



The Space-time (In)consistency of the System of National Accounts: Causes and Cures

Nicholas Oulton (London School of Economics, UK)

Paper Prepared for the IARIW 33rd General Conference

Rotterdam, the Netherlands, August 24-30, 2014

Session 8C

Time: Friday, August 29, Afternoon

**The space-time (in)consistency of the
System of National Accounts:
causes and cures**

Nicholas Oulton

**Centre for Economic Performance,
London School of Economics**

July 2014

Table of contents

ABSTRACT	2
1. Introduction	3
2. The size of the inconsistency problem.....	5
2.1 Inconsistency in practice: OECD data	7
2.2 Inconsistency in practice: data from the Penn World Table	9
3. The inconsistency issue: insights from Divisia index numbers.....	11
3.1 Divisia index numbers: theory and derivation	11
3.2 The relationship between Divisia and Konüs index numbers.....	14
3.3 Compensated versus actual budget shares	15
4. The Divisia approach applied to cross-country comparisons.....	16
4.1 A discrete analogue to Proposition 2	19
5. Chain indices compared to the GEKS multilateral price index (PPP)	21
6. How much of measured inconsistency is due to path-dependence or chain index approximation?	24
6.1 Inconsistency due to path-dependence.....	24
6.2 Inconsistency due to chain index approximation.....	26
6.3 Conclusion on empirical importance of path-dependence and chain index approximation	27
7. Conclusions	27
TABLES	29
CHARTS AND FIGURES	34
References	37

ABSTRACT

Suppose you estimate the standard of living in a group of countries for some past year. You do this by converting each country's GDP per capita into a common currency using Purchasing Power Parities (PPPs) for that year. Now you want to see whether the gaps today are narrower than they were in the past. So you extrapolate each country's GDP per capita forward to the present using the growth rate of real GDP per capita from each country's national accounts. That will give you one answer to your question. Alternatively you could take each country's *current* level of GDP per capita, measured in its own currency, and convert these levels into a common currency using PPPs for the *current* year. This will give you a second and most probably different answer. So which answer is right? Or are both right? Or both wrong? This is what I call the problem of space-time inconsistency in the System of National Accounts (SNA) which I argue is of significant size empirically. I show that the SNA is in principle space-time *consistent* if the consumer's utility function is homothetic (or the revenue (GDP) function exhibits constant returns to scale) and if Divisia price indices are used to deflate nominal GDP or consumption, both over time and across countries. It follows that any observed inconsistency must be due to either (a) non-homotheticity in consumption (non-constant returns in production); (b) approximation error when chain indices are used instead of Divisia indices; or (c) errors in price indices. I develop indicators of the size of non-homotheticity and chain index approximation error. I conclude that errors in price indices are most likely the major cause of inconsistency.

JEL codes: C43, O47, I31, F43

Key words: SNA, ICP, PPP, Divisia, Konüs, path-dependence, consistency

1. Introduction

Suppose you estimate the standard of living in a group of countries for some past year. You do this by converting each country's GDP per capita into a common currency using Purchasing Power Parities (PPPs) for that year. Now you want to see whether the gaps today are narrower than they were in the past. So you extrapolate each country's GDP per capita forward to the present using the growth rate of real GDP per capita from each country's national accounts. That will give you one answer to your question. Alternatively you could take each country's *current* level of GDP per capita, measured in its own currency, and convert these levels into a common currency using PPPs for the *current* year. This will give you a second and most probably different answer. So which answer is right? Or are both right? Or both wrong? This is what I call the problem of space-time inconsistency in the System of National Accounts (SNA).

The existence of space-time inconsistency has long been recognised. But there is disagreement over its cause or whether it is a problem or just a fact of life. The viewpoint of the World Bank up to and including the 2005 round of the International Comparison Program (ICP) was that only the estimates in the latest round should be employed; all earlier rounds are to a greater or lesser extent unreliable (World Bank 2008). If adopted this advice would ensure that no inconsistency would ever be observed since only one set of PPPs would be used. (This advice has been criticised by Deaton (2010) and Deaton and Heston (2010)). The issue has been raised again by the release of the overall results of the 2011 ICP which are very different from what would be expected on the basis of extrapolating from the 2005 round (Deaton and Aten 2014).

