EUROMOD WORKING PAPER SERIES

EM 16/17

Welfare Analysis and Redistributive Policies

Olivier Bargain

October 2017

Welfare Analysis and Redistributive Policies*

Olivier Bargain

Bordeaux University and Institut Universitaire de France

Abstract

Applied welfare analyses of redistributive systems nowadays benefit from powerful tax-benefit microsimulation programs combined with administrative data. Arguably, most of the distributional studies of that kind focus on social welfare defined as a function – typically inequality or poverty indices – of household equivalized income. In parallel, economic research has made considerable progress in the measurement of welfare along several dimensions. Distinct but related branches of the literature have attempted (i) to model different behavior (in a way that matter for incidence and redistribution of tax-benefit policies), (ii) to go beyond income, (iii) to better define and estimate equivalence scales, (iv) to open the household black box and measure welfare at the individual level. I suggest a general framework to critically review these streams of literatures and to discuss whether recent advances in each of these fields have been or could be readily operationalized in welfare analyses and policy simulations.

JEL: C81; D31; H21; H23; H31

Keywords: welfare analysis, redistributive systems, microsimulation, equivalence scales, collective models

Contact:

olivier.bargain@u-bordeaux.fr

* This paper is based on a keynote presentation at the *20th EUROMOD anniversary conference*, ISER, University of Essex, 5th of September 2016. It is forthcoming in the *Journal of Economic Inequality* special issue in honor of Tony Atkinson. I am very grateful to Frank Cowell for a thorough reading and excellent comments on the text, as well as to Marc Fleurbaey for the point regarding the social welfare aggregation of money-metric utilities.

1 Introduction

Welfare analysis is an active field of study, especially since the rebirth of income distribution studies in the 1990s (Atkinson, 1997) and the resurgence of interest in inequality following the Great Recession. In particular, the welfare and distributional impact of redistributive policies has been assessed more and more systematically by policy analysts and researchers using tax-benefit microsimulation programs. These computer-based simulations of social and fiscal policies, combined with microdata, represent a powerful tool for policy analysis and policy design because they directly link real-world policies to their effects on household income and household decisions (Bourguignon and Spadaro, 2006; Figari et al., 2017). The development of multi-country models has also offered the possibility to perform welfare analysis in a comparative way across European countries (Atkinson et al., 1988; Immervoll et al., 2011) and monitor the evolution of social policies and their welfare impact in Europe (Atkinson, 2005.). Overall, simulation studies represent a lively field of research in terms of applied welfare analysis, at the junction of several disciplines including economics, social policy and political sciences.

Arguably, most of the studies analyzing policy effects – and microsimulation studies in particular – focus on social welfare defined as a function of household equivalized income, be it a level of inequality, poverty, mean income or more sophisticated combinations of these different measures. While we cannot deny the 'distributional' orientation of these studies, it is more difficult to give them a broad label as 'welfare' analyses, primarily because of the limitations surrounding the concept of household equivalized income as a measure of welfare. In fact, it is striking to witness that the considerable advances achieved in applied economic research – notably the constant progress in the definition and measurement of more complete notions of individual welfare – are rarely incorporated in routine distributional analyses of real-world policies. In this paper, I build bridges between these two worlds with a critical review of the possible improvements or extensions that could be made when performing the welfare analysis of redistributive systems. I will focus on distinct but related branches of the literature that attempt to model behavior, to go "beyond income", to better define and estimate equivalizing scales, and to open the black box of the household in order to measure welfare at the individual level. While reviewing this huge literature seems an ambitious task, my goal is to provide a sketch of the recent progress in these fields and critically question which advances have been or could be operationalized in policy studies to perform more complete welfare analyses.

I will start by reviewing briefly the "standard practice" of policy studies (section 2), highlighting the contribution of microsimulation as a way to design interesting counterfactual distributions of disposable income to improve welfare analysis. Then I will focus on the modelling of behavioral responses that may enrich these studies (section 3). Modelling behavior does not only help to predict the potential responses – in terms of economic decision like consumption or labor supply – to policy changes, but also directly connects to individual preferences and broader concepts of welfare than income. Next, I will specifically address the way welfare measures can go beyond income in normative analyses (section 4). Finally, I will review the conceptual and empirical improvements regarding equivalizing scales and the measure of individual welfare as opposed to household welfare (section 5). Without necessarily recommending that all of these dimensions should be introduced at once in policy analyses, I will attempt to clarify which domains are particularly relevant for (and could feed more systematically into) microsimulation-based welfare analysis and what are the remaining difficulties. In the conclusion, I will discuss the opposite direction of research: how microsimulation studies can contribute to behavioral and normative research, notably by their ability to realistically capture heterogeneity in the way policies shape financial incentives.

2 Welfare Analysis of Redistributive Policies: Standard Practice

2.1 Basic Setting

Let me first describe the social welfare (SW) of a population of N households as follows:

$$SW = W(\frac{c(y_1, z_1)}{e(z_1)}, ..., \frac{c(y_N, z_N)}{e(z_N)}). \tag{1}$$

In this definition, the basic ingredients consist of equivalized disposable incomes c_h/e_h of households $h = 1, ..., N$. The aggregation function W may correspond to various ways to add up household income. It could be the formula of a poverty or inequality index, mean income, or other functions of the income distribution. For instance, nonlinear poverty gap measures combine the notion of poverty headcount and the principle of transfer among the poor. Measures based on the Generalized Lorenz curve also combine the notions of inequality and overall wealth (mean income) of a nation. This general setting is broadly the basis for welfare analysis as performed in policy studies all over the world, notably those based on tax-benefit microsimulation. In these, disposable income

$$c_h = c(y_h, z_h) \tag{2}$$

corresponds to household income after the deduction of taxes and social contributions and the payment of family and social transfers. Hence, it is itself a function of household gross incomes captured in vector y_h and of household characteristics in z_h. In this basic setting, the function $c(\cdot, \cdot)$ that transform income and characteristics into disposable income is usually approximated by tax-benefit microsimulation programs.[1] While various concepts of income compose y_h, vector z_h comprises the characteristics that matter for the calculation of taxes and benefit entitlements (for instance the number and age of children, the region or the size of the dwelling). Disposable income is necessarily deflated by a scale $e_h = e(z_h)$, commonly known as equivalence scale. It takes a value that is specific to household h as it depends again on (at least some of) the household characteristics contained in z_h, in particular the family composition.

2.2 Advantages

Standard practice adopts a limited definition of welfare, as described above, notably the focus on household income and the use of specific equivalizing scales. Referring to a simple monetary concept as c_h/e_h also has many advantages. First, it is easy to explain (to policy makers, journalists or the public). Second, it directly links policy simulations and welfare implications. Indeed, a simple welfare measure as equivalized income c_h/e_h is directly affected by changes in the tax-benefit system through changes in function c. Third, income is a welfare concept that allows aggregation and comparability (once a proper equivalence scale is defined, which is not without difficulties as discussed in section 4), while this is not the case with other welfare measures. Attempts to measure welfare more broadly have punctuated economic history but usually face the same difficulty of interpersonal comparison of preferences, as we shall see. Four, the synthetic equivalence scales commonly used in policy analyses – like the "modified OECD scale" – are simple to calculate and allow international comparability in cross-country studies.

At the same time, referring to function W as a more complete measure of social welfare that account for both inequality and mean income is perhaps a good rendering of what matters in most policy circles. Far from complex notions of welfare, social objectives based on poverty reduction – overall or among specific groups like families with children – and maximum tolerable levels of inequality seem reasonable and pragmatic. They may actually be sufficient to delimit the degree of redistribution consistent with a population's broad preferences and (i) guide the design of social and fiscal instruments for this country, (ii) explain the difference in tax-benefit schedules across countries. Also, under restric-

[1] Throughout the paper, I will focus on microsimulation as a tool to approximate the impact of tax-benefit policies on household budget constraints. I do not discuss other microsimulation methods, notably those using a regression approach to decompose changes in income distribution into the contribution of factors like fertility, labor supply and productivity (see Bourguignon et al., 2004, or a recent account by Herault and Azpitarte, 2014).

tive assumptions about function W, it is possible to rank different income distributions, notably before and after a policy change, on welfare grounds. This is the possibility offered by dominance results, including versions where broad differences in needs are taken into account (Atkinson-Bourguignon's theorem). However, restrictive notions of household/individual welfare are required, and real-world heterogeneity in preferences cannot easily be incorporated in these settings, which will motivate the revival of money-metric utility measures respecting individual preferences (section 4).

2.3 Examples of Applications for Policy Analysis

Let $W(c_{k,t}(y_{k,t}, z_{k,t}), z_{k,t})$ denote the poverty or inequality level characterizing a country k at year t. It is a function of the vector of disposable incomes $c_{k,t}$ and of the vector of household characteristics $z_{k,t}$. The stream of disposable incomes itself depends on the simulated tax-benefit function $c_{k,t}(\cdot,\cdot)$, on the vector of market incomes $y_{k,t}$ and on the vector of demographics $z_{k,t}$. In this way, we can easily summarize most of the production accomplished in national and international tax-benefit microsimulation studies. The bulk of the literature focuses on distributional analyses of policy simulations, which formally correspond to changes in function $c_{k,t}$. In particular, actual reforms in a country k can be seen as the effect of moving from system $c_{k,t}$ to $c_{k,t+1}$, i.e. the change in parameters of this function over two years (see examples in Figari et al., 2017).

A crucial aspect, ignored for a long time in most policy studies, is that the policy change occurs at the same time as a possible change in income distribution $y_{k,t}$ (or household characteristics $z_{k,t}$). Nevertheless, microsimulation allows creating counterfactuals to properly isolate the effect of policy changes over time. The complete change between years t and t' can be written as:

$$\begin{aligned} & W(c_{k,t'}(y_{k,t'}, z_{k,t'}), z_{k,t'}) - W(c_{k,t}(y_{k,t}, z_{k,t}), z_{k,t}) \\ = \; & [W(c_{k,t'}(y_{k,t'}, z_{k,t'}), z_{k,t'}) - W(c_{k,t}(y_{k,t'}, z_{k,t'}), z_{k,t'})] \\ & +[W(c_{k,t}(y_{k,t'}, z_{k,t'}), z_{k,t'}) - W(c_{k,t}(y_{k,t}, z_{k,t}), z_{k,t})]. \end{aligned} \tag{3}$$

This decomposition allows extraction of the policy change contribution (first component of the r.h.s.) versus the contribution of a change in market income or in demographics over time (second component). Note that the former is captured for a particular income/demographic situation (that of the end period), but the symmetrical decomposition could be made (policy changes evaluated at base-period data). It is possible to examine the sensitivity of the results to the choice of the decomposition, but also to use the mean of the policy effects over the two alternative decompositions (a Shapley measure, as suggested by Shorrocks, 2013). Examples of such decompositions have been provided for several countries at key periods for welfare state reforms, showing what would have happened in terms of poverty or inequality if these reforms had not been implemented (Bargain, 2015). This assessment of policy effects is also related to the literature that attempts to decompose changes in income inequality in various factors including changes in income sources or in demographics (Brewer and Wren-Lewis, 2016).

