

16. International Comparisons of GDP: Issues of Theory and Practice¹

Ian Castles and David Henderson

Introduction

This article combines a general theme, which though not new remains a subject of lively debate, with specific instances, illustrations and proposals for change. Our chief concern is with how international comparisons of real GDP and GDP per head are best made. We set out the case for using purchasing power parity (PPP) converters for this purpose, rather than exchange rates, and give reasons for rejecting various arguments that are still widely made to the contrary.² In doing so, we give instances of the differing current practices of international agencies, argue the case for greater uniformity and consistency, and make suggestions for improvement.

In developing the general theme, we start with a number of preliminary points, some of them familiar and elementary but pertinent nonetheless, relating to the measurement of output, and of changes in output over time, in individual countries. We then turn to consider the international comparative dimension, where much though not all of the argument proceeds on parallel lines. Under both headings, a basic and invaluable source is Chapter 16 of the 1993 System of National Accounts (SNA), and we have also drawn on an illuminating recently published paper by William Nordhaus.³ Following these general expository sections, we turn to specific and topical cases. First, we list and comment on some mistaken but still-continuing ways of presenting and interpreting international comparative data, by a range of individuals and official agencies. Second, we review briefly the extent to which the consistent use of PPP converters has

1 First published in *World Economics*, volume 6:1, Jan - Mar, 2005: 55 - 66.

2 The choice is often presented as being between PPP converters and 'market exchange rates' (MERs); but since it is debatable how far past and present exchange rates can be viewed as 'market-determined' we use here a more neutral term. Later in the text, when quoting or referring to other authors who speak of MERs, we place the term in inverted commas.

3 *System of National Accounts 1993*, prepared and published by five international agencies – the European Commission, the IMF, the OECD, the United Nations and the World Bank – and approved by the statistical authorities of their member governments. William Nordhaus, 'Alternative Measures of Output in Global Environmental Models: Purchasing Power Parity or Market Exchange Rates?'; revised version of a paper prepared for an Experts Meeting of the Intergovernmental Panel on Climate Change (IPCC) in January 2005. It is available at [ppp_02085.doc](#). Another valuable reference is Irving B Kravis, 'Comparative Studies of National Incomes and Prices'. *Journal of Economic Literature*, Vol. XXII, March 1984.

been incorporated into the work of the leading international agencies that are responsible for issuing international comparative data. Third, we summarise the actions that governments still need to take in order to ensure more accurate and more consistent treatment and presentation of these data. In a short annex we reproduce an earlier note by one of us which outlines a specific proposal for improvement on a broad front.

Among the international organisations whose current practice is open to question, one that we note in particular, because of its topical aspects, is the Intergovernmental Panel on Climate Change (IPCC). We have presented elsewhere a critique of the Panel's handling of economic issues; and within this critique, one element concerns the way in which inter-country comparisons of real GDP, and hence GDP per head, have been made in IPCC documents. Mistaken procedures have been adopted, and the IPCC has shown itself resistant to changes in these and unclear as to what is at stake. Such mistakes are to be found in particular, though not only, in the *Special Report on Emissions Scenarios* (SRES), which was published in 2000 as one of the documents that entered into the Panel's Third Assessment Report. The SRES provided, as its main single product, a range of projections of greenhouse gas emissions covering the period from 1990 to 2100.⁴

The Panel is now well into the preparation of its Fourth Assessment Report (AR4), which may well come to 3,000 pages and is due for completion in 2007. It has dismissed both our critique and our suggestions for wider participation in its economic work;⁵ and as part of this reaffirmation of the status quo, it has determined that 'the SRES scenarios provide a credible and sound set of projections, appropriate for use in the AR4'. Hence this particular instance of the case for reform on the international scene is especially topical: IPCC member governments have to act promptly if the decision to retain the SRES as the point of departure for AR4 is to be reconsidered – and, more broadly, if the economic aspects of this coming report to governments are to be handled in a more informed and professionally representative way than is now the case. Because of this immediate relevance, the work and procedures of the IPCC and one of its twin parent agencies, the United Nations Environment Program (UNEP), are featured in the two final sections of the paper.⁶

4 While the main single target of our critique has been the SRES, our concerns extend to the IPCC process and milieu *as a whole*, including the Panel's parent agencies and its sponsoring departments in member governments. The concerns are summarised in an article by one of us (Henderson) in the quarterly *Newsletter* of the Royal Economic Society for January 2005. This piece has now been reprinted in *Energy and Environment*, Vol 16 No 2 (2005).

5 A formal IPCC press release of December 2003, now posted on the Panel's website, says that 'In recent months some disinformation has been spread questioning the scenarios used by the IPCC', and refers to us as 'so called "two independent commentators"'.
6 The IPCC was established in 1988 as the joint creation of the UNEP and the World Meteorological Organisation.

Prices and quantities: Measuring and interpreting changes within an economy

Nominal and real changes in output and expenditure

For any country or individual economy for a specific period of time (normally a year), its GDP is defined in terms of the value of the output of goods and services produced within the economy over that period. For any given year, past or present, the outputs that enter into GDP are measured and valued at the prices prevailing in that year: the starting point is an estimate of the value of output at current prices, or nominal GDP.

To define and measure changes in aggregate output, or real GDP,⁷ such initial current-price estimates have to be corrected for year-to-year changes in the average price level of the goods and services concerned: for GDP, as for other economic time series, 'Changes in the values of flows of goods and services can be directly factored into two components reflecting changes in the prices of the goods and services concerned and changes in their volumes' (SNA 1993: 379). Only by eliminating price effects, and valuing each year's GDP accordingly, is it possible to derive a consistent measure of changes in output.

Such a measure is indispensable for establishing a true record of the past, for analysing and interpreting economic events and relationships, and for assessing current macroeconomic policy choices. As the SNA notes (para.16.1) one objective is

to assemble a set of interdependent measures which make it possible to carry out systematic and detailed analyses of inflation and economic growth and fluctuations.

For all the above purposes, a necessary first step is to separate out price and quantity elements.

Both nominal and real GDP can be expressed either as an output total, where the component parts comprise value added in different industries or sectors of or the economy, or as an expenditure total, made up of final expenditures plus exports less spending on imports. Although the items that make up the two aggregates are thus different, the totals are constrained to be the same: both emerge from the same matrix of transactions in goods and services for the period in question. Changes over time in real GDP are defined as equal whether particular estimates relate to output or to expenditure.

⁷ The SNA suggests (para.16.72) that the term 'volume GDP' is preferable to 'real GDP'. We use 'real' here because of its greater familiarity.

In separating price changes from quantity changes, both practical and analytical problems necessarily arise, and in neither case are there unique and agreed solutions.

Problems of estimation

The chief practical problems relate to the collection and interpretation of data. The process of separating the price and quantity components of economic time series is not costless. Resources have to be devoted to the collection, processing, interpretation and publication of data relating to prices and quantities over the relevant period or periods. Again, improvements in the quality of these data do not come free of charge. There is an inescapable trade-off between, on the one hand, wider and more detailed coverage and more firmly based results, and, on the other, keeping costs within reasonable limits. Inevitably, there is room for disagreement about how the balance should be struck. But no one doubts that the derivation of estimates of real GDP is necessary, or that improvements in the reliability of such estimates are to be welcomed as such.

