

Supplementing GDP: Some Recent Contributions from Official Social Statistics

Jérôme Accardo*

Abstract – In its efforts to take the criticisms levelled against GDP into consideration, national accounting benefits from the contribution made by social statistics. In the last decade, it has developed major innovations, such as accounts by household category, and it has provided other useful empirical and methodological tools to address the problem of evaluating domestic production and of taking into account non-monetary dimensions. It has also been very active in the new field of “the economics of happiness” (with the measurement of subjective well-being). This paper offers a critical discussion of the work carried out in these areas by official statistics.

JEL Classification: E01, I31

Keywords: accounts by household category, social indicators, subjective well-being

Reminder: The opinions and analyses in this article are those of the author(s) and do not necessarily reflect their institution's or Insee's views.

* Insee (jerome.accardo@insee.fr)

This article benefited from comments provided by several readers and two anonymous reviewers. I would particularly like to thank Didier Blanchet, Fabrice Lengart and André Vanoli for their criticisms and suggestions on an initial version of the text. I am solely responsible for any errors that may remain.

Received 28 June 2018, accepted after revisions on 15 April 2019.

Translated from: “Compléter le PIB : quelques contributions récentes de la statistique sociale”

Citation: Accardo, J. (2020). Supplementing GDP: Some Recent Contributions from Official Social Statistics. *Economie et Statistique / Economics and Statistics*, 517-518-519, 25–39. <https://doi.org/10.24187/ecostat.2020.517t.2016>

As an indicator of the level of socio-economic development of a country, GDP has many limitations, criticism of which dates back almost as far as the indicator itself (Vanoli, 2002, Chap. 7). The main criticisms essentially focus on three issues: *i*) aggregation: as an aggregate indicator, GDP is not able to reflect phenomena associated with the distribution of flows or stocks among economic units; *ii*) scope: it only takes into account certain transactions, while excluding others, even though their economic nature and significance seem indisputable; *iii*) relevance: it raises very complex questions when the figure is interpreted in terms of social well-being; in turn, these questions raise the further question of its ability to inform public policy.

The report by Stiglitz, Sen and Fitoussi (2009) provided both a summary of these analyses and new impetus for research aimed at going “beyond GDP”.

The aim of this paper is to present, among the recent developments made in this area by official statistics, the contribution of social statistics, understood as the collection of data describing the living conditions of individuals at the microeconomic level. The article is not intended as a survey of these questions or to cover all social statistics. It attempts only to provide a detailed presentation of some of the current approaches adopted by national statistical institutes to address the three categories of criticism by explaining the objectives identified, the methods and sources used and the obstacles encountered. Most of the practical illustrations are drawn from work carried out by Insee, with which the author is quite familiar.

Thus, this article will examine, in turn, the work devoted by social statistics over the last decade to taking into account the distribution of household account transactions, the domestic production of services, to incorporate non-monetary dimensions (health status, quality of the natural environment, security, social capital, etc.); and to the direct measurement of well-being (“happiness”, “life satisfaction”, etc.).

The first part seeks to address the criticism of the aggregate nature of GDP. The second applies to the criticism relating to scope. The last two address issues of interpretability.

1. Beyond the Analysis of Aggregates Alone: Distribution and Decomposition of Accounts by Household Category

In what precise way can the national accounts follow the recommendation of the Stiglitz-Sen-Fitoussi report to “take distributions into account” as far as the household account is concerned? As is well known, national accounts use only aggregates to describe the various transactions in the household account. However, thanks to the ever-growing number of microeconomic sources, the distribution of the corresponding variables is often known. Is it not possible to use these distributions to produce “distributional accounts”? That is, accounts which – in addition to establishing a set of monetary aggregates made consistent, according to internationally standardised methods – could also show their distribution across the population, thereby enabling the accounts to be interpreted in terms of inequality. The idea is particularly appealing: including the analysis of inequalities (inequalities in income, consumption and savings, a redistributive assessment of the tax and benefit systems, effect of public policies, etc.) within the theoretical framework of national accounts guarantees the consistency of the analysis and its exhaustive nature. It also appears very early on in academic thinking concerning National accounts (see Online Appendix C1¹). We start below by explaining why the complete fulfilment of this objective remains unattainable, then we present the solutions that have nonetheless been explored by national accountants with a view to integrating differences between households into the accounting framework.

1.1. The Practical Impossibility of “Distributional Accounts”

To facilitate the exposition of the problem, a minimal version can be considered, with a household account limited to three (aggregated) transactions: gross (adjusted) income, actual final consumption and savings. Is it possible to create a distributional version of such an account – in other words, to produce this account for each household in a representative sample? Currently, the answer is no.

The obstacle lies in the fact that knowledge of the respective distributions of two variables in the population (in this case, income and consumption) does not make it possible to

1. [Link to the Online appendices at the end of the article.](#)

calculate the distribution of their sum or of their difference. It is necessary to know their joint distribution. In other words, it is necessary to have the following information available, for at least each household of a representative sample:

- the income of the household members (income from employment, including social taxes, replacement income, social benefits, capital income, etc.);
- transfers in kind from public health bodies that can be individualised (hospital care, reimbursements for healthcare, etc.);
- the educational situation of the household members (level of education, course of study, etc.) to allow the calculation of the public education expenditure from which they directly benefited during the year (based on micro-economic data from the National Education department on costs per pupil according to the type of education);
- a sufficiently precise description of that residence for the households in the sample that own their main residence, so as to be able to apply a satisfactory imputed rent model;
- the total consumption of market services and goods of each household.

All of this represents a great deal of information and collecting it all directly in a single survey would constitute an excessive burden on respondents. The information does exist, however, scattered across household surveys on one side – the Labour Force Survey (*Enquête Emploi en Continu*, or EEC), the European Union Statistics on Income and Living Conditions (EU-SILC), the National Housing Survey (*Enquête Nationale Logement*, or ENL), the Household Income and Expenditure Survey (*Enquête Budget de Famille*, or BDF), the Health Survey (*Enquête Santé*), etc. – and administrative files on the other side – income tax and housing tax files, files on claimants from social security organisations, Annual Declaration of Social Data (*Déclarations Annuelles de Données Sociales*, or DADS), health insurance files, etc.

If all these sources were matched together, the objective would be achieved: we would then have the adjusted gross disposable income (AGDI) and the actual final consumption, and therefore the saving rate², of each household in the sample. Such matching is currently only partially feasible, its full implementation being currently hampered by legal obstacles, which

themselves reflect the political and philosophical problems raised by this type of “panoptic” project.

Work in progress in France, in a rapidly changing legal context (digital act and act on health data), certainly allows us to hope for significant progress in the coming years. Nevertheless, it should be stressed that these obstacles exist in most countries and, from this point of view, the situation concerning French official statistics is quite favourable compared to that of many countries in which the possibilities of matching survey data with administrative sources are much more limited. However, were distributional analyses possible in only a small number of countries, this “improved” accounting would lack international comparability, which is one of the major strengths of the system of national accounts.