Hypothetical reforms can also be simulated. They may correspond to announced reforms, topical reforms or more radical changes that can serve as a benchmark for policy debates, for instance the introduction of a flat tax, of a basic income scheme, of a tax harmonization at the scale of several countries, etc. They correspond to the switch from $c_{k,t}$ to $c_{\widetilde{k},t}$ with a virtual country $\widetilde{k}$, similar to k but with variation in one or several policy instruments. In this spirit, a particular motivation behind international comparisons is the possibility to analyze the redistributive potential of applying a policy from country k' on country k (ex: applying the British child benefit to Spain). At the international level, this type of analysis represents the majority of studies stemming from the EUROMOD project on tax-benefit policies in Europe (Sutherland, 2015). An even more radical exercise consists in complete system swaps. That is, the system of country k' would be applied to data from country k by simulating $c_{k',t}(y_{k,t}, z_{k,t})$, or the opposite, i.e. $c_{k,t}(y_{k',t}, z_{k',t})$. Again, a decomposition of the type

$$\begin{aligned} & W(c_{k',t}(y_{k',t}, z_{k',t}), z_{k',t}) - W(c_{k,t}(y_{k,t}, z_{k,t}), z_{k,t}) \\ = & [W(c_{k',t}(y_{k',t}, z_{k',t}), z_{k',t}) - W(c_{k,t}(y_{k',t}, z_{k',t}), z_{k',t})] \\ & +[W(c_{k,t}(y_{k',t}, z_{k',t}), z_{k',t}) - W(c_{k,t}(y_{k,t}, z_{k,t}), z_{k,t})] \end{aligned} \tag{4}$$

can be written and would indicate how much of the difference in aggregate welfare between the two countries is due to different tax-benefit systems (first component on the r.h.s.). It is assessed on the basis of incomes and demographics of country k', yet the symmetrical decomposition can also be performed. This type of analysis is limited – further research should provide more systematical swaps (notably at the European level, as promoted by the founders of EUROMOD, cf. Atkinson et al., 1988).

2.4 Limitations

There are many limitations to this framework. A primary concern is that this approach ignores behavioral responses. For instance, reforms of the tax-benefit system have direct impact on disposable income distribution, as characterized before, but also affect earnings distribution y_k through changes in work incentives – both in terms of work hours and labor market participation. The second set of limitation pertains to the definition of welfare as "household equivalized income". Why should it depend on income only? Why should the household be the basis of assessment? Why do we measure equivalized income using equivalence scales that are "synthetic" and typically unrealistic (like the modified OECD scale)?

The rest of the paper will address these points. Before doing so, I will mention other crucial limitations. Consider a slightly more general set-up that the one of equation (1) above:

$$SW = W(u_1^1(c_1^1, x_1^1), ..., u_1^{n_1}(c_1^{n_1}, x_1^{n_1}), ..., u_N^1(c_N^1, x_N^1), ..., u_N^{n_N}(c_N^{n_N}, x_N^{n_N})) \quad (5)$$

with u_h^i the "welfare function" of individual i in household h of size n_h (we ignore time for simplicity), defined over consumption c_h^i and a vector x_h^i of all other relevant dimensions. This setting seems to reflect most of our problems: it splits households into individuals (allowing the possibility of unequal sharing of resources within families), it embeds multiple dimensions beyond income/consumption, etc. Even if we could empirically identify the components of this model, so as to measure individual welfare, this framework would still be problematic and incomplete. The essential limitation is that it is static. It does not account for savings, income smoothing, dynamic effects of policies, etc., and does not account for an intertemporal measure of welfare. Recent developments in dynamic programming provide an interesting setting to study this kind of question (Keane et al. 2011). Another aspect is the nature of W. Assume that governments (or citizens who vote for them) evaluate policies with a specific notion of social welfare in mind. Even with the restriction of a welfarist world, we have no information on the shape of this welfare function (for instance the degree of inequality aversion). There is a vast literature attempting to reveal social preferences, for instance through experiments (Gaertner and Schokkaert, 2012) or through the inversion of well-defined optimal redistribution models (Bourguignon and Spadaro, 2012).

3 Modelling Behavior

While the recent years have seen applied economic research focus mainly on reduced-form approaches based on experiments, there is a long tradition of modelling human behavior using structural, micro models. The advantage for welfare analysis is that these models directly link behavior and responses to policy changes to welfare measures. Different types of behavioral modules could potentially augment tax-benefit analyses based on microsimulation tools: consumption, saving, labor supply, benefit take-up, tax compliance, migration, etc. Thus, behavioral microsimulation could inform us about the incidence and the changes in the tax base following a policy reform, which may affect the distributional and budgetary impacts of the reform.

In practice, behavioral modules are not systematically implemented because: (i) they sometimes require quite technical or heavy estimation procedures; (ii) their validity hinges on a set of additional assumptions regarding behavior, model specification and the distribution of stochastic terms that policy analysts are not ready to defend when presenting their analyses to policy makers; (iii) even among researchers, the internal validity of the structural approach is put into question when behavioral parameters are not identified thanks to experimental approaches (see the discussion and references in Bargain and Doorley, 2017). Nonetheless, I summarize here the basic principles of modelling consumption and labor supply behavior in policy analyses. Consumption is relatively easy to incorporate in the analysis of indirect taxation while labor supply (and direct taxation of labor income) is more complicated, even if extremely important for redistribution studies.[2] I will also review aspects regarding the targeting and coverage of redistributive policies, which include both the take-up of social transfers and tax compliance. Finally, while the "structure" of behavioral models directly relates to the preference functions used for welfare analyses, such a disaggregated modelling level is not always necessary to draw the welfare implication of policy changes. This calls for a short review of the recent literature on "sufficient statistics".

3.1 Consumption

The behavioral effects of indirect taxtion on consumption can be modelled and incorporated relatively easily in policy simulations. It requires expenditure data and a simple model of consumption over K goods. The latter is written as follows for a given household h with demographic characteristics z_h:

$$\begin{aligned} & \max_{q_h^1,\dots,q_h^{K-1}} u_h\left(q_h^1,...,q_h^K;z_h\right) \\ \text{s.t. } c_h &= p^1 q_h^1 + ... + p^K q_h^K. \end{aligned}$$

It leads to a demand system $q_h^k = q_h^k(p^1,...,p^K,c_h;z_h)$, $k = 1,...K-1$, that depends on household disposable income c_h and prices $p = (p^1,...,p^K)$ assumed to be net of indirect taxes. Hence, reforms of the differentiated rates of VAT (or excise taxes) across types of good can be simulated straightforwardly once the demand system is estimated. Estimation techniques based on flexible specifications like the QUAIDS have become standard (see Banks et al., 1996). Provided that many years of data are available and that price variation over time (or time $\times$ regions) can be treated as exogenous, the model is relatively well

[2] Indeed, labor income is the main income source in our societies (and almost the only one in poor countries) and the relevant basis for redistribution (unequal distribution in market productivities justifies distortive redistributive policies and a possible tradeoff between equity and efficiency, as studied in the huge literature on optimal taxation).

identified. Different price elasticities at different income levels will orientate the incidence of a commodity tax reform and its redistributive impact. For instance, raising taxes on tobacco is regressive if elasticities of cigarette consumption are lower at low income levels.

If distributional analyses are based on disposable income c_h, then the impact of price variations (due to indirect tax reforms) is trivial since the value of total consumption does not change. Very clearly, the structure of the model should be used to infer more complete welfare impacts of a change in the consumption bundle. Standard integration results can readily be applied to derive indirect utility or cost functions and, from there, to assess changes in money metric utility (e.g. equivalent variation) due to a price change or a policy reform. A few microsimulation studies have adopted this approach. In particular, an interesting application is the distributional and budgetary impact of a VAT harmonization at the EU level (Copenhagen Economics, 2007). The difficulty is that microsimulation of both direct and indirect taxation should rely on expenditure data, which is not systematically the case. In fact, administrative data on income and labor are most often used as best practice for direct taxation. In this case, in the absence of expenditure data, the redistributive effects of indirect taxes must rely on imputation techniques (Decoster et al., 2011).

3.2 Labor Supply

The case of labor supply is slightly more complicated because it relates to direct taxation, which is often nonlinear. Denoting labor supply (worked hours) l_h and consumption c_h, it starts with a basic static model of the type:

$$\begin{aligned} & \max_{l_h} u_h(c_h, l_h; z_h) \\ s.t. \quad c_h &= c(y_h, m_h, z_h) \text{ and } y_h = w_h l_h, \end{aligned}$$

for a household of composition z_h, labor income y_h (hourly wage w_h $\times$ hours worked l_h) and unearned income m_h. These three ingredients are transformed into disposable income through the function c, which aggregates income sources, adds benefits and withdraws taxes and social contributions by means of tax-benefit microsimulation. The latter is combined with an econometric model that usually treats the household as a single decision maker, represented by the household welfare function u_h. The estimation of such a "unitary" models of labor supply in the presence of taxation has become relatively standard. In particular, discrete choice models have been used broadly to analyze policy reforms (van Soest, 1995, Aaberge et al., 1995). By discretizing the choice set in a finite number of work hours, we can directly account for corner solutions (non-participation) and for complex budget constraints.[3] Once the model is estimated on pre-reform data, a new function c' can be simulated in place of c and the estimated utility function used to predict the new optimal labor supply choice for each household in the population. The behavioral response can be taken into account in the distributional analysis. In this way, the incidence of a tax or benefit reform can be assessed while taking into account the different elasticities of labor supply (worked hours or participation) at different income levels, i.e. which groups respond to a reform and by how much.[4]

The model above is extremely simple. We assume only one potential worker in the household, there is no childcare cost, no fixed cost of work, no benefit take-up. Most these refinements can be incorporated relatively easily (Blundell et al., 2001). In particular, it is straightforward to account for the joint decision of several workers in the households.[5] It is also possible to include penalties for work costs (not without some identification issues) and benefit claim costs (see below). Because of the flexibility and advantages of this setting, many studies have emerged over the past twenty years, which attempt to measure behavioral responses to policy reforms using structural model estimations with detailed tax-benefit microsimulation. Early surveys focus on various techniques and evidence for the US and the UK (Blundell and MaCurdy, 1999); more recent ones review international evidence (Bargain and Peichl, 2016, Aaberge and Colombino, 2015). Arguably, the main issue in most of these studies is the fact that preference parameters are not necessarily well identified. Indeed, possible confounding factors (e.g. being a hard worker) affect both work preference in u_h and wage rates w_h. A now standard solution is to use exogenous variation in *net* wages $(1-\tau_h)w_h$, with $\tau_i = 1 - \partial c/\partial l_h$, as caused by tax-benefit reforms affecting the function c (see Blundell et al., 1998, or Bargain et al., 2014). Another option is to combine structural estimations with (quasi)experiments for identification (Bargain and Doorley, 2017). Critically, this identification issue is largely ignored in the policy literature using behavioral microsimulation, casting doubt on the validity of the predictions of policy impacts.

