

MANCHESTER  
1824

The University  
of Manchester

**Economics**  
**Discussion Paper Series**  
**EDP-1506**

**Talking about the National Accounts: Statistics and the Democratic  
Conversation**

Diane Coyle

May 2015

Economics  
School of Social Sciences  
The University of Manchester  
Manchester M13 9PL

# Talking about the National Accounts: Statistics and the Democratic Conversation<sup>1</sup>

Diane Coyle, University of Manchester

May 2015

## 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, not just statistics.

Politicians use the published statistics to tell their story, to persuade voters and ‘stakeholders’ that some things and not others are true about the state of the economy, or the state of the nation. They are telling the story with the intention of promoting either their electoral prospects or an ideological perspective. Newspapers or news organisations or 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 indeed misinterpreted, to tell different stories. There are countless examples. To give just one from the UK General Election campaign in 2015, the third set of figures for 2014Q4 GDP, published on 30 March 2015, generated two contrasting stories: “UK economy grew at fastest rate for nine years in 2014,” and “Data shows slowest recovery since 1920s.”<sup>2</sup>

News about the latest statistics from all of these sources washes over a citizenry that is increasingly cynical about many of these mediators of economic statistics, when not simply overcome with indifference. 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.<sup>3</sup> The independence of

---

<sup>1</sup> Paper prepared for the conference of the IARIW at the OECD, Paris, April 2015. Forthcoming in the *Review of Income and Wealth*. I am grateful to Dave Giles, James Grant, Magnus Henrekson, Johannes Hirata, Alice Nakamura and Geoff Tily for comments on an early draft, and to participants at the conference. All responsibility is of course mine. Please send comments to Diane.Coyle@manchester.ac.uk

<sup>2</sup> <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.

<sup>3</sup> See for example: for the US <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->

national statistical offices is an important bulwark against cynicism about the statistics themselves, and highly desirable for those not at present independent from the political process (although accountability is important too). There are also new mediators emerging online, some affiliated to traditional media organisations and others independent organisations, ‘fact checking’ the claims or stories being told in political debate.

It is likely that many people notice the battle of the political narratives over official statistics, given that this is played out in mass media. Perhaps this contributes to the general distrust of politicians and media; perhaps it has also unfairly caused people to be cynical sometimes about the statistics themselves. Cynicism is not new – one example that leaps to mind is the deep distrust British citizens had for unemployment figures during the 1980s and 1990s, or some Americans for the Boskin Commission’s 1996 view that inflation was overstated because of a failure to take into account quality improvements in computer equipment.<sup>4</sup> Indeed, Oskar Morgenstern noted the same phenomenon of distrust in 1950.<sup>5</sup> But just as the importance of monetary policy autonomy has come to be recognised, so too there is a growing emphasis on the need for the independence of statistical offices from the political process – examples include the UK, since the establishment in 2008 of the UK Statistics Authority as a non-ministerial body with the Office for National Statistics (ONS) as its executive arm, or the status of Mexico’s INEGI as an autonomous agency since 2007.

Statistics have always been of huge interest to governments, often as one of the soft weapons of warfare as well as a tool of economic and social policy.<sup>6</sup> They should be seen now as a birthright of citizens, enabling them to hold politicians and officials to account. Independent and reliable official statistics are a public good in democratic, information-based economies. The move to give statistical offices independence in some countries recognises the role they play as one of the checks and balances of a healthy polity.<sup>7</sup>

---

[mori.com/researchpublications/researcharchive/3504/Politicians-trusted-less-than-estate-agents-bankers-and-journalists.aspx](http://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>.

<sup>4</sup> <http://www.ssa.gov/history/reports/boskinrpt.html>

<sup>5</sup> In *On The Accuracy of Economic Observations*, Princeton University Press 1950, 1963: “The professional users of economic and social statistics, strangely enough, often seem to be less sceptical than the public” (1963 edition, p 12)

<sup>6</sup> It is regrettable when political interference occurs, as in Greece in the early 2000s, or as in the removal of the long-standing independence of Statistics Canada in 2010. On Greece, see Diane Coyle, *GDP: A Brief but affectionate history*, Princeton University Press 2014.

