International Institute of Social Studies

Erasmus

Working Paper
No. 632

Measurement error of global production

Peter A.G. van Bergeijk

August 2017

ISSN 0921-0210

The International Institute of Social Studies (ISS) is Europe's longest-established centre of higher education and research in development studies. On 1 July 2009, it became a University Institute of the Erasmus University Rotterdam (EUR). Post-graduate teaching programmes range from six-week diploma courses to the PhD programme. Research at ISS is fundamental in the sense of laying a scientific basis for the formulation of appropriate development policies. The academic work of ISS is disseminated in the form of books, journal articles, teaching texts, monographs and working papers. The Working Paper series provides a forum for work in progress which seeks to elicit comments and generate discussion. The series includes academic research by staff, PhD participants and visiting fellows, and award-winning research papers by graduate students.

Working Papers are available in electronic format at www.iss.nl

Please address comments and/or queries for information to:

International Institute
of Social Studies
P.O. Box 29776
2502 LT The Hague
The Netherlands

Or

E-mail:
wpapers@iss.nl

Table of Contents

Abstract

This working paper discusses the need and possibility to report measurement error together with key (macroeconomic) statistics as shown by a case study of the real rate of growth of world GDP (Gross Planet Product). The IMF estimates for individual years since 1980 and continue to change when a new vintage of the World Economic Outlook data base is published (each year in October). The different vintages provide an indication of the extent of measurement error. According to two measures for measurement error the IMF data for Gross Planet product on average have an implicit minimal measurement error (IMME) of four percent and maximum ratio (MR) of eighteen percent. Even for long-term growth rates that are calculated over two decades growth rates have a substantial measurement error, namely an IMME of 1.7% and an MR of 8.0%. Measurement error of Gross Planet Product is thus economically and statistically significant and needs to be addressed in studies that analyse or use global production data. Measurement error in economics currently is significant, is not showing improvement over time and could be reported transparently without technical or budgetary problems.

Keywords

Measurement error, IMF, GDP, world production, implicit minimal measurement error, maximum ratio.

Measurement error of global production[1]

1 Introduction

This article starts from an important observation in Oskar Morgenstern's seminal *On the accuracy of economic observations*. Morgenstern (1950) pointed out that – just as in the natural sciences – absolute precision and certainty are impossible to obtain in economic observations. Indeed, economic observations are often hampered by significant measurement error. Unlike the natural sciences, economics in general does not report measurement errors for the key concepts (such as consumer and producer prices, value and production) that it seeks to measure and explain. Economics is not a natural science and its methodology must be different from the natural sciences as it lacks the benefit of repeated experiments under laboratory conditions, but the non-reporting of measurement error is unscientific by any standard and carries an important cost for society in terms of decision-making. Indeed, using data without consideration of the extent of measurement error has potentially significant implications for evidence-based policy-making (Reiss, 2016).

Recent studies have consistently demonstrated the importance of measurement error for analysis, behavior and policy-making in the fields of monetary policy (Cavallo et al. 2016, Eichenbaum et al. 2015, Salter and Smith, 2017), economic and fiscal policy (Andriessen et al. 1995, Ferrera 2013), development policy (Carletto et al. 2015, Jerven 2013, Wolff et al., 2011), international trade and investment policy (van Bergeijk 2009, Egger and Průša, 2016), population, health and education (Burch 2015, Cawley et al.2015, O'Neill and Sweetman 2013, Wolff et al., 2011) and even the detection of illegal activities for law enforcement and prevention of terrorism (Blattman et al. 2016; Chambers et al.2010).

Despite this increasing body of evidence, Morgenstern's contribution continues to be neglected by mainstream economics (Bagus, 2011, Boumans 2012, 2015): data producers do not report measurement error for most of the aggregated data and estimated concepts (Manski, 2015). Typical examples of variables for which we do not know the extent of error include inflation, unemployment, Gross Domestic Product and its components, such as consumption, investment and exports.[2] Non-reporting of measurement error is a highly relevant issue in view of the fact that research consistently shows significant measurement errors for these concepts, both in a domestic context (e.g. Andriessen et al. 1995, Cañal-Fernández, 2012, Sinclair and Stekler, 2013)

---

[1] Peter van Bergeijk is professor of international economics and macroeconomics, Erasmus University, International Institute of Social Studies, The Hague, The Netherlands. Peter A.G. van Bergeijk bergeijk@iss.nl / orcid: 0000-0002-4098-0483.

[2] The situation appears to be better for survey data for which well-recognized statistical procedures are available and applied, but is not without problems (Abowd and Stinson 2013, Gibson et al.2015, Groves and Lyberg, 2010, Kim and Tamborini, 2014, Meyer et al.2015, Tasciotti and Wagner, 2017).

and in an international and often comparative setting (van Bergeijk 1995 and 1998, Benita et al.2016, Jerven, 2014).