[3] Function c embeds the many nonlinearity and discontinuity generated by real-world tax-benefit systems. Yet it is only necessary to evaluate budget constraints at finite points in the set while tangency conditions need not be derived as in older models based on marshallian labor supply functions.

[4] Labor supply responses can also be used to enrich the aforedescribed counterfactual analysis used to extract the role of policy reforms on income distribution over time. The contribution of the policy effect is split in a direct effect and the behavioral effect (Bargain, 2015).

[5] The utility function becomes $u_h(c(w_h^1 l_h^1, w_h^2 l_h^2, m_h, z_h), l_h^1, l_h^2; z_h)$ in the presence of workers 1 and 2, say, husband and wife. Discrete choices correspond to the various combinations (l_h^1, l_h^2).

3.3 Benefit Take-up

Other dimensions are also potentially important for welfare analyses, especially in the case of benefit take-up. Leakages in the targeting of social transfers may have serious consequences in terms of poverty reduction. The bulk of the literature, surveyed in Currie (2006), attempts to disentangle the different channels explaining non-take-up, and particularly stigma (Moffitt, 1983), transaction costs (Anderson and Meyer, 1997) and a low perceived utility from transfers (Pudney et al., 2007; Bhargava and Manoli, 2015). Take-up in the context of working age households can be modelled simultaneously with labor supply, as a simple extension of the previous framework (Moffitt, 1983; Brewer et al., 2005). I just rewrite the budget constraint by decomposing c into its components: gross incomes, taxes t and benefits b:

$$\begin{aligned} & \max_{l,P} u_h(c, l; z_h) - \phi_z P \\ s.t. \quad c \ = & \ w_h l + m_h - t(w_h l, m_h, z_h) + P.b(w_h l, m_h, z_h) \end{aligned}$$

with P a dummy indicating whether household h, eligible for benefit b, takes it up. Eligibility will depend on earnings, other income and family characteristics. The hassle cost and stigma is given by ϕ_h, which may depend on household characteristics and be incorporated as an additional element in welfare analyses (Pudney et al. 2007). If b is an out-of-work benefit, then the participation decision ($l > 0$ or $l = 0$) will be jointly taken with the claim decision in case of inactivity. Symmetrically, if b is an in-work transfer for which a claim is necessary, then the range of hour choices l that may yield eligibility will also be associated with the claiming decision. The main difficulty consists in measuring precisely entitlement in the data, i.e. not to attribute it by mistake due to some eligibility condition based on unobserved information (e.g. a wealth test) or misreporting (e.g. underreporting of income in benefit means-tests).

Many refinements and ways to address measurement errors in this context have been suggested by Duclos (1995), Terracol (2002), Hernandez and Pudney (2007), among others. They have rarely been implemented in practice. Several studies agree on the fact that countries characterized by generous social assistance are affected by large rates of non-take-up, which considerably weaken the anti-poverty potential of the redistributive system (for instance in Germany or France, see Riphahn, 2001, and Terracol, 2002, respectively). Despite of this, most microsimulation studies assume full take-up, i.e. they are interpreted as showing the intended effects of the system. Some attempts have been made to estimate the extent and incidence of non-take-up in comparative European simulations (Matsaganis et al., 2010), yet they have not been operationalized.[6]

[6] A simple correction for benefit non-take-up is sometimes added as a random draw of beneficiary

3.4 Tax Avoidance and Informal Work

A complementary issue is that of tax compliance. Leakages in the collection of tax money may seriously affect public finances and reduce the efficiency of progressive tax systems. Jointly with labor supply, the extensive margin of tax compliance (i.e. fully avoiding tax t) can be modelled in a similar way as take-up: there is a disutility from legal tax avoidance (legal costs, tax mobility) or tax evasion (fear of being caught, warm glow effect, etc.) while there is the gain from not paying tax in the budget constraint. In the case of tax evasion, some parameters must also reflect the probability of being caught and the cost of punishment. It is also possible to model the allocation of time between formal and informal work. The literature has started with models of tax evasion models, like Allingham et Sandmo (1972) and Cowell (1985), which actually combined conventional labor supply theory with a standard portfolio model, i.e. how much labor to supply all in all and how much in each sector. Empirical applications rely on interviews where people are induced to give some information about time (or money) spent on untaxed activities (expenditure). This information is used to estimate structural models of labor supply and formal-informal sector choice in emerging economies (Gong and van Soest, 2002; de Hoyos, 2012) or rich countries (Lacroix and Fortin, 1992; Lemieux et al., 1994).

Structural models of that sort have rarely been implemented in welfare analyses based on microsimulation, for the same reason as what we mentioned above for labor supply and, also, because of data limitation on the degree of avoidance. An interesting stream of research attempts to uncover indirect evidence of evasion, which Slemrod and Weber (2012) call "traces of evasion". An early example is the consumption-based approach, whereby it is assumed that income is understated in household surveys while expenditure is correctly measured (Pissarides and Weber, 1989; Lyssiotou et al., 2004). Other approaches infer evasion by comparing the marginal increase in charitable donations with respect to sources of income subject to different third-party reporting requirements (Feldman and Slemrod, 2007). Recently, attempts have been made to model tax avoidance by comparing the responses in income surveys and in tax returns of identical or similar individuals. In this way, these studies estimate the extent and distribution of income under-reporting for the purposes of tax evasion (Benedek and Lelkes, 2011; Matsaganis et al., 2012) and assess the implications for distributional analyses based on microsimulation.[7] In countries where

based on the take-up proportions reported in national statistics. For benefits, a reduced-form estimate of the claiming behavior, including the level of benefit entitlement, could be at least estimated and used in microsimulation.

[7] The approach relies on the assumption that respondents to income surveys truthfully reveal their income as they have no incentive to conceal it, while the opposite is the case when individuals fill in their tax return. While the former assumption is opposite to that made in the consumption-based approach cited above, these studies provide reasonably intuitive results and notably show that the deviation of

tax evasion is a widespread phenomenon, a correction adopted in some simulations consist in splitting income sources (generally employment and self-employment income) recorded in the survey into a first component which is assumed to be reported to the tax authority and in a residual component which is assumed to be evaded (Ceriani et al. 2013), the share of each being drawn from national statistics on the extent of the shadow economy. Overall, however, most microsimulation studies assume full compliance and ignore untaxed labor. Admittedly, given the different types of tax avoidance one may consider, the different populations it involves (poor informal workers, high-skill tax evaders, etc.), and the large variety of methods and empirical results, it is difficult to suggest a particular path to take – and this lack of guidance explains the unfortunate result that microsimulation studies tend to assume full tax compliance and ignore informal employment and informal consumption.[8] It is nonetheless tempting to recommend the inclusion of a probability of tax evasion by income levels in order to enrich distributional analyses based on disposable income, even if this is difficult without country-specific estimates of the magnitude of such effects.

3.5 Taxable Income and Sufficient Statistics

Welfare analyses suggested above rely on structural models that allow one to model and estimate behavior, identify household welfare and simulate the effect of policy reforms on households' income and utility. Recent years have witnessed a decline in the use of structural models. It is often seen as challenging to identify all the primitive parameters in an empirically compelling manner because of selection effects, simultaneity bias and omitted variables. The recent literature tends to favor reduced-form approaches based on randomized or natural experiments, much stronger in terms of causal inference of actual policy effects. Yet, it is not excluded using them to validate structural models (see Thoresen and Vattø, 2015, and reference therein) or combining them (as suggested by Blundell, 2012), a direction increasingly taken in the literature (see Attanasio et al., 2011, Bargain and Doorley, 2017, and Low and Meghir, 2017, for examples, discussions and references).

It is also noticeable that the literature on behavioral responses to tax policies has recently concentrated on (quasi)experimental approaches that are supposed to capture a broad

incomes between the two data sources is greater in the case of farming and self-employment income.

[8] Another issue is that survey-based studies may be accused of lacking internal validity. In fact, recent research relies of field experiments (for instance randomized audit experiments, in Kleven et al., 2011) or natural experiments (for instance changes in tax rules that give opportunities to save on taxes by overstating charitable donation, in Landais and Fack, 2016, or by claiming additional dependents, in LaLumia and Sallee, 2013).

set of potential responses of the tax base y_h, including labor supply, work effort, benefit take-up, legal or illegal tax evasion, etc. In particular, there is now a large literature on the elasticity of taxable income (ETI) that estimates the responsiveness of taxpayers to income tax changes (see Saez et al. 2012 for an overview). A related literature estimates elasticities of the tax base using the degree of bunching at kinks (provided by piecewise linear tax system) or notches (provided by sharp discontinuities in budget curves due to eligibility thresholds), as survey by Kleven (2016). It could be relatively easy for microsimulation studies to rely on such estimates to incorporate basic responses of the overall tax base in policy simulation. A possible issue is the relatively high sensitivity of estimates to the model specification or the method used.[9] For these reasons, and because of the absence of precise and systematic estimates in all countries and time periods, a pragmatic approach would consist in enriching policy simulations by behavioral responses of different magnitudes. One could use the usual no-response scenario (as a lower bound), a mean estimate (for instance an average value of 0.25 for the US, cf. Saez et al., 2012) and a reasonable upper bound (for instance a value of 1), overall providing a broad confidence interval of the likely responses of policy changes. Several studies have actually performed welfare analyses of redistributive policies while using this type of pragmatic approach and reducing most of the heterogeneity in household preferences to the variation in elasticities across discrete income groups (see Immervoll et al., 2007, 2011).

Since the seminal contribution of Feldstein (1999), the ETI literature has led to the question of whether welfare analyses of policy changes could be performed without structural models. This literature indeed suggests using ETI estimates as a "sufficient statistic" to estimate the deadweight loss of income taxes. Yet, as put by Chetty (2009), this should be seen as a complement rather than a replacement of the structural framework. Indeed, the sufficient statistic approach must derive new formula for each specific question, and possibly deals with specific types of social welfare functions. Moreover, while modelling assumptions are weaker than in the structural approach, out of sample predictions may also be less reliable. Finally, it assumes that all types of behavioral responses that affect taxable income can boil down to one single number (e.g. the ETI). In fact, Chetty (2009) shows that the channel along which taxable income responses occur may make a difference for the efficiency losses induced by income taxes. For instance, losses are lower when behavioral adjustments generate externalities such as transfers (charitable donations), compared to cases where they are associated to real resource costs (Dörrenberg et al., 2016).

[9] For instance, Burns and Ziliack (2016) show that estimates obtained with a grouping instrumental variables estimator à la Blundell et al. (1998) are 2-3 times larger than those from the standard individual-level synthetic tax instrument used in much of the literature.