1.2. The Principle of Decomposition of the Household Account

In the absence of broad matching allowing the creation of a complete account for each household in a representative sample, the solution lies in statistical imputation: this method (which is often referred to as “bottom-up”) consists of selecting a survey in which the information collected, at household level, on the account’s transactions is as extensive as possible. This information is then supplemented by imputing a value for each missing transaction for each household. Imputation is performed using models estimated in the other available sources.

One possible way to complete this process, in practice, is as follows. The primary source is the BDF survey: it provides an estimation of the annual consumption (at a fine level of the product nomenclature) and annual income of each household in the sample, obtained by matching with socio-fiscal administrative sources. This income represents only part of the income taken into consideration by the national accounts. Therefore, social security contributions, resources in kind, imputed rents, etc. must be added to it. These additions are obtained by applying the value predicted by models estimated based on the Tax and Social Incomes Survey (*Enquête revenus fiscaux et sociaux*, or ERFS) and the EU-SILC for social security contributions and income from self-employment

² This would, of course, be a saving rate that is not rigorously dated, as several sources are only available on a multi-year basis and their availability is not synchronised.

and on the ENL survey for the imputed rents to each household. The same is carried out using models relating to public health and education expenditure to obtain an adjusted income at the household level. The crucial point here is that the explanatory variables in the different models used are also collected in the BDF survey. This condition is clearly essential for applying these models to each household in the BDF survey to estimate the most likely value (given the characteristics taken into account) of the missing components.

The end result is a representative sample in which, formally, resources and expenditure (and thus savings) are fully known for each household in the sample.

However, this file does not allow distributional analysis of the accounts. Indeed, if the sample is large enough, the distribution in the file of a given account transaction, whether expenditure or resources, adequately represents its actual distribution. In contrast, the imputation procedure does not make it possible to ascertain the true joint distribution of the various account transactions, only the joint distribution that is conditional on the explanatory variables used in the models. This makes it impossible to determine the distribution of sums and balances, starting with the distribution of savings.³ The measures of inequality in adjusted gross disposable income or in actual consumption carried out using this microeconomic file will be biased.⁴

An example may shed some light on the difficulty: household out-of-pocket health expenditure (i.e. what they spend beyond what is covered by the health care system) is difficult to quantify in a household survey. Respondents often find it very difficult to estimate what they have spent on medical care, and they find it even more difficult to determine the portion that was not reimbursed. Here, the Health Surveys constitute the source of reference: as health expenditure is one of their main variables of interest, they devote questioning time to it and, if necessary, they carry out matching between their sample and the health insurance data. This is not the case with the BDF surveys which, consequently, provide an unreliable estimate of this expenditure. The solution is then to impute it for the households in the BDF survey sample based on a model estimated in the Health Survey.

The health expenditure of a household depends on its socio-demographic characteristics (age, income, social category and level of educational

attainment of its members) and health-specific variables: health status, medical history and health cover of its members. Of all these factors, those specific to health are of course by far the most explanatory. Information on those factors is collected in the Health Survey, not the BDF survey. The imputation model will therefore have to make do with the usual socio-demographic variables; it will then only be able to explain a fairly small part of the dispersion of health expenditure. In fact, between two households with the same usual socio-demographic characteristics (age, income, qualification, social category, etc.), expenditure can differ greatly if the factors most directly related to health are different. Imputation then amounts to assigning to each of the two households a value selected at random in the Health Survey from among the expenditures of households with the same socio-demographic characteristics.

On average, this procedure is unbiased: it provides, for any given group of households of fixed age, income, etc., that household's true average level of health expenditure. In contrast, it is incorrect in terms of distribution since it assumes that, once these characteristics have been fixed, health expenditure is randomly distributed among households, regardless of the rest of their consumption, in particular. However, with other given characteristics, a very sick person will have a lower final consumption than a person in good health, but a higher health expenditure. Random imputation lacks this correlation and will tend to assign too low a health expenditure to that person and, therefore, an underestimated total consumption.⁵ This limitation is inherent in the very principle of imputation. Only the actual collection of the variables for each household makes it possible to obtain their joint distribution.

Although it is not, strictly speaking, possible to determine the true joint distribution (income and consumption) in the population based on the distribution of income on the one hand, and the distribution of consumption on the other, but only an approximation, obtained under the at best rather crude assumption that they are

3. To provide a very simplified example: knowing the distribution of consumption C on the one hand, and that of income R, on the other, does not make it possible to determine that of savings R - C while it is not known whether the two covariate in the same way (the wealthier a person is, the higher their consumption) or whether, on the contrary, they tend to compensate each other (the wealthier a person is, the more they save).

4. The direction of the bias has not, a priori, been determined.

5. The imputation of health consumption, as with that of health expenditure by public health bodies that can be individualised, raises the exact same difficulty.

independent of each other⁶, imputation is nevertheless a method in line with good statistical practice, even if it remains a little cumbersome to implement.

However, it is not the method that has been used in the various studies aimed at decomposing the household account. The practical problem is that the imputations depend on the explanatory variables used in the model. The international comparison of distributions, which is clearly a major objective (an enrichment of the national accounts that would be doomed to lose international comparability would be of limited interest), is reliable only between countries that have rigorously followed the same imputation methodology, i.e. the same models, with the same variables. However, it is virtually impossible to display a core of variables common to all sources used in the different countries that is sufficiently large to be useful.

Consequently, international work on the decomposition of the accounts has fallen back on a much simpler pseudo-matching method (often referred to as “top-down”): it consists of dividing households in each source into groups according to a particular criterion that is present in all sources. Each aggregate of the account is then distributed (using the relevant source) between these different groups.

For example, households can be classified by age group and, for each group, the average value of the missing components in the BDF survey can be calculated using the appropriate sources. A complete account is thus calculated for each age group.⁷ The method entails matching average (or “representative”) households in a group between the different sources. This is what is referred to as pseudo-matching of sources. It may also be seen as an elementary case of the imputation method, in which the imputation model is reduced to a single explanatory variable, namely the criterion used (in this case, age); incidentally, this confirms that it shares the same limitations as the imputation method. This method has been followed since the beginning by the international working group, coordinated by the OECD, dedicated to the development of accounts by household category. Online Appendix C1 briefly traces the history of the attempts and efforts to decompose national accounts aggregates by category of households.

