the review of income and wealth

Review of Income and Wealth
Series 63, Supplement 2, December 2017
DOI: 10.1111/roiw.12333

THE FUTURE OF THE NATIONAL ACCOUNTS: STATISTICS AND THE DEMOCRATIC CONVERSATION[1]

BY DIANE COYLE

University of Manchester

1. TELLING A STORY

Public debate about the main economic indicators, including GDP and the rest of the National Accounts, is a matter of politics as well as statistics. Politicians use the published statistics to tell their story, to persuade voters of some view about the state of the economy. They are telling the story with the intention of promoting their electoral prospects or an ideological perspective. Newspapers, news organizations and blogs with their own political or commercial interests mediate the politicians' stories. Economic commentators who appear on the news or are quoted in the papers are also in the business of narratives and marketing. The same statistics can be interpreted, or misinterpreted, to tell different stories. There are countless examples. To give just one from the U.K. General Election campaign in 2015, the third set of figures for 2014Q4 GDP, published on March 30, 2015, generated two contrasting interpretations: "U.K. economy grew at fastest rate for nine years in 2014," and, "Data shows slowest recovery since 1920s."[2]

News about the latest statistics from all of these sources washes over a citizenry that is somewhat cynical about the deployment of economic statistics, when not simply indifferent. A number of polls indicate that trust in both the media and in politicians has been on a downward trend over a long period and stands at or near all-time lows.[3] The independence of most national statistical offices is an important bulwark against cynicism. There are also new mediators emerging online, some affiliated to traditional media organizations and others independent organizations, 'fact checking' the claims or stories being told in political debate. Yet a high degree of

*Correspondence to: Diane Coyle, Department of Economics, University of Manchester, Oxford Road, Manchester M13 9PL, UK (Diane.Coyle@manchester.ac.uk).

[1]My thanks to Dave Giles, James Grant, Magnus Henrekson, Johannes Hirata, Alice Nakamura, Peter Sinclair and Geoff Tily for comments on an early draft, and to participants at the conference of the IARIW at the OECD, Paris, April 2015. I would also like to thank the referees for their careful attention and comments. All responsibility is of course mine. Diane.Coyle@manchester.ac.uk

[2]http://www.telegraph.co.uk/finance/economics/11505763/UK-economy-grew-at-fastest-rate-for-nine-years-in-2014.html and http://www.theguardian.com/business/2015/mar/31/uk-gdp-growth-revised-up-to-06. Accessed 04/04/15.

[3]See for example: for the U.S. http://www.gallup.com/poll/176042/trust-mass-media-returns-time-low.aspx; http://www.gallup.com/poll/5392/trust-government.aspx; http://www.people-press.org/2014/11/13/public-trust-in-government/; for the EU and UK https://www.ipsos-mori.com/researchpublications/researcharchive/3504/Politicians-trusted-less-than-estate-agents-bankers-and-journalists.aspx; http://www.gfk.com/news-and-events/press-room/press-releases/pages/gfk-verein-global-study-on-trust-in-professions-.aspx.

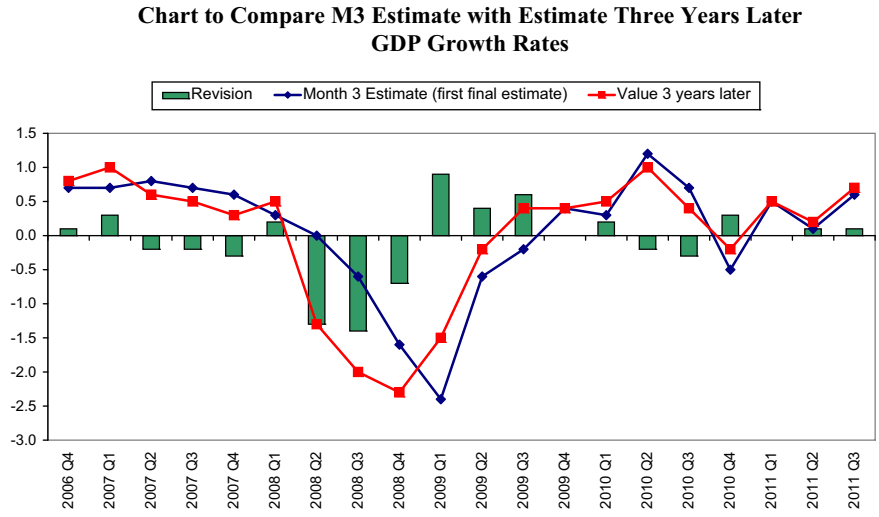

Figure 1. GDP quarterly growth revisions, between 3rd month estimate and 3 year estimate [Colour figure can be viewed at wileyonlinelibrary.com]

Source: ONS (https://www.ons.gov.uk/economy/grossdomesticproductgdp/datasets/revisionstrianglesforukgdpabmi).

skepticism about economic statistics is reflected in a renewed debate about whether they are painting an accurate picture of a changing economy, or an adequate portrait of the economic health of the nation. For example, recent recommendations in a Dutch parliamentary report and in an influential independent review of the U.K.'s economic statistics by Sir Charles Bean, have highlighted both the inadequate capture of the digital economy in existing statistics, and the debate about going 'Beyond GDP' (Bean, 2016; Tweede Kamer, 2016.)