This of course does not mean that measurement error is not a concern amongst economists. Typically the existence of measurement error has given rise to two reactions: firstly, to improve the quality of data collection (see, e.g., Keller et al. 2017) and, secondly, to develop and strengthen advanced statistical techniques to estimate 'true' parameters after correction for measurement error (see Schennach 2016 for an overview of the state of the art). The point that this article wants to make is thus not that economics ignores measurement error related problems; the point is that measurement error is not reported and that this endemic worst practice needs to be cured. Yes, data quality improvement is important, but does not answer the issue of the accuracy of an observation and, importantly, in order to monitor the efforts (and to set priorities) an indicator of the final outcome variable (that is: measurement error) is necessary. Yes, correcting for measurement error in econometric analysis is important, but does not solve the problem that policy-making is frequently by necessity based on primary data and often over a short period of time so that econometric techniques cannot be applied meaningfully. Therefore reporting measurement error is important in order to prevent misinterpretation of the primary data. Treating data as if there is no measurement error generates wrong diagnoses and thus may give rise to type I and type II errors in economic policy. For example, a government that erroneously believes that a recession has ended may engage in fiscal austerity thus increasing the period of below-equilibrium activity. In sum, while economic science addresses measurement issues, it neglects its scientific duty by not reporting measurement error.

The next section discusses how measurement error can be assessed and reported. Section 3 deals with the apparent measurement error in a well-known data series for the real rate of growth of world production that is produced by the International Monetary Fund (IMF). This is an interesting case because the IMF is a top class data producer so that the extent of error in this series gives a good idea into the minimum manner of error to be expected in statistics on the growth rate of an economy. Section 4 draws some conclusions.

2 Assessing and reporting measurement error

In order to establish the extent of measurement error one needs multiple measurements M_i of the same item either by using the same measurement tool repeatedly (so that the variation of the measurement data can provide an indication of the inaccuracy of the tool and/or measurement procedure) or by using different measurement tools and/or points (perspectives) of observation. Measurement error of relatively inaccurate measurement tools can also be established with the use of highly accurate measurement tools. As with measurement itself, measurement errors can only be established with some inaccuracy. So neither the item that is measured nor the measurement error can be established with absolute precision.

So as to establish measurement error several measurements need to be available. In economics, often the impression is that only one observation or series of observations exist: data producers typically report their preferred single estimate for a variable and data users typically use their preferred single estimate secondary data. Very frequently, however, two or more independent series (observations) of the same concept exist:

a) **The same variable is measured by different institutions, perspectives and/or observers.** This typically is the case for all international economic exchanges (flows). Exports from country A to country B (registered by A's statistics) should match imports by country B from country A (registered by B's statistics) and similar for capital flows (and sometimes stocks) such as lending, investment, remittances and official transfers (van Bergeijk 1995; Beita et al. 2016). Also for typical national data in many cases different institutions report data on the same phenomenon: household surveys can be compared to census data (Abowd and Stinson, 2013; Tasciotti and Wagner 2017); micro data can be compared to administrative data (Kim and Tamborini, 2014). A complicating factor in these cases is that measurement procedures need to be evaluated that differ due to different approaches in different disciplines (e.g. economics and physics; so a transdisciplinary approach is necessary) or sub-disciplines (e.g. microeconomics and macroeconomics; which may give rise to the Fallacy of Composition). With this caveat in mind, it still seems quite possible to use the heterogeneity of measurements to arrive at meaningfull estimates of measurement error.

b) **The same concept is measured by different measurement tools/procedures/perspectives**. The clearest example is provided by the National Accounts estimates for Gross Domestic Product (GDP). This estimate is based on three different approaches that measure production, expenditure and income, respectively. Typically the information of the three approaches is bundled by the producers of primary data to provide the best possible estimate of GDP, but it is also possible to use these separate estimates to establish measurement error (Aruoba et al, 2016). Technological innovations and increased computing power have greatly enhanced the opportunities for alternative measurements. Observations from mobile devices, drones and even outer space provide new and alternative perspectives on economic activities (Nordhaus and Chen 2014, Pinkovskiy and Sala-i-Martin 2014). Microdata from a great many individual stores could be compared to the samples typically used by central bureaus of statistics and central banks to analyze inflation (Eichenbaum et al. 2014).