These considerations call for a number of comments:

- The use of such pseudo-matching to introduce a decomposition of the household account is, in the case of age groups, an old and proven procedure: it is the method used by the generational accounting developed in the 1980s and 1990s by Auerbach & Kotlikoff, an objective that has been taken up again since the 2000s by the promoters of the *National Transfer Accounts Project* (see Online Appendix C2). However, the method can decompose the accounts according to any household classification criterion (gender of the reference person, household size, level of educational attainment of the reference person, etc.), provided that, for each transaction and for all countries, a microeconomic source is available that identifies the households according to that criterion, in a homogeneous manner across sources and countries;

- Breaking down the household account requires taking into account monetary transfers between households (support, donations, etc.), as well as exchanges of market goods and services between them (sales of second-hand vehicles, rentals, etc.);

- The decomposition exercise for the household account aims to enrich the economic description provided by the national accounts. However, one of its associated results is improving the quality of household surveys. Rigorous collation⁸ with the accounting aggregates makes it possible to accurately assess the lack of coverage of these surveys in order to try to remedy it or, at least, to take it into account in the analyses;

- The availability of microeconomic sources, which are often only available on a multi-year basis, means that, in principle, it is not possible to perform decomposition of the account each year. Nevertheless, work is currently carried out to address this shortcoming, at least in part (see hereinafter).

1.3. Methodological Issues and Avenues for Progress

Even with the agreed simplifications in relation to the unattainable goal of a complete account at the level of each household, in practice, the decomposition of the household account

6. Strictly speaking, this is their independence conditional on the household description variables used in the imputation models. This is much more plausible than unconditional independence.

7. The total of a transaction across the different age groups should give the aggregate of the account. Otherwise, it is sufficient to recalibrate the source used using the aggregate. The sole function of the source is to provide the profile, not the level.

8. i.e. by ensuring that we are working on the same scope and with the same concepts.

raises several technical difficulties, exposed in Bellamy *et al.* (2009). In this article, the choice has been made to focus on one of them; it gives an idea of the work that still needs to be done to overcome the obstacles to the implementation of a system of accounts by household category, which has the same properties of reliability and international comparability as those under the central framework. It is also an interesting illustration of the differences, in both the objectives and the approach, between the decomposition of the accounts and two important recent approaches that also aim to link the distribution of resources and consumption with the corresponding accounting aggregates: the *National Transfer Accounts*, at the initiative of R. Lee and A. Mason and the *World Income Database* (WID.World), developed by researchers led by T. Piketty. Online Appendix C2 provides a presentation of these two approaches and details their similarities and their discrepancies in comparison with the breakdown of the household account.

The accounts by quintile of living standard show substantial dissaving by the poorest households in every country, except France. At the root of this discrepancy is the excess, in the consumption surveys, in the level of consumption over the level of income for a significant number of households.⁹ The BDF survey is no exception in this respect. In this instance, however, the decomposition of the French account by quintile of living standard was based on a specific BDF variable that makes it possible to identify and adjust the responses of households showing aberrant consumption-income discrepancies. The effect of this treatment is considerable. Without it, the lowest quintile would have a dissaving rate of around 20% in France. However, this variable that allows for adjustment is not present in household income and expenditure surveys in most other countries. In addition, this adjustment is merely a practical method, which has the sole merit of simplicity and plausibility. The assumptions on which it is based are open to discussion and, therefore, the results published also include a version with a different, less selective, adjustment method. With this method, the dissaving rate stands around 13% in the first quintile.

That is to say that the use of microeconomic information is not always an immediate operation. This information must be analysed, discussed, arbitrated and without any guarantee to find a satisfactory solution for the problems encountered. Moreover, the solutions possible in a particular information system cannot

necessarily be generalised. Undoubtedly, the most satisfactory solution requires a significant improvement in the accuracy of the microeconomic measurement of consumption. But this is an objective that will be difficult to achieve, even in the long term.

The difficulty outlined above is just one example of the problems to be solved. One could also mention the multi-year frequency (at least in the majority of countries) of certain microeconomic sources, such as the consumption survey; can an annual publication of accounts by category nevertheless be envisaged? In what manner?¹⁰

Another important issue is the accuracy of the accounts. Traditionally, the statistical accuracy of the aggregates of the central account is not considered. As there is no alternative, these aggregates are assumed to be “accurate”. In contrast, it is known that microeconomic data from surveys are marked (at least) by a sampling risk, which can be estimated. Is it possible to take this risk into account in order to assess confidence intervals for the differences established between household categories? The work of the OECD Expert Group (see Online Appendix C1) on these methodological issues and others is continuing, with the challenge to provide answers that are not only conceptually and practically satisfactory but also common, so as to arrive at a process for the production of accounts by household category that is stabilised and standardised as that for the aggregate account.

2. Expansion of the Scope of GDP

2.1. Time Spent on Domestic Work

Of all the expansions of the scope of GDP, the inclusion of the domestic production of services is probably the one that is most in line with the logic governing the indicator: first, the domestic production of goods is already taken into account (self-consumption); second, GDP includes the value of the housing service that owners-occupiers render to themselves (imputed

9. This is a classic finding, both at household level and at the level of groups of households. Consumption econometricians readily explain this by poor measurement of income, assuming that households tend to under-report their resources to the survey (hence the traditional practice in econometric study models of instrumenting income). The problem is, in reality, more profound and more complex, as the excess of consumption over income also appears implausibly widespread when the data on income is of administrative origin, as in the case of the BDF 2010 survey.

10. One possibility that has been explored by Insee recently (Accardo *et al.*, 2017) is to fix the disparities between households, as observed in the surveys, but to change the aggregates annually as indicated in the national accounts.

rents). This is clearly domestic production of services. Furthermore, it generally constitutes a major item in the household account (in France, for example, it accounts for around 13% of final consumption expenditure). Another is that ignoring the domestic production of services can bias international comparisons (this is also one of the justifications for including imputed rents). As the Stiglitz-Sen-Fitoussi report points out, a country in which the level of household production for self-consumption is significant may have a lower GDP than another, in which more goods and services go through the market, while households have the same level of consumption, if their own production is taken into account. For example, Alesina & Ichino (2009) calculate that when all domestic production is taken into account, Italy's GDP per capita rises from 56% to 79% of US GDP; finally, ignoring this (non market) production can lead to an overestimation of GDP growth, as households turn to the market for activities they used to do themselves.¹¹

In practice, however, there are many unresolved difficulties in measuring the value of these activities, despite the efforts that have been devoted to it for several decades now:

- the precise scope of the activities to be taken into consideration remains a subject of debate. In principle, the criterion of “delegability” (or third party) is agreed upon. However, its application is often problematic¹² (Gershuny, 2011; Roy, 2012);

- various valuation options are available: at opportunity cost or at the observed market wage for an equivalent task. The latter option is the one most often used, as the former raises quite a number of objections; but it is not necessarily more realistic;¹³

- in the absence of accurate information on the characteristics of the task and the resulting product, their valuation is probably fairly biased.¹⁴

The estimated value of domestic work not only varies considerably depending on the scope and the valuation option chosen (in a ratio of 1 to more than 3), but in all cases it also represents a substantial sum (up to 50% of GDP according to Roy, 2012). This makes it difficult to include it in the central framework (and suggests instead that it be processed in a satellite account).