Cynicism is not new—older examples include the deep distrust British citizens had for unemployment figures during the 1980s and 1990s, or some Americans for the Boskin *et al.* (1996) view that inflation was overstated because of a failure to take into account quality improvements in computer equipment (Boskin *et al.*, 1996). Indeed, Oskar Morgenstern much earlier noted the same phenomenon of distrust: "The professional users of economic and social statistics, strangely enough, often seem to be less skeptical than the public" (Morgenstern (1950, 1963)). Statistics have always been of huge interest to governments, considered in the past as one of the soft weapons of warfare, and now as a sign of the success or failure of economic policies. Independent and reliable official statistics are a public good in democratic, information-based economies.

However, the weaving of statistics such as GDP growth rates into political or otherwise slanted stories distracts attention from the inevitable margin of uncertainty on any official statistics—and this includes most economists. The already-mentioned ONS bulletin on the third estimate of UK GDP in the final quarter of 2014 makes it clear the figures will be revised later:

> "All estimates, by definition, are subject to statistical uncertainty and for many well-established statistics, ONS measures and publishes the sampling

> error associated with the estimate, using this as an indicator of accuracy. The estimate of GDP, however, is constructed from a wide variety of data sources, some of which are not based on random samples and as such it is very difficult to measure the sampling error. While development work continues in this area, ONS like all other G7 national statistical institutes does not publish a measure of the sampling error associated with GDP."[4]

The typical revisions of 0.1 or 0.2 percentage points sound reassuringly small—although even such small changes are enough to generate dramatic headlines and stories about the health of the economy. The actual absolute revisions can be much larger, especially at turning points in the business cycle. Figure 1 shows for the period of the financial crisis the ONS revisions to quarterly GDP growth rates between the third month's estimate and the estimate three years later.

There was substantial debate in the 1950s about unavoidably judgmental treatment of sectors such as government and finance (Studenski, 1958). Oskar Morgenstern wrote:

> "Statements concerning month-to-month changes in the growth rate of the nation are nothing but absurd and even year-to-year comparisons are not much better. The same applies to variations in price levels, costs of living and many other items. It is for the economists to reject and criticize such statements which are devoid of all scientific value, but it is even more important for them not to participate in their fabrication."

The issues of consistency and reliability are also addressed by André Vanoli in his history of national accounting (Vanoli, 2005).

However, economists and commentators do not act on the knowledge of the uncertainty and scale of revisions to GDP figures. Charles Manski has pointed out that sampling error—albeit not published with the GDP data anyway—is not the only source of potential error in national accounts statistics (Manski, 2015). There are potentially many others, processing errors, faulty assumptions in data gathering, variability in the data, incorrect or incomplete or misleading survey responses, and so on. Manski categorises these errors as transitory statistical uncertainty (due to the fact that data collection takes time and will at first be incomplete); permanent statistical uncertainty (due to finite samples, or provision of inaccurate data by respondents); and conceptual uncertainty, (due to the fact that the statistics do not mean what users think—seasonal adjustment is his example here, chain weighted price deflators could be another). He favors publishing ranges rather than point estimates in an effort to educate users about the high degree of uncertainty:

> "In the absence of agency guidance, some users of official statistics may naively assume that errors are small and inconsequential. Persons who understand that the statistics are subject to error must fend for themselves and conjecture the error magnitudes. Thus, users of official statistics may misinterpret the information that the statistics provide."

[4]http://www.ons.gov.uk/ons/rel/naa2/quarterly-national-accounts/q4–2014/index.html accessed April 13 2015.

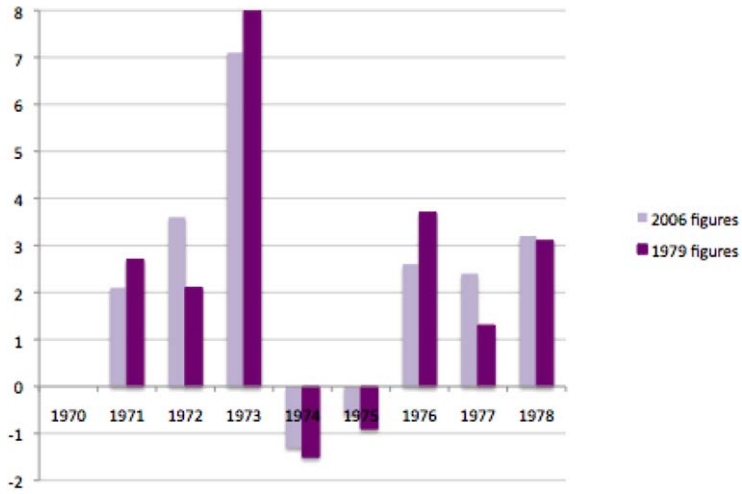

Figure 2. Annual % change in real GDP, 1979 and 2006 vintages [Colour figure can be viewed at wileyonlinelibrary.com]

Source: http://www.measuringworth.com/datasets/UKdata/UKGDPs.pdf

Another example of the source of uncertainty is time aggregation, the fact that the data are quarterly when both the underlying statistical series and the economic decision-making could be monthly or weekly. David Giles has shown that the dynamic behavior of national accounts time series apparently revealed by regressions can often be the result of aggregation into discrete time periods, which also affects the way hypothesis and specification tests can be interpreted (Giles, 2014).