4 Going Beyond Income

In the basic version of our social welfare measure, in equation (1), we defined individual or household welfare solely as a function of disposable income (or consumption). Let us introduce some separability in household income levels:

$$SW = W(u(c_1), ..., u(c_N)).$$

With a uniform function u, the theoretical literature on poverty and inequality has provided interesting results. Under some assumptions on u, it is possible to make welfare comparisons on the basis of income distribution data. The well-known theorem of Atkinson (1970) shows that dominance in terms of inequality between two populations is equivalent to dominance in terms of welfare (Shorrocks, 1983, extends the result to inequality and average income). Yet, this result applies to a world without heterogeneity regarding other dimensions than income, and in particular no heterogeneity in function u, so that the latter is often interpreted as part of the social welfare function. Heterogeneity in household types can still be introduced. With further regularity assumptions on the way households of different size differ in needs, sequential dominance criteria have been obtained for the comparison of joint distributions of income and needs (Atkinson and Bourguignon, 1987). Another approach is to use equivalence scales and apply standard dominance results on artificial populations of 'equivalent adults' (Ebert, 1997). These approaches may be reconciled (Lambert and Ramos, 2002), but the question is whether they can easily be made operational in empirical work and policy analysis. The sequential approach has been used in some applications (e.g. Chambaz and Maurin, 1998) while the identification of equivalence scale poses serious challenges that are discussed in the next section. Another fundamental limitation with this framework is the fact that welfare depends on income only. While the different behavioral models suggested in the previous section show ways to incorporate other dimensions (leisure, disutility from claiming behavior, etc.), two main problems appear. First, it is not clear how to compare utility for aggregation. Second, these models suggest a still limited set of dimensions while individual well-being may depend on a very large range of factors. Hence, I review two approaches at the forefront of welfare measurements: the use of subjective well-being information and the development of money metric utility aimed at providing comparable and aggregatable welfare measures (a considerably more extensive discussion of these issues is suggested in Fleurbaey and Blanchet, 2013).

4.1 Subjective Well-being

The resurgence of interest for the measurement of welfare has taken the form of different initiatives (the Human Development Index by the United Nations, subjective measures in OECD reports, expert views after the Stiglitz-Sen-Fitoussi Commission, etc.). Among the different measures that go "beyond GDP", subjective well-being (SWB) has been successful among the public and policy makers, for it provides single-valued summary indices of well-being. This approach relies on answers about happiness or life satisfaction on a numeric scale or on an explicit scale (happy, very happy, etc.). This approach is fully welfarist, since it is assumes that self-reported levels of well-being can be directly used as cardinal and interpersonally comparable measures of welfare. In that sense, it is close to the early attempts by utilitarians who searched for welfare measures associated with pleasure and pain following an experience (see Kahneman and Sugden, 2005). For the same reason as what led to the ordinal(ist) revolution a century ago, SWB is still not accepted by a large part of the profession (see Fleurbaey, 2009). The bone of contention is essentially the question of interpersonal comparability with SWB measures but also the type of information and welfare definition they may provide.

On the informational content, a rapidly growing amount of evidence collected by economists and psychologists over the recent years has shown that SWB is not pure noise and can be validated in numerous ways (see critical reviews in Clark et al., 2008, Kahneman and Krueger, 2006, or Fleurbaey and Blanchet, 2013), notably against behavior and more objective measures of well-being (see Krueger and Schkade, 2008, and Oswald and Wu, 2010). In this sense, it may come as an interesting complement to standard measures based on ordinal preferences. Indeed, an old critique of the money metric approach pertains to the fact that ordinal measures obtained by revealed preferences may not incorporate sufficient information about subjective welfare.[10] In principle, elements of subjective welfare may be incorporated in utility functions and help to provide the complement of information about one's well-being (see Fleurbaey, 2008). In fact, SWB studies push the logic to its end and beyond. For them, individual choices are irrelevant so that revealed preferences must be replaced by self-reported well-being to evaluate social welfare. Individual SWB is regressed on consumption c and other relevant determinants of welfare x so that the estimated function $u_h^i(c_h^i, x_h^i)$, for individual i in household h, can be used to evaluate national welfare as in equation (5).

The most problematic aspect of this approach is somewhat the opposite of the above criticism against ordinal preferences. It cannot be taken for granted that SWB directly provides interpersonally comparable measures of welfare. The main issue is that self-reported well-being contains a huge degree of individual heterogeneity in self-perception about one's situation, both in relation to others (relative concerns) and to one's self over time (adaptation), and in adaptation to this situation. For instance, the resilient poor

[10] As recalled by Sen (1979) about ordinal preferences: "A variation of one's intensities of pleasure or welfare cannot find any reflection in this numbering system as long as the ordering remains unchanged".

may report high well-being levels while the demanding rich may declare experiencing low satisfaction. In both cases, it cannot justify a policy redistributing to the latter or failing to address the poor conditions of the former.[11] At least, SWB information should be cleaned from individual relative concerns and aspirations for it to be used in welfare analysis. Recent studies have suggested ways to do so in order to construct money-metric measures based on many dimensions of the "good life" (Decanq et al., 2015, and other references below). Attempts to bring SWB data together with policy simulations are also rare (see Jara and Schokkaert, 2017).

4.2 Money Metrics

To go beyond income and construct multidimensional welfare measures requires data on the different relevant dimensions of well-being and a system of individual preferences to weight these dimensions. This can be done for instance on the basis of structural models, as those presented above, if the utility function can be retrieved thanks to the revealed preferences approach. The way to summarize the different dimensions in a single index is the standard money metric utility approach. The latter only requires information about ordinal (non-comparable) preferences, but nonetheless provides welfare metrics that are cardinal and comparable, just like ordinary income. Hence, these metrics can be used for distributional analyses and potentially for social welfare aggregation. Yet, while estimations and the revealed preferences approach can provide information about preference heterogeneity, the latter is usually ignored during the normative analysis (i.e. the construction of money metric utility). In labor supply modelling, for instance, the standard approach first introduced by King (1983) relies on money-metric utilities that are evaluated by inserting chosen bundles into a fixed reference preference ordering – that of a reference household chosen by the social planner – and using fixed reference prices (see applications for tax reform in Aaberge and Colombino, 2004, for instance).

Progress have been made with the 'fair allocation' theory, which suggests ways of ranking individual situations when preferences differ while escaping from most of the standard criticisms about money metrics (Fleurbaey and Maniquet 2006). In particular, the choice of a reference set is less arbitrary, or at least supported by explicit normative principles. Arrow's impossibility is overcome by relaxing the Independence of Irrelevant Alternative

[11] This was already the standpoint of opponents to subjective welfarism 30 years ago. It was notably Sen (1985) who argue that utility does not sufficiently take into account the real physical conditions of the person (physical-condition neglect): "A person who is ill-fed, undernourished, unsheltered and ill can still be high up in the scale of happiness or desire-fulfillment if he or she has learned to have 'realistic' desires and to take pleasure in small mercies". Moreover, valuing a life is a reflective activity in a way that 'being happy' need not be (valuation neglect).

and making use of all the information about an individual's ordinal preferences (*non-paternalism*). An ordinal equity concept of *egalitarian-equivalence* is used: it consists in retrieving a configuration where the actual allocation of individual bundles is Pareto equivalent to an egalitarian allocation – the reference set – that is chosen according to explicit fairness criteria. Interpersonal comparisons (*dominance principle*) are then made in the specific regions of the indifference set that are justified by these ethical choices (Fleurbaey, 2008). Let me exemplify this approach in the case of labor supply models. Ordinal preferences are respected and summarized by a person's indifference curves in the income-leisure space. We can define well-being at the optimal choice as

$$u_h^* = \max_l u_h(c(w_h l, m_h), l),$$

then use tax-free linearized budget constraints $c = \widetilde{w}_h l + \widetilde{m}_h$ (with virtual wage $\widetilde{w}_h$ and nonlabor income $\widetilde{m}_h$) to define various possible money metrics. Take for instance one defined as nonlabor income m_h^*, which leads to the same welfare level as above:

$$u_h^* = \max_l u_h(\widetilde{w} l + m_h^*, l),$$

for an egalitarian reference wage $\widetilde{w}$. In this context, inequalities arising from endowed circumstances (like productivity) should be removed but those from other factors (like preferences) may remain, depending on the level of $\widetilde{w}$. At one extreme ($\widetilde{w} = 0$), the money metric m_h^* will be equal for all those with the same unearned income, whatever the shape of their preferences.[12] In this case, inequality due to preference heterogeneity is fully neutralized, i.e. people are not held responsible for their work aversion. The degree of responsibility will increase with $\widetilde{w}$ and might depend on whether preferences is seen as a responsibility factors (for instance, if underemployment is involuntary or if work aversion is due to unpleasant or risky jobs for the unskilled, it seems charitable to choose a low value of $\widetilde{w}$, see Fleurbaey and Maniquet 2006).

Empirical implementations of this approach have been suggested using revealed preferences, for instance in Bargain et al. (2013), Decoster and Haan (2014) or Carpantier and Sapata (2016). These applications have focused on the income-leisure domain, as illustrated above. That is, money metrics is used to rank individuals when accounting for non-market time in addition to income in labor supply models – a modest departure from income but nonetheless important for tax-benefit policy studies. In general, it is difficult to address many dimensions of welfare beyond income since 'price' variation is required for each of the 'goods' added to the welfare measure. In fact, SWB regressions provide an interesting way to put a weight on many dimensions at once. Some studies have shown how ordinal preferences could be inferred from SWB data. Decanq et al.

[12] A related idea can be found in Kolm's 'leisurely equivalent income' (Kolm, 1969).

(2015) originally suggest the construction of money-metric evaluation of multidimensional welfare, while Schokkaert et al. (2011) focus on income and job satisfaction. Decancq and Schokkaert (2013) and Decancq et al. (2015) follow similar approaches while focusing on social progress and poverty respectively.

Aggregation of welfare money metrics in social welfare functions is possible, in principle. Yet, a well-known issue remains: equivalent income measures are not necessarily concave in income and, hence, may induce antiegalitarian policy implications (see Blackorby and Donaldson, 1988). A way around this problem is to restrict the social objective to positional welfare functions, like the maximin social objective. Arguably, that money metrics do not satisfy the Pigou-Dalton principle everywhere is not necessarily a strong argument against using them to construct a social welfare function that is less extreme than the maximin. Indeed, the violation of the Pigou-Dalton principle occur only when indifference curves change shape as utility increases, in a way that makes the violation of the principle not so shocking (see Bosmans et al. 2017 for a discussion and a suggestion on how to use aggregation more systematically).

5 From Households to Individuals

The last set of improvements I would like to discuss pertain to the unit of welfare analysis and the comparison of households for welfare measurement. Common practice consists in comparing households of different size and composition by adjusting incomes using "equivalence scales". The main reason for taking households as the reference unit is that it is difficult to observe the intrahousehold decision-making process. Complications arise from the presence of goods consumed jointly (a degree of 'publicness' in consumption). Another related issue pertains to data availability: It is invoked that the relevant information is usually recorded at the household level, notably the expenditure of most goods. As a result, distributional analyses traditionally rely on "equivalized" household income, i.e. before aggregation (into poverty, inequality or social welfare measures), household income is deflated by equivalence scales that reflect the overall needs of the household (depending on the number and type of persons composing the household) and the possibility of economies of scale (that potentially improve the welfare of multi-person households). This approach suffers from a fundamental identification problem and broadly ignores the possibility of unequal resource sharing within the household. This double issue has been addressed by new advances in the literature on family economics, which suggest shifting the unit of analysis to the individual and offer new ways to model and estimate scale economies.