The main source of information on domestic activities is the Time Use Survey (*Enquête Emploi du Temps*). The results of the valuation

are closely dependent on the information gathered by these surveys and the methods used to collect it. The standard method consists of having a sample of respondents complete a daily diary as they go about their activities.¹⁵ The retrospective survey method¹⁶, which is less costly and cruder, can give results that differ significantly in their level and distribution (Kan, 2008), with a tendency to overestimate the time spent on domestic activities.¹⁷ More elaborate and costly methods (such as the experience-sampling method or continuous observation) also exist, which could be developed in the future thanks to technological developments (online surveys, sensors installed on respondents' mobile phones, etc.). These methods still result in different estimates. Furthermore, the level of accuracy of the information collected is crucial in characterising domestic activities. Knowing all the secondary activities carried out at the same time as the main activity represents a much heavier survey burden for the respondent but is the only way to gain an understanding of all the domestic activities.¹⁸

11. However, this bias is undoubtedly more limited than the previous one. For example, in France, the average time spent producing domestic services (cooking, cleaning, childcare, etc.) per person (aged 18 or over) per day fell by 28 minutes between 1974 and 2010 (Brousse, 2015, p. 84). Valued using the super-gross minimum wage (SMIC super brut) used by Roy (2012), this reduction in domestic production represents, under the maximum hypothesis in which it is entirely externalised in the monetary sphere, a contribution to GDP of €91 billion in 2010. The annual GDP growth, 2% for the period 1974-2010, would then be overestimated by a maximum of about 0.13 percentage points.

12. Do playing with one's children, DIY and shopping qualify as domestic production, or are they done for the personal pleasure derived from them? Depending on the answer given, time spent on domestic production varies by 50% (Roy, 2012). Similarly, excluding care of one's own body, as done by Roy (2012), rather than including it, as done by Alesina & Ichino (2009), has a very significant impact (a reduction of one hour in the time spent on domestic production per person per day).

13. In particular, its reference to the observed market price is questionable, as there is generally no precisely defined market price for these activities. For example, there is no market, for reasons that are easy to understand, where it is possible to buy 15 minutes of cooking time to prepare the children's ham and mashed potatoes in the evening, the two and a half minutes of time to wash the dishes and the 18 minutes of time to read them a bedtime story. In addition, the fact that the parent(s) who performed these tasks that evening produced a value of exactly 35.5 minutes x the super-gross hourly minimum wage is far from obvious.

14. In practice, in fact, the studies resolve to value these tasks uniformly at or around the minimum wage. However, there is no guarantee that an activity carried out by the household is of a quality comparable with that of activities performed professionally. It should also be noted that imputed rents, the only domestic service currently included in the national accounts, are not set equal to the average actual rent but are determined by taking into account the characteristics of the stock of principal residences occupied by their owners.

15. The time-slot of the diary is variable: in the French survey, it was five minutes until 1998, when it increased to ten minutes. It is 15 minutes in many surveys. Some, such as the Australian survey, leave it blank.

16. It is also known by the name “Stylised time-use items”. It is the one adopted by the Labour Force Surveys in most countries.

17. However, this point, which has been observed on British data, is debated; Schulz & Grunow (2011), in contrast, find fairly good consistency between the two methods on German data.

18. For example, the statement “I watch TV” in a daily diary results in that time not being coded as a domestic production activity. However, if the diary also gathers information on secondary activities and if one of them indicates the presence of children under the respondent's care, then the time will be counted, at least in part, as a domestic activity (“Childcare”).

Obtaining comparable estimates of the domestic production of services across countries or over time therefore requires a high degree of harmonisation between the surveys used for measurement. This harmonisation is still only partial. Many countries carry out Time Use Surveys with variable, but generally quite widely spaced, frequencies (in France the survey is carried out about every ten years), as those surveys are considered most expensive. At present, Eurostat has managed to coordinate the European countries by getting them to use a common methodology: collection of information using a daily diary, taking secondary activities into account and using a nomenclature of activities. Japan has chosen to collect information via a diary completed gradually. However, the United States, Canada, Australia and New Zealand have chosen (at least in the most recent surveys) to stick to the retrospective method.¹⁹ Generally speaking, despite the many international efforts to standardise nomenclatures²⁰, they do not always coincide, with risk of different classifications for the same activity.

Clearly, a major effort to harmonise the scope, valuation and measurement methods is still needed before domestic work can be integrated into the preparation of the accounts, with the figure produced having a status comparable to that of the aggregates in the standard account.

2.2. Non-Monetary Dimensions: Health, Safety, Social Capital, Human Capital, etc.

Criticising GDP (and national accounting more generally) for ignoring many dimensions of life that have a value for individuals raises three questions: *i*) Is a quantitative measurement of these values necessary? *ii*) Can such measures be designed and determined? *iii*) How can this information be linked with that provided by GDP?

Statisticians, economists and national accountants are undoubtedly (is it a professional bias?) inclined to answer the first in the affirmative, but this view is not as straightforward as we may think. Just consider Robert Kennedy's famous speech during the 1968 American presidential election campaign²¹ in which, clearly, most of the values mentioned did not, in his mind, require quantification. It should also be recalled that economic theory itself stresses "the rather loose nature of the link between overall income and social well-being" (Fleurbay & Blanchet, 2013, p. 115²²), which may put into question the utility of quantifying the unquantified.

If, however, the decision is taken to measure these values, questions (*ii*) and (*iii*) can be addressed in two ways: by juxtaposing complementary indicator tables, possibly summarised into composite indicators, with national accounting aggregates or by calculating a monetary equivalent for the non-monetary dimensions that can be directly measured against GDP and other accounting variables.

2.2.1. Synthetic Indicators and Dashboards

The first approach is limited to identifying indicators (in principle, non-monetary indicators) capable of describing the situation of individuals in the dimension (health, safety, democracy, social cohesion, etc.) under consideration. They provide information that complements that provided by the major accounting aggregates (GDP, gross disposable income, consumption, savings, etc.). This approach has been developed since the 1970s. An intuitive approach and GDP limits that have become commonplace explain the ever-increasing demand from decision-makers or the public for these indicators. An ever-increasing abundance of economic and social information that is easier to process explains why supply has been able to keep up. The result has been a flurry of initiatives that (based on pre-existing statistics) constitute sets of indicators meant to compensate for the shortcomings of traditional macroeconomic variables.²³ Among the most recent and significant examples are the European Sustainable Development Indicators (2005), the Sustainable Development Indicators for France (2010), the European Union's 2020 Strategy Indicators (2010), the OECD's Better Life Indicator (2011), the French government's new wealth indicators (2015) and the indicators

19. A surveyor questions a household member by telephone about their activities the previous day.

20. ICATUS nomenclature (UN), HETUS nomenclature (Eurostat), UNECE guidelines in 2013 as well as the work of Gershuny's team (the MTUS project of the CTUR in Oxford).