   It is also worth to point out that an identity, for example,

   sold products = produced products – change in inventories

   provides two items (left hand side and right hand side of the equation) that can be independently observed and compared to detect measurement error. Economics actually compares favorably to other sciences because these identities are well-known, do not depend on conditions of time and place and refer to concepts that are solidly grounded in theory.

c) **Different estimates at different moments *T+n* are produced for the same variable at a specific moment *T* are produced by the same institution** Often unnoticed, economic history is constantly being rewritten when new information arrives at the offices of the data producers and previous estimates are reviewed in the light of the evidence that has just become available (Croushore, 2011, Manski 2016). Early preliminary estimates of GDP are provided by many countries after 6 weeks, 3 months, 2 quarters and 1 and 2 years.[3] The latter is often referred to as the 'final estimate' which means that the estimate is not expected to be changed unless a structural change in the procedures of National Accounting occurs. Structural change of the procedures includes changes in the UN System of National Accounting and the inclusion of previously non-included economic activities (for example, sectors that were negligible or non-existent before but started to grow so that they needed to be considered). A structural change will typically give rise to a change of previous 'final estimates' over a longer period of time. As a consequence of new information (feeding into the first and second year) and structural changes in procedures and samples, time series produced at time T and time $T+n$ can differ significantly; these time series are called the different vintages of a dataset or database. In Section 3 we will study different vintages of the IMF *World Economic Outlook* data base.

In view of the above it can be ascertained that for many if not most macroeconomic concepts at least one alternative measurement is available so that establishing minimal levels of measurement error is not only possible, but actually quite doable. Two approaches with clear intuitive interpretation are available: the Implicit Minimal Measurement Error (IMME, Van Bergeijk 1995) and the Maximum Ratio (MR, Tsao and Wright 1983). IMME measures the average distance from the average estimate and provide the lowest estimate of measurement error. IMME was developed and has been used in the context of variables and rates of growth that could both be positive and negative and for two estimates M_1 and M_2 (but can be generalized to situations where more than two estimates are available) and is defined as

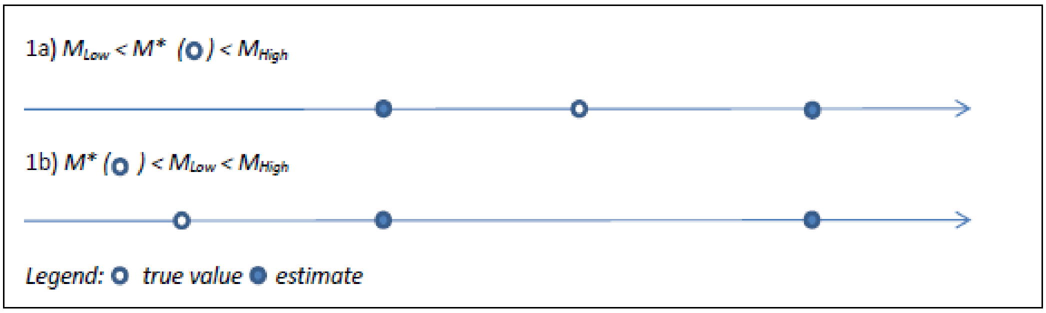
$$IMME = \frac{|M_1 - M_2|}{|M_1| + |M_2|} \times 100\% \qquad (1)$$

[3] An important line of research is to see how preliminary early-on estimates (that are by definition most inaccurate) can be corrected or evaluated so as to guide policy. See, for example, Jacobs and Van Norden (2011)

MR provides the maximum measurement error consistent with a given set of estimates (assuming that the true value lies within this range). MR was developed in the context of two or more estimates that typically are strictly positive (but the formula can easily be adjusted for situations in which both negative and positive estimates of the same variable exist) and is defined as

$$MR = \frac{max(M_i) - min(M_i)}{min\,(M_i)} \times 100\% \qquad (2)$$

Diagram 1
True value as assumed by IMME and MR (1a) and alternative where IMME and MR by definition underestimate measurement error because true value lies outside interval of estimates (1b)



While IMME assumes that both estimates are wrong and focusses attention on the minimal kind of measurement error that data users should take into account, MR offers another conservative approach by expressing the largest difference between the available estimates in terms of the smallest estimate. The originators of the approaches do recognize that it is impossible to determine which of the available estimates is correct and also that the true value M^* of the object that is measured may be outside the range in which the estimates M_i are located, but evaluate error in over the domain of reported estimates (Diagram 1). In this sense both IMME and MR could be underestimating the problem at hand. With this caveat in mind, the merits of IMME and MR are clear because the indicators do not require computing power, have an intuitive and clear interpretation and can be reported transparently, as illustrated by the case study in the next section.

3 Case study: IMF *World Economic Outlook* Data for Global Growth

The choice for the IMF's data series on the growth rate of GPP (Gross Planet Product[4]; see van Bergeijk 2013) is motivated by the fact that the IMF is a highly professional organization with well-trained staff and thus stands out as a top notch data producer. The extent of measurement error can therefore be expected to be comparatively low. The case study is relevant for academics, policy-makers and the public at large because the IMF's global growth statistic is a key feature of its flagship publication *World Economic Outlook* (published every year in April and October), that gets wide media coverage and is an important input for research, policy and business decisions.