5.1 From Equivalence to Indifference Scales

Comparing households of different size and composition is a difficult task. The only certainty regarding the difference in needs across household types is that children, at least under a certain age, cost less than adults. Many other circumstances may deserve attention, for instance that of adults with specific conditions (like disability or old age care expenditure), who may have larger needs than other adults – but these aspects are rarely accounted for. Synthetic equivalence scales rely on very crude weighting systems that focus almost exclusively on the child-adult differentiation. For instance with the modified OECD scale, a weight of 0.3 is assigned to a "child" – defined as a person under 14 years of age – while it is 0.5 for every "adult" beyond the first adult. Hence, the need of a person under 14 is deemed $3/5 = 60\%$ that of a person above 14. This is the scale used almost systematically for internationally comparative distributional analyses in Western countries (notably for microsimulation-based welfare analyses) and also frequently used at national levels. As for economies of scale, the only safe bet is that a couple is going to cost less than two single individuals due to its joint consumption of public goods like housing costs. The modified OECD scale assumes that the weight on the first adult is 1 while it is 0.5 for every additional adult, i.e. a couple requires $1.5/2 = 75\%$ of the expenditure of two singles to achieve the same welfare level.

Clearly, synthetic scales of that type are not very realistic. Country-specific scales should be estimated that better reflect the demographic characteristics of a population. To do so, let us first formally define equivalence scales. Denote v the household's indirect utility as a function of market prices p and total income c. The traditional definition of an equivalence scale e_h for a household of type h is written:

$$v_h(p, c) = v_0\left(p, c/e_h\right)$$

with reference to a childless single individual (household of type 0). The scale is defined as the answer to the question: "how much income is needed for the adult living alone to be as well off as the family of composition h". For instance, a scale of $e_h = 2$ means that a single adult needs half of the household resources c to be as well off as household h. If the latter increases in size (new children) or starts to consume less public goods (drop in scale economies), the equivalence scale increases so that the single individual needs less than before to be as well off as the family, i.e. the standard of living of the latter, at a constant c, have decreased.

There are several problems with this definition, well summarized by Chiappori (2016). First, equivalence scales are not (ordinally) identified, as they depend on how utility is cardinalized (Pollak and Wales, 1979). Positive transformations of v_h yield new indirect utilities that are observationally equivalent, i.e. ordinal preferences are unchanged so that observable marshallian demands will be the same, but give different numerical values of e_h compared to the original definition using v_h. Identifying a single equivalence scale would require much more information about households. Interpersonal comparability would require one to know the household decision function and to perform comparison on the shaky ground of a ill-defined concept of 'household welfare'. Another problem is the fact that equivalence scales ignore possibly unequal sharing within families. These issues are related: the comparison of households of different types is difficult precisely because we ignore the collective decision making process – and particularly the allocation of resources within the household – underlying function v_h.

The traditional literature has suggested adding identifying assumptions when attempting to estimate equivalence scales on expenditure data. For example, the Engel scale assumes that household welfare varies with the share of expenditure dedicated to food: it suffers from several biases, the main one being the overstatement of child scales because households with children spend disproportionally more on food. The Rothbarth scale assumes separability between adult and child goods in $u_h(c)$, as well as the stability of adult preferences across family statuses, i.e. the presence of children is assumed not to change adult preferences regarding adult goods. Under these assumptions, it is possible to measure the cost of children as the level of resources shifted from adult consumption (in a childless household) to child consumption (after the arrival of a child in this household). Given the observability of adult goods in standard data, this old idea remains central in the attempt to measure economies of scale and has recently been incorporated in the collective model framework in order to identify both the individual welfare function and the household consumption technology (Browning et al. 2013).

The idea is to compare individuals in different family situations. Let v_i be the indirect utility of adult i. Assume that life with others can be summarized by resource sharing and joint consumption: adult i will receive a share $\eta_{i,h}$ of total resources when she lives in household h. She also jointly consumes with other members so that market prices are altered by a set of coefficients $\alpha_{i,h}$ (Barten scales). For instance, if she lives alone with her husband and they always drive the car together, the car is consumed twice, as if the market price was multiplied by a factor $\alpha_{i,h}^{car} = 0.5$. The identification problem regarding equivalence scales is solved by changing the question above into: "how much income is needed by individual i living in a family of composition h to be as well off as when she lives alone?". The answer is a so-called "indifference scale" $I_{i,h}$ (Lewbel, 2003), defined as follows:

$$v_i(\alpha_{i,h}p, \eta_{i,h}c) = v_i\left(p, c/I_{i,h}\right). \tag{6}$$

In this formula, we compare the utility of a person i in different family contexts: living with others (left-hand side) and living alone (right-hand side). The formula reflects the double effect of living with others: scale economies (summarized by Barten scales $\alpha_{i,h}$) and sharing (summarized by the sharing rule $\eta_{i,h}$). The income effect equivalent to this, the indifference scale, is expressed here as a deflator of total income, just like the equivalence scale. Yet it is uniquely defined and hence empirically identifiable, i.e. a positive transformation of v_i would leave $I_{i,h}$ unchanged. For instance, an indifference scale $I_{i,h} = 2$ means that a person, if living alone, needs half of the household h's resources to reach the same indifference curve as when she shares and consumes jointly with others. Browning et al. (2013) have formalized this approach in a context with price variation over time. With mild additional assumptions and in a setting with constant prices, Lewbel and Pendakur (2009) simplify the problem as follows:

$$v_i(p, \frac{\eta_{i,h}}{\sigma_{i,h}}c) = v_i\left(p, c/I_{i,h}\right).$$

Scale economies are now summarized by a single deflator $\sigma_{i,h}$. The indifference scale combines this scaling factor deflated by the resource share. Note that $c/I_{i,h}$ is a money metric utility, with specific reference prices. A more general normative characterization of sharing rules as equivalent income is explored by Chiappori and Meghir (2015).

5.2 From Household to Individual Welfare

Welfare analysis usually focuses on the household as the core unit of analysis. Yet, there may be inequality (and different levels of poverty) within a family. This is intuitive when one think of poor regions of the world, as evidenced in anthropometric studies (Haddad and Kanbur, 1990). Microsimulation studies – and welfare analysis more generally – cannot ignore the question of within-household distribution in this context, especially when development policies aim to target specific individuals like women or children (Alderman et al., 1995). The motivation to place the individual at the center of the analysis naturally goes beyond poor regions. For instance, gender inequality may prevail to some extent in rich countries too, partly driven by gender wage gaps, social norms and mating dynamics (Lise and Seitz, 2010).

The motivation is not only driven by policy analysis on individual-based inequality. In fact, the economic literature has clearly demonstrated that unitary models of household behavior are not well theoretically grounded when applied to multi-person households, and are empirically rejected in this case (Chiappori and Donni, 2011). Attempts to depart from the unitary framework date back at least to Becker's "rotten kid" model and Samuelson's "consensus" model, which respect individualism and represent households as a collection of individual utility functions. Yet these models are not empirically distinguishable from the unitary one. Cooperative models have been applied to the household agency problem, suggesting specific structures to explain household decisions like the Nash cooperative models (for instance in McElroy and Horney, 1981). This approach relied on ad hoc assumptions about threat points, which may be viewed as an external situation (utility of each spouse in case of divorce) or an internal equilibrium (e.g. from a non-cooperative game). The most encompassing framework has become the collective model (Chiappori, 1988, 1992), which assumes only the Pareto efficiency of household decisions while remaining agnostic about the underlying cooperative mechanism. The justification for efficiency is that household decisions can be seen as a repeated game with perfect information, which eventually lead to efficiency.[13] Hence, at least in the context of rich countries and for repeated decisions like consumption of non-durables, efficiency and the collective model have been retained as the leading applied theory.

Consider a collective model focusing on consumption and representing how household h spends total resources c on K types of good given a set of prices, a sharing rule (each member i receives a share $\eta_{i,h}$ of total resources) and scale economies (each good k is consumed according to a degree of publicness summarized by the Barten scale α_h^k):

$$\max_{q_{1,h}^1,...,q_{1,h}^K,...,q_{n_h,h}^1,...,q_{n_h,h}^K} \sum_{i=1}^{n_h} \rho_{i,h}(p^1,...,p^K,\zeta) \cdot u_i\left(q_{i,h}^1,...,q_{i,h}^K\right)$$

$$\text{s.t. } \eta_{i,h} c = \alpha_h^1 p^1 q_{i,h}^1 + ... + \alpha_h^K p^K q_{i,h}^K \text{ for } i =, 1..., n_h.$$

The maximization of a sum of individual utilities guarantees the efficiency of consumption decision. Pareto weights $\rho_{i,h}$ determine the balance of power in the household and depend on prices p and a set ζ of distribution factors.[14] Efficiency makes that this program can be represented as a two-stage budgeting problem: household decisions are as if the household chose individual shares $\eta_{i,h}$ for member $i = 1,...,n_h$, and then each member solved her/his own decentralized program. Using this decentralized 'sharing rule' interpretation, the early literature has focused on testing the efficiency assumption. Bourguignon et al. (2009) suggest a unifying framework to present the different ways to perform efficiency tests, notably tests based on nonlinearity in demand or on distribution factors.[15]

[13]It is of course possible to argue against these assumptions for at least three reasons: information needs not be symmetrical; some decisions are rarely repeated and lead to strategic behavior; people live in a dynamic world whereby credible commitments cannot always be enforced (see the discussion in Lundberg and Pollak, 2003). Dynamic versions of the collective consumption, the insurance role of the household as well as the representation of the household to be adopted in models with adverse shocks are discussed in the survey of Chiappori and Mazzocco (2016).

[14]The latter are household characteristics or environmental factors that influence the intrahousehold decision rule but do not directly affect preferences nor the budget set (Bourguignon et al., 1994).

[15]It turns out that tests on consumption and labor supply decisions in a static framework are rarely

Identification results have also been derived, but most of them allow retrieving only the "marginal" sharing rule $\partial\eta/\partial c$, i.e. how an additional dollar of income is shared between members (see surveys by Vermeulen, 2002, Chiappori and Donni, 2011, Browning et al., 2013).