Yet while the statistics are frequently revised, and some users show they are aware of this, more often the stories economists (and commentators and politicians) tell display unwarranted certainty. This is certainly true of the overarching political narratives. Enrico Berkes and Samuel Williamson have created a database of U.K. GDP statistics consisting of the contemporary figures describing the path of growth from successive ONS publications (Berkes and Williamson, 2015). Their aim is to understand the lens through which the economic situation was interpreted at the time. The contrast between the contemporary statistics and the latest figures can be startling. For example, the 2012Q4 vintage of national accounts data reveal 7 recessions (defined as two consecutive quarters of negative growth) between 1955 and 1995, whereas the 1996Q4 vintage data show there were ten recessions (28 quarters as against 20 quarters.)

The revisions to headline GDP growth figures have some potential to be politically significant. Using the Berkes/Williamson data set, Figure 2 compares annual real GDP growth rates for the 1970s as published in 1979 and 2006. While this is a period of relatively small revisions between vintages of data, the contemporary figures show a more extreme boom-bust cycle than the later figures. Of course, people were at the same time experiencing high inflation, rising unemployment and large-scale public sector strikes, but the path of real growth in GDP over the decade was less variable than it seemed at the time, ahead of the 1979

General Election. The paper concludes that: "Elections that were won by the Labour party are characterized by real-time GDP growth figures that are skewed towards the maximum of the vintage distribution. On the other hand, elections won by Conservatives are characterized by real-time growth figures that are skewed towards the minimum," although this holds only up to 1997.

The Berkes/Williamson paper considers another example, looking at contemporary economic research, namely the "five tests" in 1997 of the U.K.'s readiness or otherwise to join the Euro. The Treasury-commissioned research at the time looked at the correlations in business cycles between the U.K. and other countries. Berkes and Williamson repeat the method with the latest vintage of data and conclude:

> "[A] researcher [who] would study the synchronization of French and British business cycles between 1960Q1 to 1997Q4 using the 1999Q4 vintage would conclude that the U.K. and the U.S. show a much higher degree of synchronization than France and the U.K. However, the same researcher performing the same analysis using the 2012Q4 vintages would end up concluding that business cycles in the U.K. are as synchronized with the French business cycles as much as they are with the U.S. ones."

This discussion has referred only to revisions (and rebasing) but changes in methodology—especially the changes to price indices used to create the real GDP growth figures—have also redrawn the broad contours of history, and by implication the stories economists tell about the way economies work. Angus Maddison made this point about the introduction of chain weighting: "Acceptance of the new measure for this period [i.e. applying chain weights to pre-1950 data] would involve a major reinterpretation of American history," he wrote (Maddison, 2001). The statistics have changed but the rewriting has not and does not occur. In the vast amount of empirical economic research, based on the national accounts data that can so easily be downloaded now, how much has been re-done after significant revisions or changes in methodology applied to historical data? No doubt there are examples, but it is not the habit of economic researchers to repeat earlier work even when the data change, part of the wider failure of the profession to pay much attention to measurement problems or replicate significant results, in contrast to some other sciences.

There are of course exceptions. For example in his AER Presidential Address, discussing the economy's productivity performance, the main point made by Zvi Griliches was: "Our understanding of what is happening in our economy (and in the world economy) is constrained by the extent and quality of the available data" (Griliches, 1994). Some sophisticated users of the statistics do also take statistical uncertainty into account. One example of this is the Bank of England's "fan chart" for GDP growth, which includes forecast uncertainty in the forward projection, and uncertainty about revisions in the historical series, showing 30 percent, 60 percent and 90 percent confidence intervals. In its latest (February 2017) version of the chart, the Bank is 90 percent confident that year-on-year real GDP growth is somewhere between about 0.5 percent and 4.5 percent.[5] This

[5]Bank of England Inflation Report February 2015. Another way of grasping the range is to consider that it implies living standards will double either every 70 years or every 14 years.

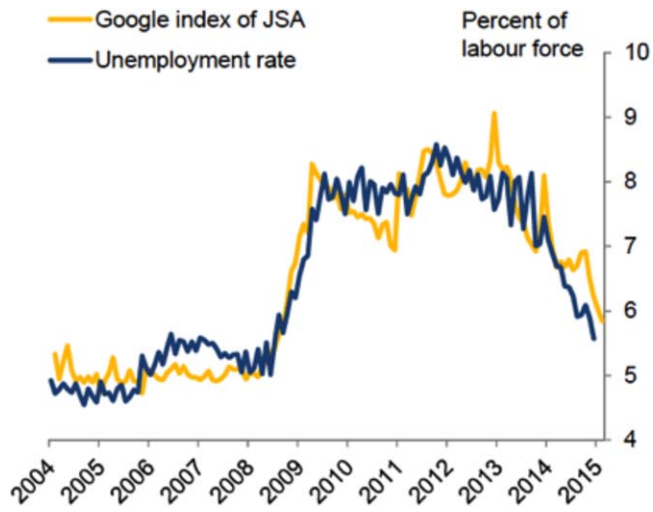

Figure 3. Google search versus official statistics [Colour figure can be viewed at wileyonlinelibrary.com]
Source: Bank of England Quarterly Bulletin (2011, Q2, pp. 134–140)

is of course a very large range for a growth rate estimate, the difference between living standards doubling in approximately 140 rather than 16 years.

Few economic researchers invest the time required to truly get to grips with the national accounts. Both Allin and Hand (2015) and Osterwald-Lenum (2015) point out that these statistics have become too complicated, and that their compilation is a producer-driven activity rather than a user-driven one. This is not to absolve economists. The demand for certainty about the statistics—from economists themselves and still more from the final users of statistics in the public policy debate—is strong, even when that certainty is non-credible. As noted, the role of national accounts statistics in the public debate is to enable policy makers to tell their story and be held to account. It is perhaps not surprising that a supply of certainty emerges to meet the demand, although it is disappointing.