21. "Yet the gross national product does not allow for the health of our children, the quality of their education or the joy of their play. It does not include the beauty of our poetry or the strength of our marriages, the intelligence of our public debate or the integrity of our public officials. It measures neither our wit nor our courage, neither our wisdom nor our learning, neither our compassion nor our devotion to our country, it measures everything in short, except that which makes life worthwhile".

22. Chapter 4 of their work provides an in-depth analysis of how to give a monetary expression to preferences and how to use it in a normative analysis of well-being.

23. It should be noted that this article does not propose a general epistemological and historical analysis of the indicators (see Noll, 2002, for example), but only presents the main characteristics of the indicator compilations that aim to go beyond the description through only the aggregates of the national economic and social development accounts, and to quantitatively establish a notion of quality of life or quality of growth (for example, sustainability, inclusiveness, etc.).

for the Sustainable Development Goals adopted by the UN in 2017.

The indicators are practically always selected from the vast set of indicators published (or at least publishable) by the various public and private producers of economic and social information, generally in a very pragmatic way, outside of any theoretical framework, after more or less long and complex negotiations between political, administrative, scientific, expert and community representatives. Various authors or organisations have indeed tried to identify general principles for selecting an indicator, but these principles are primarily pragmatic (and sometimes *ad hoc*)²⁴ and do not provide a theoretical justification for the indicators selected.²⁵ This explains, at least in part, why the sets of indicators produced are often very disparate.²⁶ Once the indicators have been identified and collected, the question of how they relate to the usual accounting aggregates is likely to be resolved in two ways. The most simple solution is to make the information available as it is, in the form of a dashboard. It is left up to the user to consider the various messages in front of their eyes and to draw the conclusions they can.

During the 1980s and especially in the 1990s, the advocates of an indicator-based approach readily nurtured the ambition of creating an indicator capable of replacing GDP. In the belief that the place of GDP in the public debate is largely explained by the fact that it is a single figure (therefore easy to remember, easy to quote and allowing countries to be classified), they have tried to summarise sets of indicators into a single so-called “synthetic” index²⁷: this has led to the design of the Index of Social Health (Miringoff, 1987; Miringoff & Miringoff, 1998), the Human Development Index (Haq, 1990), the Advanced Quality of Life Index (Diener, 1996), the Weighted Index of Social Progress (Estes, 1997), the CSLS’s Index of Economic Well-Being (Osberg & Sharpe, 1998), the Index of Living Standards (Sarlo, 1998), the BIP40 (Inequality Observatory, 2004), etc. The OECD’s Better Life Index (2011) is among the most recent.

To allow aggregation of variables describing very heterogeneous phenomena²⁸ into a single figure, these indices project them linearly (the minimum observed for the variable on 0 and its maximum observed on 100) and then create a simple (like the UN HDI) or weighted average for them. The weights are then either chosen at the discretion of the designer of the index,

determined by factor analysis techniques (ISP), left to the user’s initiative (the CSLS provides an Excel macro to vary the weight of its IEW and the OECD provides an online application for the Better Life Index²⁹).

The procedure is convenient, but *ad hoc* and, technically, it is not without flaws (Gadrey & Jany-Catrice, 2012, p. 41; Accardo & Chevalier, 2005). Above all, the summary produced is problematic to interpret and it is generally recommended not to stick to the index (which is tantamount to acknowledging that it is basically just a convenient artefact) but to take into consideration the information provided by its components.³⁰

As the issue of the weighting of the components of a synthetic index is without solution other than conventional, the designers of alternative indicators to national accounts aggregates currently tend to abandon the objective of a single index competing with GDP. The recent initiatives mentioned above are all either of the “dashboard” type or allow the user to choose their weighting preference.

2.2.2. Monetisation

Monetising the non-monetary dimensions is the alternative to tables of indicators and synthetic indicators. This time, the stage of aggregation to GDP is immediate (or almost immediate³¹) and it is obviously the first stage, which consists in assigning a price to things that are readily

24. The Indicators Sub-Group of the European Social Protection Committee has thus proposed the following principles: “An indicator should 1) capture the essence of the problem and have a clear and accepted normative interpretation, 2) be robust and statistically validated, 3) provide a sufficient level of cross countries comparability, 4) be built on available underlying data, and be timely and susceptible to revision, 5) be responsive to policy interventions but not subject to manipulation”. See also Atkinson et al. (2002).

25. To clarify the criticism made here: this lack of theory can be contrasted with the construction of notions such as poverty in living conditions (Townsend, 1979), in which indicators are selected within an explicit conceptual framework. This, at least, makes it possible to question the ability of the indicators to measure what needs to be measured.

26. For example, the ten key European Sustainable Development Indicators (of 130 indicators) produced by Eurostat since 2007 include monetary aggregates (e.g. GDP per capita), counts of people (e.g. number of poor people) or years (e.g. life expectancy) or animal species (e.g. common birds) and tonnes of CO₂.

27. Sen (2003) describes this reasoning in detail, in the case of the HDI.

28. In effect, this involves aggregating rates of infant mortality, unemployment, youth suicide, housing access indicators (ISH) or even the number of doctors per 1,000 inhabitants, the saving rate, an income inequality index and the number of environmental treaties ratified (AQoLI), etc.

29. <http://www.oecdbetterlifeindex.org/fr/#/11131111111>

30. If only to (try to) understand why the different indices do not match up. For example: “Hence, the probable reason for Canada’s fall from first (HDI) to 31st (WISP) in international ranking is the greater breadth of coverage of the WISP – but the complexity of the WISP calculation prevents a clear comparison” (Osberg & Sharpe, 2001).

31. It is not that immediate, insofar as the monetisation is based on a stock rather than an annual flow.

said to be priceless, which is the tricky part of this solution.³²

Two methods are used to assign a value to a non-monetary asset.

(i) Stated preferences: this method is based on directly questioning a sample of individuals (in principle representative of the population). The questioning can be more or less sophisticated. Individuals may simply be asked: “What do you think this asset is worth?”. They may also undergo complex questioning protocols, involving detailed experimental simulations and questions on binary choices or rankings of assets or scenarios, designed to allow formal choice models to be estimated.

(ii) Revealed preferences: here, the method is not based on statements but on observed behaviour. There are two main techniques. The first is implicit costs: the most commonly cited example is transport expenditure incurred by visitors to a nature reserve. This transport cost itself is a lower bound of the value that the public places on this environmental good. In random utility models, it can be used to estimate the value itself (although, it must be said, at the expense of quite a number of other additional assumptions). The second is hedonic prices: this uses as inputs the observed variations in the market price of a good depending on its characteristics. Thus, by observing the different prices of cars, depending on the models and ranges, it is possible to identify, econometrically, the value of a particular vehicle characteristic (speed, driving comfort, fuel economy, etc.) despite there being no specific market for that characteristic. The different methods have been in common use for many years in the field of cost-benefit analysis for the selection of public investments.