The complete resource allocation between spouses has been recently identified in Browning et al. (2013), followed by the simplified approach of Lewbel and Pendakur (2009) and the extension to children in Bargain and Donni (2012) and Dunbar et al. (2014). These methods generalize the Rothbarth approach embedded in the collective framework: the identifying assumption in these contributions pertains to the existence of assignable consumption and the stability of individual preferences across demographic groups. This is best illustrated in the case of parents versus children, as in the Rothbarth application. Denote k_a the index of a good consumed only by adults and ignore price variation. Engel curves $q^{k_a} = \delta_{k_a} + \beta_{k_a} c$ can be estimated on a sample of adults living alone. The separability assumption and the stability of preferences guarantee that parameters δ_{k_a} and β_{k_a} can also be used for adult consumption when living with children. Hence, data on adults with children can be used to estimate Engel curves $q^{k_a} = \delta_{k_a} + \beta_{k_a}(\eta_a c)$, so that the share of resources accruing to adults (and, residually, to children) is directly obtained. The approach can equally be used to obtain the share of a specific adult. For instance, denote k_f the index of a good consumed only by adult females, then Engel curves $q^{k_f} = \delta_{k_f} + \beta_{k_f} c$ on single women and $q^{k_f} = \delta_{k_f} + \beta_{k_f}(\eta_f c)$ on women living in families can be jointly estimated to retrieve women's shares in families, η_f.

This schematic representation excludes economies of scale, but the logic can easily be extended as in Browning et al. (2013), who use many years of expenditure data and price variation to identify Barten scales α. Lewbel and Pendakur (2009) posit a single function representing the economies from joint consumption – the function $\sigma_{i,h}$ previously introduced – and assume it is independent of total expenditure ('independence of base'). With this assumption, they can identify both resource sharing and economies of scales without observing price variation, i.e. the demand system reduces to a mildly nonlinear system of Engel curves estimated on cross-sectional data. This very tractable approach can be easily implemented to recover indifference scales, as suggested above. More specifically, it may be used to infer intrahousehold resource allocation in microsimulation studies based on expenditure data. The main challenge is the availability of identifying goods. Typically, the recent literature has relied on clothing as a good assignable to either the children, the mother or the father (Bargain and Donni, 2012).

rejected. In contrast, production decisions in the context of poor regions lend themselves to rejection of efficiency (for instance in Udry, 1996).

Finally, the framework presented above has focused on consumption for simplicity. Consider the budget constraint of a couple, $c = c(y_f, y_m)$: it depends on female and male earnings respectively ($y_i = w_i h_i$ for $i = f, m$), transformed into disposable income by the microsimulation function c. Labor supply is more difficult to handle than consumption. First, it entails an additional wealth effect. Second, tax-benefit schedules are usually nonlinear. It is nonetheless possible to extend our framework to labor supply decisions by linearizing budget constraints, accounting for the fact that *net* wage rates $(1 - \tau_i)w_i$, with $\tau_i = 1 - \partial c(y_f, y_m)/l_i$, depend on labor supply, as characterized in Donni (2003). Collective models of labor supply in the presence of corner solutions and nonlinear budget constraints are also estimated in Moreau and Donni (2002) and Bloemen (2009), who focus on testable constraints and identification of the marginal sharing rule. Extending the complete sharing rule identification as in Browning et al. (2013) remains to be done in the case of labor supply and time allocation more generally.[16] Another issue is the lack of guidelines regarding how taxation should enter Pareto weights $\rho_{i,h}$ or the sharing rule $\eta_{i,h}$.

5.3 Applications and New Perspectives in Policy Analyses

With these developments, welfare analyses take a new perspective. Let me give a few examples. First, it becomes possible to decompose total inequality across *individuals* of a country in its within-household and between-household components. The within component can for instance tell us something about how the gender wage gap translates into unequal sharing of time and/or resources within households. Lise and Seitz (2010) estimate a collective model on British data over the long period using single and couple data for identification. They find that an increase in marital sorting explains both the rise in consumption inequality between households and the fall in inequality within households since the 1970s. In the recent period, they report that within-couple inequality contributes to $10 - 20\%$ of total inequality, depending on the dispersion measure and model specification.

A second example is the analysis of poverty at the individual level, notably child poverty. The approach of Lewbel and Pendakur, in particular, has been extended to resource allocation between parents and children in studies by Bargain and Donni (2012), Bargain et al. (2015) and Dunbar et al. (2013). In the context of Côte d'Ivoire and Malawi, the two latter studies show whether all household members are equally poor or whether some individuals (e.g., the children) are disproportionately poor. They suggest original poverty

[16] Recent advances suggest to model economies of scale in time allocation using time use survey. In particular, Couprie and Ferrand (2015) model scale economies in time use following Browning et al. (2013).

analyses by computing a direct measure of individual poverty, whereby poor persons are poor because the resources they receive in the household are below some poverty line. This is innovative compared to usual measures where a poor child is simply defined as one living in a poor household. Bargain et al. (2015) also find that adults in couples, who are apparently poorer than singles using traditional per capita expenditures, are in fact compensated by the gains from economies of scale. Similar studies can be performed in rich countries with relatively large rates of child poverty.

Finally, new types of question can be asked when evaluating the redistributive effect of social and fiscal policies. At present, welfarist analyses of optimal policy designs, and above all the vast literature on optimal taxation, only consider the distribution of welfare across households. Yet, a social welfare function will generally be misspecified, except in the unlikely event that a household puts relative weights on its members in accordance with the redistributive views of the social planner (Apps and Rees, 1988). From this, one of the many new questions pertains to the implication of basing redistributive policies on the household. For instance, if rich households are also those with more equal internal sharing, Pigou-Dalton transfers necessarily reduce between-household inequality but mechanically increase within-household inequality. This question has received some attention both theoretically and when using a collective model to assess how within-household inequality varies along the income distribution (Couprie et al., 2010).

5.4 Implementation and the Bargaining Effect of Policy Reforms

How can model estimations and tax-benefit microsimulations be combined in order to simulate the welfare impact of policies in a way that benefits from these advances? How to use collective models to study the intrahousehold redistribution incurred by policy reforms? To investigate these questions, it is useful to distinguish the different types of effects taxation or redistribution may have on individual welfare. The first one is a change in c, the total household resources net of taxes and transfers, as predicted by microsimulation models. If one rules out behavioral responses, sharing rules estimated prior to a reform can be directly applied to the new resources available to the household to calculate the change in individual allocations. The second effect pertains to household responses following a change in prices or net wages that affect the budget constraint. This is the same effect as previously analyzed with a unitary model in section 3. Microsimulation is particularly useful in this case since it accurately defines the impact of the reform on household work or consumption incentives. Redistribution can take place within households following the standard price and income effects of a policy reform, even with Pareto weights kept constant, simply because spouses may not be equally rich before the reform and hence be located on different parts of their utility function of income. The third effect is precisely the within-household distributional effect of reforms, i.e. one that corresponds to the shift in Pareto weights ρ, and hence the sharing rule η, following the changes in prices, wages or distribution factors following the reform.

This last effect is specific to the collective model but might be difficult to specify. The collective consumption model presented above can readily be used to analyze VAT reforms and their potential effect on individual welfare. This is the case with the approach of Browning et al. (2013), which relies on price variation for complete identification (Barten scale and sharing rule). In a decentralized version of the model, the sharing rule $\eta_{i,h}$ can be specified as a function of (net of tax) prices, among other things. Hence, it is possible to use the estimated parameters to predict the welfare impact of a differentiated change in VAT across types of good. The model would fully account for the usual substitution and income effects on expenditure patterns as well as this new bargaining effect. In particular, the framework could be extended to the allocation of resources between children and parents to study the impact of actual or hypothetical VAT reforms on child poverty (e.g. subsidies on child goods) and empirically compare their effect to other forms of redistribution (e.g. a change in child benefit).

Consider now the effect of taxes and transfers that affect nonlabor income separately from labor income. If Pareto weights ρ can depend on individual levels of unearned income, it is possible to model the effect of who brings nonlabor money to the household. This is important since it is shown to significantly influence consumption patterns, as in the "wallet to purse" reform that changed the identity of the spouse receiving family benefits in the UK in the 1970s (Lundberg et al., 1997). This was a rejection of the (unitary) assumption that household members pool their resources. It is possible to directly introduce levels of nonlabor income controlled by specific family members in the sharing rule $\eta_{i,h}$ of a collective consumption or labor supply model identified à la Browning et al. (2013). Yet, nonlabor income is not only the amount of benefits to which specific members are entitled to, but also some capital income. This poses two issues for modelling choices. First, capital income results from past individual and collective decisions about consumption, saving and investment, which is much less exogenous than prices. Second, the taxation of income may not be separable between labor and nonlabor income, so that changes in capital income may affect labor supply.

Things get even more complicated when focusing on the direct taxation of earnings in collective labor supply models. Take the simple introduction of wage rates w_f and w_m in the sharing rule (for instance in the case of no taxation or linear taxation). This is itself an empirical problem: gross wage rates are possibly endogenous to past and present household decisions.[17] Then, when tax schedules are nonlinear, the sharing rule may depend on net wages $(1-\tau_i)w_i$ and becomes endogenous to actual labor supply choices, as seen above. Donni (2003) suggest ways to identify the sharing rule at the margin in this case. Yet it is not clear how his results transpose to the more general framework à la Browning et al. (2013) with labor supply and nonlinear taxation. Finally, one would like to go beyond wage rates and model the effect of individual earnings y_f and y_m on the sharing function. It is possible to make $\eta_{i,h}$ a function of individual labor incomes in a pure consumption model, but only by positing separability between consumption and labor decisions (and facing the same difficulties as emphasized above on the endogeneity to past household decisions). Moreover, it is clearly not possible to do so in labor supply models, since efficiency would collapse in this case (as in Basu, 2006). Several attempts have been made to circumvent this difficulty, notably using the relative financial contribution of each spouse at a fixed labor supply level (for instance at full time). This is the approach followed in Laisney (eds, 2006): a collective model of labor supply is calibrated on single and couple data to retrieve individual utility functions of men and women and the value of a power index respectively. The power index is then regressed on relevant bargaining factors including a set of variables retracing the spouses' *potential*, rather than actual, relative contributions to household disposable income. This approach is fairly ad hoc and illustrates the fact that models relying solely on the efficiency assumption are 'semi-structural' in the sense that they cannot explain the form of the sharing rule – we lack guidance concerning the way taxation could affect negotiation. Cooperative models in which outside options are specified help to better characterize the role of tax policies but are also ad hoc regarding the choice of a given threat point among many (Bargain and Moreau, 2012).

6 Concluding Remarks

Tax-benefit microsimulation studies and applied welfare economics coexist but do not cross-fertilize much. Policy analyses that rely on household equivalized income can benefit from research in welfare measurement and/or household modelling. New developments could indeed help to provide much richer analyses of how redistributive systems shape individual welfare. Some improvements are already on the agenda – for instance, accounting for behavioral responses to policy reforms. But much remains to be done, including a more extensive use of money metrics, the estimation of sharing rules to shift the analysis

[17] We have seen the effect of confounding factors on both wage and work preferences which applies here as much as in the unitary model. Omitted variables may also affect both wages and household preferences (i.e. the decision process). For instance, low wages may reflect past decisions (inactivity and human capital depreciation) and hence the past balance of power, that may affect the current one.

to the individual rather than the household, the estimation of more reliable and realistic equivalizing functions drawing from the indifference scale concept. However, some advances cannot readily be operationalized. In particular, collective labor supply requires more research effort before implementation. Furthermore, complexity may quickly compound, for instance if one wants to derive money metric measures in a multidimensional context while doing so at the individual level. This will necessarily lead to interesting questions that are already on the research agenda – for instance, how sharing among spouses depends on fairness principles applied by spouses themselves.