<sup>7</sup> Report On Greek Government Deficit And Debt Statistics Brussels, 8.1.2010 COM(2010) 1 final; <http://www.savestatcan.ca/>

However, the weaving of statistics such as GDP growth rates into political or otherwise slanted stories is only one element of the state of uncertainty about the economy; most people do not notice the health warnings that come with 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 is clear (on its second page) that the figures could be revised later, saying the typical revision to the quarter on quarter percentage change between first and third estimates is 0.1 to 0.2 percentage points. It adds:

“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.”<sup>8</sup>

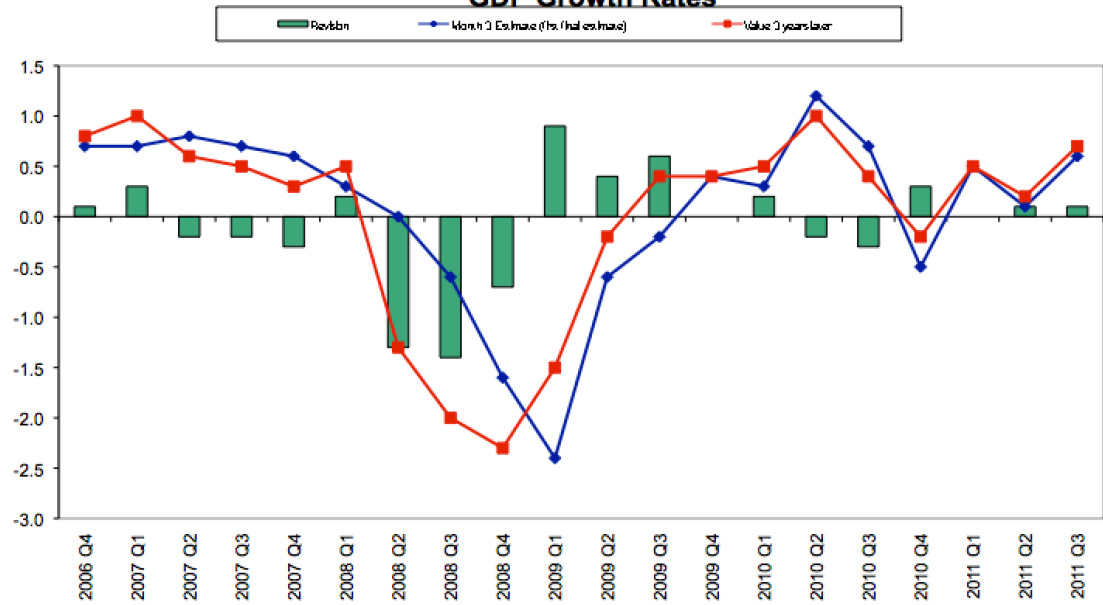
Revisions of 0.1 or 0.2 percentage points sound reassuringly small – although even such small changes are enough to generate dramatic headlines and stories about the health of the economy. The actual absolute revisions can be much larger, especially at turning points in the business cycle. The figure below 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.

**Figure 1: GDP quarterly growth revisions, between 3<sup>rd</sup> month estimate and 3 year estimate**

---

<sup>8</sup> <http://www.ons.gov.uk/ons/rel/naa2/quarterly-national-accounts/q4-2014/index.html> accessed 13 April 2015.

**Chart to Compare M3 Estimate with Estimate Three Years Later  
GDP Growth Rates**



Source: ONS

In a recent paper 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.<sup>9</sup> 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 suggests it is made too easy for users of statistics to claim “incredible certitude”, and favours at a minimum 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.”

<sup>9</sup> *Communicating Uncertainty In Official Economic Statistics: An Appraisal Fifty Years after Morgenstern* Charles F. Manski, Forthcoming, *Journal of Economic Literature*. Accessed 03/04/15.

The ranges needed might be large, but so be it, he argues. He quotes Oskar Morgenstern on GDP specifically:

“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.”

Sophisticated users of the statistics are well aware of the need to 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%, 60% and 90% confidence intervals. In its latest version of the chart, the Bank is 90% confident that year-on-year real GDP growth is somewhere between about 1% and 5%.<sup>10</sup> This takes account of past experience with data revisions.

I am not aware of significant effort to understand or remedy the data collection hurdles that lie behind sometimes large revisions. Perhaps they are occurring behind the scenes. Given the challenges revisions pose to all the users of SNA statistics, it would seem a sensible avenue to explore. No big business could operate with the scale of the uncertainty around its basic metrics. (I return later to conceptual questions.)

While the statistics are frequently revised, and occasionally users show they are aware of this, more often the stories are not. This is certainly true of the overarching political narratives. Enrico Berkes and Samuel Williamson have created a database of UK GDP statistics consisting of the contemporary figures describing the path of growth from successive ONS publications.<sup>11</sup> Their aim is to understand the lens through which the economic situation was interpreted at the time. The change in understanding required is 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).