The revealed preference method is regularly used in the national accounts: for example, in the valuation of services provided by government bodies at their production cost, which is the implicit cost technique; hedonic methods are commonly used for the valuation of the housing service produced to themselves by owner-occupiers or for the determination of price indices (vehicles, computers, household appliances, etc.) at constant quality. By contrast, national accounts do not use the stated preference method for which, in fact, there is no reference procedure. This method is based on hypothetical choices, raising the crucial question of under what conditions these reported

estimates constitute truly relevant information on individuals’ preferences.

Thus, there is no consensus on the assessment of non-market dimensions, such as environmental capital, on the statistical value of life or social capital (in the meaning of Putnam, see OECD 2001) in a country. Correlatively, the available studies are rarely comparable and international comparisons are impossible most of the time.

3. Another Paradigm: Subjective Well-Being, Satisfaction and Happiness

The current wave of interest in a direct measure of well-being as perceived by individuals can probably be traced back to the mid-1990s. Easterlin’s article in *Journal of Economic Behaviour and Organization* (Easterlin, 1995) is thought to be the starting point. It was essentially a continuation of an article from twenty years earlier (Easterlin, 1973) little noticed at the time. In 1995, however, economists, who are traditionally more inclined to rely on what individuals do rather than on what they say, had become influenced by work on behavioural economics, developed in particular by Thaler, Kahneman and Tversky, who were much more familiar with approaches questioning the canonical model of a *Homo Economicus* and more willing to consider the perception that individuals report about their economic situation.

3.1. The Good Fortune of the Happiness Paradox

Easterlin’s idea is to use Happiness Surveys: since 1946, at least, surveys have asked respondents directly whether they consider themselves “happy” (phrased this way or similarly). Looking at average satisfaction calculated based on the responses collected, it seems to have remained stationary over the post-war decades even though, over the same period, GDP per capita has increased by a factor of two, three or even more, depending on the country. This result is the “Easterlin paradox”.

As he points out in his 1995 article (p. 37), this paradox was fairly well documented as early as the late 1970s. However, it was not until the 1990s that the interpretation of it as an index to

32. Problems related to the link between monetary value and well-being are left aside here (see Fleurbaey & Blanchet, 2013, chap. 4). They also concern the usual monetary dimensions, those for which the national accounts can use existing market prices.

use against GDP became evident. Previously, the same observation seemed to reinforce economists in their bias against the use of subjective data, a material that the majority of them deemed to be acceptable, at a push, for “soft” disciplines such as sociology and psychology, but not for positive economic science for which only actions can reveal preferences.

This is in stark contrast to the current situation: for more than twenty years now, more and more supporters of the measurement of subjective well-being have been stressing that “the ways in which people value their lives [...] should be an integral part of the concept of human well-being”³³, an idea that is also supported by the Stiglitz-Sen-Fitoussi report.

At present, regular survey data on perceived well-being are available for many countries, feeding into the abundant work on an “economics of happiness” that gained momentum in the academic field during the 2000s and in which decision-makers, media and the general public take a keen interest. The OECD, which includes a measure of reported satisfaction in its Better Life Index, has also published guidelines to that end, aimed in particular at national statistical institutes (OECD, 2013). In 2013, a secondary module of the EU-SILC panel, a survey under European regulation, collected the answers of respondents to questions on their well-being and Eurostat published the results.

Several NSIs have taken an interest in this type of indicators; Insee has produced an annual measurement of life satisfaction since 2011 and the British ONS since 2015. The French indicator is also included in the table that the law on new wealth indicators (or “*Loi Sas*”), adopted in April 2015, obliges the government to publish annually.

3.2. Measuring Happiness

The economic theory of happiness actually distinguishes three notions of subjective well-being (OECD, 2013):

- respondents’ satisfaction with the life they lead: here, the respondents must produce a global judgement on their life overall or over a more limited period (most often the current period). The issue then boils down to a question. The respondents are generally asked to rate their life on a scale (known as the Cantrill scale) that goes from 0 (very poor rating) to 10 (very good rating);

- “affects”: the term is used in psychology to refer to the emotional states of an individual at a given time. The reference measurement method is then the Experience Sampling Method (ESM), in which the participants must keep an accurate account of the following states (“happy”, “peaceful”, “irritated”, “angry”, etc.);

- the “eudaimonic” approach: the aim is to take a step back from the hedonism underlying the two previous methods and to have the respondents assess the degree to which they feel fulfilled in their life and the degree to which they think they have effectively developed their physical, intellectual and moral potential.

In practice, the first two approaches are dominant. The measurement of satisfaction is certainly the most widespread, due to the simplicity of its implementation. Nevertheless, recent technological developments offer researchers new options for measuring affects, from applications installed on the respondents’ smartphones reminding them of the times of day when they should send information on their emotional state, to medical-type devices that continuously measure their blood pressure, stress, etc., information from which the researcher is supposed to be able to derive an assessment of the respondent’s affects throughout the day. It should be noted here that, with this method, the respondent’s subjectivity is set aside: this is a physical type of measurement, through use of a measurement tool. Naturally, the question then arises as to whether the usual subjective assessment and this much more objective measurement relate to the same thing.

Judging by its public success and its integration among the instruments guiding public action, “life satisfaction” appears to be particularly well placed to compete with GDP as a measurement of well-being. However, it raises major difficulties, the resolution of which seems all the more difficult since they are probably not taken into account sufficiently by the various users of this indicator. First of all, there is a fundamental theoretical question: what conclusion should be drawn from the level of satisfaction reported? Is it to be understood that maximising satisfaction should be the goal of public policy? On this point, one may object, firstly, that it is by no means obvious that this is the objective of the individuals themselves and, secondly, even

³³ Taken from the *Recommendations for Measuring Sustainable Development of the joint UNECE/Eurostat/OECD Task Force established by the Conference of European Statisticians in 2014.*

if that were the case, such a choice is likely to be criticised from an ethical point of view (Fleurbaey & Blanchet, 2013, p. 169 *et seq.*).

3.3. The Meaning of the Figures

In addition, many problems arise in the practical construction of this measure and in its use, most of which do not seem to be resolved, nor are they in the process of being resolved. They are not due to the subjective nature of the respondents' response, which in itself does not exclude the possibility of developing indicators of proven usefulness, but to the lack of clarity over what the response covers.

Firstly, the wording of the question in statistical surveys is necessarily brief and does not make it possible to specify with what a respondent is satisfied with sufficient precision: his/her life at present, their life since birth, or their life over a shorter period of time? Their own life, their own life and the lives of those they care about or their life as a member of a wider community? And how can one be sure which dimensions of life the respondent takes into account in their assessment?

Similarly, how is it possible to control the respondent's reference point: in other words, to what does the "0" (or the "10") on the scale refer? The worst life ever lived in the history of the world? Or just in the respondent's country at present? Or the worst life that the respondent has a reasonable chance of experiencing personally? Or the worst life that the respondent has actually experienced? Etc.