Finally, let me summarize examples of how microsimulation, in return, could support applied welfare research. First, tax-benefit simulation of existing systems may help to better characterize redistributive views embedded in existing policies, either regarding vertical equity (see for instance Bourguignon and Spadaro, 2012) or horizontal equity (e.g. implicit equivalent scales embedded in tax-benefit systems). Then, there is a striking imbalance between how little we know about individual/household preferences versus how precisely we can characterize household financial incentives. Indeed, it is hard to capture much heterogeneity in preferences beyond the basic variables provided in microdata (like age, education, etc.). In contrast, microsimulation studies provide very detailed and accurate evaluation of real-world budget sets. This information could be better used by researchers when trying to understand and model behavior. For instance, this is useful for natural experiments based on nonlinearities and discontinuities in budget constraints (cf. Saez et al., 2012, Kleven, 2016), which may themselves help to identify behavioral models. It can also provide thin variation in net-of-tax prices or wages across households to be used in revealed preferences approaches to elicit bounds of individual welfare functions (for instance to extend Cherchye et al., 2012).

References

[1] Aaberge R., Dagsvik J., Strøm S. (1995): "Labor supply responses and welfare effects of tax reforms", *Scandinavian Journal of Economics* 97(4):635–659

[2] Aaberge, R., Colombino, U. and Strøm, S. (2004): "Do more equal slices shrink the cake? An empirical investigation of tax-transfer reform proposals in Italy", *Journal of Population Economics* 17, 767–785

[3] Aaberge, R., and Colombino, U. (2015): "Labour Supply", Chap. 5, *Handbook of Microsimulation Modelling* (C. O'Donoghue editor)

[4] Alderman H, P.A. Chiappori, L. Haddad, J. Hoddinott and R. Kanbur (1995): "Unitary versus Collective Models of the Household: Is It Time to Shift the

Burden of the Proof", *World Bank Research Observer*, 10, 1–19

[5] Allingham, M.G. and A. Sandmo (1972): "Income tax evasion : A theoretical analysis", *Journal of Public Economics*, 1(3-4), 323-338

[6] Anderson, P. M. and B. D. Meyer (1997): "Unemployment insurance takeup rates and the after-tax value of benefits", *Quarterly Journal of Economics* 112, 913–937

[7] Apps P.F. and R. Rees (1988): "Taxation and the household", *Journal of Public Economics*, 35, 355–369

[8] Atkinson, A.B. (1970): "On the measurement of inequality", *Journal of Economic Theory*, 2, 244-263

[9] Atkinson, A.B. and F. Bourguignon (1987): "Income distribution and differences in needs", Chapter 12 in G.R. Feiwel (ed.) Arrow and the Foundations of the Theory of Economic Policy. London: Macmillan.

[10] Atkinson, A. B. (1997): "Bringing income distribution in from the cold", *Economic Journal* 107(441), 297–321

[11] Atkinson, A.B. (2005): "EUROMOD and the development of EU Social Policy", EUROMOD working paper 1/05.

[12] Atkinson, A., F. Bourguignon, P.A. Chiappori (1988): "What Do We Learn About Tax Reforms from International Comparisons? France and Britain", *European Economic Review*, 32, 32, 343-52

[13] Attanasio, O., C. Meghir and A. Santiago (2011): "Education choices in Mexico: Using a structural model and a randomized experiment to evaluate PROGRESA", *Review of Economic Studies*.

[14] Banks, J., Blundell, R. and Lewbel, A. (1996): "Tax reform and welfare measurement: do we need demand system estimation?", Economic Journal, 106(438), 127-124

[15] Bargain, O. (2015): "Decomposing Inequality and Poverty Trends and Quantifying the Role of Tax-Benefit Policy Reforms", Chap. 5, *Handbook of Microsimulation Modelling* (C. O'Donoghue editor)

[16] Bargain, O. and N. Moreau (2007): "Does taxation affect intrahousehold distribution? A Simulation Approach", *Research on Economic Inequality* 15

[17] Bargain, O. and Donni, O. (2012): "The Measurement of Child Costs: A Rothbarth-Type Method Consistent with Scale Economies and Parents Bargaining". *European Economic Review*, 56(4), 792-813

[18] Bargain, O., and A. Peichl (2016): "Steady-State Labor Supply Elasticities: An International Comparison", *IZA Journal of Labor Economics*, 5,10, 1-31

[19] Bargain, O. and K. Doorley (2017): "The Effect of Social Benefits on Youth Employment: Combining RD and a Behavioral Model", forthcoming in *Journal of Human Resources*

[20] Bargain, O., A. Decoster, M. Dolls, A. Peichl, D. Neumann, S. Siegloch (2013): "Welfare, heterogeneity and responsibility: Evidence from Europe", *Social Choice and Welfare*, 41(4), 789-817

[21] Bargain, O., K. Orsini, and A. Peichl (2014): "Comparing Labor Supply Elasticities in Europe and the US: New Results", *Journal of Human Resources* 49(3), 723-838

[22] Bargain, O., O. Donni, and P. Kwenda (2015):. "Intrahousehold distribution and poverty: Evidence from Côte d'Ivoire," *Journal of Development Economics*, 107(C), 262-276

[23] Basu K. (2006): ""Gender and Say: A Model of Household Behavior with Endogenously-determined Balance of Power", *Economic Journal*, 116, 558-580

[24] Benedek, D. and Lelkes, O. (2011): "The distributional implications of income under-reporting in Hungary", *Fiscal Studies*, 32 (4): 539–60

[25] Bhargava, S. and D. Manoli (2015): "Psychological frictions and the incomplete take-up of social benefits : Evidence from an IRS field experiment", *American Economic Review* 105, 3489–3529

[26] Blackorby, C. and Donaldson, D. (1993): "Adult-equivalence scales and the economic implementation of interpersonal comparisons of well-being", *Social Choice and Welfare* 10, 335–361

[27] Bloemen, H. (2009): "An Empirical Model of Collective Household Labour Supply with Nonparticipation", *Economic Journal*, 120, 543, 183-214

[28] Blundell, R. (2012) "Tax policy reform: The role of empirical evidence", *Journal of the European Economic Association*, 10, 43–77.

[29] Blundell, R. and McCurdy T. (1999): "Labor Supply: a Review of Alternative Approaches", in: Ashenfelter O. and Card D. (eds.), *Handbook of Labour Economics*, Vol. 3A., Elsevier Science Publishers

[30] Blundell R., Duncan A., McCrae J., Meghir C. (2001): "The labour market impact of the working families' tax credit", *Fiscal Studies* 21(1):75–104

[31] Bosmans, K., K. Decancq and E. Ooghe (2017): "Who's afraid of aggregating money metrics?", mimeo

[32] Bourguignon, F. and A. Spadaro (2006): "Microsimulation as a tool for evaluating redistribution policies", *Journal of Economic Inequality,* 4:2, 77-106

[33] Bourguignon, F. and A. Spadaro (2012): "Tax–benefit revealed social prefences", *Journal of Economic Inequality,* 10:75–108

[34] Bourguignon, F., Browning, M., Chiappori, P.-A., and Lechene, V. (1994). "Income and Outcomes: A Structural Model of Intrahousehold Allocation", *Journal of Political Economy*, 102(6), 1067-1096.

[35] Bourguignon, F., M. Browning and PA Chiappori (2009): "Efficient Intra-Household Allocations and Distribution Factors: Implications and Identification," *Review of Economic Studies*, 76(2), 503-528.

[36] Bourguignon, F., Ferreira, F. and Lustig, N. (2004). *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*, World Bank, Washington, DC

[37] Brewer, M., Duncan, A., Shephard, A., and Suarez, M.-J. (2005), "Did the WFTC work? Analysing the impact of in-work benefits on labour supply and programme participation", *Labour Economics*

[38] Brewer, M. and L. Wren-Lewis (2016): "Accounting for Changes in Income Inequality: Decomposition Analyses for the UK, 1978–2008", *Oxford Bulletin of Economics and Statistics*, 78 (3), 289–322

[39] Browning, M., P-A. Chiappori and A. Lewbel (2013): "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power", *Review of Economic Studies*, 80(4), 1267-1303

[40] Burns, SK. and J.P. Ziliak (2016): "Identifying the Elasticity of Taxable Income" *Economic Journal*, Doi: 10.1111/ecoj.12299.

[41] Ceriani L., Fiorio C., and C. Gigliarano (2013): "The importance of choosing the data set for tax-benefit analysis", *International Journal of Microsimulation* 6(1): 86-121

[42] Chambaz, C. and E. Maurin (1998): "Atkinson and Bourguignon's dominance criteria: extended and applied to the measurement of poverty in France", *Review of Income and Wealth*, 44, 497-513

[43] Cherchye, L., B. De Rock, F. Vermeulen (2012): "Married with Children: A Collective Labor Supply Model with Detailed Time Use and Intrahousehold Expenditure Information", *American Economic Review*, 102 (17), 3377-3405

[44] Chetty, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 1(1), 451-488

[45] Chiappori, P.-A. (1988): "Rational Household Labor Supply", *Econometrica*, 56, 63-89.

[46] Chiappori, P.-A. (1992): "Collective Labor Supply and Welfare", *Journal of Political Economy*, 100, 437-467.

[47] Chiappori, P.A. and C. Meghir (2014): "Intra-household Welfare", NBER Working Paper No. 20189

[48] Chiappori, P-A. and O. Donni (2011): "Non-Unitary Models of Household Behavior: A Survey of the Literature". In: A. Molina (eds), *Household Economic Behaviors*, Berlin: Springer.

[49] Chiappori, P-A. (2016): "Equivalence versus Indifference Scales", (Lecture at the Royal Economic Society conference, 2015), forthcoming in the *Economic Journal*

[50] Chiappori, P.-A. and Mazzocco, M. (2016): "Static and Intertemporal Household Decisions", forthcoming in the *Journal of Economic Literature*

[51] Copenhagen Economics (2007): "Study on reduced VAT applied to goods and services in the member states of the European Union", Research Report on behalf of the EU Commission, 6503 DG TAXUD

[52] Couprie, H., Peluso, E. and A. Trannoy (2010): "Is power more evenly balanced in poor households?", *Journal of Public Economics*, 94, 7-8, 493-507

[53] Couprie, H. and G. Ferrant (2015): "Welfare Comparisons, Economies of Scales and Equivalence Scale in Time Use", *Annales of Economics and Statistics*, 117-118. 185-210

[54] Cowell, F. (1985): "Tax evasion with labor income", *Journal of Public Economics* 26, 19–35

[55] Currie, J. (2006). The Take-up of Social Benefits. In Public Policy and the Income Distribution (New York : Russell Sage Foundation ed.)., pp. 80–148. Alan J. Auerbach, David Card, and John M.Quigley.