However, neither the degree of underlying uncertainty nor the everyday practice of ignoring it seems sustainable, even though this situation has lasted for decades. There are two forces driving change. One is the way digital technology appears to be changing the character of the economy in ways that make it increasingly difficult to measure according to existing national accounts conventions (Coyle, 2015). The other is the revived interest in taking better account of sustainability and broader well-being, captured in the "Beyond GDP" discussions taking place in the European Commission and OECD as well as many individual countries. To consider how to ensure the statistics can serve their purpose in the democratic debate, it is necessary to revisit the question of the purpose of the national accounts—what are they supposed to measure?

2. The Purpose of the National Accounts

The fundamental purpose of economic statistics has shifted during the capitalist era, the present standards being the product initially of the needs of the

Great Depression and Second World War, then co-evolving with post-war Keynesian macroeconomics (Coyle, 2014). The original purpose of official statistics was to serve the needs of the state, hence the etymology of the word. But in the modern era, the needs of the people must guide the work of statisticians. Official statistics are for citizens, not for officials.

Of course, one use of economic statistics is tracking the macroeconomic conjuncture. It is hard to imagine how macroeconomic policy could be implemented without aggregate statistics such as GDP. However, as the Bean Report pointed out, gathering suitable statistics for that purpose might be on the verge of becoming faster and cheaper. One example explored by the Bank of England is the correlation between official unemployment statistics and the Google Trends series for searches for the term "Job Seeker's Allowance," shown in Figure 3 (McLaren and Shanbhogue, 2011.) The correlation is high, and although not perfect, a cost-benefit assessment might favor the use of Google searches. In practice, the commercial imperatives of a search company are completely different from the public service requirements of official statistics, including unbiased sampling, and open access to the data. However, the potential productivity gain in the gathering of economic statistics from the use of new technology alternatives is high indeed, including in terms of increased accuracy and timeliness. This is surely also attractive in the face of budget cuts; although like all large improvements in productivity it would require substantial and possibly uncomfortable change in how things are done. More important, the ability of statistical offices to use online and scanner data sources would require an updated legal framework to ensure companies do not manipulate data, and to ensure official statisticians have access to raw data and adequate publication rights. The full ramifications of the public goods character of information in a world of private information monopolies have yet to unfold; the use of data of this kind for official statistical purposes is one important aspect.

For macroeconomic policy and research, it is hard to envisage doing without GDP or a similar aggregate. However, it is well known that GDP measures activity at transacted prices and therefore does not include the consumer and producer surpluses familiar from the microeconomic analysis of welfare. Yet in normal usage, and often in professional debate, real GDP growth is taken as shorthand for progress or an improvement in social welfare. Sometimes, economists strongly caution against doing so, pointing out that GDP should not be taken as a measure of economic welfare, but simply a measure of economic activity at market prices. There was a significant debate in the late 1930s and early 1940s about whether the aggregate measure of the economy then being developed should explicitly account for aspects of welfare or not; Simon Kuznets thought it should, arguing for example for removing "bads" such as spending on advertising. He lost the debate, although of course there were many judgments to be made about where exactly to locate the production boundary.

Still, GDP growth is widely used as a progress indicator. There is indeed a strong correlation over time between GDP growth and a range of indicators of welfare such as health, longevity and education (van Zanden *et al*., 2014). As soon as we move from the nominal GDP figures to real GDP, and especially with the use of hedonic price indices, or as soon as we introduce PPP conversions, we are clearly interested in living standards or social welfare in some sense. The aim,

no matter how implicit, must be to get to a measure of purchasing power, command over the use of resources (Hirata, 2015).

There is the separate question of whether it is the level or rate of growth of GDP that either measures or correlates with welfare. Geoff Tily (2015) argues that the national accounts were developed:

> "To support policy: to resolve the unemployment crisis of the Great Depression and to aid the deployment of natural resources to their fullest possible extent for the conduct of the Second World War. … It is fundamental to recognise that these theoretical and practical initiatives were aimed at the *level* of activity."

The attention of policymakers turned to growth year after year only from the late 1950s on, with a milestone in 1961 when the OEEC become the OECD and agreed a target of 50 percent GDP growth for 1960–70 (Schmelzer, 2016).

The distinction between setting a policy target in terms of levels or growth rate is interesting because some environmentalists advocate zero GDP growth on sustainability grounds. Advocates of happiness metrics point to some evidence of the breakdown in the (level of) per capita GDP-life satisfaction correlation at some income level, while later work has noted there is nevertheless a correlation between GDP per capita *growth* and life satisfaction. This finding of a positive relationship has been challenged in turn by the originator of the happiness literature, Richard Easterlin and others (2010). The question seems to be this: is it the case that how people feel about their life has nothing to do with the goods and services measured by GDP once certain basic needs have been satisfied; or rather does life satisfaction increase, albeit much less than proportionately, as the things measured by GDP continue to grow? Addressing this, one needs to bear in mind that most of GDP (and all of its increment in advanced economies) is now non-material (although energy-consuming) and so hard to measure, and that it includes new products and services whose unmeasured consumer surplus is likely to be high. More time series regressions on existing data will not resolve the debate.