No matter how detailed and meticulous the inquiries may be that yield estimates of changes in real GDP, there will remain room for doubt and for differences of view. Three factors in particular are involved here. One is changes in the quality of goods and services that may not be fully reflected in their prices. A second is the appearance of new products and the disappearance of older ones, so that the lists of goods and services entering into output over the period under review may not be identical. Both these factors loom larger if the GDP estimates in question relate to changes over a substantial period. A third problem arises from the presence of (to quote para.16.4 of the SNA) 'non-market goods and services whose valuation is difficult at current as well as constant prices'. However, the existence of these significant and unavoidable problems does not put in question the need for reliable estimates of changes in real GDP, nor does it make long-run intertemporal comparisons inadmissible.

Questions of interpretation

As to analytical aspects, there is admittedly no single and unique formula for measuring changes in a country's real GDP, since different sets of price weights can be used to value the respective outputs. If for example a comparison is being made between two years, outputs can be valued at the prices of the earlier year, thus yielding a Laspeyres quantity index (expressed as $\frac{\sum p_1 q_2}{\sum p_1 q_1}$); at those of the later year, yielding a Paasche index ($\frac{\sum p_2 q_2}{\sum p_2 q_1}$); or at an average of the two, as in the Fisher index. If the series is to extend over several years, the range of choice is wider, the possibilities of chaining come into play and the choice of reference year may become an issue. Once again, however,

these inherent features of the exercise do not put in question its purpose or its rationale. For one thing, there is now broad agreement on index number choices and procedures, as reflected in the recommendations set out in para. 16.73 of SNA 1993. More fundamentally, the fact that index number problems are inescapable, and that different ways of treating them can be defended, does not affect the need to derive estimates of real GDP as an essential measure of the course of economic change.

The choice of index number formula can be linked to a second and possibly more contentious analytical issue. This concerns the significance to be attached to estimates of real GDP over time: just what is it that is being measured, or should ideally be measured?

Although these are deep waters, we believe that a contrast can be drawn between two approaches or modes of interpretation. In the evolution of economists' thinking on the subject, issues of the definition and measurement of real GDP have often been considered in relation to its status as a measure or indicator of an economy's performance: thus in a classic and influential article published in 1940, John Hicks referred to, and joined himself with, 'a long line of economists who have sought in the Social Income an index of economic welfare, of the wealth of nations'.⁸ This orientation, or way of thinking, characterises what the SNA refers to (paras 16.21-30) as 'the economic theoretic approach to index numbers', in which 'the observed quantities may be assumed to be functions of the prices, as specified in some utility or production function'. Thus measures of real GDP are (or can be) interpreted as reflecting underlying shifts in individual utilities of consumers (where expenditure is under consideration) or in the economy's production possibility frontier (where the focus is on output); and a lot of work has been done to define precisely, first, the conditions under which this relationship would hold good, and second, the price deflators that it would then be appropriate to use.

In our view, an alternative (or complementary) approach is possible, in which the movement in real GDP is viewed and interpreted, at least initially, in more neutral terms. It can be defined, in the first instance at any rate, as no more than the estimated volume component of a series that is originally given in value terms, i.e., in current prices. As such, its status is descriptive only: in itself, it does not pretend to convey information about changes in welfare or production possibilities, any more than an index of the volume of exports or of industrial production does. The relationship of real GDP – or, more strictly, real GDP per head – to economic welfare represents a second and separate stage of inquiry, in which various other influences have to be taken into account,

⁸ JR Hicks, 'Valuation of Social Income', originally published in 1940 in *Economica*, reprinted as Chapter 3 of *Wealth and Welfare: Collected Essays on Economic Theory*. Volume 1, Oxford, Blackwell, 1981.

starting with changes in the terms of trade over the period in question. To be sure, this positivist approach leaves open the choice of price deflator. However, adopting it does not prejudice the choice, nor does it affect the broad agreement on index numbers referred to above. Moreover, and as will be seen below, such a neutral interpretation of real GDP estimates may be seen as more appropriate in the context of many international comparisons.

The irrelevance of exchange rate changes

In estimating current-price GDP series, as also in correcting these for price changes so as to derive a series for changes in real GDP, only the prices at which the outputs for each year were actually transacted are relevant. Just as the real or volume component of current-price GDP is the output produced, or the real final expenditure incurred, within the country or economy concerned, so the price component is derived from the prices at which the goods and services in question were bought and sold within it. Hence in constructing measures of price and output changes, *exchange rate changes do not enter in*: to 'correct' a real GDP series for such changes is wrong.

A recent instance of such an error, which itself follows in the footsteps of others, is to be found in Jagdish Bhagwati's book *In Defence of Globalization* (New York, Oxford University Press, 2004). In Chapter 13 of the book Bhagwati refers to 'the perils of gung-ho international financial capitalism'. He states there (199), by way of illustrating those perils, that in the financial crises which affected five East Asian economies in the 1990s:

Per capita incomes tumbled to almost one-third of their 1996 level in Indonesia, with the other crisis-stricken Asian countries showing declines ranging from a quarter to nearly half of the 1996 levels.

From the source that Bhagwati quotes, which itself gives another source, it becomes apparent that these figures refer to changes in GDP per head as between calendar years. However, actual GDP per head for Indonesia, according to standard published data, fell by only about one-twelfth, as opposed to two-thirds, between 1996 and its lowest subsequent point in 1998. Again, the corresponding largest falls for four other 'crisis-stricken' countries over the period 1996-99 ranged, not from 'a quarter to nearly half of the 1996 levels', but from $7\frac{1}{2}$ per cent in the case of South Korea to just under 14 per cent for Thailand.⁹ The alarmist numbers that Bhagwati quotes, and which he believes to carry implications for the external policies that developing countries should

⁹ The figures that we quote here, and others that appear below, are taken from Angus Maddison, *The World Economy: Historical Statistics* (Paris OECD Development Centre, 2004).

follow, appear to reflect some process of revaluing GDP totals so as to reflect the big falls in exchange rates that occurred during the crises. This is not how changes in real GDP are defined and measured.¹⁰

The argument here can be extended from the record of the past to projections of the future. Just as with past changes, the prospective growth of real GDP in an economy is defined and measured with reference to output and expenditure at constant domestic prices. It is wrong either to adjust the results of such a projection by building in assumptions about possible future changes in the exchange rate, or to rest the projection itself on such assumptions. When assessing or modelling the possible future growth of output, the exchange rate does not directly enter in.

Prices and quantities: Cross-country comparisons

Nominal and real differences in output and expenditure

Where comparisons between two countries are in question, whether for a given year or over a series of years, there is a close analogy with the procedures that are called for, and the problems that arise, when making estimates of changes over time in real GDP for a particular economy.