[56] Decancq, K., M. Fleurbaey and E. Schokkaert (2015): “Happiness, equivalent incomes, and respect for individual preferences”, *Economica*, 82, 1082-1106

[57] Decancq, K., M. Fleurbaey & E. Schokkaert (2015): "Inequality, income, and well-being". in: Anthony B. Atkinson & François Bourguignon (eds.) Handbook on Income Distribution, Volume 2A. Elsevier. 67-140

[58] Decoster A, Verwerft D, Loughrey J and O'Donoghue C (2011): "Microsimulation of indirect taxes", International Journal of Microsimulation, 4(2), 41–56

[59] De Hoyos, R. (2012): "Female Labour Participation and Occupation Decisions in Post-NAFTA Mexico", *Research in Labor Economics*, volume 33, Emerald Group Publishing Limited, 85-127

[60] Donni O. and N. Moreau (2007): "Collective Labor Supply: A Single-equation Model and Some Evidence from French Data", *Journal of Human Resources*, 42, 214-246.

[61] Donni O. (2003): "Collective Household Labor Supply: Non-participation and Income Taxation", *Journal of Public Economics*, 87, 1179–1198.

[62] Dörrenberg, P., A.Peichl and S. Siegloch (2016): "The Elasticity of Taxable Income in the Presence of Deduction Possibilities", forthcoming *Journal of Public Economics*

[63] Duclos, J.-Y. (1995): "Modelling the take-up of state support", *Journal of Public Economics* 58, 391–415

[64] Dunbar, G., A. Lewbel and K. Pendakur (2013): "Children's Resources in Collective Households: Identification, Estimation and an Application to Child Poverty in Malawi", *American Economic Review*, 103, 438-471.

[65] Ebert, U. (1997): "Social welfare when needs differ: an axiomatic approach", *Economica*, 64, 233-244

[66] Feldman, N. and Slemrod, J. (2007): "Estimating tax noncompliance with evidence from unaudited tax returns", *Economic Journal*, 117(518):327–352

[67] Feldstein, M. (1999): "Tax avoidance and the deadweight loss of the income tax", *Review of Economics and Statistics*, 81(4), 674-680

[68] Figari, F., A. Paulus, and H. Sutherland (2017): "Microsimulation and policy analysis", *Handbook of Income Distribution*: Volume 2B, Ch. 24 - 2015

[69] Fleurbaey, M. (2008), *Fairness, Responsability, and Welfare*, Oxford University Press: Oxford, UK.

[70] Fleurbaey, M. (2009): "Beyond GDP: The Quest for a Measure of Social Welfare", *Journal of Economic Literature* 47: 1029–1075

[71] Fleurbaey, M., and Maniquet, F. (2006): "Fair income tax", *Review of Economic Studies* 73(1), 55-83

[72] Fleurbaey, M., & Blanchet, D. (2013), *Measuring Welfare and Assessing Sustainability*, Oxford University Press, USA

[73] Frederiksen, A., E. K. Graversen, and N. Smith (2005): "Tax evasion and work in the underground sector", *Labour Economics* 12, 613–628

[74] Gaertner W. and E. Schokkaert E (2012), *Empirical Social Choice: Questionnaire-experimental Studies on Distributive Justice*, Cambridge University Press, Cambridge

[75] Gong, X. and A. van Soest (2002): "Family Structure and female Labor Supply in Mexico City", *Journal of Human Resources*, 37(1), 163-191

[76] Haddad L. and R. Kanbur (1992): "Is There an Intrahousehold Kuznets Curve? Some Evidence from the Philippines", *Public Finance*, 47 (supplement), 77–93.

[77] Herault, N. and Azpitarte, F. (2014). 'Understanding changes in the distribution and redistribution of income: a unifying decomposition framework', Review of Income and Wealth. doi: 10.1111/roiw.12160

[78] Hernandez, M. and S. Pudney (2007): "Measurement error in models of welfare participation", *Journal of Public Economics* 91, 327–341.

[79] Immervoll, H. (2005): "Falling up the stairs. An exploration of the effects of "bracket creep" on household incomes", *Review of Income and Wealth*, 51 (1), pp. 37-62.

[80] Immervoll, H., H.J. Kleven, C.T. Kreiner, E. Saez (2007): "Welfare Reform in European Countries: A Microsimulation Analysis", Economic Journal, 118 (516), 1-44

[81] Immervoll, H., H.J. Kleven, C.T. Kreiner, N. Verdelin (2011): "Optimal tax and transfer programs for couples with extensive labor supply responses", *Journal of Public Economics*, 95 (11-12), 1485-1500

[82] Jara, X.H. and E. Schokkaert (2016): "Putting subjective well-being to use for ex-ante policy evaluation," KU Leuven.

[83] Kahneman, D., P. Wakker, & R. Sarin (1997): "Back to Bentham? Explorations of Experienced Utility", *Quarterly Journal of Economics*, 112(2): 375-405

[84] Kleven, H., M. Knudsen, C. Kreiner, S. Pedersen, and E. Saez (2011): "Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark", *Econometrica* 79 (3), 651-692

[85] Kleven, H. J. (2016): "Bunching", *Annual Review of Economics*, 8(1).

[86] Kolm, S. C. (1969): "The Optimal Production of Social Justice", in J. Margolis and H. Guitton (eds.), Public Economics. Macmillan, London

[87] Lacroix, G. and B. Fortin (1992): "Utility-Based Estimation of Labour Supply Functions in the Regular and Irregular Sectors", *Economic Journal*, Vol. 102, No. 415, pp. 1407-1422

[88] Laisney F. (2006 ed.): "Welfare Analysis of Fiscal and Social Security Reforms in Europe: Does the Representation of Family Decision Process Matter?", *Review of Economics of the Household*, 4, 99-180

[89] LaLumia S. and J. Sallee (2013): "The value of honesty: empirical estimates from the case of the missing children," *International Tax and Public Finance*, 20(2), 192-224

[90] Lambert, P.J, and X. Ramos (2002): "Welfare Comparisons: Sequential Procedures for Heterogeneous Populations", *Economica*, 69, 276, 549–562

[91] Landais, C. and G. Fack (2016): "The Effect of Tax Enforcement on Tax Elasticities: Evidence from Charitable Contributions in France", forthcoming in *Journal of Public Economics*

[92] Lemieux, T., Fortin, B., Fréchette, P., 1994. The effect of taxes on labour supply in the underground economy. *American Economic Review* 84 (1), 231– 254

[93] Lewbel, A. (2003): "Calculating Compensation in Cases of Wrongful Death," *Journal of Econometrics*, 113, 115-128.

[94] Lewbel, A. and K. Pendakur (2008): "Estimation of Collective Household Models with Engel Curves", *Journal of Econometrics*, 148, 350-358.

[95] Lise, J. and S. Seitz (2011): "Consumption Inequality and Intra-Household Allocations," *Review of Economic Studies*

[96] Low, H. and C. Meghir (2017): "The Use of Structural Models in Econometrics" *Journal of Economic Perspectives*, 31/2, 33–58

[97] Lundberg S. and R.A. Pollak (2003): "Efficiency in Marriage", *Review of Economics of the Household*, 1, 153–167

[98] Lyssiotou P., Pashardes P. and Stengos T. (2004): "Estimates of the black economy based on consumer demand approaches", *Economic Journal* 114 622-40

[99] Matsaganis M., Leventi, C. & Flevotomou, M. (2012): "The crisis and tax evasion in Greece: what are the distributional implications?", *CESifo Forum*, 13 (2): 26-32

[100] Matsaganis M., Levy H., and M. Flevotomou (2010): "Non take up of social benefits in Greece and Spain", *Social Policy & Administration* 44(7): 827-844

[101] McElroy M.B. and M.J. Horney (1981): "Nash-Bargained Decisions: Toward a Generalisation of the Theory of Demand", *International Economic Review*, 22, 333-349.

[102] Meghir, C. and Phillips, D. (2008): "Labour supply and taxes", IZA Discussion Paper No. 3405.

[103] Moffitt, R. (1983): "An economic model of welfare stigma", *American Economic Review*, 73, 5, 1023-1035

[104] Moreau N. and O. Donni O. (2002): "Un modèle collectif d'offre de travail avec taxation", Annales d'économie et de statistique, 65, 55–81.

[105] Pollak, R. and T. J. Wales (1979): "Welfare Comparisons and Equivalence Scales", *American Economic Review*, 69, 216–221.

[106] Pissarides C.A. and Weber G. (1989): "An expenditure-based estimate of Britain's black economy", *Journal of Public Economics* 39 17-32

[107] Pudney, S., M. Hernandez, and R. Hancock (2007): "The welfare cost of means-testing : pensioner participation in income support", *Journal of Applied Econometrics* 22, 581–598

[108] Riphahn, R.T. (2001): "Rational Poverty or Poor Rationality? The Take-Up of Social Assistance Benefits", *Review of Income and Wealth*, 47(3), 379-98

[109] Rothbarth, E. (1943): "Note on a method of determining equivalent income for families of different composition", in C. Madge (ed) *War-time Pattern of Saving and Spending*, Cambridge University Press, Cambridge.

[110] Saez, E., J. Slemrod, and S. H. Giertz (2012): "The elasticity of taxable income with respect to marginal tax rates: A critical review", *Journal of Economic Literature* 50 (1), 3-50

[111] Schneider, F., and Enste, D.H. (2000): "Shadow economies: Size, causes, and consequences". *Journal of Economic Literature* 38, 77–114

[112] Sen, A. (1979): "The welfare basis of real income comparisons: A survey" *Journal of Economic Literature* 17: 1–45

[113] Sen, A. (1985), *Commodities and Capabilities*. Amsterdam: North-Holland

[114] Shorrocks, A. F. (1983): "

[115] Shorrocks, A. F. (2013): "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value", *Journal of Economic Inequality*, 11(1), 99-126

[116] Slemrod, J. and C. Weber (2012): "Evidence of the Invisible: Toward a Credibility Revolution in the Empirical Analysis of Tax Evasion and the Informal Economy," *International Tax and Public Finance*, 19, 1, 25–53

[117] Stiglitz, J., Sen, A. and Fitoussi, J.-P. (2009). Report by the Comission on the Measurement of Economic Performance and Social Progress, Technical Report.

[118] Sutherland, H. (2015): "Multi-country microsimulation", Chap. 4, *Handbook of Microsimulation Modelling* (C. O'Donoghue editor)

[119] Terracol, A. (2002): "Analysing the Take-Up of Means-Tested Benefits in France", mimeo, Université Paris I-Panthéon-Sorbonne

[120] Thoresen, T.O. and T.E. Vattø (2015): "Validation of the discrete choice labor supply model by methods of the new tax responsiveness literature, *Labour Economics*, 37, 38–53

[121] Udry C. (1996): "Gender, Agricultural Production, and the Theory of the Household", *Journal of Political Economy*, 104, 1010–1046.

[122] van Soest A. (1995): "Structural models of family labor supply: a discrete choice approach", *Journal of Human Resources*, 30(1):63–88

[123] Vermeulen, F. (2002): "Collective Household Models: Principles and Main Results", *Journal of Economic Surveys*, 16(4): 533-564