



Micro-Simulating the Impact of Public Policies on Households: Why, How and Which Ones?

Les notes du conseil d'analyse économique, no 74, September 2022

Public policymaking requires instruments that make it possible to anticipate the impact of existing measures or proposed reforms in fiscal terms on the one hand and in distributional terms on the other. “Micro-simulation” models attempt to address this twofold concern by applying to each household in a representative sample of the population the changes in the rules governing monetary redistribution (taxes and benefits), on the basis of the economic situation of the members of the household and its composition.

There are currently three such models on the French landscape: this diversity is desirable, if only for the sake of transparency and neutrality. However, the various models are not always based on the same data sources, do not offer the same level of detail and do not necessarily simulate the same set of measures. As a consequence, the published results on redistributive outcomes are not always completely convergent. This creates a risk of confusion in public debate and may cast doubt on the accuracy and neutrality of the models, all the more so as the media barely bother with methodological distinctions. This *Note* takes stock of current micro-simulation tools in France, makes a few proposals to avoid confusion between the various results published, and draws up some prospects for development.

Firstly, in terms of presentation of results, it is a good idea for each micro-simulation organization to use a common reading grid in its main publication or in an online appendix.

The comparison exercise undertaken in this *Note* also shows the interest of a more systematic approach to confronting their practices at the initiative of modelers and/or users. We recommend institutionalizing exchanges between modelers, users and data producers within an annual seminar. Secondly, granularity is a key dimension of micro-simulation: more comprehensive detailed data would in particular allow for a better understanding of the tails of the distribution (the poorest and the best-off) and of the variation in ‘horizontal’ inequalities. It is therefore recommended to explore the use of exhaustive or quasi-exhaustive administrative data in the double dimension of income and of demographic and socio-professional characteristics of households. Finally, we recommend maintaining a strict separation between exercises focusing solely on mechanical effects, which make it possible to understand the redistributive effects, and exercises incorporating behavioral responses in addition, which make it possible to evaluate the budgetary impact.

More systematic consideration of the distributional and redistributive effects of any public policy in its design and implementation is highly desirable, yet with analytical challenges. To push forward the frontier of micro-simulation, calls for research projects would be a possible way to enrich the evaluation of public policies, beyond the classical field of static micro-simulation of the socio-fiscal system.

This Note is published under the sole responsibility of its authors

Objectives and structure of micro-simulation models of redistributive systems

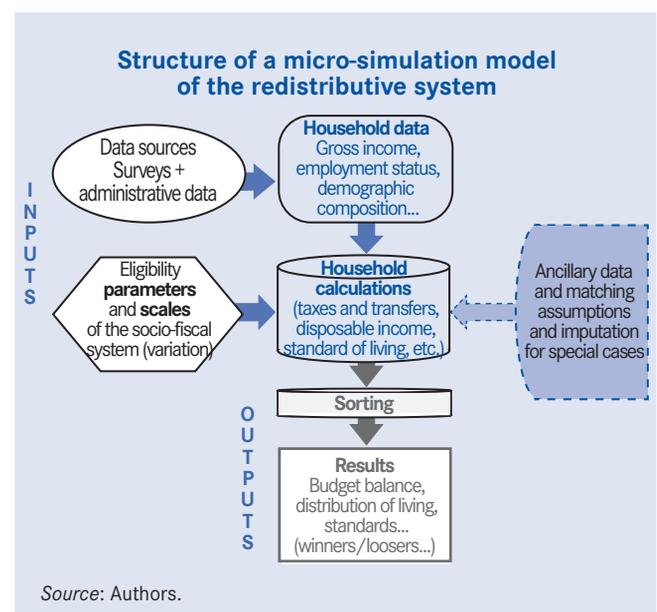
The development of macroeconomics models makes it possible to assess the aggregate impact of public policy measures such as changes in taxes or social transfers, with the advantage of taking into account looping effects (a tax cut may, for example, stimulate consumption and have a positive effect on employment). Nonetheless, they are not able to assess the differentiated effects in the population, especially when these measures depend on necessarily heterogeneous individual characteristics. While analysis of model cases can give an initial idea of the effects of a reform at the individual level, they are insufficient to describe this heterogeneity or even the fiscal impact.

However, prior to any redistribution reform, it is essential to assess its budgetary consequences as well as its effects along the living standards scale. The same information is also necessary *ex post* for the sake of transparency, especially as delays and corrections may have occurred during the implementation of the reform. This is what micro-simulation models allow. Their limitation is that they propose partial equilibrium analyses: the simulated reforms are assumed to have no effect on economic equilibria. However, they allow precise evaluations at the individual level of these so-called ‘first round’ effects, i.e. before taking into account behavioral reactions that could lead to a modification of economic equilibria. In practice, micro-simulation is thus a useful tool both for *ex ante* evaluation, by simulating reform projects ‘all other things being equal’, and for *ex post* evaluation, by describing in particular the contribution of past fiscal or social measures to the variation in the standard of living of the various categories of population.

If a full description of the socio-economic situation of a household is available, including its various sources of income, the labor force status of its adult members, its demographic composition, and some additional information, the application of official scales and, for benefits, eligibility rules allow the calculation of taxes, contributions, and benefits paid or received by the household. If this calculation is done for each household in a representative sample of the national population, it is then possible to assess the budgetary and redistributive impact of a reform of the existing system. This simulation is purely arithmetical in nature: this means that it assumes that the economic situation of the household is not affected by the reform. In other words, the assumption is that the fiscal change does not affect decisions of the household in terms of activity, sources of income, use of the benefits...

This hypothesis is sometimes questionable because behavior can change. A benefit that was not used before is claimed if its amount becomes greater, an improvement in the conditions of parental leave leads to more frequent use of this option, a tax exemption for overtime increases the number of hours worked, etc. *Ex ante*, these reactions are either ignored or simulated on the basis of assumptions based on econometric knowledge of these behaviors. This takes us away from the purely accounting approach.

At this point, it is important to stress the “static” nature of the micro-simulation models presented and discussed in this *Note*, in the sense that they neglect the inter-temporal effects of a given measure on a household’s welfare. It may be restrictive in several cases.¹ The following figure represents the common structure of micro-simulation models of redistributive systems according to the logic just described. At the same time, it shows the various potential sources of divergence between models. These are the blocks that feed into the central vertical axis, which is solely concerned with numerical calculations: data sources, taking into account all the parameters of the socio-fiscal system (scales and eligibility criteria) and, above all, the dotted line showing the part of the measures that cannot be simulated from the basic data and requires additional data or imputation hypotheses. This same block also includes the assumptions necessary to take account of possible behavioral reactions.