Finally, what metric is the respondent using? In other words, is the respondent harsh or indulgent in rating their life? What is a "5", a "7" or a "10" actually worth to them? This is what Fleurbaey & Blanchet (2013) refer to as the "calibration problem".³⁴

In the absence of a minimum degree of clarification of this set of ambiguities affecting the responses collected, it seems doubtful that it would be possible to give a reliable meaning to the aggregation of the satisfaction ratings reported by a sample of individuals. There is indeed good reason to believe that these ambiguities are of very real practical importance.

Let us thus consider the calibration problem: it is possible to seek to assess (and attempt to control) its importance using a "vignette methodology" (Kapteyn *et al.*, 2009; Angelini *et al.*, 2014). This involves brief descriptions of individual

situations that the respondent is asked to rate. The distribution of the ratings collected for a single vignette makes it possible to measure calibration differences within the population.

The self-questionnaire that has been included in the EU-SILC panel since 2011 includes eight different vignettes of this kind.³⁵ They are presented to the panel respondents (more specifically, to those in the second re-interview). Whatever the year of the survey and whatever the vignette considered, it is found that the ratings are about as dispersed (sometimes even significantly more so) as the ratings given by respondents concerning their own lives.

This is quite a remarkable phenomenon: if, for the same situation, the respondents give such divergent ratings, it becomes very risky to give any substantial interpretation to the rating they give to their own situation. Seeing that Respondent A gives their life a satisfaction rating of 5, while Respondent B gives a satisfaction rating of 8, what can be concluded about their actual situation and how they perceive it? It should be noted that the average life satisfaction indicator included in the alternative indicators to GDP provided for by the law of 2015 is not corrected for calibration differences, nor is the indicator calculated for each European Union country by Eurostat, based on the EU-SILC module carried out in 2013.

It is important to underline that even a modest correction can have a highly visible impact, due to the smallness of the differences generally observed between the average satisfaction levels in the various countries: according to data from the EU-SILC 2013³⁶, the average rating for 30 countries are between 6.2 and 8; a correction of +0.5 therefore represents about 10 places gained in the ranking. Under these conditions, and to use an example, France's mediocre rating

34. In psychometrics, the problem is known as "Differential Item Functioning". See for example Osterlind & Everson (2009).

35. Here are two examples of these vignettes:

(No 7) Maria is a veterinarian aged 58. She lives with her husband in a house with a garden. She has three children and five grandchildren who visit her regularly. She plays tennis every weekend. How would you rate Maria's situation in respect of the life she is currently living?

(No 8) Anne is 40 years old. She works as a nursery assistant. She lives with her husband and their three children in a small apartment they rent on the outskirts of the city. The neighbours are quite noisy. Her husband has been unemployed for two years, it is not always easy to make ends meet and this creates tensions in their relationship. She suffers with back pain and has trouble sleeping because, this year, she is working in a difficult class. How would you rate Anne's situation in respect of the life she is currently living?

36. The European Union Statistics on Income and Living Conditions (EU-SILC) is a panel survey carried out annually in Europe, within the framework of a European regulation. The French part of the EU-SILC is referred to as SRCV in French.

in terms of life satisfaction (with an average rating of 7.0, measured in the 2013 EU-SILC module, the 16th place out of 32 countries), a phenomenon that is regularly observed, is easily interpreted either as an indication that life in France is less pleasant than in its neighbouring countries, or as the mark of a national mood that is more gloomy than elsewhere.³⁷ In the absence of rigorous calibration, these interpretations appear fragile, to say the least.

Of course, the interpretation of this dispersion as an issue of calibration can be questioned. The available data do not make it possible to rule out that, in reality, individuals use the scale in exactly the same way and that it is simply their conceptions of what is pleasant, bearable, painful or intolerable in life that differ.³⁸ However this objection does not answer questions about the relevance of measurements based on self-assessment of subjective well-being.³⁹ The heterogeneity of preferences can even be viewed as further calling into question their usefulness, as substantially different preferences from one individual to the next make it very difficult to interpret the collected ratings, for which it is not clear not only to which exact situations they relate but also what, in these situations, is judged positively and what is judged negatively by each respondent.

For the promoters of the economics of happiness, measuring well-being is measuring what people think of their own happiness.⁴⁰ This approach has become quite widely accepted in recent years, resulting in high demand for data. Social statistics has acted with remarkable responsiveness, but it cannot confine itself to producing figures. It must also enable users to understand their nature, their scope and their limitations. From this point of view, the measurement of subjective well-being still requires significant clarification efforts. While going beyond GDP is clearly a necessity, it is still necessary to know exactly where we stand once we have gone beyond.

* *
*

Looking “beyond GDP” may simply mean not limiting oneself to that indicator but taking other socio-economic indicators into consideration in the economic analysis. However, one could also set a more ambitious goal of “going beyond GDP”, i.e. developing a conceptual and operational framework that integrates other information in a coherent way without losing

that provided by the existing framework. Each of the three avenues of research described in the article illustrates not only the benefit of such an approach, but also its difficulty.

Disaggregation of GDP is fully in line with the accounts approach. Admittedly, conceptual questions arise, particularly with regard to the scope concerned: should all the components of GDP be disaggregated? Or only the household account (but including consumption)? There is also the issue of the statistical unit: should the distribution be measured at the level of the individual or at the level of the household? The answers depend on the objectives set. However, the more complex issue remains a practical one. It is the issue of the sources of information on income and consumption distributions: how reliable are they, what is their availability, how comparable are they over time and between countries. The use of sources of social statistics (surveys and data from government departments) is on the rise, particularly over the last ten years or so, and undeniable progress has already been made. This progress is expected to speed up. It can be expected that within ten years, various distributions (income, consumption, wealth and savings) that are fully consistent with the framework of the accounts will be available in many countries.

Taking into account socio-economic dimensions not included in GDP but deemed necessary to make judgements on well-being is more difficult to combine with the traditional accounts approach and the conception and construction of GDP. Monetisation is certainly the approach most directly consistent with them, as it makes it possible to produce an expanded GDP, in the continuity of the traditional GDP and subject to the same analyses. However, this raises conceptual and technical issues (for example, the dependence of the valuation obtained on the method chosen and the unavoidable and numerous additional assumptions), the solution to which does not appear to be forthcoming.

The subjective well-being approach is surely the one that poses the greatest difficulties: how can the national accounts integrate the information

37. Algan et al. (2018) thus speak of the “exception of the French malaise”.

38. It is not possible, however, with the existing data, to estimate the respective weights of calibrations and preferences in the heterogeneity of the vignette evaluations.

39. For more general information on these methodological issues, see the OECD manual (OECD, 2013, op. cit.) which provides detail on them.