As in the inter-temporal case, the point of departure for a cross-border comparison is the respective estimates of GDP in value terms; that is, at current prices. These current price estimates are given in the respective currencies of the two countries. The GDP of the one economy can thus be translated into that of the other, often though not always straightforwardly, by conversion at the average exchange rate for the period in question. But just as the current price estimates for a single country for two different years take no account of price *changes* as between the two, so an exchange-rate-based conversion of the money GDP of two countries in a particular year takes no account of price *differences* between them. *It therefore does not yield a measure of comparative output.* Only by eliminating price effects, and thus valuing each country's GDP at a common set of prices, it is possible to derive a valid measure of differences in real GDPs. For cross-border comparisons as for inter-temporal estimates of real GDP, price

10 To be sure, in such episodes changes in GDP per head are not the whole story. Because of the very large turn-round in their current account balances, the crisis-affected countries, as Martin Wolf has noted, had to 'reduce spending in relation to national incomes by an amount equal to 15.5 per cent of their combined GDP' – a huge adjustment in so short a time. Further, real incomes were reduced by adverse changes in the terms of trade – which, however, are not correctly measured by the fall in nominal exchange rates.

and quantity components have to be separated out, so that the respective GDPs are expressed in a common set of prices. This is achieved by the use of PPP (purchasing power parity) ratios or converters.

The SNA gives clear guidance on this matter. In its opening chapter, it specifies (para.1.38) that:

When the objective is to compare the volumes of goods or services produced or consumed per head, data in national currencies must be converted into a common currency by means of purchasing power parities and not exchange rates ... Exchange rate converted data must not ... be interpreted as measures of the relative volumes of goods and services concerned.

Just as with inter-temporal comparisons, attention may be focused on differences in either real output or real expenditure; and here again, the two totals are defined as being equal. It is true that most cross-country PPP comparisons have so far been made with reference to final expenditures rather than output totals, but this is largely for reasons of data availability.

Again as with inter-temporal comparisons, and largely for the same reasons, both practical and analytical problems arise in making the separation between prices and quantities.

Problems of estimation

The practical problems are the same in kind, though they are often worse in degree. Here again, resources have to be devoted to the collection, processing, interpretation and publication of data relating to the prices of comparable goods and services in the two countries; and in doing so, it is necessary to make judgements as to what is genuinely comparable. Again, improvements in the quality of such estimates are not costless, and there is the same trade-off between achieving wider and more detailed coverage and keeping costs within acceptable bounds. Again, similar problems may arise because differences in quality are not reflected in price differences, because the respective product lists are not identical, and from the presence of non-marketed goods and services. All these problems can be more serious in cross-country comparisons than when changes in a single economy are in question: the SNA notes (para.16.81) that 'there is little doubt that it is more difficult to compile reliable international than inter-temporal price indices'. But as with intertemporal comparisons, the existence of the problems does not undermine the case for making comparative estimates which relate to real GDP, rather than using actual exchange rates on no better grounds than that they are readily available, and despite the fact that they do not yield measures of comparative output.

Questions of interpretation

The analytical issues that arise are likewise similar. As in the inter-temporal case, the results that emerge will depend on the choice of price weights: index number problems arise in the same way and for the same reasons. Thus for a particular year the respective outputs of goods and services (or real expenditures) can be valued at the prices of Country A ($\sum A_p B_q / \sum A_p A_q$) or of Country B ($\sum B_p B_q / \sum B_p A_q$): the choice is precisely analogous to that between a Laspeyres and a Paasche index for a comparison in a single country as between two years. In this case also, there may be good arguments for using an average such as the Fisher index. Again as with intertemporal comparisons, the fact that these index number problems have to be faced does not undermine the case for the deflating the value series in national currencies by a suitable price index, so as to derive a measure of differences in real GDP.

In the international context also, the question arises as to how such differences are to be interpreted; and here again, the same argument as for inter-temporal comparisons can be made for adopting a neutral or 'positive' interpretation. Indeed, the argument is stronger, since cross-country differences, where the economies in question may have widely different properties, are more liable to put in question the assumptions that enter into the construction of 'economic-theoretic' measures of comparative GDP. Thus PPP-based estimates of cross-country differences in real GDP are best viewed, at any rate initially, as no more than that. They are measures of comparative outputs, neither more nor less. *As such, they do not measure comparative living standards, productivity or welfare, though they can be used, along with other evidence and on clearly specified assumptions, to throw light on all of these.*

The irrelevance of exchange rate changes

Since exchange rate changes, past or prospective, do not enter into measures of changes in the real GDP of a single economy, they are not relevant when comparing differences in real GDP growth across countries. Recent comparative data relating to the US and the Euro Area can serve to illustrate the point. From recent OECD published data, it appears that between 2001 and 2004 the aggregate GDP of the Euro Area rose by an estimated 3.3 per cent, as compared with 9.6 per cent for the US. Over the same period the euro rose in relation to the US dollar by some 38 per cent; and this occurred despite the fact that there was little difference in the respective increases in the two domestic price levels.¹¹ Hence a series for Euro Area GDP at constant prices, if converted from

11 These figures are taken from OECD Economic Outlook, No. 76, December 2004. The estimated rise in the consumer price index for the Euro Area between 2001 and 2004 is 6.5 per cent, which compares with 6.6 per cent for the US. The corresponding figures for the GDP deflator are 6.5 per cent for the Euro Area and 5.6 per cent for the US.

euros to US dollars at the rates that prevailed, would show a rise of over 40 per cent for the period. Conversely, the corresponding series for the US, if converted from dollars to euros, would show a fall of over 20 per cent. But neither of these two figures has any economic meaning, nor do they offer alternative estimates for changes in real GDP. These latter estimates are what they are: they are derived, without reference to exchange rate changes, by correcting each of the respective value series for GDP for changes in domestic prices. In this particular comparison, they would not be affected if the movement of exchange rates over the period had been in the opposite direction. Since exchange rate variations have no bearing on how the growth of output is defined and measured, they have no bearing on comparative growth rates. In this context also, they do not enter in.

The above example also serves to illustrate a related point. It is well recognised that making cross-country conversions of GDP at actual exchange rates, rather than using PPP converters, causes the gap between rich and poor countries to appear as much greater, chiefly because the prices of non-traded goods are typically lower in poor countries. This overstatement of the gap in real GDP and GDP per head is often viewed, with good reason, as an argument in favour of using PPP-based estimates when comparisons between these two groups of countries are in question. But the argument against bringing in exchange rates is a general one, which applies to other comparisons also. As is shown by the above example of the Euro Area and the US – and many other past instances could be cited¹² – market exchange rates may have a momentum of their own, which is unrelated to changes in the outputs or price levels of the countries concerned. In such cases, actual exchange rates give seriously misleading results if used to make comparisons of output as between countries even with similar levels of real GDP per head. It is not only as between rich and poor countries that the *Purchasing Power Parity theory* (or hypothesis), according to which exchange rates reflect and move in line with comparative national price levels, has not been borne out by actual events.