In addition to the differences in the various components of their model, the modelers’ choices concerning the definition of the “counterfactual” (i.e. the reference used to evaluate

The authors would like to thank Étienne Fize and Madeleine Péron and Paul Bachoffer and Pierre Rousseaux for their assistance in the preparation of this *Note*. They would also like to thank the INSEE, the French Treasury, the OFCE and the IPP for all the exchanges and cooperation that were essential for the completion of this work.

¹ For an overview of dynamic micro-simulation models, see Blanchet D. (2020): “Des modèles de micro-simulation dans un institut statistique: pourquoi, comment, jusqu’où ?”, *Courrier des Statistiques*, INSEE, June. See also Cheloudko P. and H. Martin (2020): “Une décennie de modélisation du système de retraite : la genèse du modèle de micro-simulation TRAJECTOIRE”, *Courrier des Statistiques*, INSEE, June.

the variations in living standards and the budget balance due to the set of reforms studied), the scope of the measures studied, and the behavioral reactions taken into account may contribute to different results. Consequently, reading the results of the models requires multiple precautions, and therefore calls for the greatest transparency from the point of view of communication.

Micro-simulation of redistribution in France

The actors, their objectives and their models

In France, the first micro-simulation model was the MIR model and concerned only income tax. The transfer and social benefits dimension was missing in order to assess the impact of the entire redistributive system. This integration was carried out by a team of researchers at the end of the 1980s at the Delta, the predecessor of the Paris School of Economics (PSE). Quite soon, the INSEE (the French Institute of statistics and economic studies) developed a similar instrument, but based on better quality data, under the name of INES in the mid-1990s, now managed with DREES (Department of research, studies, assessments and statistics of the Ministry of Health and solidarity) and CNAF (national family allowances fund). Other institutions followed: the Direction Générale du Trésor (DGT – the administrative body of the Ministry of Economy and Finances) has had its own “Saphir” model since 2008 and, more recently, the Institut des Politiques Publiques (IPP: Institute of Public Policies is a research institute attached to the Paris School of Economics) created the TAXIPP model in 2012.²

The number of users can go beyond the modelers alone insofar as these are freely accessible in the form of an open license. This has been the case since 2016 for the INES model, of which the OFCE (French economic observatory, a research institute attached to Sciences Po Paris) is one of the first “external” users. The IPP chose open access from the outset by using a free and collaborative socio-fiscal calculator (OpenFisca) for its model, while, following a request from the national commission for the access to administrative documents (CADA), the Direction générale du Trésor opened publicly the code of the Saphir model in 2018.

Today, four institutions regularly publish the results of their estimates of the redistributive effects of the reforms of the socio-fiscal system undertaken during the year or planned in the draft budgets. The DGT publishes its results at the time of the presentation of the Budget Bill (PLF) and the Social Security Financing Bill (PLFSS) in the Economic, Social and Financial Report (RESF).³ These are therefore prospective results for the year N or beyond, as published at the autumn N – 1. They may cover a single year or several years in the context of a presidential term of office. The IPP and the OFCE now publish an annual assessment of the redistributive impact of the reforms included in the budget for year N, as voted, in the first quarter of year N. Most often, one can also find cumulative results over several years, again with a view to the term of office.⁴ Finally, the INSEE, together with the DREES and the CNAF,⁵ generally proposes an *a posteriori* redistributive assessment for year N, in the autumn of N + 1:⁶ this assessment is thus guaranteed to take into account all the measures actually implemented in a given year because, even if OFCE and IPP rely on voted budgets, changes may occur during the year through amending laws or decrees.

Although the objectives to draw up the budgetary and redistributive balance sheets of the reforms are similar, these estimates sometimes tend to differ from one institution to another, for several reasons which are due to certain characteristics of the models and simulations.⁷

Characteristics of simulation models

The data and details of the results

Both the Ines and the Saphir models use the *Enquête revenus fiscaux et sociaux* (ERFS) as their main data source. This annual survey is based on data from the Continuous Employment Survey matched with data from income tax returns and data on social benefits. This database includes approximately 50,000 households representative of the French population in metropolitan France. When simulating the legislation (or draft legislation) of year N + 1, the most recent ERFS informs about the income of year N – 3 (for example, ERFS 2016 to simulate the 2020 Budget bill). To remedy this drawback, each institution “ages” the data using more recent aggregate information. Thus, observations are re-weighted to take into account changes in demographic

² For a history of micro-simulation in France, see Legendre F. (2019): “L'émergence et la consolidation des méthodes de micro-simulation en France”, *Économie et Statistique*, no 510-511-512, pp. 201-217.

³ See for example the RESF annexed to the PLF 2020.

⁴ See for example Fabre B., A. Guillouzouic, C. Lallemand and C. Leroy (2020) : “Budget 2020: quels effets pour les ménages?”, Notes IPP, no 49, February and Ducoudre B., P. Madec, M. Plane, M. and R. Sampognaro (2020) : “Budget 2020: du pouvoir d'achat au travail”, Policy Brief OFCE, no 64, February on the impact of the 2020 budget and the cumulation over 2018-2020.

⁵ The INSEE, the DREES and the CNAF are co-developers of the Ines model. In the remainder of this Note, to simplify the writing, only the INSEE will be cited to refer to these three organizations.

⁶ See for example France, Social Portrait, 2020 Edition (coll. INSEE Référence) with a review of 2019.

⁷ On the same subject, see André M. and M. Sicsic (2020): “Évaluation des effets redistributifs des réformes socio-fiscales: comment s'y retrouver?”, INSEE Blog, 3 December, <https://blog.insee.fr/evaluation-des-effets-redistributifs-des-reformes-socio-fiscales-comment-sy-retrouver/>

structures or unemployment, while incomes are updated according to their nature. Given their forward-looking approach, the DGT and the OFCE carry out an aging over four years, whereas INSEE carries out an aging over two years only because of its *ex post* analysis (ERFS 2018 for the 2020 analysis carried out in the summer of 2021).⁸

The current version of the TAXIPP model (TAXIPP 2.1) relies mainly on FIDELI, (Fichier démographique sur les logements et les individus), which gathers exhaustive administrative data from the registers of housing tax, built-up properties, personal income tax, and tax returns. These data are statistically matched with FELIN (a sample of 500,000 tax households from income tax data with exhaustive representation of the top 4% of income earners), DADS the annual declarations of social data for employees (Déclarations annuelles des données sociales pour les salariés) and BNS, a self-employed database (Base non-salariés). The IPP also “ages” its data to perform its simulations.