40. “Self-reported happiness has turned out to be the best indicator of happiness” (Frey & Stutzer, 2002).

it produces? For, whatever the merits of the “economics of happiness”, the advantage of the research it generates and the insights it can shed on socio-economic behaviour and the fundamental problems of the economics of well-being⁴¹, it is clear that its purpose is not of the same nature as those developed in the national accounts. Indeed, the collection of subjective preferences is an interesting tool for the monetisation of non-monetary dimensions and can thus contribute to their inclusion in the

expanded GDP. However, notions of happiness, subjective well-being, life satisfaction, etc. still raise, in terms of the clarity of concepts as well as comparability and traceability of measurements, too many issues that are difficult to resolve and which, whether they can be resolved or not, will always clearly fall outside the scope of national accounts. □

41. See, for example, Layard (2005), one of the main representatives of the field, and Clark (2018) for a presentation of research developments.

Link to the Online Appendices: https://www.insee.fr/en/statistiques/fichier/4770148/ES-517-518-519_Accardo_Online_Appendices.pdf

BIBLIOGRAPHY

Accardo, J. & Chevalier, P. (2005). Les indicateurs synthétiques. *In*: Verger, D., Bas revenus, consommation restreinte ou faible bien être : les approches statistiques de la pauvreté à l'épreuve des comparaisons internationales. *Économie et Statistique*, 383-384-385, 33–35. <https://doi.org/10.3406/estat.2005.7192>

Accardo, J., Billot, S. & Buron, M. L. (2017). Les revenus, la consommation et l'épargne des ménages par grande catégorie entre 2011 et 2015. *Insee Références – L'économie française - Comptes et dossier - édition 2017*. <https://www.insee.fr/fr/statistiques/2894028?sommaire=2894036>

Alesina, A. & Ichino, A. (2009). *L'Italia fatta in casa. Indagine sulla vera ricchezza degli italiani*. Milano: Mondadori.

Algan, Y., Beasley, E. & Senik, C. (2018). *Les Français, le bonheur et l'argent*. Paris: Éditions Rue d'Ulm.

Angelini, V., Cavapozzi, D., Corrazzin, L. & Paccagnella, O. (2014). Do Danes and Italians rate life satisfaction in the same way? Using vignettes to correct for individual-specific scale biases. *Oxford Bulletin Economics and Statistics*, 76(5), 643–666. <https://doi.org/10.1111/obes.12039>

Atkinson, T., Cantillon, B., Marlier, E. & Nolan, B. (2002). *Social Indicators: The EU and Social Inclusion*. Oxford: Oxford University Press.

Bellamy, V., Consalès G., Fesseau, M., Le Laidier, S. & Raynaud, E. (2009). Une décomposition du compte des ménages de la comptabilité nationale par catégorie de ménage en 2003. *Insee, Document de travail* N° G2009/11. <https://www.insee.fr/fr/statistiques/1380884>

Brousse, C. (2015). La vie quotidienne en France depuis 1974. Les enseignements de l'enquête Emploi du temps. *Économie et Statistique*, 478-479-480, 79–117. <https://doi.org/10.3406/estat.2015.10559>

Clark, A. E. (2018). Four Decades of the Economics of Happiness: Where Next? *Review of Income and Wealth*, 64(2), 245–269. <https://doi.org/10.1111/roiw.12369>

Conference of European Statisticians (2014). *Recommendations on Measuring Sustainable Development*. UNECE. https://www.unecce.org/publications/ces_sust_development.html

Easterlin, R. A. (1973). Does money buy happiness? *The Public Interest*, 30, 3–10.

Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior and Organization*, 27(1), 35–47. [https://doi.org/10.1016/0167-2681\(95\)00003-B](https://doi.org/10.1016/0167-2681(95)00003-B)

Fleurbaey, M. & Blanchet, D. (2013). *Beyond GDP: Measuring Welfare and Assessing Sustainability*. Oxford: Oxford University Press.

Frey, B. & Stutzer, A. (2002). *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*. Princeton: Princeton University Press.

Gadrey, J. & Jany-Catrice, F. (2012). *Les nouveaux indicateurs de richesse*. Paris: La Découverte.

Gershuny, J. (2011). Time-Use Surveys and the Measurement of National Well-Being, mimeo, Centre for Time-use Research, Department of Sociology, University of Oxford, september.

- Kan, M. Y. (2008).** Measuring housework participation: The gap between « stylised » questionnaire estimates and diary-based estimates. *Social Indicators Research*, 86, 381–400. <https://doi.org/10.1007/s11205-007-9184-5>
- Kapteyn, A., Smith, J. P. & Van Soest, A. (2009).** Comparing life satisfaction. Rand Corporation Publications Department, *Working Paper* N° 623-1. https://www.rand.org/pubs/working_papers/WR623-1.html
- Layard, R. (2005).** *Happiness: Lessons from a New Science*. London: Penguin Books.
- Noll, H. H. (2002).** Social Indicators and quality of life research: background, achievements and current trends. In: Genov, N. (Ed.), *Advances in Sociological Knowledge over Half a Century*. Paris: International Social Science Council.
- OECD (2001).** *The Well-Being of Nations. The Role of Human and Social Capital*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264189515-en>
- OECD (2013).** *OECD Guidelines on Measuring Subjective Well-being*. Paris: OECD Publishing.
- Osberg, L. & Sharpe, A. (1998).** An Index of Economic Well-Being for Canada. CSLS Conference on the State of Living Standards and Quality of Life in Canada, October 30-31, Ottawa, Ontario.
- Osberg, L. & Sharpe, A. (2001).** Trends in Economic Well-Being in Canada in the 1990s. *The Review of Economic Performance and Social Progress*, 1, 233–247. <https://econpapers.repec.org/RePEc:sls:repsls:v:1:y:2001:aslo>
- Osterlind, S. J. & Everson, H. T. (2009).** *Differential item functioning*. Thousand Oaks, CA: Sage Publishing.
- Roy, D. (2012).** Le travail domestique : 60 milliards d’heures en 2010. *Insee Première* N° 1423. <https://www.insee.fr/fr/statistiques/2123967>
- Sen, A. (2003).** L’indice de développement humain. *Revue du MAUSS*, 21(1), 259–260. <https://doi.org/10.3917/rdm.021.0259>
- Schulz, F. & Grunow, D. (2011).** Comparing Diary and Survey Estimates on Time Use. *European Sociological Review*, 28(5), 622–632. <https://doi.org/10.1093/esr/jcr030>
- Stiglitz, J.E., Sen, A., Fitoussi, J.-P. (2009).** Rapport de la Commission sur la mesure des performances économiques et du progrès social. <https://www.vie-publique.fr/rapport/30513-mesure-performances-economiques-et-progres-social>
- Townsend, P. (1979).** *Poverty in the United Kingdom*. London: Allen Lane and Penguin Books.
- Vanoli, A. (2002).** *Une histoire de la comptabilité nationale*. Paris: La Découverte.
-