The argument here can be extended to the measurement of combined GDP totals, with the above figures again taken as an illustration. A measure of the change in the combined GDP of the Euro Area and the US, as between 2001 and 2004, has to be a weighted average of the 3.3 per cent increase for the one and the 9.6 per cent for the other, and the weights should reflect the relative size of the two economies. Since the US economy grew faster over the period, its relative weight for 2004 should be higher than for 2001. If however the respective GDPs

12 For example, Irving Kravis, in the article mentioned above, noted (2) that 'exchange rate conversions indicate that Japan's per capita GNP was 47 per cent higher than that of the United Kingdom in 1978 and 5 percent lower than the UK level in 1980 ... The Japanese constant price series for GNP shows an increase of about 8 per cent on a per capita basis while the UK constant price series shows an approximate decrease of one per cent'.

are converted at actual exchange rates, the reverse change occurs: because of the appreciation of the euro: the Euro Area GDP appears as having increased in relation to that of the US. Such a result does not make economic sense. In measuring the growth of output for two or more countries grouped together, or for the world as a whole, the appropriate weights are the comparative real sizes of the economies concerned in some agreed base period, and these can be derived only from a PPP-based comparison. The correct procedure, as summarised by Nordhaus in the abstract of the article cited above, is to combine 'cross-sectional PPP measures for relative incomes and outputs' with 'national accounts price and quantity indexes for time-series extrapolations' (and also, we may add, for estimates of past growth).¹³

Exchange rate-related confusions: Projecting the closure of an imagined gap

In some model-based projections, such as those that enter into the SRES, it has been the practice to start with base-year cross-country (or cross-region) GDP data valued at actual exchange rates. These are treated, wrongly, as yielding measures of real GDP per head; and as a result, the initial gap in GDP per head between rich and poor countries is greatly overstated: thus in the SRES the ratio of GDP per head in 1990 in the OECD group to that of the Asian developing countries is put at close to 40 to 1, as compared with a figure of around 9 to 1 that emerges from Maddison's estimates. Building in this overstated gap increases the apparent scope for future convergence in GDP per head: convergence is defined in nominal rather than real terms, though such a notion has little if any economic meaning. If such 'nominal convergence' is built into a model, it gives rise to projections of output and GDP per head for developing regions which are higher than they would have been if the base year figures had been correctly derived, from PPP-based comparisons, and the gap had accordingly been smaller. In such inflated projections for the growth of GDP in poor countries, therefore, there are two elements, one genuine and the other imaginary. The first reflects higher growth in real GDP per head, correctly defined, while the second represents no more than the assumed closure of an imaginary initial gap.

In some models, a further element, which should likewise have no place, enters into 'nominal convergence'. The assumption is made, and reflected in projections, that in the case of initially poor countries the gap between 'MERS' and PPP converters, which arises from relative poverty, will be gradually closed over time. By tracing out the projected path of such a change, it is possible to express

¹³ Where cross-border comparisons extend to more than two countries, the analogy with intertemporal comparisons no longer holds good, and a question arises as to how best to 'multilateralise' comparisons. The issues which then arise are reviewed in all three of the sources cited in footnote 2 above.

the GDP of the groupings of poor countries either in the original 'MER' units or in terms of exchange rates that are projected to converge on what are taken to be PPP values. The bizarre result of such a procedure is that, to quote a recent article by Richard Tol, 'Developing countries grow slower with a purchasing power exchange rate than with a market exchange rate'.¹⁴ But as already noted, the projected growth of output for any country or group of countries has to be derived from national price and quantity measures, suitably weighted by relative outputs where more than one country is involved. Assumptions relating to the possible future course of exchange rates are not relevant to the choice of a unit of account for valuing projected GDP, whether for a single economy or for a group of countries. Exchange rates, whether past or projected, do not enter into the measurement of changes in real GDP. Where output is in question, there is no such thing as 'MER growth rate' or a 'PPP growth rate'. These are no more than figments of modellers' imaginations. They acquire meaning only in relation to a process of assumed 'nominal convergence' which itself has no economic meaning.

A puzzle

If the above arguments hold good, a question at once arises. If exchange-rate-based cross-country comparisons of GDP are analogous to current-price inter-temporal comparisons of GDP for a single economy, in that neither yields a valid measure of comparative real GDP, how is it that the need to correct for price changes is universally accepted in the latter case but widely rejected in the former? Answering this question throws light both on the rationale of PPP-based measures and on some reasons currently given for not accepting them for what they are.

PPP and 'MER': Some still-prevalent mistakes

Misconstruing PPPs

The meaning and rationale of PPP measures remain subject to various ill-founded notions. One such notion is that they do not arise from, or reflect, actual transactions. For example, Professor Lord Desai, in an intervention last year in the House of Lords (21 April 2004), expressed the view that

nobody pays their bills in purchasing power parity: they pay it in real money, which is based on market exchange rates.

¹⁴ Richard Tol, 'Exchange Rates and Climate Change: an Application of *FUND*'. Within the SRES, a so-called PPP-based set of projections is given as a variant of the MESSAGE scenario group.

The fact is that PPP converters, as is clear from the index number formulae quoted above, are built up from detailed price comparisons that are based on actual purchases made in the respective domestic currencies in each of the countries concerned. For any pair of countries, and for all market transactions covered, 'real money' is involved.

A related argument is that while exchange rates are really existing entities, the values of which are known, PPP converters are no more than artificial constructs. Thus a group of authors associated with the SRES have argued, in responding to the critique of the Report that we had made, that

...the main appeal of market exchange rates is that they can actually be *observed* in the market transactions. In contrast, PPP need to be estimated by statistical offices and international organizations.¹⁵

But as just noted, PPP converters are derived from information relating to actual observed market transactions. It is true that they have to be estimated, and that different weighting procedures will yield different results: but the same is equally true of the price indices that are used to deflate current price series in order to derive estimates of real GDP for a single country. Even where 'MERs' can be readily observed, they do not yield a measure of differences in output.

On similar lines, some commentators view PPP converters as inferior *ersatz* exchange rates. In contrast to actual observed exchange rates, PPP estimates are portrayed as a costly and questionable rival product, built up from the obscure labours of statisticians round the world whose time could have been better spent. But PPP converters are not, and do not pretend to be, exchange rates: they are price index numbers that result from systematic inter-country price comparisons. Their purpose and rationale is to enable cross-country comparisons of output to be made, while at the same time yielding cross-country price comparisons which are of interest in their own right. The fact that they may be subject to uncomfortably wide error margins – as a result of data limitations, problems of comparability, and the relatively modest resources that often go into their compilation – is not a reason for giving up the task of producing and improving them, still less is it an argument for using exchange-rate-based conversions. As Jacob Ryten has correctly observed, 'The only viable alternative to the use of inadequate PPP-based estimates is better PPP-based estimates'.¹⁶

15 The quotation is taken from p.191 of 'IPCC SRES Revisited: A Response', by a group of 15 authors, published in *Energy and Environment*, Vol. 14, Nos 2 and 3, 2003.

16 Jacob Ryten, 'MERs, PPPs and IPCC: illusion and reality'. *Energy and Environment*, Vol. 15. No. 3, 2004.

Misdefining the rationale of PPP-based comparisons

A widely held mistaken belief is that PPP converters are appropriate for making international comparisons of GDP aggregates only when those comparisons relate to 'living standards': when comparing outputs and shares in world GDP, and in computing total world GDP, a 'MER'-based measure of some kind should be used. For example, Richard Cooper has argued, in a letter to *The Economist* (18 June 2004), that:

... how best to measure global output depends on the purpose of the measurement. Some variant of PPP is surely needed to compare standards of living across countries but for that we do not need global output. If we are interested in the vigour of global demand, national demand (and output) must be added by using market exchange rates, not PPP rates.