It is important to first consider the implications of my choice for and specific use of this variable by means of an *ex ante* measurement error prognosis analysis. From this perspective it is noteworthy that studying a real annual growth rate of a key concept at a very high level of aggregation with data that are at least three years old and produced by an international institution can be expected to reduce measurement error compared to an alternative analysis of the level of a nominal component of GPP such as export, investment or consumption including the most recent data and produced by a national government institute, because it avoids five important problems that have been identified in the literature:

a. **the choice of the numeraire:**

Obstfeld et al. (2015) report significant differences in world GPP growth in 2015 based on different currencies, pointing out, for example, that the 2015 nominal GPP growth rate in US dollars is –4.9%, but + 13.6% in euro. This is an important reason to focus on a real (constant price) rather than a nominal variable.

b. **constant measurement error**:

First differencing gets around constant measurement error that may impact for example the trend of productivity levels as shown by Bils et al.2017 with firm level data for the US over the years 1978–2007 where a 50% difference over the period was observed

c. **seasonal adjustment:**

Manski (2017) singles out seasonal adjustment at the most important conceptual measurement issue of GDP statistics; using annual data avoids this problem

d. **classification error:**

Components of GDP can get easily misclassified. A consumer good, for example, according to National Accounting principles is an investment good until it has been sold to the final consumer, but teller data do not report if goods are really for final consumption or will be used for further

[4] GPP is often referred to in academic publications, policy documents and the popular media as world GDP, global production or world output.

sale; likewise goods classified as intermediate goods may be used by small business owners for consumption. Classification error also regularly occurs in product classification, country classification and with regard to regional data (van Bergeijk 2010, appendix A1).

e. strategic over- and/or underreporting:

Fariss et al. (2017) point out that governments may have strategic incentives to under or over report their GDP, for example, in order to avoid inconvenient truths, to meet the criteria for development aid or the requirements for inclusion in an organization or institution. The Fund may have strategic reasons for under- or overreporting in order to avoid self-fulfilling prophecies regarding its forecasts that tend to be positively biased (Independent Evaluation Office 2014), but not regarding historic data.

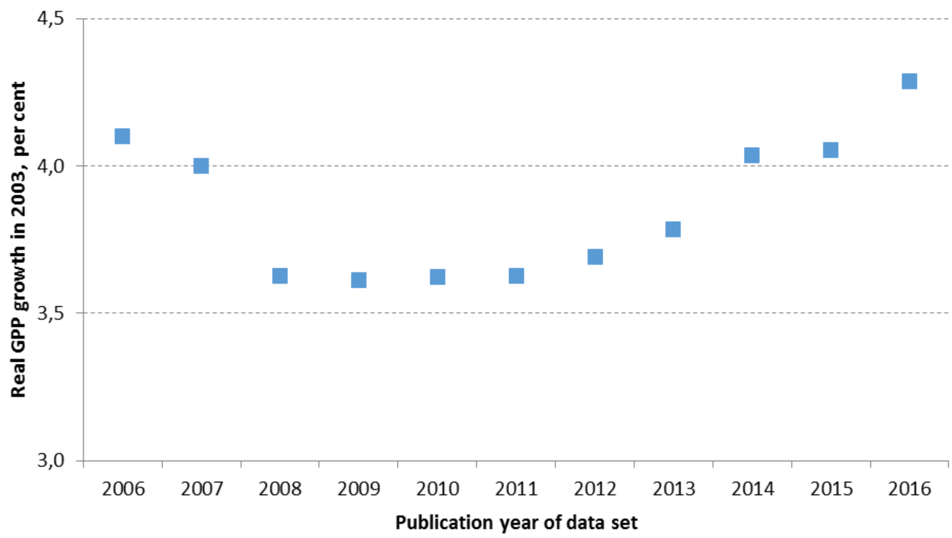
It is, moreover, important to note that the series for GPP growth are constructed in a highly professional context. The Fund has detailed knowledge about the quality of the underlying official statistics produced by its Members amongst others due to the fact that IMF staff regularly visits Member States in the context of surveillance (Article IV consultations). The IMF staff is well aware of differences between the vintages of its series for GPP and occasionally reports the most recent vintage in comparison to earlier data versions. Uncertainty intervals are provided, but only regarding forecasts; not regarding historic data. Finally it is noteworthy that underlying data are obtained from a large number of countries (the IMF currently has 189 Members); the five largest country shares in current 2016 US dollars are the United States (23%), China (14%), Japan (6%), Germany (5%) and the United Kingdom (4%).[5] The large number of countries suggests that the law of large numbers could ensure that errors across countries cancel out; the large share of especially the United States and China, however, implies the risk that country-specific errors could dominate global measurement.

We start by comparing different estimates for GPP growth in the year 2003. Figure 1 reports the annual real rate of growth of GPP as reported in 11 October versions of the *World Economic Outlook*. The 11 data versions (vintages) have been reported in the period 2006-2016 alongside the IMF flagship publication *World Economic Outlook* (the October version of each year in this period). The lowest number reported for GPP growth in the year 2003 was published in the 2009 October data set (3.61%). The highest value for the 2003 growth rate (4.29%) was published in 2016. Over the years the reported estimate for 2003 varies considerably (the standard deviation of the estimates is 0.24; the coefficient of determination is 0.06; IMME is 5.6%; MR=18.7%).