---

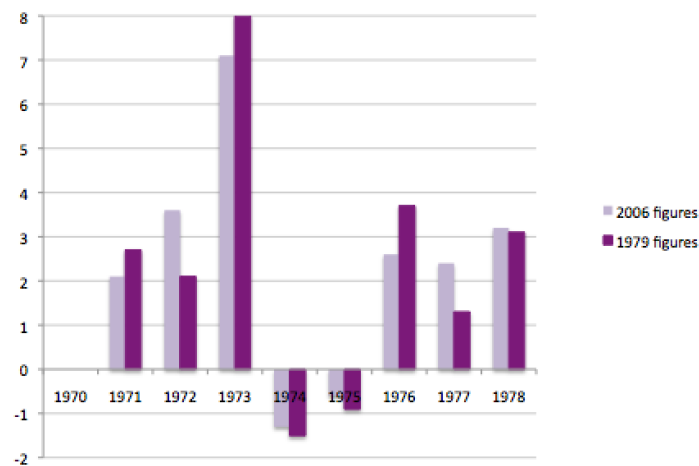
<sup>10</sup> Bank of England Inflation Report February 2015. Another way of grasping the range is to consider that it implies living standards will double either every 70 years or every 14 years.

<sup>11</sup> *The Impact and Interpretation of Post War Revisions in the Official Estimates of GDP for the United Kingdom* by Enrico Berkes & Samuel H. Williamson,

<http://www.measuringworth.com/datasets/UKdata/UKGDPs.pdf> accessed 18/04/15.

The revisions to headline GDP growth figures have some potential to be politically significant. Using the Berkes/Williamson data set, the figure below 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 conservative are characterized by real-time growth figures that are skewed towards the minimum,” but this holds only up to 1997. Nevertheless, the economic narratives are greatly affected by the contemporaneous GDP figures, even if the hold these narratives have over voting outcomes has diminished.

**Figure 2: Annual % change in real GDP, 1979 and 2006 vintages**



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

So much for the political story-telling. What about economic research, which one would expect to be more sophisticated and cautious? The Berkes/Williamson paper considers another example, looking at contemporary economic research, namely the ‘five tests’ in 1997 of the UK’s readiness or otherwise to join the Euro. The Treasury-commissioned research at the time looked at the correlations in business cycles between the UK 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 UK and the US show a much higher degree of synchronization than France and the UK. However, the same researcher performing the same analysis using the 2012Q4 vintages would end up concluding that business cycles in the UK are as synchronized with the French business cycles as much as they are with the US 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 [ie. applying chain weights to pre-1950 data] would involve a major reinterpretation of American history,” he wrote.<sup>12</sup> The statistics have changed but the rewriting has not and does not occur. I am not aware of any of the vast amount of empirical economic research, based on the national accounts data that can so easily be downloaded now, that 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.

In his recent book, Anthony Atkinson wrote: *“All too often economists race ahead, drawing conclusions from figures that happen to be there, without asking why the data are suitable.”*<sup>13</sup> He was writing about data on income inequality, but the point applies much more broadly. Economists should consider the character of the national accounts data far more carefully than they ever do. One example is 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 behaviour 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.<sup>14</sup>

Is such cavalier disregard for the need for care using the SNA data entirely the fault of economists? Largely, yes, and it is a carelessness encouraged by the ease with which statistics can be downloaded and fed into convenient software packages. However, I do have great sympathy with the point made by both Allen and Hand and also by Osterwald-Lenum that the national accounts statistics have become too complicated, and that it is now a producer-driven activity rather than a user-driven one.<sup>15</sup> It is not surprising that so few economic researchers invest the time required to truly get to

---

<sup>12</sup> Angus Maddison, *The World Economy: A Millennial Perspective*, page 79, OECD, Paris 2001.

<sup>13</sup> Anthony Atkinson, *Inequality: What can be done?*, Harvard University Press, 2015.

<sup>14</sup> *The Econometrics of Temporal Aggregation 1956-2014*, AWH Phillips Memorial Lecture, New Zealand Association of Economists Annual Conference July 2014.

<sup>15</sup> Paul Allin (Imperial College London) and David J. Hand (Imperial College London), ["From a System of](#)



grips with the national accounts. This is not to absolve economists at all. As Manski points out, the demand for certainty about the statistics – from economists and still more from the final users of statistics in the public policy debate – is strong, even when that certainty is non-credible. It is not surprising that supply emerges to meet the demand.