A similar distinction enters into the long established and still continuing practice of the World Bank, as expressed in its *World Bank Atlas*. The Director of the Bank's Development Data Group, in writing to one of us recently, expressed the view that 'we all agree that comparisons of welfare between countries are better stated using purchasing power parities'. When it comes to output, however, the latest edition of the *World Bank Atlas*, which is subtitled *Measuring Development*, states (6) that:

While more than 80 percent of the world's people live in developing countries, their economies in 2003 produced goods and services worth \$7.1 trillion, about one-fifth of the world's total output ... To make comparisons between countries, local currencies must be converted to a common value ... Valuations based on exchange rates better measure the tradable value of a country's output and a country's relative importance in the global economy.

As against such ways of thinking, the following points are relevant.

- Global output is the sum total of the outputs of the countries that comprise the world economy. These can be compared one with another only by correcting for differences in the prices of the goods and services that enter into output - that is through the use of PPP converters.
- It makes no sense to argue, as the World Bank does, that 'exchange rates better measure the tradable value of a country's output', since there is no basis for estimating the comparative tradable value of non-tradable goods.
- In a world where exchange rates are often volatile, and often do not track changes in national price levels, there is no satisfactory basis for choosing a particular rate, or constellation of such rates, as valid or representative for cross-country comparisons of output. In this connection, it is worth noting

that even trade volumes, for country groupings as well as for individual economies, are computed without reference to the exchange rates that prevailed: export and import values are deflated by their own price or unit value indices.

- It is unclear what economy-wide national accounts aggregate, to be distinguished from real GDP and potentially divergent from it, is supposed to be the appropriate measure, with PPP converters as a basis, for making comparisons of 'living standards'. As noted above, cross-country PPP-based comparative figures for GDP do not measure, and do not claim to measure, differences in living standards.
- In the letter quoted above, Cooper refers to the problems of making accurate cross-country price comparisons, 'reflecting both the difficulty of finding truly comparable goods and services in different countries, then deciding how to weight them'. But the problems are no different in kind from those that arise in making long-run comparisons for individual economies.

A variant of the 'living standards' line of thinking finds expression, among other places, in the article by Richard Tol already quoted. The notion here is that the function of PPPs is to measure and compare inter-country differences in the cost of buying 'a standard basket of goods'. This too is a misapprehension. PPP converters are not defined with reference to a specific 'basket': their purpose (to repeat) is to enable estimates to be made of cross-country differences in real GDP, and following from this, of movements in the real GDP of country groupings and the world as a whole. Each economy has its own 'basket', which comprises its own aggregate output of goods and services.

In one more misdefining comparative output in this latest edition of its *Atlas*, the World Bank gives further currency, in the excerpt quoted above, to a widely repeated but seriously misleading statement. The statement is to the effect that the developing countries, with 80 per cent of the world's population, produce (or, on some interpretations, 'receive') only 20 per cent of total world GDP, while the rest produce (or, on some interpretations, 'take' or control) the remaining 80 per cent. This comparison is a favourite one with the President of the Bank, James Wolfensohn; and indeed it occurs in the most recently published issue of *World Economics*, in an article written by two World Bank staff members.¹⁷ A more soundly based estimate, derived from PPP-based comparative data, is that the current share of developing countries in world GDP is well over 40 per cent.¹⁸

17 Kirk Hamilton and Ian Johnson, 'Responsible Growth to 2050', *World Economics*, Vol 5, no 4, October-December 2004. Three outrageous variations on the '80/20' theme, one of them from Wolfensohn, are quoted in David Henderson, *The Role of Business in the Modern World: Progress, Pressures and Prospects for the Market Economy*, London, Institute of Economic Affairs, 2004 (83-4) and Washington, DC, Competitive Enterprise Institute, 2004 (91-2).

18 Angus Maddison's estimates for 2001 would suggest a share of 42.5 per cent, while rapid growth in China and India since that year may well have pushed this figure up a bit. IMF projections for 2005 suggest a figure of close to 43 per cent.

Treating the issue as trivial or unimportant

In some quarters, the choice between exchange-rate-based and PPP-based conversions is wrongly seen as of little or no consequence. Thus in the IPCC press release of December 2003, referred to in footnote 4 above, the statement is made that 'the economy does not change by using a different metrics (PPP or MEX), in the same way that the temperature does not change if you switch from degrees Celsius to Fahrenheit'. A different variation on this theme of inconsequence was provided last April by Lord Desai, in his remarks in the House of Lords already quoted. He advanced then the view that, '... whether one takes one measure of GDP or another is very much a matter of taste'.

In weighing such dismissive assertions, the case of China today is instructive (though many other instances could be given). Last year one of us (Henderson) made the following comment, to no visible effect, on a speech by the present Leader of Her Majesty's Opposition:

Michael Howard is quoted in today's *Times* (29 April) as saying that China's GDP is now close to that of Britain, and that the UK is the fourth largest economy in the world. Howard should know, and certainly his advisers should, that these are highly questionable statements. In my opinion, they are seriously misleading... Taking Angus Maddison's most recent estimates, which are for 2001, the GDP of China in that year appears as 3.8 times that of the UK. China is now easily the second largest economy in the world, while the UK comes seventh.

Over the period 2001-04 the gap has widened, since Chinese GDP appears to have grown by well over one quarter, as compared with an increase in the UK of some 7.5 per cent. Hence the ratio of Chinese to British GDP might now be put at around four and a half to one – and indeed, a figure of approximately 4.4 for 2005 is indicated by recent IMF data.

Whether the GDP of China today is roughly equal to that of the UK or well over four times as large is neither a mere matter of taste nor a trivial issue of choosing between two equivalent 'metrics'. The same is true of broader contrasting estimates of the gap in GDP per head between rich and poor countries. Thus in the recent report of the 'World Commission on the Social Dimension of Globalization', commissioned by the International Labour Office (ILO), the relative gap for the period 2000-02 between the twenty richest and the twenty poorest countries in the world is put at 120 to 1, whereas a PPP-based estimate would suggest a figure of around 40 to 1.¹⁹

19 *A Fair Globalization: Creating Opportunities for All*, Geneva, International Labour Office, 2004: 37. The figure of 40 to 1 is based on Maddison's estimates. The World Commission report was reviewed by one of us (Henderson) in *World Economics*, Vol 5 No 3, October-December 2004. Among the members of the Commission were two prominent academic economists, Professors Deepak Nayyar and Joseph Stiglitz.

Misleading figures for GDP can go together with correspondingly misleading figures for GDP-related ratios. A recent instance is contained in a study published last year by the International Energy Agency (IEA), entitled 'Analysis of the Impact of High Oil Prices on the Global Economy'. In reviewing the impact across countries, the study provides comparable measures of oil intensity, defined as 'primary oil consumed per unit of GDP'. In deriving this figure for different countries and country groups, the GDP of the countries covered is expressed in a common unit of measurement through conversion of national currencies into US dollars at actual exchange rates. As a result of this procedure, it is wrongly made to appear that (to quote a leading instance from the IEA text) India 'uses more than two and half times as much oil as developed countries per unit of GDP'. This again is not an instance merely of 'Celsius versus Fahrenheit'. It is not a trivial matter, nor a matter of indifference in judging the impact of higher oil prices on the Indian economy, whether that economy uses well over twice as much oil per unit of GDP as developed countries, as asserted by the IEA, or (as PPP-based estimates would suggest) much the same amount.