In order to extend the scope of the measures studied, the modelers use additional data sources. For example, consumption data from the Family Budget survey are used to estimate the VAT paid by households, as well as duties on tobacco or alcohol. Similarly, additional sources can be used to simulate the taxation of wealth (income and wealth tax declarations, wealth survey).

By using larger databases, the level of detail is higher in the IPP’s model than with the Ines or Saphir models where, due to their very high dispersion, it is not possible to measure precisely income differences within the richest 5%, or perhaps even 10%, of households. With an exhaustive representation, TAXIPP allows an analysis by hundredths of the population, and is therefore much more precise in simulating certain reforms.

The scope of the measures studied and the choice of the counterfactual

Modelers may also differ in the scope of the measures studied, which may lead to discrepancies in the evaluation of a Budget bill or the reforms of a mandate. Furthermore, the counterfactual scenario of a simulation may depend on its temporal specification. For example, taking stock of the reforms of the five-year period that has just ended requires a counterfactual based on the pre-2017 socio-fiscal system. To assess the 2022 Budget bill, the counterfactual would be the system in force in 2021.

As regards the scope of the measures simulated, each institution has its own logic. The IPP covers legislative measures relating to taxes and benefits affecting households when they are decided by the government in power, as well as measures taken by the social partners, when they can be simulated, for a given year. Thus, the scope covered corresponds to reforms voted during the period of analysis (whether their entry into force during the period is immediate or deferred). Measures passed before the analysis period are therefore not covered by the IPP.⁹ For their part, the OFCE and the DGT take into account, for a given year, all the new measures having an impact on the public accounts of the year, at least those that can be quantified. In principle, the OFCE and the DGT do not include measures taken by the social partners.¹⁰ However, the Treasury’s approach in the RESF differs because the measures studied are measures that have not yet been voted (PLF and PLFSS), unlike the OFCE that focuses on voted measures only.

In the publication *France portrait social*, in autumn N + 1, the INSEE finally presents an *a posteriori* assessment of year N and only takes into account in its main analysis the measures falling within the scope of monetary redistribution.¹¹ The INSEE only reports indirect taxation measures such as the TICPE (domestic tax on energy products) or tobacco taxes in an annexed analysis of the same publication. In its socio-fiscal evaluation, the INSEE distinguishes itself by favoring the presentation of simulated measures with a full-year effect, regardless of the timetable for implementing the reforms. This approach can be a source of discrepancies with other publications. This is particularly the case for 2018 due to the CSG (general social contribution) increase in January and the two successive cuts in social security contributions (in January and October). On average over the year, the CSG/ social security contribution switch resulted in an overall increase in taxes on households for the IPP, the OFCE, and the DGT, while a full-year assessment leads to an overall neutral effect for the INSEE.

Simulation or imputation?

Another source of difference between micro-simulators may be the simulation method itself. The three models certainly allow for the simulation of a large part of the socio-fiscal system: income tax, social security contributions, family benefits, minimum income benefits, in-work benefits, housing benefits, housing tax rebates. However, the simulation of indirect or assets taxation requires data that are not available in the databases used by the models. The effect of reforms

⁸ In this respect, it may be noted that little information is available on how to age the databases and, *ex post*, on their accuracy. However, one suspects that the longer the time lag, the less accurate the aging.

⁹ This convention has been applied since the redistributive assessment of the 2017-2022 five-year period published by the IPP in March 2022: Dutronc-Postel P., B. Fabre, C. Lallemand, N. Loisel and L. Puschnig (2022) : “Effets redistributifs des mesures socio-fiscales du quinquennat 2017-2022 à destination des ménages”, *Note IPP*, no 81, March.

¹⁰ It should be noted, however, that the projected effects of the unemployment insurance reform are included by both the Treasury and the OFCE.

¹¹ Primary income measures such as a reform of unemployment benefits or pensions are not taken into account. Nor are the levies allocated to the financing of these benefits included in the redistributive balance.

is then ‘imputed’: the modeler computes the average tax or transfer using an annexed database for different categories of population and then, in the main database, recreates those categories and attributes those average tax and transfers to the individual or household according to its assumed category.¹² In practice, the OFCE and the Treasury carry out imputations to assess the redistributive impact of the introduction of the real estate wealth tax (IFI) and the single flat-rate tax (PFU), while the INSEE and the IPP carry out real simulations by matching their database with wealth tax registers.

Taking into account behavioral reactions

The purely arithmetical logic of the micro-simulation presented above focuses on the ‘mechanical’ effects of public policies in terms of redistribution. We will see below that the presentation of these mechanical effects is of interest conceptually, in addition to bringing the advantage of simplicity and transparency. However, this arithmetical logic ignores the expected behavioral reactions to the socio-fiscal measures analyzed. Yet they are important for at least two reasons: firstly, they are decisive for the evaluation of the budgetary impact of the measures analyzed, and secondly, in certain cases, they can affect the redistributive impact beyond the simple mechanical effects.

A good example of the difficulties involved is the increase in tobacco taxation. The arithmetical logic would be to apply the new tax under the hypothesis that consumption will remain the same. However, some smokers will stop smoking or limit their consumption as a consequence of the policy, responding to the price effect, while others won’t reduce their consumption and will devote a larger share of their budget to smoking in comparison to any counterfactual. This will affect tax revenues. To take this into account, behavioral assumptions must be introduced. These can be based on existing research, but as it is generally difficult to avoid arbitrariness, they are likely to produce substantial differences between micro-simulation exercises. For tobacco, the four modelers all assumed identical behavior across deciles but used different price elasticities for average tobacco consumption.

Another example of behavior to consider is the non-take up of benefits by eligible beneficiaries. The factors explaining non-take-up are numerous and generally depend on the social benefit concerned.¹³ They are also heterogeneous within the population. *A posteriori*, we can get an idea of the average rate of non-take-up by comparing the public accounts with the disbursement that would be involved if the eligibility scales and rules were strictly applied. It is more difficult

to infer *a priori* how this rate would vary with a change in the latter while taking into account the heterogeneity of this reaction in the population. We know, for example, that the *Prime d’activité* (PA- in-work benefit) saw its number of beneficiaries increase sharply in 2019 following its major revaluation in response to the Yellow Vests crisis. By 2020, not all modelers had changed their assumptions about the take-up rate of the PA.

Behavioral responses to labor supply are often highlighted in the economic literature on the impact of public policies. They are ignored by French micro-simulators. Their estimation is certainly complex and the accuracy of existing econometric models is low. Despite these difficulties, it should be noted that several foreign modelers have integrated this dimension into their micro-simulation models (see box).