This line of argument does, though, point to the conclusion that the next steps for statistical offices do not involve ever more baroque evolutions of the SNA. We have already arguably reached a point of unusable complexity. If what we have now is more abused than used, even when statisticians point out the uncertainties, it is time to go in a new direction.

## 2. The purpose of the national accounts

This argument takes us back to the question of the purpose of the SNA. The fundamental purpose of economic statistics has changed several times during the capitalist era, the present approach being the product initially of the needs of the Great Depression and Second World War, then co-evolving with post-war Keynesian macroeconomics.<sup>16</sup> In each phase of history, the purpose of official statistics has been to serve the needs of the state, hence the etymology of the word. In the modern era of mass democracy in countries whose economies have grown well beyond global poverty levels, the needs of the people should guide the work of statisticians. Official statistics are no longer for officials, they are for citizens.

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. This is certainly true at present, when macroeconomic conditions in many countries remain distressed, and indeed uncomfortably reminiscent in some ways of conditions in the 1930s, which

---

[National Accounts to a Process of National Wellbeing Accounting](#)"; Michael Osterwald-Lenum (Statistics Denmark), "[Sketch of Elements of a Measurement Theory of Economics as an Extension of the Current Sequence of SNA Manuals](#)" Papers prepared for IARIW-OECD conference April 2015, accessed 03/04/15.

<sup>16</sup> See Diane Coyle, *GDP: A brief but affectionate history*, Princeton University Press 2014

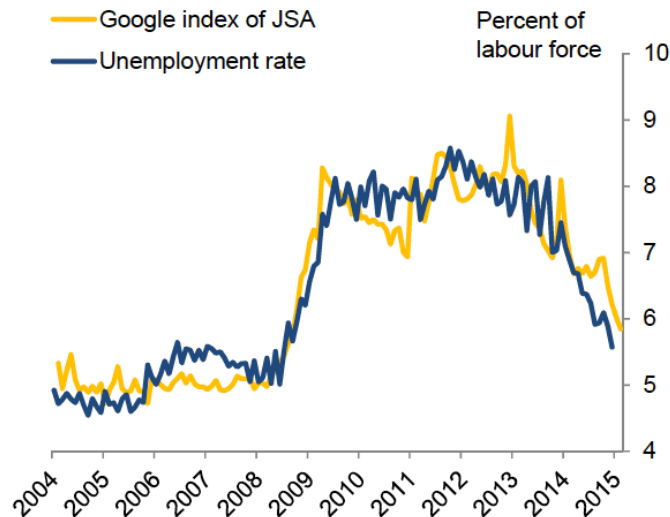
ultimately gave birth to the present SNA. The macroeconomic aggregates are essential to unravelling what is happening now.

However, as the Bank of England has pointed out, gathering suitable statistics for that purpose might be on the verge of becoming faster and cheaper.<sup>17</sup> One of the examples given 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 below. The correlation is high, and although not perfect, a cost-benefit assessment would apparently favour the use of Google searches, at least in this case (which shows the highest correlation of the examples presented in the Bank's research). In reality, 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 might require a 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.

**Figure 3: Google search versus official statistics**

---

<sup>17</sup> McLaren, N and Shanbhogue, R (2011), 'Using internet search data as economic indicators', *Bank of England Quarterly Bulletin*, Vol. 51, No. 2, pages 134–40



Source: Bank of England Quarterly Bulletin

So for macroeconomic policy and some macro research, it is hard to envisage doing without GDP. However, in normal usage, and often in economics too, real GDP growth is taken as shorthand for progress or an improvement in social welfare. In theory, economists strongly caution against doing so. It is often said that GDP should not be taken as a measure of social welfare, but simply a measure of economic activity at market prices – so often that I have repeated this point myself. 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 judgements to be made about where exactly to locate the production boundary.

Yet not only is the practice different – with GDP growth widely used as a progress indicator – but the claim that GDP is not in theory a welfare measure is also misleading. This is not just because in everyday political and policy conversation the latest growth figures are used to evaluate policy performance, although that is the case of course. Nor is because there is a strong correlation over time between GDP growth and a range of indicators of progress such as health, longevity and education.<sup>18</sup> In fact, we constantly make welfare judgements on the basis of real GDP. As soon as we move from the nominal GDP figures to real GDP, and especially with the move to hedonic price indices, or as soon as we introduce PPP conversions, we are clearly interested in living standards or

<sup>18</sup> "How Was Life? Global Well-Being Since 1820," edited by Jan Luiten van Zanden, Joerg Baten, Marco Mira d'Ercole, Auke Rijpma, Conal Smith and Marcel Timmer, OECD, 2014.