PPPs are a critical building block in the widely-used Penn World Table. Prior to version 8, successive versions of the Penn World Table (PWT) used weights for the final expenditure components of GDP which derived from the latest round of the ICP. The problem here was that countries which were growth stars on one round turned into growth dogs on a later round (Johnson *et al.* 2013). Version 8 of the Penn World Table has adopted the opposite approach. In the main part of the table growth rates of GDP and its components are based on successive sets of PPPs, interpolated or extrapolated where necessary so that for each country in the tables estimates can be presented for every year from 1955 to 2005 (Feenstra *et al.*, 2013).

The advantage of this approach is that the addition of new, more recent PPPs will not change any of the earlier growth rates. But a completely different set of time series for GDP and its components based on national accounts are also made available in PWT 8.0. These growth rates are often strikingly different from the PPP-based ones (see e.g. Figures 1 and 2 in Feenstra *et al.* 2013). Feenstra *et al.* (2013) take space-time inconsistency as a fact of life; they argue that the national accounts are just measuring something different from the measures based on PPPs.

Throughout this paper I adopt the economic approach to index numbers.¹ This means that household and firms are assumed to be trying to maximise something. More specifically, I assume that households are trying to maximise utility and firms are trying to maximise a revenue function.² Of course households and firms might be trying to maximise something else or maybe they are not trying to maximise anything at all. But these possibilities will not be further explored here. Under these assumptions plus the further assumption that the utility function is homothetic (the revenue function exhibits constant returns to scale) I show that in principle the SNA is space-time consistent provided that Divisia (continuous) price indices are employed to measure PPPs and domestic prices. So if we observe space-time inconsistency it must be due to one or more of the following causes:

1. Non-homotheticity of the expenditure function (or non-constant returns to scale in the revenue (GDP) function).
2. Approximation errors due to replacing continuous Divisia price indices by discrete approximations, such as chained Törnqvist or chained Fisher indices.
3. Data errors in the PPPs or in the national accounts.

Based on these theoretical results, the paper seeks to assess the empirical size of these three causes of space-time inconsistency.

The plan of the paper is as follows. In Section 2 I set out the size of the inconsistency problem by looking first at OECD countries over an 18 year period and next at 53 countries which participated in both the 1980 and the 2005 rounds of the ICP, whose data are reported in the latest Penn World Table. Section 3 reviews the main properties of Divisia indices and

¹ See Balk (2009) for a review of different approaches to international comparisons.

² My approach is therefore in the same spirit as that of Neary (2004) and Feenstra, Ma and Rao (2009).

also discusses the relationship between Divisia and Konüs (true) price indices, the latter being the theoretically preferred ones. Section 4 analyses how Divisia index numbers can be applied to measure PPPs and proves a number of propositions about them. In particular, I show that under homotheticity the SNA is space-time consistent (Proposition 2). In practice PPPs are constructed as multilateral price indices rather than as chain indices (which are the discrete counterpart of Divisia indices). In the International Comparison Program (ICP) the Gini-Eltetö-Köves-Szulc (GEKS) procedure is widely used and is the preferred method; it is also the method employed for OECD-Eurostat PPPs. Section 5 accordingly compares chain index with GEKS index numbers. These theoretical results are then used in Section 6 to estimate how much of the measured inconsistency in the SNA between 1980 and 2005 is due to non-homotheticity (non-constant returns) and how much to the other causes listed above (approximation and data errors). These results employ estimated income elasticities derived from my earlier work on the 2005 ICP (Oulton 2012c). Section 7 concludes.