Modelers on four case studies simulations

In order to test the respective properties of the models and simulation methods, the four French modelers were invited to a joint simulation exercise on four recent measurements following a precise set of specifications:

- The additional reduction of the tax d’habitation (TH – housing tax) from 30% to 65% in 2019;
- The 2020 income tax (IR) reform, including the lowering of the first bracket of the tax scale;
- The exceptional revaluation and extension of eligibility for the in-work benefit (PA) in 2019;
- The introduction of the single flat rate tax (PFU) on capital income in 2018.

The purpose of this exercise was twofold: to check the concordance of the simulations and to assess the weight of different assumptions about behavioral responses or the inclusion of ancillary data on the conclusions. In order to make the comparison, it was necessary for the modelers to adhere to a common specification of the simulations and presentation of the results, which may differ from the format of their publications.

Counterfactuals

Before comparing the results, it seems useful to check the concordance of the counterfactuals on the basis of which the effects of the simulated measures will be evaluated. Here, it takes the form of the distribution of ‘living standards’ in the base year. In fact, the presentation of the results of the micro-simulations in terms of standard of living¹⁴ constitutes

¹² The term ‘decile’ is used in this *Note* in the Anglo-Saxon sense of the term and refers to a tenth of the population classified by standard of living, rather than the income level between two deciles.

¹³ Among the main causes of non-use are the lack of information, the complexity of the procedures and the stigmatizing nature of certain measures.

¹⁴ The INSEE defines the standard of living as follows: “The standard of living corresponds to the disposable income of the household divided by the number of consumption units (CU). It is therefore the same for all persons in the same household. Consumption units are calculated according to the so-called ‘modified OECD’ equivalence scale, which assigns 1 CU to the first adult in the household, 0.5 CU to other persons aged 14 or over and 0.3 CU to children under 14”. Other equivalence scales, such as the number of members or the number of adults, are however possible, as well as the absence of an equivalence scale, which amounts to retaining the household as the statistical unit. Clearly, these choices are not neutral as regards the distributional effect of the simulated reforms.

Micro-simulation of redistribution in advanced economies

Without presuming to provide an exhaustive overview, the *Focus* associated with this *Note*^a describes micro-simulation activity in a small sample of countries, where there is a wide range of current practices and experiences: the Netherlands, the UK, Germany, and the USA. The following lessons can be drawn from this exercise:

- With the exception of the Netherlands, there is generally, as in France, a plurality of actors in micro-simulation, combining administrations and research institutes;
- Dialogue or partnerships between the players vary: nonexistent in the Netherlands (the Central Planning Bureau being in a monopoly situation), it mainly takes the form of outsourcing in the United Kingdom, where the administration can place 'orders' with the Institute for Fiscal Studies (IFS), for example, or exchanges about costing and results both in Germany (within the Arbeitskreis Steuerschätzungen) and in the United States (especially between the Treasury and the two Congressional bodies, the Joint Committee on Taxation and the Congressional Budget Office);
- Simulation exercises differ according to whether or not they incorporate behavioral responses. In this respect, the two German research institutes (DIW Berlin and IFO) include labor supply responses to redistributive reforms and their aggregate effects on employment and wages in the simulations of the social tax system,

which the Finance Ministry model does not. The IFS in the UK has also developed a detailed econometric model of household labor supply to feed back into its main TAXBEN model where appropriate. In the Netherlands, the CPB has also introduced a structural supply model within its micro-simulation model. The US models all incorporate various behavioral responses to tax measures in their models to estimate the fiscal impact of a reform;

- By convention, however, the Tax Policy Center, as well as the US Treasury, retain only the 'static' impacts of tax changes in their measures of redistributive impact, i.e. without any behavioral reaction. In doing so, they follow a well-known result of the economic theory of well-being, according to which the variation in the well-being of a consumer in the presence of a modification of his budgetary constraint corresponds, to a first approximation, to the sole variation in the value of his consumption basket net of his activity income at constant labor supply.

Compared to France, it seems that the main difference stems from a greater propensity in advanced countries to include behavioral reactions in the micro-simulation, particularly with regard to labor supply, and sometimes even a complete closure (employment, prices) of the labor market.

^a Bourguignon F. and C. Landais (2022): "La micro-simulation de la redistribution dans les économies avancées", *Focus du CAE*, no 087-2022, July.

in itself a first kind of standardization between modelers. The presentation of the results in percentage of the standard of living per quantile or for the whole population also calls for a remark. The four modelers use the ratio of averages (and not the average of the ratios): for example, within a quantile, the average variation in euros in relation to the average standard of living. Welfare economic theory would recommend also presenting the average of individual variations in percentage of the standard of living. The difference is minor for the middle deciles but can be significant for the extreme quantiles, let alone for the total population.¹⁵

Since the INSEE, the OFCE and the Treasury all use the ERFS, the average living standards per decile can only be very close to each other. This is broadly the case, with differences of no more than 2%, which cannot significantly affect changes in living standards, which rarely exceed a few percentage points. The IPP uses a different database and works on a

broader scope, taking into account the French overseas departments and territories and non-ordinary (or collective) households: the sample covers a population of 67 million individuals, compared with around 63.5 million for the other three organizations. The particularities of the additional population covered mean that the average living standards calculated by the IPP are lower, particularly at the bottom of the distribution. This difference may be important because it means, for example, that for identical impacts in euros of a measure, the percentage impact on the standard of living in the first deciles would be systematically higher for the IPP.

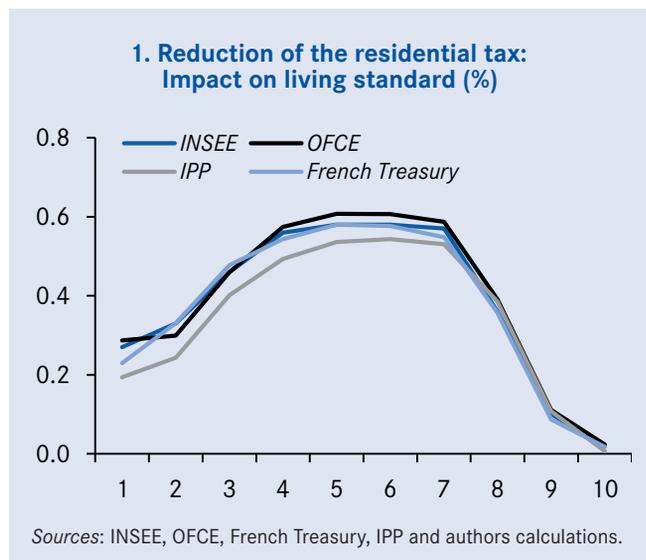
The additional reduction of the housing tax

Modelers first simulated a measure implemented in 2019, consisting in an increase in the rate of rebate from 35% to 60%. In terms of budgetary impact (of the order of 3.7 billion euros), of the proportion of households gaining¹⁶ (between

¹⁵ Technically, the ratio of averages presupposes a social utility function giving the same weight to all individuals, while the average of individual ratios assigns to each individual a decreasing weight with her standard of living, according to the principle of decreasing marginal utility.