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.<sup>19</sup>

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 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.”<sup>20</sup>

The attention of policymakers turned to growth year after year only from the late 1950s on, he argues, with a milestone in 1961 when the OEEC became the OECD and agreed a target of 50% GDP growth for 1960-1970.

The distinction between setting a policy target in terms of levels or growth rate is interesting because it is growth that bothers many critics of the current policy focus. Some environmentalists favour zero GDP growth on sustainability grounds. Advocates of happiness metrics point to the breakdown in the (level of) per capita GDP-life satisfaction correlation at some point, 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.<sup>21</sup> Thinking about the concepts rather than the regressions, 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 that life satisfaction increases, albeit much less than proportionately, as the things measured by GDP continue to grow? Addressing this question, one needs to bear in mind that most of GDP (and all of its increment in advanced economies) is now non-material (although energy-

---

<sup>19</sup> I am grateful to Johannes Hirata for this. *How Should We Measure GDP? The origin and nature of a contested concept: Book review of Diane Coyle 'GDP: A Brief but Affectionate History'*, Johannes Hirata, International Review of Economics. <http://dx.doi.org/10.1007/s12232-014-0222-8>

<sup>20</sup> *The National Accounts, GDP and the 'Growthmen'*, Geoff Tily, January 2015. [http://static1.squarespace.com/static/541ff5f5e4b02b7c37f31ed6/t/54b3afeae4b0d2480d43d760/1421062122575/CoyleReview\\_Tily.pdf](http://static1.squarespace.com/static/541ff5f5e4b02b7c37f31ed6/t/54b3afeae4b0d2480d43d760/1421062122575/CoyleReview_Tily.pdf) Accessed 13/4/15.

<sup>21</sup> The happiness–income paradox revisited, Richard Easterlin et al, Proc Natl Acad Sci U S A. 2010 Dec 28; 107(52): 22463–22468. Published online 2010 Dec 13. doi: [10.1073/pnas.1015962107](https://doi.org/10.1073/pnas.1015962107)

consuming), and that it includes new products and services. More time series regressions will not resolve the debate.

GDP is an unsatisfactory measure of social welfare, nevertheless, and not only because of the increasingly complicated methodology needed to calculate a real chain weighted measure, or to make hedonic adjustments. 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. Economic historians such as Brad DeLong have previously pointed out that GDP under-accounts for the benefits of innovation, and 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. Others have pointed out that the basic data collection anyway heavily emphasises manufacturing rather than the services that now constitute the bulk of developed economy GDP.<sup>22</sup> This bias towards tangibility also manifests itself in the absence of quality adjustments in services data, yet the emphasis on hedonic price calculations for manufactured goods could give a misleading picture of which sectors of the economy are the most dynamic.

What's more, it seems the public is well aware that the use of GDP statistics is freighted with normative meaning, and is increasingly sceptical. There have long been objections to the GDP-centricity of economic policy, and quite a number of suggested alternatives – dating at least as far back as the Club of Rome. These alternatives are gaining significant traction in the media and in policy debates, and – if my experience is typical – also among students and members of the public. There is huge interest in 'happiness' or well-being as a policy aim and metric. And of course the Sen-Stiglitz-Fitoussi work and its follow-up, along with the European Commission's 'GDP and Beyond' and the OECD's Better Life Index are clear evidence of official interest in a different approach to measuring the economy, one very explicitly based on social welfare.<sup>23</sup>

### **3. Taking the measurement of welfare seriously**

---

<sup>22</sup> J Bradford DeLong, *Cornucopia: The Pace of Economic Growth in the Twentieth Century*, NBER Working Paper 7602, March 2000. <http://www.nber.org/papers/w7602> Accessed 13/04/15

See also Coyle, *GDP*, Chapter Six. On manufacturing and services, see <https://growthecon.wordpress.com/2015/04/29/there-is-more-to-life-than-manufacturing/> accessed 13/05/15

<sup>23</sup> [Report of the commission on the measurement of economic performance et social progress](#), 2009; European Commission, [GDP and beyond: Measuring progress in a changing world \(COM\(2009\) 433 final\)](#), 2009; <http://www.oecd.org/statistics/better-life-initiative.htm>

This is surely the right aim. 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.<sup>24</sup> In modern democracies we surely do want statistics that enable citizens to hold policymakers to account for social welfare rather than simply aggregate economic activity. 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.<sup>25</sup> 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.<sup>26</sup> 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.