There is no clear theoretical link between GDP as currently defined and the consumer (and producer) surplus created by innovation, although the empirical link—at least in the long run—is intuitive and clear. I have argued elsewhere that there is a growing wedge between GDP and consumer surplus because of the increase in the variety of goods and services, and because of the economic characteristics (non-rivalry and zero-marginal cost) of the important and growing category of new digital goods and services (Coyle, 2014). Others have pointed out that the basic data collection anyway heavily emphasizes manufacturing rather than the services that now constitute the bulk of developed economy GDP (Delong, 2000). This bias towards tangibility also manifests itself in the absence of quality adjustments in many services, logistics and construction.

The conflicting needs of macro and micro measurement are not new but arguably more acute due to the likelihood that technical innovation is not being fully measured in GDP (although the extent of the mismeasurement is hotly debated) and the possibility that the wedge between GDP and a welfare-based approach has been growing, due to the increasing (private and social) social returns character of digital goods and services (Bean, 2016; Byrne *et al*., 2016).

3. Taking the Conversation Forward

In modern democracies we surely do want national statistics that enable citizens to hold policymakers to account for a wider perspective on economic welfare, rather than simply aggregate economic activity. Statistics shape the boundaries of what is politically possible. They originated in the development of the modern, administrative nation state, and have a strongly performative character (Porter, 1988; Desrosieres, 2002.) The ideal indicators should have the following characteristics: they should be linked to the kinds of levers available to policymakers or to outcomes policy can plausibly affect; they should be available as consistent time series and in a timely enough manner that there is some meaningful attribution of outcomes to policy decisions; they should be not-too-complicated and reasonably intuitive.

One much-debated question is whether or not it is preferable to have a single index rather than a suite of indicators, of dashboard. There is clearly some desire for a single indicator, given the number of GDP-alternatives that have been produced from time to time, often as an adjusted version of GDP. Alternatives such as the 'Genuine Progress Indicator' or the well-known ISEW deduct "regrettables" from GDP such as inequality, crime, pollution and so on. These alternatives invariably show progress halting in about 1973. This has been taken as a sign of the diminishing "welfare productivity" of GDP (Offer, 2000). However, that welfare has stagnated or even declined since the 1970s seems wholly implausible, not only because of the extraordinary innovations that have occurred since, but also because of improvements in the quality of housing and many everyday goods and services. The extant alternative indices might measure sustainability in some sense but do not measure social welfare.

The best argument for going down this single alternative index route is the public salience of a headline indicator going up or down. While this political economy argument certainly has some appeal, the strong counter-argument is that summing all the dimensions of social welfare to single index will be unsatisfactory (Hirata, 2015). It is not only that there are several incommensurate dimensions, but also that there are difficult trade-offs between them. The obvious one, submerged within GDP as in all other single indicator alternatives, is the trade-off between present and future. Without being explicit about this, it will never be possible to assess the sustainability of current economic activity.

The problem with dashboards, apart from the question of how much public traction they might have, is that there is a strong temptation to pile more and more indicators into them. The Better Life Index has 11 topic headings. A new entrant to the dashboard field, the Social Progress Indicator (which does not even include any economic categories such as employment or income), has 54 components. They are all plausible as elements of social welfare, but the sheer number highlights the difficulty of creating a parsimonious dashboard, preferably consisting of indicators on which many countries could agree so that international standards might be developed.

Parsimony requires some structure or theory. Alternative social welfare approaches—and Amartya Sen's concept of capabilities is probably the most-often advocated—lead anyway to the need for a range of indicators to capture the

incommensurable dimensions of well-being. One way of determining or limiting the number of potential indicators in this range would be to look at the empirical evidence in the well-being literature, which provides some apparently robust results about the contributors to well-being at the individual level. Not all of the factors are or ought to be the subject of economic policies, but others clearly are. Employment is one of these, beyond the income it provides. Other candidates with policy implications would be health (especially mental health), the local environment, commuting time (Frey and Stutzer, 2010). Another solid part of the empirical evidence is the importance of relative status and positional goods in determining individuals' happiness, which points to the need for distributional indicators in any dashboard (Frank, 2008).

This approach would be pragmatic and might help determine a reasonably parsimonious set of dashboard components. However, it does not help address some key sets of questions. Here are (at least) seven—some previously raised, for example in the Sen-Stiglitz-Fitoussi Commission (European Commission, 2009):

a) Utilitarianism and methodological individualism are the philosophical underpinnings of the standard approach to welfare evaluation in economics. The fundamental welfare theorems derive from the aggregation of individuals' utility maximization problems, albeit that economists gloss over two decisive objections (even accepting the empirically doubtful assumption that preferences are fixed): the fact that aggregate welfare and distribution cannot in fact be separated as the basic textbooks claim (Scitovsky, 1976); and the fact that complete markets (over all future goods and states of the world) do not exist. The "happiness" economics advocated by Richard Layard and others are even more explicitly utilitarian (Layard, 2005). The advantage of the standard methodology from the perspective of aggregate economic statistics is that it provides a theoretical basis for the calculation of a single number by adding up individuals, whether that is Gross Domestic Product or Gross National Happiness. The disadvantage is the way it has rooted the concept of social welfare so profoundly in methodological individualism (with individual utilities aggregated by—someone, a benign but external entity) (Coyle, 2012a; 2012b), when the concept of prosperity built on specialization and the division of labor means individuals are inescapably mutually dependent. Adam Smith's assertion that people benefit society by acting in their individual self-interest is at the heart of the modern machinery of welfare economics; yet the resulting total social economic welfare is more than the sum of the parts. For centuries, prophets and poets have told us that, "No man is an island/Entire of itself/Every man is a piece of the continent,/A part of the main." Now psychologists and biologists have supplemented the poetry with empirical evidence. And the more complex our advanced economies based on extended global supply chains become, the more the aggregate social welfare outcome will depend on the degree of *interdependence*. There is an interesting apparent tension between the summing of individual utilities and the gains from

specialization, which perhaps comes to a point in the question of how to aggregate. In their Atlas of Economic Complexity, Ricardo Hausman and Cesar Hidalgo have documented the correlation between the variety of goods and services a country trades, and its trading links, and the level of GDP per capita. It would be useful to measure an economy's degree of specialization in a way explicitly linked to the social welfare benefits that specialization delivers; the complexity index does so implicitly.