---

[5] Note that the calculations of shares are based on current US dollars and thus critically depend on the choice of the numeraire (see Oldersma and van Bergeijk 1993, and on IMF data Obstfeld et al. 2015).

Figure 1
Real GPP growth for the year 2003 by vintage (year of publication) of IMF WEO data set



The reported growth rate for the year 2003 varies by a maximum of 0.68 percentage points between the different data versions – the key word here is fluctuates as there is neither a clear strictly positive or negative trend, nor an obvious mean reversion tendency. Clearly new vintages change IMME and MR. Before the 2016 estimate of GPP growth in 2003 became available IMME was 4.9% and MR was 12.6%. Therefore the 2016 estimate increases IMME by 0.7 and MR by 6.1 percentage points (note that this estimate for 2003 is published 13 years later). All in all the established indicators for measurement error for 2003 is economically and statistically significant and therefore should be an important reporting item.

Figure 2 shows that this variation between vintages for historical data is a regular phenomenon in the *IMF World Economic Outlook* data base. The figure reports for the same 11 vintages that were used before in the construction of Figure 1 the minimum and maximum (indicated by the grey area) and the median growth rate (dotted line) for each year. Even for 1991 (so a year that lies at least 15 years before the year of publication of the vintages that I investigate) changes occur that amount to a maximum of 1.1 percentage points (IMME = 18%; MR = 79%). This is the largest absolute difference, but for a number of other years the 'revisions' are also economically important. If the measurement error in percentage points is related to the level of the growth rate (as both IMME and MR do) then the year 1983 where real GPP growth is below 1% is also highly problematic (IMME = 17%; MR = 76%). Table 1 reports the difference between maximum and minimum estimates reported in the 11 vintages for a specific year, as well as IMME and MR for individual years.

Figure 2
Real GPP growth (median, maximum, minimum) for specific years (1980-2005) according to IMF WEO database vintages 2006-2016