---

<sup>24</sup> *The Rise of Statistical Thinking 1820-1900*, Theodore Porter, Princeton University Press 1988. *The Politics of Large Numbers: A History of Statistical Reasoning*, Alain Desrosieres, Harvard University Press 2002.

<sup>25</sup> *Economic Welfare Measurements And Human Wellbeing* Avner Offer, University of Oxford Discussion paper, January 2000, accessed 03/04/15.

<sup>26</sup> See Hirata, *ibid*.

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.<sup>27</sup> 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.<sup>28</sup>

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:

- 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 maximisation 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;<sup>29</sup> 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.<sup>30</sup> The advantage of the standard methodology from the perspective of aggregate economic

---

<sup>27</sup> B.S.Frey and A.Stutzer, *Happiness and economics: How the economy and institutions affect human well-being*, Princeton University Press 2010.

<sup>28</sup> Frank, Robert H., 2008. "[Should public policy respond to positional externalities?](#)" *Journal of Public Economics*, Elsevier, vol. 92(8-9), pages 1777-1786, August.

<sup>29</sup> Scitovsky (1976). *The Joyless Economy: An Inquiry into Human Satisfaction and Consumer Dissatisfaction*. Oxford.

<sup>30</sup> Richard Layard, *Happiness: Lessons from a new science*, Allen Lane 2005.

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),<sup>31</sup> when the concept of prosperity built on specialisation and the division of labour 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.<sup>32</sup> It would be useful to measure an economy's degree of specialisation in a way explicitly linked to the social welfare benefits that specialisation 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.<sup>33</sup> As already noted, it is not clear

---

<sup>31</sup> See *The Social Responsibilities of the Economist*, Diane Coyle, Tanner Lectures, Oxford 2012; and *The Economist As Outsider*, Diane Coyle, Pro Bono Economics Lecture, 2013.

<sup>32</sup> R Hausman, C Hidalgo, *Atlas of Economic Complexity*, <http://atlas.cid.harvard.edu/>

<sup>33</sup> Hans Rosling, The Magic Washing Machine,

[http://www.ted.com/talks/hans\\_rosling\\_and\\_the\\_magic\\_washing\\_machine?language=en](http://www.ted.com/talks/hans_rosling_and_the_magic_washing_machine?language=en)



what link (if any) there is between current national accounts statistics and the consumer and producer surplus created by the cornucopia of innovations that characterises the market economies of the past 250 years.<sup>34</sup> 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.<sup>35</sup> 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'.<sup>36</sup> 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.<sup>37</sup> 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."<sup>38</sup> A dashboard can avoid this potentially crippling doubt by not trying to evaluate all

---

<sup>34</sup> DeLong *ibid.*

<sup>35</sup> For two examples see: Brynjolfsson, Erik, Smith, Michael and Hu, Yu (November, 2003) "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science*, Vol 49, No. 11; S Greenstein and R McDevitt, , OECD Digital Economy papers 2012, DOI 10.1787/5k9bcwkg3hwhf-en

accessed 07/04/15

<sup>36</sup> Barry Schwartz, *The Paradox of Choice: Why More Is Less*, Harper Collins, 2004.

<sup>37</sup> Simon, Herbert A. 1991. "Organizations and Markets." *Journal of Economic Perspectives*, 5(2): 25-44. DOI: 10.1257/jep.5.2.25

<sup>38</sup> *Economic Welfare Measurements And Human Wellbeing* Avner Offer, University of Oxford Discussion paper, January 2000, accessed 03/04/15.

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 specialisation 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.<sup>39</sup> 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 trade off 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?

---

<sup>39</sup> <http://www.project-syndicate.org/commentary/inclusiveness-key-strategy-for-growth-by-ricardo-hausmann-2014-11> Accessed 07/04/15

- f) Nobody ever took to the barricades for the sake of liberty, equality and efficiency. Yet efficiency – in a tightly defined sense – is the narrow lens through which economists evaluate social welfare. This is a straitjacket self-imposed by the discipline as a result of the utilitarian and individualist turn. Is it still the right lens for statistics intended to evaluate economic progress and hold policymakers to account for delivery? The answer is yes, probably; there are plenty of social statistics too. However, we should be mindful of the gap between the public and economists on some issues that most people clearly see through a fairness lens instead. One example is the resale of tickets for sports events or concerts, which economists see as an efficient mechanism for making sure the people who most value the event are the ones who attend, while normal people see it as scalping. We should also acknowledge that for at least 30 years economics has been far more open to arguments from freedom than from fairness, although that might have changed, post-Piketty.
- g) 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 specialisation in global supply chains be measured? Or the dependence of sustainability at the national level on global environmental developments?