b) Is it possible to account for the contribution of innovation to social welfare? It is clear that over time innovations both large and small have made the biggest contributions, from new medicines and public health discoveries, to the sequence of general purpose technologies, to everyday and incremental innovations that make life easier and pleasanter. Hans Rosling has nominated the humble washing machine as the single most important innovation of modern capitalism because of the amount of women's time it freed up. As already noted, it is not clear what link (if any) there is between current national accounts statistics and the consumer and producer surplus created by the cornucopia of innovations that characterizes the market economies of the past 250 years (Delong, 2000). It is particularly difficult to do so for the zero marginal cost, free to the consumer, digital innovations of recent times. There is some evidence that the consumer surplus created by new digital services is large (see for example Brynjolfsson *et al*., 2003; Greenstein and McDevitt, 2012). Nor do we know what to make in terms of aggregate measurement of the evidence that at the level of individual decision-making people can experience a "paradox of choice" (Schwartz, 2004). The proliferation of variety in the advanced economies needs explanation if people (in the aggregate) do not want it, even if as individuals we do not want "too much" choice. This might be another question mark over aggregation by adding up individual choices.

C) If the practice in current use of national accounts data and the aim in potential future use of dashboards is to measure social welfare, what are statisticians to make of the fact that so much welfare is created outside the market? Herbert Simon once famously said that if a Martian were to observe society, market transactions would clearly be a minority of activities, with most occurring within non-market institutions—firms, societies, households, not to mention leisure (Simon, 1991). It is possible of course to calculate imputed values for leisure or household production. But if more than half of social welfare arises from non-market activities, then as Avner Offer has commented: "This salience of non-commodities casts doubt on the welfarist assumption that all well-being can be priced" (Offer, 2000). A dashboard can avoid this potentially crippling doubt by not trying to evaluate all contributors to social welfare in the same metric (in which case a price metric is as good as any). However, this approach does introduce additional trade-offs to be considered and made part of the democratic conversation, not least that between leisure and/or household production and income.

d) If rooting a dashboard in a capabilities approach to social welfare, how is mutual dependence via specialization in general, and access to public or collective goods in particular to be accounted for? As well as the pure public good arguments, Ricardo Hausmann has argued recently that people who are poor face high fixed costs of access to the networks essential for economic betterment, including social networks, but also infrastructure (Hausmann, 2014). So there is also an important distributional aspect to this. If you do not have any or much private capital, the welfare value of public capital is likely to be greater. An implicit recognition of this underlies the principle of universality. An explicit accounting for public goods benefits would be desirable. This is all the more desirable because so many new digital goods have the public good characteristics of non-rivalry and high fixed costs or network costs, even if technically excludable.
e) How should sustainability be captured in a dashboard? Specifically, do we know enough about the depletion of natural capital and its relationship to the national accounts data that exist at present? Is it sensible to develop and use a single, aggregate natural capital measure, or would it be better to disaggregate to some degree? Can a dashboard approach present the tradeoff between current consumption or well-being and future states, or does that simply depend too much on a range of normative assumptions—in which case, would it be preferable to include some statistics already available but to do so in a more straightforward and accessible way? Should sustainability indicators be confined to environmental ones or also include more scope for scrutiny of other balance sheets, especially the state's, perhaps widely defined to include contingent liabilities? Can the existing environmental "satellite" accounts combined with the national accounts form the basis for a more satisfactory sustainability accounting? (Fleurbaey and Blanchet, 2012; Karacaoglu, 2015.)
f) Is the nation state going to remain the best basis on which to build a dashboard of statistical indicators? For example, how can the societal implications of the specialization in global supply chains be measured? Or the role of global online services whose location is difficult to pin down? Or the dependence of sustainability at the national level on global environmental developments?

This is an array of extremely difficult issues. It took decades for the System of National Accounts to be developed to its present level of sophistication. So we can expect it to take a long time to reach anything like a settled professional consensus on what an economic dashboard should include, especially if some of these questions do turn out to raise rather fundamental issues about welfare economics. What then are the next steps the path toward statistics that will help citizens to monitor and judge the performance of policymakers in our democracies?

Both the political stories given as examples in the introduction—both the increasing momentum in the U.K. economy at the start of the General Election campaign and the fact that the recovery from the financial crisis has been slow and lackluster—will tally with many people's sense of reality. Personal circumstances will determine which version speaks more closely to any individual's

experience. There is nevertheless a steady and perhaps growing chorus of discontent with the reliance on real GDP growth as the thermometer of economic progress. Explanations might be found in some of the indicators already collected (for example in the new Economic Well-being publication from the Office for National Statistics), such as income distribution indicators, and Net National Disposable Income per capita as well as GDP per capita.