In all cross-country economy-wide comparisons, as also in relating such series as energy use or CO₂ emissions to GDP, it is real GDP – in other words, output – that is relevant. Only conversion of nominal GDPs through PPP converters yields a measure of cross-country differences in output: and the substitution of PPPs for 'MERs' in defining and measuring such differences can have a significant or even dramatic effect on how the world appears and how events and relationships are interpreted.

Disregarding evidence

In weighing the twin rival criticisms of PPP converters just outlined – that using them makes no difference to anything that matters, and that they have their uses but these do not extend to measuring differences in real GDP – facts and figures can shed some light. Again, the case of China today provides a good example. On an exchange-rate-based comparison, the share of China in world GDP in 2003 was close to 4 per cent, while for a PPP-based computation the figure can be put at around one-eighth. In considering which of these two widely divergent estimates may be nearer to the mark, the following illustrative data are of interest:

- *Agriculture*: According to the FAO database (FAOSTAT), in 2004 China's estimated share of the global output of selected agricultural commodities was: cereals, 19 per cent; fibre crops, 25 per cent; fruit, 15 per cent; vegetables, 50 per cent; roots and tubers, 26 per cent; meat, 28 per cent; eggs, 45 per cent; wool, 15 per cent; and milk 4 per cent.

- *Industrial commodities*: According to the *Industrial Commodities Yearbook 2000*, produced by the Statistics Division of the United Nations Department of Economic and Social Affairs, in 2000 China's share of the global output of major industrial commodities was: iron ore, 18 per cent; wheat flour, 31 per cent; cotton yarn, 40 per cent; cotton woven fabrics, 48 per cent; woollen woven fabrics, 29 per cent; sulphuric acid, 25 per cent; cement, 37 per cent; crude steel ingots, 17 per cent; refrigerators for household use, 18 per cent; washing machines for household use, 25 per cent; colour television receivers, 39 per cent; telephones, 49 per cent; and watches, 26 per cent.
- *Communications*: From information in the *Human Development Report 2004*, published by the United Nations Development Program (UNDP), it can be calculated that in 2002 China accounted for 20 per cent of the world total of telephone mainlines, 18 per cent of the world's cellular telephone subscribers and 10 per cent of the world internet subscribers.
- *Energy*: According to *Energy Balances of Non-OECD Countries 2001-2002*, published by the IEA, China in 2002 produced 12 per cent of the world's energy and accounted for 12 per cent of the world's total primary energy supply. China was responsible for 10 per cent of the total generation of electricity, and also for 10 per cent of the final consumption of electricity for all uses.

In the light of such evidence, it is absurd to argue that the choice between exchange rate-based and PPP-based estimates of the comparative size of the Chinese economy is arbitrary, unimportant, or no more than a matter of personal taste. Further, it is not credible to maintain, which as just noted the World Bank has recently done, that 'the tradable value of China's output', and its relative importance in the global economy, are better indicated by the country's share of global GNI measured at actual exchange rates (4 per cent) than by the corresponding share measured at purchasing power parities (12 per cent).

PPP and 'MER': The thinking and practice of international agencies

Official responsibility for the production and diffusion of international comparative economic data rests primarily with international agencies, though those agencies are subject to direction and supervision by their member governments. Over the past 15 years or so, there has been what Angus Maddison has termed a 'creeping acceptance' among the agencies of the case for basing comparisons of real GDP on PPP estimates, rather than relying on exchange

rates, and for modifying published official statistics accordingly. However, as the following brief notes on agency practice suggest, the creeping process still has a long way to go.

IMF

A major step was taken in 1993 by the International Monetary Fund (IMF), when PPP-based weights were adopted as the basis for computing changes in world GDP in the Fund's *World Economic Outlook* (WEO). Since then the Fund's analyses of the world economy have relied entirely on PPP-based measures for inter-country comparisons of income and output and estimates of changes in world GDP. At the same time, however, the Fund's *International Financial Statistics* database, which presents 'MER' conversions, is a widely used source of GDP estimates from 1948 onwards; and by maintaining these series for all countries, the IMF is helping to perpetuate the inappropriate use of data that are based on such conversions.

OECD

Chronologically, the first displacement of exchange-rate-based estimates by a PPP-based alternative may well have been in the OECD, where the decision was made in 1990 to use estimates of GDP (PPP) to determine the country weights in the OECD-wide consumer price index. A similar reweighting was later introduced for the totals presented in the Organisation's *Economic Outlook*. However, in the Organisation's annual *National Accounts of OECD Countries: Main Aggregates* exchange-rate-based GDP comparisons are published in parallel with their PPP-based counterparts; and the main tabular presentation of comparative levels of GDP in *Monthly Economic Indicators* is done (a) in current prices and exchange rates and (b) in the prices and exchange rates of 2000.²⁰ Again, in its new publication, *OECD Factbook: Economic Environmental and Social Statistics*, the Organisation has incorporated uncritically exchange-rate-based data provided by the IEA.

IEA

Although one of the IEA's lapses has been noted already, a further word is in order here. The Agency's current practice is inconsistent. Its flagship publication is the biennial *World Energy Outlook*, which presents itself, on the back cover of its 2002 edition, as 'the authoritative source for projections of

²⁰ In an OECD publication entitled *OECD Environment Outlook 2001*, it was stated (66) that the OECD countries accounted for 80 per cent of world GDP, which of course was an exchange-rate-based (and misleading) figure. However, this report appears to have had no successors.

global trends in energy supply and demand, trade and investment and carbon dioxide emissions'. Since the 1998 volume, all of the projections of output in successive editions of the *Outlook* have been PPP-based, and none has been expressed in terms of 'MERs'. The 2002 edition includes the statement (40, note 4) that:

All GDP data in this report are expressed in 1995 dollars using purchasing power parities (PPPs) rather than market exchange rates...*This is important in analysing the main drivers of energy demand or comparing energy intensities among countries* (italics added).

Notwithstanding this firm (and correct) statement of doctrine on how to measure energy intensities, the Agency published last year (as noted above) comparative figures for oil intensities which, by contrast, incorporate exchange-rate-based conversions and hence give seriously misleading results. Again, its *Energy Balances of OECD Countries 2000-01:2003* edition presented an analysis of the relative importance of the OECD countries collectively in the world economy, and of their relative energy intensity as a group, which relied entirely on 'MER'-based estimates of GDP.²¹ And the discussion in the text of the 2004 edition of the companion document (*Energy Balances of Non-OECD Countries*) assumes without explanation that 'world GDP' is the aggregate of country GDPs using exchange rates. On p. 1.50 it is stated that:

In 2002, the world GDP experienced a 1.9% growth explaining in part a 2.2% growth in Total Primary Energy Supply. From 1971 to 2002 as an annual average, the growth of GDP and in TPES were respectively, 2.9% and 2.2%.

The GDP increases quoted here are incorrect. As one of the tables in the publication shows, estimated PPP-based world GDP increased between 2001 and 2002 by 2.9 per cent, not 1.9 per cent – so that it was higher, not lower, than the growth in TPES. Further, it can be seen from the same table that the annual average growth of world GDP from 1971 to 2002, correctly measured, was not 2.9 per cent but 3.3 per cent.

