgium (BE). 2019-2022. EUROMOD version I5.0+*. Seville: JRC.

Assal, E.-M., Derboven, J., Rongé, Z., Vanderkelen, J., Van Houtven, S., & Vanheukelom, T. (2020). *EUROMOD country report. Belgium (BE) 2017 - 2020. EUROMOD version I3.0* Seville: JRC.

Brewer, M., & Tasseva, I. V. (2020). *Did the UK policy response to Covid-19 protect household incomes?* (EUROMOD Working Paper EM 12/20). Colchester: University of Essex.

Capéau, B., Decoster, A., Vanderkelen, J., & Van Houtven, S. (2021). *De impact van de COVID-19 schok voor loon- en weddetrekkenden in 2020* (COVIVAT Beleidsnota 9). Antwerpen/Leuven: COVIVAT.

Capéau, B., Decoster, A., Vanderkelen, J., & Van Houtven, S. (2022). *Details matter for policy evaluation: the case of Covid-19 in Belgium1* (COVIVAT Working Paper 5). Antwerpen/Leuven: COVIVAT.

Christl, M., De Poli, S., Figari, F., Hufkens, T., Papini, A., & Tumino, A. (2021). *The cushioning effect of fiscal policy in the EU during the covid-19 pandemic* (JRC Working PApers on Taxation and Structural Reform 2/2021).

Dekkers, G. (2012). *The simulation properties of microsimulation models with static and dynamic ageing. A guide into choosing one type of model over the other.* . Mimeo.

Dekkers, G., & Liègeois, P. (2012). *The (dis)advantages of dynamic and static microsimulation*. Mimeo.

Derboven, J., Neelen, W., Vanderkelen, J., Verbist, G., & Marchal, S. (2021). *Maandinkomens in COVID-jaar 2020* (COVIVAT Beleidsnota 10). Antwerpen/Leuven: COVIVAT.

Eurostat. (2021). Income inequality and poverty statistics. Retrieved from https://ec.europa.eu/eurostat/web/experimental-statistics/income-inequality-poverty-indicators

Eurostat. (2022). *Flash estimates of income inequalities and poverty indicators for 2021 (FE 2021). Experimental results.* Luxembourg: Eurostat.

Eurostat. (2023). *Flash estimates of income inequalities and poverty indicators for 2022 (FE 2022). Experimental results*. Luxembourg: Eurostat.

Federaal Planbureau. (2018). *Beschrijving en gebruik van het model EXPEDITION. Document opgesteld in eht kader van de voorbereidende werkzaamheden van de doorrekening van de verkiezingsprogramma's van 2019.* (Working Paper 3 DC2019).

Goedemé, T., Van den Bosch, K., Salanauskaite, L., & Verbist, G. (2013). Testing the statistical significance of microsimulation results : a plea. *The international journal of microsimulation, 6*(3), 50-77.

Immervoll, H., Lindstrom, K., Mustonen, E., & Viitamaki, H. (2005). *Static data 'ageing' techniques: accounting for population changes in tax-benefit microsimulation models* (EUROMOD Working Papers EM7/05). Colchester: Institute for Social and Economic Research.

Leulescu, A., Agafiteit, M., Gallelli, A., Di Guglielmo, A., & Demontis, F. (2021). *Nowcasting income indicators during the pandemic (FE 2020). Methodological developments* Luxembourg: Eurostat.

Leulescu, A., Di Guglielmo, A., & Agafitei, M. (2023). *Flash estimates of income inequalities and poverty indicators for 2022 (FE 2022). Methodological note*. Luxembourg: Eurostat.

Leulescu, A., Di Guglielmo, A., Agafiteit, M., & Gallelli, A. (2022). *Flash estimates of income inequalities and poverty indicators for 2021 (FE 2021). Methodological not*. Luxembourg: Eurostat.
Leulescu, A., Lopez Vilaplana, C., Micu, B., Agafiet, & Gallelli, A. (2020). *Flash estimates of income inequalities and poverty indicators for 2019 (FE 2019). Methodological note*. Luxembourg: Eurostat.
Li, J., O'Donoghue, C., & Dekkers, G. (2014). Dynamic models. In C. O'Donoghue (Ed.), *Handbook of Microsimulation* (pp. 305-333). Bingley: Emerald Publishing.
Marchal, S., Vanderkelen, J., Cantillon, B., Decancq, K., Decoster, A., Kuypers, S., . . . Verbist, G. (2021). *The Distributional Impact of the COVID-19 Shock on Household Incomes in Belgium* (COVIVAT Working Paper 2). Leuven/Antwerpen: COVIVAT.
Navicke, J., Rastrigina, O., & Sutherland, H. (2014). Nowcasting Indicators of Poverty Risk in the European Union: A Microsimulation Approach. *Social Indicators research, 119*(1), 101-119. Retrieved from http://www.jstor.org/stable/24721071
Neelen, W., Derboven, J., & Marchal, S. (2022). *Methodological note: Nowcasting the COVID-19 impact throughout 2020* (COVIVAT deliverable). Antwerp: University of Antwerp.
O'Donoghue, C., & Loughrey, J. (2014). Nowcasting in Microsimulation Models: A Methodological Survey. *Journal of Artificial Societies and Social Simulation, 17*(4), 12. doi:10.18564/jasss.2635
O'Donoghue, C. (2001). Dynamic microsimulation: a survey. *Brazilian Electronic Journal of Economics, 4*(2).
O'Donoghue, C., Sutherland, H., & Utili, F. (2000). Integrating output in EUROMOD: an assessment of the sensitivity of multi-country microsimulation resutls. In L. Mitton, H. Sutherland, & M. Weeks (Eds.), *Microsimulation in the new millennium*. Cambridge: Cambridge University Press.
Parolin, Z., Curran, M., Matsudaira, J., Waldfogel, J., & Wimer, C. (2022). Estimating Monthly Poverty Rates in the United States. *Journal of Policy Analysis and Management, 41*(4), 1177-1203. doi:https://doi.org/10.1002/pam.22403
Piette, C. (2016). *Predicting Belgium's GDP using targeted bridge models* (NBB Working paper 290). Brussels: National Bank of Belgium.
Sutherland, H. (2001). *EUROMOD: an integrated European Benefit-tax model: Final Report* (EUROMOD Working Paper EM 9/01). University of Cambridge: Microsimulation Unit, Department of Applied Economics.
Thuy, Y., Van Camp, G., & Vandelannoote, D. (2020). *Covid-19 crisis: simulatie impact van het loonverlies bij tij-delijke werkloosheid in geval van overmacht en het inko-mensverlies in geval van een overbruggingsrecht.* Paper presented at the Webinar Federal Planning Bureau.
Vinck, J., Audenaert, B., & Van Lancker, W. (2022). *Profiel en transities van tijdelijk werklozen tijdens het eerste jaar van de coronacrisis: analyses op basis van administratieve gegevens*. Leuven/Antwerpen: COVIVAT.
Vinck, J., & Verbist, G. (2022). *Wetenschappelijk onderzoek naar een uniform inkomensbegrip en gezinsbegrip in Vlaanderen* (Rapport in opdracht van Departement Welzijn, Volksgezondheid en Gezin). Antwerpen: Centrum voor Sociaal Beleid Herman Deleeck. Retrieved from https://armoede.vlaanderen.be/onderzoek
Wizan, M., Neelen, W., & Marchal, S. (2023). *Balancing speed and effectiveness: smoothing income volatility through COVID-19 social policy responses in Belgium* (CSB Working Paper Series 23/04). Antwerp: University of Antwerp.