However, public interest in either an alternative to GDP or a more rounded "beyond GDP" view of the economy will be latent until both the professional debate and the political conversation conducted via the media reach a consensus about the need to switch to an alternative measurement standard. Even if a majority of the public, the journalists and the politicians agreed on the need to make the switch, there would be a co-ordination problem.[6] One argument for a single indicator as an alternative to GDP is precisely because of this problem, the case being that it is more realistic to switch to a single indicator than can be presented as a better version of GDP than to adopt a dashboard. If dashboards tend to accumulate indicators (as they will as long as they are atheoretic), this pragmatic argument will be a strong one. The existing examples of well-being publications or dashboards present a large amount of data in ways that are hard to interpret and do not (yet) include the time series that are necessary to hold policy to account.

A practical alternative might be to ask people what they think should be included in a small dashboard, through a public consultation. Some statistical offices such as Australia and New Zealand have undertaken important consultation exercises. Campaign groups have also concluded public consultation is the best method of selecting indicators relevant to well-being.[7] One risk in public consultation is that the salience of indicators is partly determined by previous narratives from a particular political or ideological perspective. There is some evidence that voting habits are determined by viewing habits (for example, Facchini *et al*., 2009; Martin and Ali Yurukoglu, 2014.) Nevertheless, the principle of involving the public in the selection of indicators is attractive. Indeed, it seems fundamental to devising a set of statistics that would facilitate the democratic conversation.

A conversation is two-sided, and another aspect of improving the democratic debate would be to increase statistical literacy in general. The term "GDP" is for most people an incantation they hear on the news without any real understanding of what it means or even what the letters are short for. Economists teaching in universities bear a special responsibility for ensuring none of their students graduate without the practical ability to read and understand a press release from the statistical office (as well as understanding the specific data, and its limitations, in any of the practical work they do in the course); it is a responsibility very often unfulfilled.

4. Conclusions

This paper has explored two kinds of problems with GDP and other national accounts statistics.

[6]For a discussion of the political economy challenge, see Coyle (2017).

[7]One example is the New Economics Foundation: http://www.neweconomics.org/blog/entry/labours-living-standards-index-does-it-go-far-enough Accessed April 13, 2015.

One concerns the failure of everyone—statisticians and economists as well as politicians, media and public—to acknowledge the extent of the inevitable uncertainty about the published statistics. "Stories" about the economy might be without any substance, yet are confidently told. There is a particular responsibility on economists as users and interpreters of the statistics, and as teachers, to pay more careful attention to the limitations of "data" and express due caution.

The second concerns whether GDP and the national accounts still do a good enough job in tracking economic welfare (which they are used for, albeit with long-understood limitations.) My answer is that they do not. The longstanding and now acute concern about sustainability, the implications of digital change for measuring the economy, and a growing public and political interest in going "Beyond GDP", all point to the need for a new debate about what exactly 'beyond' consists of. This raises some very difficult conceptual and practical questions, sketched above. There are some major challenges to both the production of statistics and to the analysis of welfare economics underpinning aggregate statistics. This is likely to take both considerable research effort and institutional and political co-ordination, but the effort is necessary if the national accounts of the future are to serve the public good.

References

Allin, P. and D. J. Hand, "From a ystem of National Accounts to a Process of National Wellbeing Accounting", IARIW-OECD conference April 2015, (accessed April 3, 2015).

Atlas of Economic Complexity, http://atlas.cid.harvard.edu/

Bean, C., *Independent Review of Economic Statistics*, March 2016 https://www.gov.uk/government/publications/independent-review-of-uk-economic-statistics-final-report

Berkes, E. and S. H. Williamson, *The Impact and Interpretation of Post War Revisions in the Official Estimates of GDP for the United Kingdom*, 2015, http://www.measuringworth.com/datasets/UKdata/UKGDPs.pdf (accessed April 18, 15).

Boskin, M. J. et al., "Toward a more Accurate Measure of the Cost of Living," United States Social Security Administration, https://www.ssa.gov/history/reports/boskinrpt.html#list (accessed June 28, 2017). Orignial date of Publication: December 1996.

Brynjolfsson, E., M. Smith, and Y. Hu, "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science*, 49, 11, 2003.

Byrne, D., M. John, G. Fernald, and M. B. Reinsdorf, "Does the United States have a Productivity Slowdown or a Measurement Problem?" *Brookings Papers on Economic Activity*, 2016.

Coyle, D, "The Political Economy of National Statistics", in *Wealth*, ed. K. Hamilton and C. Hepburn, forthcoming OUP, 2017.

Coyle, D., "Modernising Economic Statistics: Why It Matters," *National Institute Economic Review* No. 234, 2015.

———, *GDP: A Brief but Affectionate History*, Princeton University Press, Princeton, NJ, 2014.

———, *The Social Responsibilities of the Economist*, Diane Coyle, Tanner Lectures on Human Values, Oxford, 2012a.

———, *The Economics of Enough: How to Run the Economy as if the Future Matters*, Princeton University Press, Princeton, NJ, 2012b.

Delong, J. B., Cornucopia: The Pace of Economic Growth in the Twentieth Century, NBER Working Paper 7602, 2000. http://www.nber.org/papers/w7602 (accessed April 13, 2015).

Desrosieres, A., *The Politics of Large Numbers: A History of Statistical Reasoning*, Harvard University Press, Cambridge, MA, 2002.