World Bank

Conspicuous inconsistencies of treatment are likewise to be seen in the past and current practice of the World Bank. In the 1997 edition of its annual publication, *World Development Indicators* (WDI), the bank gave estimates for a range of countries relating energy use to GDP in which the GDP figures

21 It included the statement, which belongs with the class of groundless '80/20' assertions already referred to, that 'In 2001, about 19 per cent of the world population lived in the OECD, but more than 80 per cent of the world GDP was created in its 30 member countries'.

had been converted into US dollars using actual exchange rates: and it then said, misleadingly, that 'Per dollar of GDP, developing countries produce four times the carbon dioxide that high income economies do' (p.113). These same errors entered into the next two editions of WDI, but were corrected in the 2000 edition. However, the President of the Bank continued to use exchange rate conversions of GDP in his public speeches and in his personally signed 'Foreword' to each issue of the WDI and the *World Bank Atlas*; and in WDR 2003, published on the eve of the Johannesburg World Summit on Sustainable Development, the Bank repeated the statements that it had made in the late 1990s about the allegedly high relative consumption of energy in relation to GDP in non-OECD countries, which were said to use '3.8 times as much energy per dollar of GDP [as OECD countries]'

When it comes to comparative figures for GDP, the bank publishes, for example in its annual *World Development Report*, country estimates and group and world totals on both a 'MER' and a PPP basis: the two are shown together in the same table, as though they had equal claims to validity. At the same time, and as noted above, in its flagship publication, the *World Bank Atlas*, the Bank continues to give prominence and a special status to flawed exchange-rate-based comparative figures, for reasons which do not hold water.

This is not an impressive record, especially for an organisation that lays special claim to economic expertise. It is high time for the Bank to improve further its handling and presentation of international comparative data, so as to ensure acceptable standards of accuracy and consistency.

UNDP

A recent case of change for the better is that of the United Nations Development Program (UNDP), in the context of its annual *Human Development Report* (HDR). Although from the outset in 1990 the Report made use of PPP-based comparative figures, in the issues for 1992 and subsequent years the authors used exchange-rate-based estimates of GDP to make exaggerated and well publicised claims about the extent of global inequality in incomes and to argue that inequality was continuing to widen (although PPP-based estimates showed the opposite).

Following the release of the widely quoted 1999 *Human Development Report*, one of us (Castles) made extensive statistical criticisms of the treatment there of trends in global poverty and inequality. At the request of the 2000 meeting of the UN Statistical Commission, those criticisms were examined by a group of expert statisticians constituted as the Friends of the Chair of the Commission. The report of this group upheld the main criticisms. In particular, it took the view that HDR 1999 had made a 'material error' (i.e. one which left the reader with 'a fundamentally distorted view of the phenomenon being described') in

using national accounting aggregates converted into \$US at current exchange rates to compare output and living standards between countries. The HDR Office of the UNDP accepted the criticism, and has improved its statistical presentation and reporting in subsequent issues of the Report.

SRES, IPCC and UNEP

Among the recent reports which have given approving currency to the 'material errors' contained in the 1999 HDR are the SRES, two major IPCC reports that formed part of the Panel's Third Assessment Report of 2001, and a leading publication of the United Nations Environment Program (UNEP) which is one of the IPCC's two parent agencies (the other is the World Meteorological Organisation).

As noted already, the SRES starts from an assumption that 'MER'-based estimates and projections of GDP can be taken as a measure of relative levels of economic activity. Consistently with this point of departure, the Report presents estimates of energy intensity which are misleading in the same way as those just referred to from the IEA and (in some earlier years) the World Bank. For example, two charts in the Report (pp.97 and 125) show alleged historical trends in comparative energy intensities, but both use 'MER'-based GDP as the denominator. The SRES contains 17 double-columned pages of references, yet the SNA is not listed there: it is possible that none of the 53 authors, 4 review editors and 89 expert reviewers who took part in the preparation of the Report were aware of the System and what is said in it.

It is not only in the SRES that similarly inappropriate exchange-rate-based comparisons are to be found in recent IPCC documents. The Panel operates through three Working Groups, each of which produced its own full-scale report as part of the Third Assessment Report. WGI is concerned with scientific aspects of climate change, WGII with the prospective impacts of such change and ways of adapting to it, and WGIII with mitigation of the impacts. In the WGIII Report, the old-style HDR is taken as the source for comparisons of GDP per head between rich and poor countries which are exchange-rate-based. The report quotes an early HDR as saying that 'in 1998 the richest fifth of the world population received 82.7% of the global income, which is nearly 60 times the share of the income received by the poorest fifth (1.4%)'; and the text goes on to quote the 1999 HDR as showing that since 1988 this relative gap had widened.²² In the WGII Report also, these same UNDP comparisons are referred to, while misleading statements are made about the nature and purpose of PPP-based comparisons.

²² *Climate Change 2001: Mitigation*. A Report of Working Group III of the Intergovernmental Panel on Climate Change. Cambridge, Cambridge University Press, 2001: 97.

As just noted, the UNDP which was a prime source for these IPCC reports has now admitted its error, and has affirmed that international comparisons of real GDP should be PPP-based. This change of front on the part of the UNDP may have come too late to influence the drafting of the SRES and the two IPCC Working Group reports. One might however have expected that by 2003 news of the change would have reached the IPCC milieu, to be taken into account in a high-level conference organised by the Panel early that year in preparation for AR4. To the contrary: in the published proceedings of this conference one of the three current Vice-Chairs of the Panel, Professor Mohan Munasinghe, quoted yet again the same kinds of misleading international comparisons, long since disowned by the agency that had produced them, which had been referred to in the report of WGIII.

News of the UNDP's change of front was likewise not picked up by the UNEP, which made similarly misleading statements about 'the gap', also based on the 1999 issue of the HDR, in a major report which appeared in 2002.²³ In this flagship document it is also stated, contrary to the facts, that 'per capita incomes have risen *only marginally* in most regions [since 1972], *with the exception of Europe and North America*' (italics added).²⁴ The UNEP press release for the report says that over 1,000 people contributed to its preparation. It seems likely that, even aside from their shaky economic history, none of these persons was aware either of the admission made by the UNDP or of the existence of the 1993 SNA.

Paths to improvement

To ensure more accurate and more consistent treatment of international comparative data requires action by the member governments of responsible international agencies; but a wider understanding and agreement among economists and economic statisticians generally would be valuable in itself and would help to create and stiffen official resolve.

The main change required, a straightforward one, is agreement on the basic point that exchange-rate-based international comparisons of nominal GDP do not yield differences in output: to measure such differences, prices have to be directly compared, and PPP converters estimated accordingly. From this it follows (1) that comparative figures such as are to be found in the publications

23 UNEP, *Global Environment Outlook 3: Past, present and future perspectives*, London, Earthscan, 2002: 36.

24 On Maddison's estimates, average per capita incomes in Asia (excluding Japan), with 57 per cent of the world's population in 2001, almost trebled between 1972 and 2001. For China the increase was more than fourfold, while for the rest of developing Asia it was 129 per cent. The corresponding increase in per capita incomes in Europe and North America, with less than 12 per cent of the world's population in 2001, was 78 per cent.

of several international agencies, and are taken as the point of departure of the SRES, are misleading; (2) that to offer exchange-rate-based international series as an acceptable representation of historical events, and a data base that can be safely used by modellers, is a questionable practice; (3) that to make use of exchange-rate-based comparative GDP data in computing such ratios as energy intensities and emissions intensities is inadmissible; and (4) that in computing estimated changes in real GDP for country groupings or for the world as a whole, PPP-based weights should be used. Acceptance of these principles would open the way to major improvements in agency practice.