¹⁶ Each organization was asked to adopt the same convention for declaring a household a winner or loser, based on an absolute threshold on the standard of living in euros: ± 5 euros.

63 and 68% depending on the organization) or of the average impact as a percentage of the standard of living per decile, the results of the four organizations are very similar. At most, the average effects are slightly lower for the IPP, which is probably due to the fact that, in the larger population covered by the IPP, there are more households not subject to TH and therefore not affected by the measure.

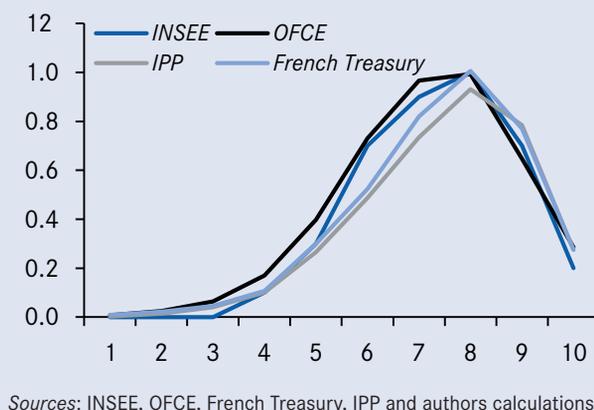


The income tax cut in 2020

The measure simulated here is the reform of the IR in 2020: the rate of the first bracket of the IR scale is lowered from 14% to 11%, with a ceiling on the gain for households in the second bracket and a neutralization of the reduction for the wealthiest households. Again, the simulations by the four organizations are fairly similar for a budgetary impact of slightly more than 5 billion euros and of the order of 15 million households gaining. As the income tax reform only concerns households that are subject to income tax, the gains are small or nil for the first deciles of the standard of living. They are also very low at the top of the scale since the gain for the wealthiest households is neutralized.

However, there are significant differences in the intermediate deciles, particularly when comparing the results of the IPP with those of the OFCE (the two most distant simulations). For example, the impact on the standard of living of the 7th decile measured by the IPP is 0.3 points lower than the estimation of the OFCE, i.e. a difference of around one third. Here again, the explanation lies in the difference in samples. A higher proportion of the additional population with a moderate standard of living covered by the IPP is not liable for income tax or benefits from mechanisms such as the discount. The gain per decile is therefore lower. As for the Treasury's results, they are lower than those of the OFCE and the INSEE, although they are based on the same population and the same initial data (before aging of the data).

2. Lowering of the income tax/ Impact on living standard (%)

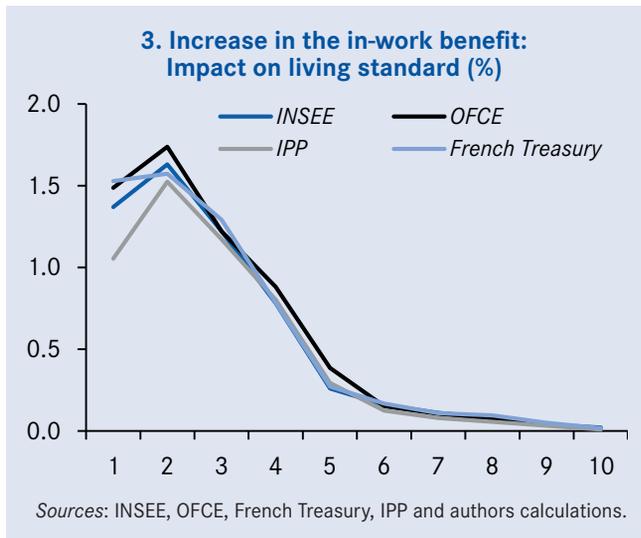


Exceptional increase and widening of eligibility for the in-work benefit in 2019

The simulated measure concerns the changes that occurred in the in-work benefit (PA) at the beginning of 2019: on the one hand, the exceptional 90-euros increase in the maximum amount of the individual bonus (from 70.49 to 160.49 euros), and on the other hand, the broadening of the eligible public (the amount of the bonus increases up to 1 SMIC, compared to 0.8 SMIC previously, and then decreases with a higher exit point). In retrospect, we know that these measures favored a strong increase in the number of households receiving the in-work benefit (+ 1.1 million beneficiaries between December 2018 and March 2019). This sharp increase can be explained both by the arrival of new eligibles who benefit from the drop in the benefit's exit point, as well as by the increase in the take-up rate among households already eligible, due in particular to the extensive communication about the revaluation of the in-work benefit and to the higher amounts of the benefit. Consequently, this measure is more complex to simulate as it involves that each modeler has to make assumptions or to choose a specific model in order to integrate the change in behavior in the benefit take-up modeler. For the present comparison exercise, the observed variation in the take-up rate was introduced exogenously, with the take-up itself initially assumed to be randomly distributed according to the net gain associated with PA. However, it took several iterations for the results produced by the different organizations to converge, which demonstrates the value of modelers maintaining a dialogue for more complex simulations.

The results of the Treasury, the INSEE, and the OFCE on the scope of ordinary households in metropolitan France are fairly close as regards the effect in percentage of the standard of living per decile, even if the budgetary impact is slightly higher for OFCE (4 billion euros vs. 3.7 billion for the INSEE and the Treasury). As far as the IPP is concerned, the broader scope of the population covered should justify a

greater impact compared to other organizations, particularly for the lower end of the living standards scale. However, the total budgetary cost obtained is lower than that of the other three organizations, and even *lower* than the official figures for the whole of France (the overall budgetary impact is estimated at 3.5 billion euros compared with 4.1 billion in the budget documents for 2019)¹⁷ and the effect on the 1st decile is much lower than that of the other modelers (a difference of 0.4 to 0.5%). The reasons for this discrepancy need to be clarified.