Easterlin, R., L. Angelescu McVey, M. Switek, O. Sawangfa, and J. Smith Zweig, "The Happiness–Income Paradox Revisited", *Proc Natl Acad Sci U.S.A.* 2010 Dec 28; 107, 52, 22463–8. Published online Dec 13, 2010. doi: 10.1073/pnas.1015962107, 2010 (accessed April 7, 2015).

European Commission, Report of the commission on the measurement of economic performance et social progress, 2009; GDP and beyond: Measuring progress in a changing world (COM(2009) 433 final), 2009.

Facchini, G., A. M. Mayda, and R. Puglisi, *Illegal Immigration And Media Exposure: Evidence On Individual Attitudes* ECARES, Université Libre de Bruxelles, Discussion Paper No. 7593 December 2009 Centre for Economic Policy Research (accessed April 4, 2015).
Fleurbaey, M. and D. Blanchet, *Beyond GDP: Measuring Welfare and Assessing Sustainability*, Oxford University Press, Oxford, 2013.
Frank, R. H., 2008. "Should public policy respond to positional externalities?" *Journal of Public Economics*, 92, 1777–86, 2008.
Frey, B. S. and A. Stutzer, *Happiness and economics: How the economy and Institutions Affect Human Well-Being*, Princeton University Press, Princeton, NJ, 2010.
Hausmann, R., 'The Economics of Inclusion', Project Syndicate, 7 November 2014, https://www.project-syndicate.org/commentary/inclusiveness-key-strategy-for-growth-by-ricardo-hausmann-2014-11
Hirata, J., How Should We Measure GDP? The Origin and Nature of a contested concept: Book review of Diane Coyle 'GDP: A Brief but Affectionate History', *International Review of Economics*. https://doi.org/10.1007/s12232-014-0222-8
Giles, D., *The Econometrics of Temporal Aggregation 1956–2014*, AWH Phillips Memorial Lecture, New Zealand Association of Economists Annual Conference July 2014.
Greenstein, S. and R. McDevitt, Measuring the Broadband Bonus in Thirty OECD Countries, OECD Digital Economy Papers, No. 197, OECD, 2012.
Griliches, Z., "Productivity, R&D, and the Data Constraint," *The American Economic Review*, 84, 1, 1–23, 1994.
Karacaoglu, G., *The New Zealand Treasury's Living Standards Framework–A Stylised Model* draft paper, April 2015
Layard, R., *Happiness: Lessons from a New Science*, Allen Lane, 2005.
Maddison, A., *The World Economy: A Millennial Perspective*, 79, OECD, Paris, 2001.
Manski, C. F., "Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years after Morgenstern," *Journal of Economic Literature*, 53, 631–53, 2015.
Martin, G. J and A. Yurukoglu, *Bias in Cable News: Real Effects and Polarization* December 22, 2014 http://polisci.emory.edu/faculty/gjmart2/ (accessed April 4, 2015).
McLaren, N. and R. Shanbhogue, "Using Internet Search Data as Economic Indicators," *Bank of England Quarterly Bulletin*, 51, 134–40, 2011.
Morgenstern, O., *On The Accuracy of Economic Observations*, Princeton University Press, Princeton, NJ, 1950; 1963.
OECD, *Measuring GDP in a Digitalised Economy*, N. Ahman, P Schreyer, Committee on Statistics and Statistical Policy, April 2016.
Offer, A., *Economic Welfare Measurements And Human Wellbeing*, University of Oxford Discussion paper, January 2000 (accessed April 3, 2015).
Osterwald-Lenum, M., "Sketch of Elements of a Measurement Theory of Economics as an Extension of the Current Sequence of SNA Manuals" Prepared for IARIW-OECD conference April 2015 (accessed April 3, 2015).
Porter, T., *The Rise of Statistical Thinking 1820–1900*, Princeton University Press, Princeton, NJ, 1988.
Schmelzer, M., *The Hegemony of Growth: The OECD and the Making of the Economic Growth Paradigm*, Cambridge University Press, Cambridge, 2016.
Schwartz, B., *The Paradox of Choice: Why More Is Less.* Harper Collins, New York, 2004.
Scitovsky, T., *The Joyless Economy: An Inquiry into Human Satisfaction and Consumer Dissatisfaction*. Oxford, 1976.
Sen, A., J. Stiglitz, and J.-P. Fitoussi, *Report of the Commission on the Measurement of Economic Performance and Social Progress*, 2009, http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report
Simon, H, A., "Organizations and Markets," *Journal of Economic Perspectives*, 5, 25–44, 1991, DOI: 10.1257/jep.5.2.25
Stevenson, B. and Wolfers, J. "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox," Brookings Papers on Economic Activity, 1–87, 2008.
Studenski, P. *The Income of Nations*, New York University Press, New York, 1958.
Tily, G. *The National Accounts, GDP and the "Growthmen*". January 2015. http://static1.squarespace.com/static/541ff5f5e4b02b7c37f31ed6/t/54b3afeae4b0d2480d43d760/1421062122575/CoyleReview_Tily.pdf (accessed April 13, 2015).
Tweede Kamer, *Mapping Welfare: Report of Temporary Committee*, April 2016 https://www.houseofrepresentatives.nl/sites/default/files/atoms/files/mapping_welfare_0.pdf (accessed May 10, 2016).
Vanoli, A., *A History of National Accounting*, IOS Press, 2005.
van Zanden, Luiten, J. J. Baten, M. Mira d'Ercole, A. Rijpma, C. Smith, and M. Timmer (eds), *How Was Life? Global Well-Being Since 1820*, OECD, 2014.