Official acceptance largely depends on actions, preferably taken in concert, within two areas of responsibility and expertise in member governments: first, national statistical offices; and second, the central economic departments of state – that is, treasuries, ministries of finance or economics, and, where they exist, agencies such as the US Council of Economic Advisers. In both areas, the basis for action has to be greater awareness of the issues, and of the scope for improving present international statistical systems and practice. A possible first step could be the setting up of a special commission of inquiry, along the lines of the UN Statistical Commission's 'Friends of the Chair' referred to above, but with wider sponsorship and terms of reference.

A special case

Among the various international mechanisms affected, the IPCC and its two supporting agencies constitute a special case. Here prompt action is called for, in the context of AR4, while resistance to change is strong and determined. We believe that a necessary first step is that national statistical offices and central economic departments of state should now become involved in the IPCC process; but the Panel, in responding to our critique, has rejected any suggestion of change. The opening paragraph of the official press release referred to above says of the IPCC that

It mobilises the best experts from all over the world, who work diligently on bringing out the various reports ... The Third Assessment Review of the IPCC was released in 2001 through the collective efforts of around 2000 experts from a diverse range of countries and disciplines. All of IPCC's reports go through a careful two stage review process by governments and experts and acceptance by the member governments composing the Panel.

In relation to economic aspects, however, there is good reason to question the claims to authority and representative status that the IPCC makes on its behalf. We do not question the numbers of those involved, their diligence, or the existence and observance of formal review processes. But we think that when it comes to the treatment of leading economic issues, the IPCC milieu is neither

fully competent nor adequately representative. We also hold that building in peer review is no safeguard against dubious assumptions, arguments and conclusions if the peers are all drawn from the same restricted professional milieu.²⁵

Fortunately a straightforward route to wider official involvement in the IPCC process exists for the taking. For the economic departments and agencies in OECD member countries, an instrument is to hand for their prompt collective engagement: it is the OECD itself. They should act now to ensure that IPCC-related economic issues are placed on the agenda of the OECD's Economic Policy Committee, where they could be reviewed by the Organisation's Economics Department, in conjunction with its Statistics Directorate and Environment Directorate.

Beyond officialdom

It is not only in official circles that greater awareness, together with closer acquaintance with generally agreed principles of national income accounting, could help to clear a path to improvement. As the preceding text has shown, there remain surprisingly wide differences of opinion among economists on how these issues should be viewed and handled: we have cited here the cases of several leading members of the profession whose current opinions, though not necessarily identical or consistent with one another, are at variance with what is laid down in SNA93. We hope that the arguments we have advanced here will help to increase awareness, to widen professional debate, and to extend the area of agreement.

ANNEX I

Improving the International Statistical System: Building on Angus Maddison's Work

The text that follows formed the basis for a presentation by one of us (Henderson) at a conference held in Canberra in 1999 by the Academy of the Social Sciences in Australia. The case that it makes is still pertinent today.

In my remarks this morning, I paid tribute to Angus Maddison's work as a source of data, analysis and ideas on the long-continuing story of economic growth and change across the world. Much could be said about his contribution, but here I want to stress three aspects of it in particular:

²⁵ It is not only in relation to economic aspects that such queries have been raised about the IPCC process and its results by critics writing about other subject areas.

- First is its *range* in both space and time. In principle, though subject of course to availability of data, every country in the world is covered. In every case, statistical series are laid out, and quantitative evidence presented, as far back in history as the often limited sources will permit.
- Second is *comparability*: estimates of GDP, as of some connected series too, are expressed in common units of measurement, so that comparisons can be made across countries and between different periods of time. Related to this is a third key aspect, which is that of
- *Continuity*: for most countries of the world, from whatever date the available evidence makes this possible, continuous annual series are presented.
- For the harassed user of figures, these three features combine to yield a store of published data which within its limits – for to be sure, not all the relevant aspects of economic change are covered – is uniquely rich, accessible and convenient to use. No other source in the world compares with it. In particular, the statistical output of the various international agencies, incorporating though it does, directly and indirectly, the work of thousands of professionals, has not to my mind yielded a comparably useful product of the same kind.

I believe that what Maddison has done offers ideas and lessons for improving the present international statistical system. In my view, one element in a program of improvement should take the form of a conscious attempt to build on Maddison's achievement and his vision of what is needed. His example should be followed, and his work continued, refined and extended, by the international agencies and the governments that support them.

Let me illustrate and clarify this notion with a concrete example to show the kinds of improvements I have in mind. Here I draw on my own experience – my own sad experience, let me add, since it relates to one of the failures of my years as Head of what was then the Economics and Statistics Department of the OECD.

Among its many statistical publications, the OECD produces an annual volume entitled *OECD Historical Statistics*.²⁶ Its coverage is restricted, quite properly, to the OECD member countries. Within this limit, care is taken to try to ensure that the figures are internationally comparable; and as compared with Maddison, the volume covers a much wider range of headings and series. So far, so good; but there are two serious limitations.

- The series go back only as far as 1960, the time of the earliest beginnings of the Organisation.
- Even for the years that have passed since then, the value of the publication as an historical source is fatally undermined by two features. First, it gives

²⁶ No longer true. Alas, since this presentation was made, the OECD has actually discontinued the publication.

data only for selected years, not continuous annual series; and second, many of the figures are presented in the form of year-to-year percentage changes, although these are useful only for short-term analysis.

During my time at the OECD I tried to remedy this situation. I decided that, rather than remodelling *OECD Historical Statistics*, we would replace it with a new, more ambitious and better-designed product. This would so far as possible take all the series back to 1900; and in every case, again subject only to availability of data, continuous annual figures would be shown: the idea was to establish a comprehensive '20th century data base' for the OECD countries. Two eminent scholars, Charles Feinstein and Angus Maddison, agreed to direct the project as consultants. With their help we drew up a project proposal. All that remained was to raise the quite modest sum of money required to get the new publication launched; but in this our best efforts, which I dare say were less skilful than they could have been, proved unavailing. None of the sources that we approached, public or private, was interested in supporting such a venture.

Despite its failure to take off, I think my idea at OECD was a good one. I would like to see a project of this kind designed and put into effect today, preferably covering the world as a whole and not just the OECD area (though an OECD venture would be an excellent start). I think that agencies and governments should now work towards establishing and maintaining a comprehensive and accessible database for the world, and for individual countries and groupings of countries within it.

I have put this suggestion in the context of long-run historical series, but it applies equally to recent developments and newly emerging statistics. My concern is not only for quantitative economic historians and historically minded economists, but also for commentators on the events of today and tomorrow. I would like to see current and newly published data made available within a similar comprehensive and accessible framework, which at present they are not.

As to roles and procedures, there are various possible lines of action, and I will not consider them today. The main thing now is to get the idea itself more fully worked out and put into circulation.