Single flat-rate tax (PFU)

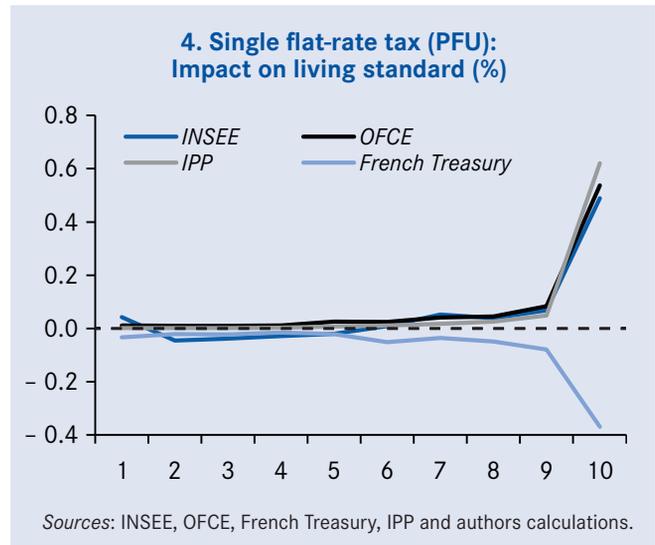
The single rate of taxation (prélèvement forfaitaire unique, PFU) was introduced in 2018 at a rate of 12.8% for income tax and 17.2% for social security contributions, to reach the single effective rate of 30%. The top marginal tax rate on taxable dividends thus falls from 40.2% in 2017¹⁸ to 30% in 2018. The PFU applies to capital gains on the sale of securities and income from movable assets (in particular dividends and interest from fixed-income investments such as bonds),¹⁹ and for interest from some ownership saving schemes opened after 1st January 2018. Tax households that would lose out by the implementation of the PFU do, however, have the option of choosing taxation on their income at the scale. Simulating the effects of such a measure requires several assumptions: on the one hand, whether or not to assume perfect optimization of households so that the PFU does not result in any losers and, on the other hand, whether or not to take into account the changes in behavior brought about by the reform, in particular the increase in dividends paid as a result of lower taxation.

¹⁷ See *Rapport économique, social et financier - PLF pour 2022* on www.tresor.economie.gouv

¹⁸ Which corresponds to the application of the marginal tax rate of 45% on the 60% of taxable dividends, taking into account an allowance of 40%, plus 15.5% of social security deductions, taking into account the deductibility of 5.1% of CSG: $40.2\% = 0.6 \times 45\% + (15.5\% - 0.45 \times 5.1\%)$ This rate must be increased by 2.2% for taxpayers subject to the exceptional tax on high incomes.

¹⁹ For income from life insurance contracts only if they relate to payments made after 27 September 2017.

²⁰ One can understand the INSEE's methodological choice to evaluate any measure once it has been fully implemented. In the present case, however, this represents for some households a jump of twelve years into the future. It should be noted that other modeling options have also been tested by INSEE (optimization, not taking into account PEL and CEL interest) in previous work.



The IPP and the OFCE seem to make similar choices in terms of modeling: assumption of optimization of households (no losers) and not taking into account behavioral effects on the volume of dividends. For the purposes of this *Note*, the INSEE has chosen to tax all income falling within the scope of the PFU and does not take into account behavioral effects either. However, it is above all another factor that leads to a substantial number of losers (19% of households) in the INSEE simulation: the reform is considered as fully implemented. Thus, interest on ownership saving schemes (whose holders are less concentrated and more numerous than for certain other savings or wealth incomes) is assumed by the INSEE to be fully subject to the PFU from 2019, whereas in the legislation only those from plans opened after 1st January 2018 or of more than 12 years are.²⁰

But it is the Treasury that stands out the most by taking into account behavioral effects to ensure consistency between the budgetary evaluations of the measure and the redistributive outcome. Thus, the Treasury associates with the measure not only its *ex ante* cost, but also its *ex post* effect due to the observed increase in dividends attributed to the introduction of the PFU. The taxes paid by households that benefited from this increase in dividends are thus associated with the reform, which mechanically leads to an overall increase in levies of around 1.3 billion euros, compared with an *ex ante* cost for public finances of 1.7 billion. Of course, the *ex ante* gain and the 'extra' tax are almost exclusively concentrated on the last decile. The difference between simulations on this decile is therefore considerable: - 0.4% for the Treasury compared to + 0.6% for INSEE, i.e. a difference of 1 point.

Such a presentation of the results is ambiguous. If the top decile pays more tax after the single flat-rate tax reform than in the counterfactual, its primary income –i.e. before redistribution– has increased, which is not visible in the result presented. The increase in tax revenue due to the rise in dividends in the budgetary balance of the transition to the PFU is effectively borne by the top decile, but not displaying the variation in primary income prevents the reader from appreciating the real variation in living standards and may lead to a reading of the reform as “progressive”. We will return to this question of taking into account changes in primary income due to the adjustment of behavior in response to a reform of the socio-fiscal system later on.

What lessons can be drawn from this comparative micro-simulation exercise? Firstly, that the results are generally consistent, even if it sometimes took one or two additional iterations between modelers to resolve certain discrepancies. Secondly, that conceptual differences may nevertheless arise between modelers, as in the case of the PFU simulation, which would be desirable to resolve. The main lesson, however, lies in the modelers’ efforts to adapt their *modus operandi* to the realization of simulations following a common specification and presentation. In a nutshell, it reveals differences in current practice that are more of form than of substance. For a given measure, these differences may originate from the specification of the simulation (counterfactuals, temporality), the population covered and the nature of the main database, the behavioral or imputation hypotheses, and finally the presentation of the results (disposable income or standard of living, households or individuals, percentiles, 20s or deciles). Hence the need for modelers to be extremely clear and precise in the presentation of their simulations so as to avoid confusion among users.

Prospects for the evolution of the micro-simulation of the French redistributive system

This overview of the micro-simulation of redistributive systems in France and the comparison with other countries gives rise to a certain number of remarks and reflections on current practices and possible developments: these are organized according to four main principles.

Diversity principle

First of all, the quality of micro-simulation work in France must be highlighted: several models are now available, leading to fairly convergent results, except sometimes, which is worth emphasizing, when it is a question of taking into account behavioral relationships or evaluating measures that require information additional to the main micro-simulation databases. Diversity is a good thing because it can be

a source of healthy emulation leading to the deepening and improvement of the models but also to the pooling of efforts (particularly with regard to data access). It also makes it possible to ensure the robustness and accuracy of simulations by comparing the results obtained by different micro-simulators. Finally, it guarantees a certain credibility in the public debate by facilitating the comparison of independent simulations

To make diversity work and to avoid confusion, however, some comparability of results must be ensured. This is not only a question of the format of the tables or graphs presented, but also of information about the databases, the counterfactual, the scope of the measures analyzed or specific assumptions. As regards the presentation of the results, it is desirable that each micro-simulation organization uses a common reading grid in its main publication or in an online annex. Incidence curves by deciles of individual living standards with a stack of simulated measurements seem the most obvious common denominator. However, this should not prevent the presentation of other formats in parallel (twenties or percentiles, disposable income per household or per adult) or the examination of redistribution according to household or individual characteristics.