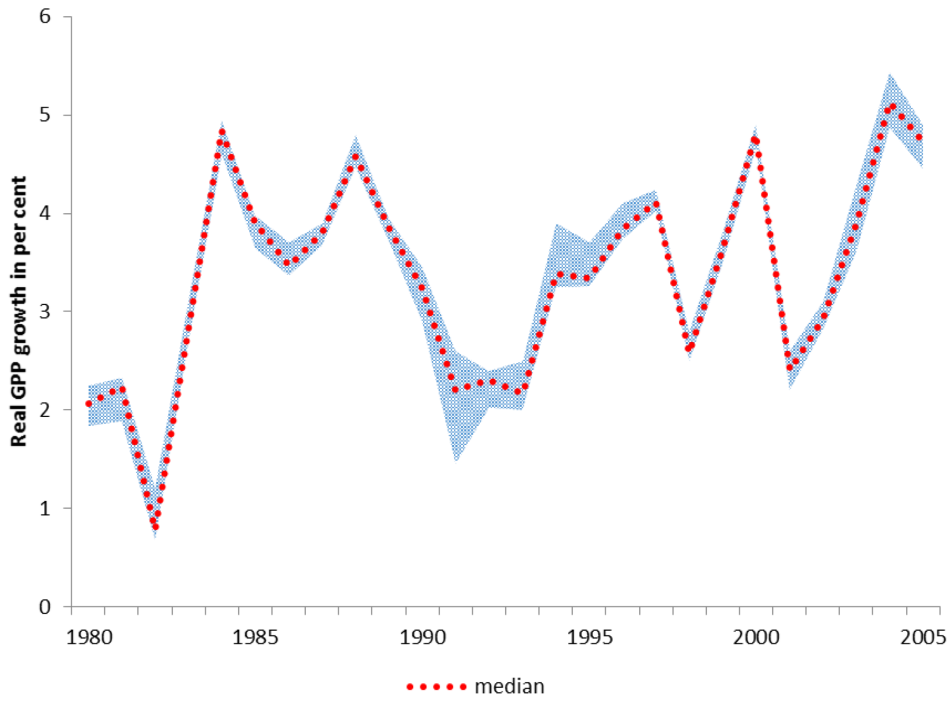


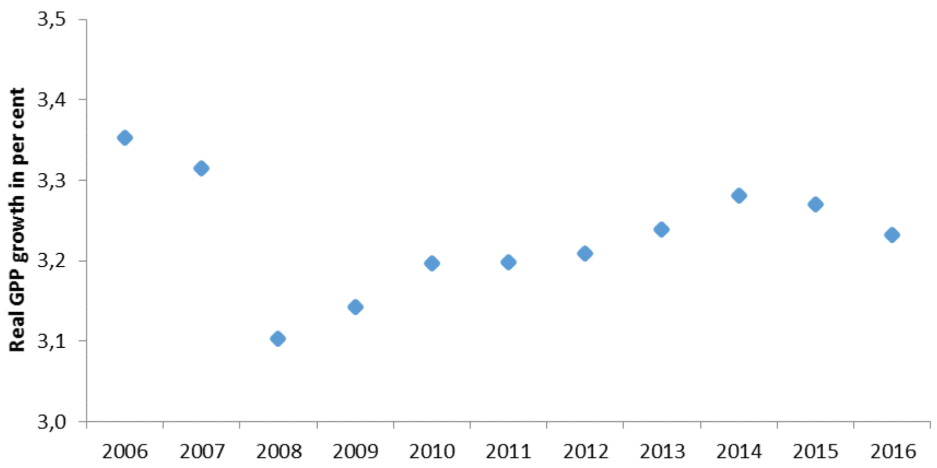
Table 1 provides three useful perspectives on measurement error in GPP according to the IMF World Economic Outlook database. First, no clear improvement over time seems to occur. If anything the downward trend in measurement error that according to the three reported indicators for measurement error appeared in the 1990s seems to have been broken around the turn of the millennium. Second, measurement error has been very significant in 1983 and 1991. Here the measurement error indicators help to identify areas for further investigation (additional research, for example, revealed that the measurement error in 1991 is mainly due to the United States). Third, on average the measurement error in percentage points is 0.4; IMME is 4% and MR is 18%. There is no reason why this could not be transparently reported in future editions of the IMF *World Economic Outlook*. This would alert users in the public and the private sector about the inaccuracy and uncertainty of the reported real growth rate of GPP. Awareness of this inaccuracy could stimulate researchers using the GPP data series to use different vintages in order to assess robustness with respect to measurement error (van Bergeijk, 2016). Note that Figures 1 and 2 and Table 1 suggest that no stable pattern in revisions could be discerned so that econometric corrections for measurement error do not offer a viable alternative for redoing (or replicating) estimated equations for different vintages.

Table 1
Maximum difference, IMME and MR for specific years (1980-2005)
for GPP according to IMF WEO database vintages 2006-2016

	Maximum – minimum estimate (percentage points)	Implicit minimal measurement error (IMME)	Maximum ratio (MR)
1980	0.4	7%	22%
1981	0.4	3%	23%
1982	0.5	17%	76%
1983	0.4	3%	15%
1984	0.3	2%	7%
1985	0.3	2%	9%
1986	0.3	3%	10%
1987	0.2	2%	6%
1988	0.3	2%	7%
1989	0.2	1%	6%
1990	0.6	5%	19%
1991	1.1	18%	79%
1992	0.4	5%	18%
1993	0.5	5%	25%
1994	0.7	4%	20%
1995	0.4	4%	13%
1996	0.4	3%	10%
1997	0.2	1%	5%
1998	0.3	2%	12%
1999	0.3	1%	9%
2000	0.2	1%	5%
2001	0.4	5%	18%
2002	0.3	3%	10%
2003	0.7	6%	19%
2004	0.6	4%	12%
2005	0.4	3%	10%
Average	***0.4***	***4%***	***18%***

It is possible that measurement errors fluctuate over time and therefore it is interesting to investigate if and how a long-term growth rate of GPP is influenced by measurement error. Figure 3 investigates a period of two decades of growth rates (1980-2000) using again the same 11 vintages as before (so now the data for the end of the period are at least 6 years old). Measurement error is substantially smaller for the long term growth rates (the indicators for measurement error are halved: the difference between the highest and lowest estimate is 0.25 percentage points, IMME is 2% and MR is 8%), but even so cannot be ignored.

Figure 3
Average real GPP growth for the period 1980-2000 by vintage (year of publication) of IMF WEO data set



All three examples investigated in this case study (estimates for a specific year, estimates over a number of years, estimates of long term growth rates for a specific period) consistently show that large measurement errors exist and persist. The examples also show the do-ability of reporting indicators for measurement error together with macroeconomic data.

4 Concluding remarks

The motivation for this paper is the necessity to take measurement error seriously in economics. First and foremost its existence should be acknowledged by requiring that measurement errors are customarily supplied and discussed. This is task for academics, data collectors – such as the bureaus of statistics and central banks – data producers and data users. To some this may seem to be a negative attitude and many may want to argue that the time and effort spent on measuring measurement error could be better invested in improvement of the data. This working paper, however, has shown that it does not require a lot of effort to report indicators of measurement error. It has substantiated that even highly professional data producers such as the IMF report data series that have significant measurement errors, for individual years, periods of observations and variables that are measured using a long term framework. Thus it should be clear that measurement error in economics currently is significant, is not showing improvement and could be reported transparently without technical or budgetary problems.

References

Abowd, J. M., and M. H. Stinson (2013) 'Estimating measurement error in annual job earnings: A comparison of survey and administrative data.' *Review of Economics and Statistics* **95** (5), pp. 1451-1467.