This is an array of extremely difficult issues. It took many years – decades – for the SNA we use now 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 position on what an economic dashboard should include, especially if some of these questions do turn out to raise rather fundamental issues about welfare economics. Meanwhile, the question remains: what is the best way now to serve citizens with statistics that will help them to monitor and judge the performance of policymakers in our democracies?

#### **4. The democratic conversation**

The national accounts statistics, especially those for the growth of real GDP, are not useless for this purpose, of course. Both the political stories based on the latest GDP figures used as examples in the introduction – both the increasing momentum in the UK economy at the start of the General Election campaign and the fact that the recovery from the financial crisis has been slow and lacklustre – 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

the economy. Part of the explanation might be found in some of the indicators collected in the new Economic Well-being publication from the Office for National Statistics, which includes for example summary income distribution indicators, and NNDI per capita as well as GDP per capita.

The new publication, alongside the second estimate of the previous quarter's GDP and some of the national accounts data, generated a small amount of polite interest in the media. Public interest in either an alternative to GDP or a more rounded 'beyond GDP' view of the economy is latent until the political conversation conducted via the media makes the switch. 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. One argument for a single indicator as an alternative to GDP is precisely because of this problem, the case being that it is a more realistic switch than adopting 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 (whether official ones such as those published by the ONS or Statistics New Zealand, or alternative examples such as the Social Progress Index) 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.<sup>40</sup> 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.<sup>41</sup> 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

---

<sup>40</sup> One example is the New Economics Foundation: <http://www.neweconomics.org/blog/entry/labours-living-standards-index-does-it-go-far-enough> Accessed 13/4/15.

<sup>41</sup> For some examples, see *Bias in Cable News: Real Effects and Polarization* Gregory J. Martin and Ali Yurukoglu December 22, 2014 <http://polisci.emory.edu/faculty/gjmart2/> accessed 04/04/15; *Illegal Immigration And Media Exposure: Evidence On Individual Attitudes* Giovanni Facchini, Erasmus University, Anna Maria Mayda, Georgetown University, Riccardo Puglisi, ECARES, Université Libre de Bruxelles, Discussion Paper No. 7593 December 2009 Centre for Economic Policy Research accessed 04/04/15.

graduate without the practical ability to read and understand a press release from the statistical office; it is a responsibility very often unfulfilled.

Finally, to return to one of the first points in this paper, citizen statistics need to have integrity and be trusted. The institutional arrangements for their collection need to deliver this, albeit always in a framework of appropriate accountability.

## 5. Conclusions

The great strength of the existing System of National Accounts, and the GDP aggregate, is its basis in Keynesian macroeconomic theory. The intellectual scaffolding was the intention to use the level of GDP as a measure of the economy's employment rate, and during the war of its ability to provide resources for national purpose. Although it is often said that GDP was never intended as a measure of social welfare – and it is certainly far from the kind of welfare aggregate advocated by Simon Kuznets – in fact it is implicitly one, and is certainly used that way in public debate. However, the statistics we have are far from those we need to enable citizens to hold their governments to account in the 21<sup>st</sup> century.

This paper has explored two sets of problems with GDP and other national accounts statistics. One concerns the failure of everyone – including many statisticians and economists – to acknowledge the extent of the uncertainty, the margins of error, in the published statistics. 'Stories' about the economy might be without any substance, yet are confidently told. The second set of problems concerns what kind of conceptual framework could be used as the basis for aggregate statistics more clearly linked to social welfare, sketching a series of very difficult questions. Social welfare has many dimensions, something recognised in the interest in dashboards. There are certainly separate measurement desiderata – aggregate activity and sustainability as well as current social welfare. All are likely to be elements of a social welfare dashboard.<sup>42</sup> Still needed though is the conceptual basis necessary for assessing whether policies are good ones.<sup>43</sup> There are in this task some major challenges to both the production of statistics and to the welfare economics underpinning aggregate statistics.

---

<sup>42</sup> The distinction is discussed in Marc Fleurbaey and Didier Blanchet, *Beyond GDP: Measuring Welfare and Assessing Sustainability*, Oxford University Press 2013. See also Diane Coyle, *The Economics of Enough: how to run the economy as if the future matters*, Princeton University Press 2012.

<sup>43</sup> One example of an administration recognising this need is *The New Zealand Treasury's Living Standards Framework - A Stylised Model* draft paper, Girol Karacaoglu, April 2015

## References

Allin, Paul, and David J. Hand, "[From a System of National Accounts to a Process of National Wellbeing Accounting](#)", IARIW-OECD conference April 2015, accessed 03/04/15.