Comparability also requires explicitly stating the elements that may differentiate the results presented from those of another micro-simulator. We have seen, for example, that the use of a France-wide sample by the IPP leads to differences with respect to the other models. Similarly, the fact that, in its socio-fiscal outlook, the INSEE favors the presentation of simulations of full-year effects may be a source of divergence in certain cases and should therefore continue to be commented on and reported. The assumptions used must also be clearly explained when imputations are made from secondary sources, or when behavioral reactions are simulated, to warn that this is not a purely accounting simulation.

Explicit comparison of a modeler with the results of another micro-simulation of the same socio-fiscal measures is also to be encouraged, as in the case of the IPP’s presentation of estimates of the distributional effects of the 2022 Budget Bill, and a comparison with the results of the DGT.²¹ This is a difficult exercise when the publications take place at roughly the same time, but briefly pointing out potential differences from other simulations would greatly enhance the transparency of the results. Indeed, the comparison between micro-simulators above on the basis of specific measurements has shown that the causes of observed differences in results are not always obvious. It would seem to be a good policy for the modelers to decipher them themselves rather than leaving them to the user.

As already mentioned, the observed differences between simulations constitute a kind of confidence interval of the

²¹ See Dutronc-Postel *et al.* (2022), *op. cit.*

results for a measurement or a field of measurements. In the same vein, it would be useful if modelers could provide, alongside their results, an estimate of their statistical accuracy, even if only in terms of sampling. The comparison exercise undertaken in this *Note* also shows the value of a more systematic approach to comparing their practices at the initiative of modelers and/or users. It may be important to also include in these exchanges the producers of public data which most often condition the nature or accuracy of simulations.

Recommendation 1. Institutionalize exchanges between modelers, users, and data producers in an annual seminar.

Principle of granularity

Granularity is a key dimension of micro-simulation. Having exhaustive or close to exhaustive data allows for a better understanding of the major role of the tails of the distribution in the average effects. This is particularly true for the top of the distribution where a small number of individuals or households can have significant effects on the aggregates due to the concentration of income and wealth. But this is also true at the bottom of the distribution, where there is a greater diversity of economic situations (pensioners, unemployed, single parents, students, etc.) and therefore greater heterogeneity of distributional effects. Reasoning in terms of income deciles or living standards does not adequately account for this heterogeneity.

Granularity is also crucial on other economic or social dimensions. Some measures have a much stronger impact on very specific groups that survey data do not necessarily capture. More detailed data, in terms of population and characteristics, would in particular allow a better understanding of the variation in 'horizontal' inequalities. In the current state of available databases –such as the ERFs– we do not observe consumption expenditure, the distance between home and work, or the availability of public services, all characteristics which mean that, for a given income or standard of living, some households are more affected than others by a reform. Due to the nature of the data used, it is not certain, for example, that micro-simulation can capture important phenomena such as the 'yellow vests', which may seem paradoxical.

In the area of income, the IPP is certainly a pioneer. It is to be hoped that other micro-simulation models use the same data sources. This does not diminish the usefulness of specific household surveys. On the contrary, they are the ones that allow the linking of various administrative sources through appropriate matching techniques. Importantly, they provide information on certain household characteristics, notably those used to simulate eligibility for certain benefits. They also allow the examination of horizontal inequalities between

groups of households defined by particular combinations of socio-economic attributes. Although it might not be a priority in the development of those models, this is now a key dimension in the public perception of inequality. In this context, the weakness of surveys, and therefore of the related micro-simulations, is obviously the limited size of the samples and the limited characterization of households. Beyond income alone, it would therefore seem desirable to explore the creation and use of much larger databases offering the same richness of cross-sectional information on households.

Recommendation 2. Explore the use of and facilitate access to exhaustive or near-exhaustive administrative data in the dual dimension of income and demographic and socio-professional characteristics of households.

Principle of simplicity and transparency

An attraction of the strictly accounting approach to micro-simulation is its simplicity and full transparency to the public. It does not rely on assumptions that might be questionable or complex calculations. *Strictly speaking*, however, in the field of redistribution, it implies ignoring behavioral reactions to a reform and their possible macroeconomic effects. It only describes the so-called "first-round effects" or "mechanical effect", i.e. modification of the economic situation of those concerned before they adjust their behavior and possibly contribute to modifying the economic equilibrium.

Not considering behavioral reactions is obviously a problem when assessing the impact of a reform on the State budget (*see above*). On the other hand, and this is a crucial point, it is not problematic for the evaluation of the distributive effects of a reform, except if it concerns non take-up of benefits or tax avoidance. Economic theory teaches that, to a first approximation, the reduction in consumer welfare due to an increase in an indirect tax is the additional amount that would be deducted from his initial consumption, i.e. the mechanical effect. The intuitive idea is that the consumer must vary his consumption of the various goods so that his total expenditure decreases by the amount of the tax increase. But since he is assumed to be optimizing, the price of each good is proportional to his satisfaction at the margin. As long as the variations are small enough, the total variation in his satisfaction can therefore be considered proportional to the tax he would have paid on his initial consumption of goods. Thus, the basis of welfare economics is that the accounting change in the cost of a household's consumption basket due to a price increase is, to a first approximation (i.e. if the increase is small) a "monetary metric of the change in its level of welfare". Being monetary in nature, this metric is comparable between individuals and can therefore be used for redistribution analysis. The argument can easily be

generalized to the case where the household's labor supply is variable and then relates to the household's disposable income. Of course, it needs to be adapted when the expected behavioral reactions are large, which is nevertheless not very frequent, or discontinuous (non-consumption choice). In all cases, it seems preferable, *at least* in a conventional manner, to maintain a strict separation between exercises focusing solely on mechanical effects, specific to the evaluation of social welfare and inequality, and exercises incorporating behavioral responses in addition and allowing the measurement of the budgetary impact. The simulation of the PFU by the Treasury is a good example of the interest of this distinction.

An important pitfall in taking into account the behavioral responses is the imprecision of the responses' modeling. This can be quite rough, consisting for example in a common elasticity of substitution and an imputed individual budget coefficient for indirect taxation, or more sophisticated as in some labor supply models used abroad (see box). In any case, these models are imprecise and this imprecision should be taken into account when it affects an important outcome of the micro-simulation, as it may be the case for labor supply and its impact on the labor market.