Andriessen, J.E., J. van Sinderen and P.A.G. van Bergeijk (1995) 'Requirements of Policy-makers', in: Z. Kenessey (ed.), *The Future of Statistics: An International Perspective*: ISI Studies and Publishing Center: Voorburg, pp. 237-251.

Aruoba, S. B., F.X. Diebold, J. Nalewaik, F. Schorfheide, and D. Song (2016) 'Improving GDP measurement: A measurement-error perspective.' *Journal of Econometrics* **191** (2), pp. 384-397.

Bagus, P. (2011) 'Morgenstern's Forgotten Contribution: A Stab to the Heart of Modern Economics.' American *Journal of Economics and Sociology* **70** (2), pp. 540-562.

Benita, F., and Urzúa, C.M. (2016) 'Mirror trade statistics between China and Latin America.' *Journal of Chinese Economic and Foreign Trade Studies*, **9**(3), pp. 177-189.

Bergeijk, P.A.G. van (1995) 'The Accuracy of International Economic Observations', *Bulletin of Economic Research* **47**(1), pp. 1-20.

Bergeijk, P.A.G. van (1998) 'Did real world per capita income really grow faster in 1870-1913 than in 1973-1992?', *De Economist* **146** (3), pp. 143-170.

Bergeijk, P.A.G. van (2009) *Economic Diplomacy and the geography of international trade*, Edward Elgar: Cheltenham.

Bergeijk, P.A.G. van (2010) *On the Brink of Deglobalization: An Alternative Perspective on the Causes of the World Trade Collapse*, Edward Elgar: Cheltenham.

Bergeijk, P.A.G. van (2013) *Earth Economics: An Introduction to Demand Management, Long-Run Growth and Global Economic Governance*, Edward Elgar: Cheltenham.

Bergeijk, P.A.G. van (2016) 'What a difference a data version makes', *The Replication Network* (November 29, 2016) http://replicationnetwork.com/2016/11/29/van-bergeijk-what-a-difference-a-data-version-makes/

Bils, M. P.J. Klenow, and C. Ruane (2017), 'Misallocation or Mismeasurement?', *Working Paper 599*, Stanford Center for International Development.

Blattman, C., J. Jamison, T. Koroknay-Palicz, K. Rodrigues and M. Sheridan (2016) 'Measuring the measurement error: A method to qualitatively validate survey data.' *Journal of Development Economics* **120**, pp. 99-112.

Boumans, M. (2012) 'Observations in a hostile environment: Morgenstern on the accuracy of economic observations.' *History of Political Economy* **44** suppl. 1, pp. 114-136.

Boumans, M. (2015) *Science outside the laboratory: measurement in field science and economics*. Oxford University Press: Oxford.

Burch, T. K. (2015) 'Error in Demographic and Other Quantitative Data and Analyses,' *Population Change and Lifecourse Strategic Knowledge Cluster Discussion Paper Series/ Un Réseau stratégique de connaissances Changements de population et parcours de vie Document de travail: Vol. 3 : Iss. 3, Article 1*. Available at: http://ir.lib.uwo.ca/pclc/vol3/iss3/1

Cañal-Fernández, V. (2012) 'Accuracy and reliability of Spanish regional accounts (CRE-95)'. *Empirical Economics* **43** (3), pp.1299–1320.

Carletto, C., S. Savastano, and A. Zezza (2013) 'Fact or artifact: The impact of measurement errors on the farm size–productivity relationship' *Journal of Development Economics* **103**, pp. 254-261.

Cavallo, A., G. Cruces, and R. Perez-Truglia (2016) Learning from potentially-biased statistics: Household inflation perceptions and expectations in Argentina. No. w22103. National Bureau of Economic Research: Cambridge MA.

Cawley, J., J.C. Maclean, M. Hammer, and N. Wintfeld (2015) 'Reporting error in weight and its implications for bias in economic models.' *Economics & Human Biology* **19**, pp. 27-44.

Chambers, D.R., D.A. Wernick., J.S. Zdanowicz. and M.A. von Glinow (2010) ' How dangerous are measurement errors to homeland security?' *Thunderbird International Business Review*, **52** (6), pp. 553-569.

Croushore, D. (2011) 'Frontiers of Real-Time Data Analysis', *Journal of Economic Literature* **49** (1), pp. 72–100.

Egger, P.H. and J. Průša, J. (2016) 'The determinants of trade costs: a random coefficient approach.' *Empirical Economics* **50**(1), pp. 51-58.

Eichenbaum, M., N. Jaimovich, S. Rebelo, and J. Smith (2014) 'How frequent are small price changes?' *American Economic Journal: Macroeconomics* **6** (2), pp. 137-155.

Fariss, C. J., C. D. Crabtree, T. Anders, Z. M. Jones, F. J. Linder, and J. N. Markowitz (2017) 'Latent Estimation of GDP, GDP per capita, and Population from Historic and Contemporary Sources.' *arXiv preprint* arXiv:1706.01099.