Atkinson, A, *Inequality: What can be done?*, Harvard University Press, 2015

Brynjolfsson, Erik, Smith, Michael and Hu, Yu (November, 2003) "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science*, Vol 49, No. 11

Coyle, D, *GDP: A Brief but affectionate history*, Princeton University Press 2014

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

Coyle, D, *The Economics of Enough: how to run the economy as if the future matters*, Princeton University Press 2012

Delong, J.B., Cornucopia: The Pace of Economic Growth in the Twentieth Century, NBER Working Paper 7602, March 2000. <http://www.nber.org/papers/w7602> Accessed 13/04/15

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

Easterlin, R., et al, The happiness–income paradox revisited, *Proc Natl Acad Sci U S A*. 2010 Dec 28; 107(52): 22463–22468. Published online 2010 Dec 13. doi: [10.1073/pnas.1015962107](https://doi.org/10.1073/pnas.1015962107)

Accessed 7/4/15.

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.

European Commission, Report On Greek Government Deficit And Debt Statistics Brussels, 8.1.2010 COM(2010) 1 final.

Facchini, G, Anna Maria Mayda, Riccardo 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 04/04/15.

Fleurbaey, M, and Didier Blanchet, *Beyond GDP: Measuring Welfare and Assessing Sustainability*, Oxford University Press 2013.

Frank, Robert H., 2008. "[Should public policy respond to positional externalities?](#)," [Journal of Public Economics](#), Elsevier, vol. 92(8-9), pages 1777-1786, August.

Frey, B.S., and A.Stutzer, [Happiness and economics: How the economy and institutions affect human well-being](#), Princeton University Press 2010

Hausman, R, and C Hidalgo, *Atlas of Economic Complexity*, <http://atlas.cid.harvard.edu/>  
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. <http://dx.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.

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*, page 79, OECD, Paris 2001

Manki, C.F., *Communicating Uncertainty In Official Economic Statistics: An Appraisal Fifty Years after Morgenstern*, forthcoming, Journal of Economic Literature.

Martin, Gregory J, and Ali Yurukoglu, *Bias in Cable News: Real Effects and Polarization* December 22, 2014 <http://polisci.emory.edu/faculty/gjmart2/> accessed 04/04/15

McLaren, N and Shanbhogue, R (2011), 'Using internet search data as economic indicators', *Bank of England Quarterly Bulletin*, Vol. 51, No. 2, pages 134–40

Morgenstern, O, *On The Accuracy of Economic Observations*, Princeton University Press 1950, 1963.

Offer, A., *Economic Welfare Measurements And Human Wellbeing*, University of Oxford Discussion paper, January 2000, accessed 03/04/15.

Osterwald-Lenum, Michael, ["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 03/04/15.

Porter, T., *The Rise of Statistical Thinking 1820-1900*, Princeton University Press 1988.

Schwartz, B, *The Paradox of Choice: Why More Is Less*, Harper Collins, 2004.

Scitovsky, T., *The Joyless Economy: An Inquiry into Human Satisfaction and Consumer Dissatisfaction*. Oxford. 1976.

Simon, Herbert A. 1991. "Organizations and Markets." *Journal of Economic Perspectives*, 5(2): 25-44. **DOI:** 10.1257/jep.5.2.25

Stevenson B, Wolfers J. Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox. *Brookings Papers on Economic Activity 2008*. 2008;(Spring): 1–87.

Tily, G, *The National Accounts, GDP and the 'Growthmen'*, January 2015.

[http://static1.squarespace.com/static/541ff5f5e4b02b7c37f31ed6/t/54b3afeae4b0d2480d43d760/1421062122575/CoyleReview\\_Tily.pdf](http://static1.squarespace.com/static/541ff5f5e4b02b7c37f31ed6/t/54b3afeae4b0d2480d43d760/1421062122575/CoyleReview_Tily.pdf) Accessed 13/4/15.

Williamson, S and Berkes, E, *The Impact and Interpretation of Post War Revisions in the Official Estimates of GDP for the United Kingdom*,

<http://www.measuringworth.com/datasets/UKdata/UKGDPS.pdf> accessed 18/04/15

Jan Luiten van Zanden, Joerg Baten, Marco Mira d'Ercole, Auke Rijpma, Conal Smith and Marcel Timmer (eds), *How Was Life? Global Well-Being Since 1820*, OECD, 2014.

PRELIMINARY VERSION