Two remarks conclude on this question of behavioral reactions. The first concerns non take-up (or avoidance) behavior, which must be taken into account in order to simulate the mechanical impact of a reform, as opposed to behavioral reactions that modify the base of a tax or benefit. Unfortunately, they are not well known, although they are clearly important.²² The second is the transparency of the models, especially the assumptions underlying the behavioral responses and the need for these to be clearly stated.

Recommendation 3. Maintain a strict separation between, on the one hand, mechanical micro-simulation, without behavioral change (apart from non take-up of benefits or tax avoidance, if applicable) and, on the other hand, micro-simulation with behavioral response.

Extension principle

There has been some discussion on how to reconcile the assessment of income distribution with national accounts, by including income not distributed to households or consumption or use of public services.²³ While the latter is not new, especially in the areas of health and education, the development of new kind of data imply a better understanding of aspects of inequality that go beyond distributed or redistributed income. Nevertheless, neither the conceptual basis, nor the data, nor the available modeling tools make it possible to extend micro-simulation to all income and expenditure flows in the National Accounts before several years. The measurement of their distribution will first have to be more firmly established. Another extension concerns the nature of the measures simulated: this is the one considered here.

Today, French micro-simulation models are primarily concerned with the social levies and benefits affecting households. Should we go further? In fact, any reform, any economic policy instrument has distributional effects on the population. It would be paradoxical if certain instruments that are potentially important from the point of view of inequality were to escape micro-simulation. What about corporate taxation and its impact on employment and wages? How should we deal with public expenditure policies whose objective is resolutely redistributive –such as, for example, the doubling of primary classes in priority education zones?

Technically, it is always possible to measure the mechanical impact of these policies as long as the beneficiaries can be accurately identified. The main problem, however, is that for many policies, the mechanical impact is often very different from the true economic impact, which requires taking into account behavioral responses and looping effects. Hence the importance of modeling to determine how these policies impact households through their primary income, including employment, and their purchasing power through the price system. This can be done, for example, through general equilibrium macroeconomic modeling. Some experience has been gained in the literature with so-called “macro-micro” models that merge micro-simulation with macroeconomic

²² However, the DREES and the CNAF have recently launched a project on this subject to improve knowledge of non-recourse: see DREES (2020): “Le non-recours aux prestations sociales. Mise en perspective et données disponibles”, *Dossiers de la DREES*, no 57, June.

²³ See Germain J-M. (dir.), M. André (rap.) and T. Blanchet (rap.) (2021): “Expert Group Report on the Measurement of Inequality and Redistribution”, *INSEE Methods*, n° 138, February; Bodier M., J. Labarthe and M. Sicsic (2021): “Réduction des inégalités : la redistribution est deux fois plus ample en intégrant les services publics”, *INSEE Références* ‘Revenus et patrimoine des ménages’, and the DINA project launched by Piketty T., E. Saez and G. Zucman (2018): “Distributional National Accounts: Methods and Estimates for the United States”, *Quarterly Journal of Economics*, vol. 133, no 2, pp. 553-609.

models.²⁴ It should also be noted that the expected effects of these policies are often delayed in time. In the case of the doubling of primary classes, for example, the real impact of this educational expenditure is the long-term impact on the wages of the pupils concerned. The conceptual and statistical extension required for the static micro-simulation model is therefore considerable.

On most of these issues, the precision that can be expected from the available models is not sufficient to meet the expectations of policy makers and observers when compared to micro-simulation models of the redistributive system *stricto sensu*. The imprecision is due not only to the macro-micro transition but also to the assumptions of the macroeconomic model, especially on the behavior of agents and the functioning of markets. It therefore seems preferable to adopt a certain flexibility in order to apprehend the redistributive and budgetary effects of such policies. The evaluation of the redistributive effects of education policies integrating behavioral responses and equilibrium effects does not require the same modeling as the evaluation of pension system reforms, for example.

Because it is highly desirable to take into account the distributional and redistributive effects of a considerably wider range of policy measures in their design and implementation, it seems useful to push forward the frontier of micro-simulation while maintaining its transparency. To this end, calls for research projects could be issued in various fields to the research community to enrich the evaluation of public policies and thus contribute to the extension of the field of micro-simulation. In particular, such work on the redistributive aspects of different pension reform scenarios would be welcome.

Recommendation 4. Launch calls for research projects to measure the redistributive impact of measures beyond the classic field of static micro-simulation

²⁴ See, for example, Bourguignon F. and M. Bussolo M. (2013): “Income distribution in computable general equilibrium modeling” Chapter 21 in *Handbook of Computable General Equilibrium Modeling*, Dixon and Jorgenson (eds), n° 1, pp. 1383-1437 for the use of computable general equilibrium models and the current progress in DSGE (Dynamic Stochastic General Equilibrium) models with heterogeneous agents –but in their endowments rather than in their behavior, which constitutes a large part of the literature (see Le Grand and Ragot, 2020).

 <p>conseil d'analyse économique</p>	<p>The French Conseil d'analyse économique (Council of Economic Analysis) is an independent, non partisan advisory body reporting to the French Prime Minister. This Council is meant to shed light upon economic policy issues, especially at an early stage, before government policy is defined.</p>		
<p>Chairman Philippe Martin Secretary General Hélène Paris Scientific Advisors Hamza Bennani, Jean Beuve, Claudine Desrieux, Thomas Renault Economists/Research Officers Étienne Fize, Madeleine Péron, Pierre-Léo Rouat</p>	<p>Members Yann Algan, Emmanuelle Auriol, Anne Épaulard, Gabrielle Fack, François Fontaine, Maria Guadalupe, Xavier Jaravel, Sébastien Jean, Camille Landais, Philippe Martin, Isabelle Méjean, Thomas Philippon, Xavier Ragot, Katheline Schubert, Claudia Senik, David Sraer, Stefanie Stantcheva, Jean Tirole Associated Members Dominique Bureau, Ludovic Subran</p>	<p>Publisher Philippe Martin Editor Hélène Paris Electronic Publishing Christine Carl Contact Press Christine Carl Ph: +33(0)1 42 75 77 47 christine.carl@cae-eco.fr</p>	
<p>TSA 20727 75334 PARIS Cedex 07 (France)</p>	<p>Ph: +33(0)1 42 75 76 84</p>	<p>Fax: +33(0)1 42 75 51 27</p>	<p>@CAEinfo www.cae-eco.fr</p>