Ferrara, L. (2013) 'Comments on:' 'Examining the quality of early GDP component estimates".' *International Journal of Forecasting* **29** (4), pp. 751-753.

Gibson, J., K. Beegle, J. de Weerdt, and J. Friedman (2015) 'What does variation in survey design reveal about the nature of measurement errors in household consumption?' *Oxford Bulletin of Economics and Statistics* **77**(3), pp. 466-474.

Independent Evaluation Office of the International Monetary Fund (2014) *IMF Forecasts: Process, Quality, and Country Perspectives*, IMF: Washington DC

Jacobs, J.P.A.M., and S. van Norden (2011) 'Modeling data revisions: Measurement error and dynamics of "true" values.' *Journal of Econometrics* **161**(2), pp. 101-109.

Jerven, M. (2013) *Poor numbers: how we are misled by African development statistics and what to do about it.* Cornell University Press: Ithaca NY.

Jerven, M. (2014) 'On the accuracy of trade and GDP statistics in Africa: Errors of commission and omission' *Journal of African Trade* **1** (1), pp. 45-52.

Groves, R.M. and Lyberg, L. (2010) 'Total survey error: Past, present, and future'. *Public opinion quarterly*, **74** (5), pp. 849-879.

Kim, C.H., and C. R. Tamborini (2014) 'Response error in earnings: An analysis of the Survey of Income and Program Participation matched with administrative data.' *Sociological Methods & Research* **43**(1), pp. 39-72.

Keller, S., G. Korkmaz, M. Orr, A. Schroeder, and S. Shipp (2017) 'The Evolution of Data Quality: Understanding the Transdisciplinary Origins of Data Quality Concepts and Approaches.' *Annual Review of Statistics and Its Application* **4** (1), pp. 85-108

Manski, C.F. (2015) 'Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years after Morgenstern.' *Journal of Economic Literature* **53** (3), pp. 631–653.

Meyer, Bruce D., Wallace KC Mok, and J. X. Sullivan (2015), 'Household surveys in crisis.' *The Journal of Economic Perspectives* **29** (4), pp. 199-226.

Morgenstern, O. (1950) *On the Accuracy of Economic Observations*. Princeton University Press: Princeton NJ.

Nordhaus, W., and X.i Chen (2014) 'A sharper image? Estimates of the precision of nighttime lights as a proxy for economic statistics.' *Journal of Economic Geography* **15** (1), pp. 217-246.

Obstfeld, M. O. Celasun, M. Hemmati, and G. M. Milesi-Ferretti (2015) The choice of numeraire matters when calculating world GDP growth *Voxeu* 11 December 2015, http://voxeu.org/article/choice-numeraire-matters-when-calculating-world-gdp-growth.

Oldersma, H., and P.A.G. van Bergeijk (1993) 'Not so constant! The constant-market-shares analysis and the exchange rate.' *De Economist* **141** (3), pp. 380-401.

O'Neill, D. and O. Sweetman (2013) 'The consequences of measurement error when estimating the impact of obesity on income.' *IZA Journal of Labor Economics* **2** (1): 3.

Pinkovskiy, M., and X. Sala-i-Martin (2016) 'Newer need not be better: evaluating the Penn World Tables and the World Development Indicators using nighttime lights. *NBR working paper 22216*. National Bureau of Economic Research: Cambridge MA.

Reiss, J. (2016) *Error in Economics: Towards a More Evidence–Based Methodology*. Routledge.

Salter, A. W., and D.J. Smith (2017) 'What You Don't Know Can Hurt You: Knowledge Problems in Monetary Policy.' *Contemporary Economic Policy* 35(3), pp. 505-517.

Schennach, S.M. (2016) 'Recent advances in the measurement error literature.' *Annual Review of Economics* **8,** pp. 341-377.

Sinclair, T.M., and H. O. Stekler (2013) 'Examining the quality of early GDP component estimates.' *International Journal of Forecasting* **29** (4), pp. 736-750.

Tasciotti, L. and N. Wagner (2017) 'How Much Should We Trust Micro-data? A Comparison of the Socio-demographic Profile of Malawian Households Using Census, LSMS and DHS data.' *The European Journal of Development Research,* pp.1-25.

Tsao, H., and T. Wright (1983) 'On the maximum ratio: a tool for assisting inaccuracy assessment.' *The American Statistician*, **37**(4a), pp. 339-342.

Wolff, H., Chong, H. and Auffhammer, M. (2011) 'Classification, detection and consequences of data error: evidence from the human development index.' *The Economic Journal*, **121**(553), pp. 843-870.

Wright T. (2013) 'A Visual Proof, a Test, and an Extension of a Simple Tool for Comparing Competing Estimates.' *Research reports series (Statistics. 2013:05)* US Census Bureau: Washington DC.