2. The size of the inconsistency problem

Consider two countries, A and B . We want to measure the economic gap between them, as measured say by GDP per capita. For each country, we have measurements of GDP per capita in volume terms, in each country's own currency, between two time periods, say from r to t . We also have GDP per capita in nominal terms. And we have purchasing power parities (PPPs) which give the value of B 's currency in terms of A 's at both time r and time t . Suppose we are interested in measuring GDP per capita at purchasing power parity in the later period t . Then there are two ways in which we could proceed. The first way is the direct one. Just take GDP per capita in nominal terms in period t in the two countries and convert B 's GDP to A 's currency using the PPP for period t . In symbols the direct index of real GDP per capita (Z) at time t , with A as numeraire, is:

$$\left[Z_t^{B/A} \right]_{direct} = \left(\frac{V_t^B}{V_t^A} \right) \left(\frac{1}{PPP_t^{B/A}} \right) \quad (1)$$

where $V_t^J = P_t^J Y_t^J$, $J = A, B$, is nominal GDP per capita at time t measured in local currency units. Here Y denotes real GDP per capita, P denotes the domestic price level (as measured typically by the GDP deflator), and PPP is the overall purchasing power parity, the PPP for

GDP, measured in units of B 's currency per unit of A 's currency, e.g. UK pounds sterling per US dollar.

The second, indirect way might be needed if a PPP is only available at time r but not at time t . Now we express B 's nominal GDP in period r in A 's currency using the PPP for that period. Then we roll forward the GDP per capita estimates using each country's real GDP and population growth rates. Finally we express GDP per capita in B relative to GDP per capita in A at time t :

$$\begin{aligned} [Z_t^{B/A}]_{indirect} &= \frac{(Y_t^B / Y_r^B) V_r^B}{(Y_t^A / Y_r^A) V_r^A} \left(\frac{1}{PPP_r^{B/A}} \right) \\ &= \left(\frac{V_t^B}{V_t^A} \right) \left(\frac{P_r^B / P_t^B}{P_r^A / P_t^A} \right) \left(\frac{1}{PPP_r^{B/A}} \right) \end{aligned} \quad (2)$$

The second line follows by applying the definition of nominal GDP per capita. Though both the direct and indirect indices purport to measure the same thing, namely the gap between the two countries at time t , there is no guarantee that they will be equal. At an empirical level there are potentially several reasons for this:

1. The PPPs are in practice derived as multilateral index numbers of expenditure on final demand in all countries included in the International Comparison Program or ICP. The weights in these index numbers are therefore not just those of the two countries involved in the present bilateral comparison but are a complicated average of the weights in all the countries. On the other hand in calculating real GDP per capita over time in any one country we use the weights of just that country and no other.
2. There are errors in the data affecting all components of the bilateral comparison (prices and quantities). It is often thought that errors in the PPPs are likely to be quite significant. Also the basket of products used to calculate PPPs is not the same as the baskets used to calculate each country's real GDP. And anyway the methodology of successive rounds of the ICP has changed.

This suggests that even if there were no errors in the data (the second reason) we would still expect the two measures to disagree (because of the first reason). But this is not a very happy conclusion. Surely, in the absence of data errors, they ought to agree? And if not, why not?

After eliminating the common elements in the direct and indirect measures, consistency across space and time can be seen to depend on the price indices and PPPs, with the criterion for consistency being:

$$PPP_t^{B/A} = PPP_r^{B/A} \left(\frac{P_t^B / P_r^B}{P_t^A / P_r^A} \right) \quad (3)$$

i.e. the more recent PPP must equal the older one after uprating the latter by inflation in the two countries. I shall refer to equation (3) as *the condition for space-time consistency in the system of national accounts (SNA)*.³ To help in interpreting this condition, note that if the right hand side (the projected PPP) is greater than the left hand side (the actual PPP) then according to the national accounts *B*'s prices relative to *A*'s have grown faster than suggested by two rounds of the ICP. So if prices have grown faster then real growth must have been slower according to the national accounts than according to the PPPs. Note too that any inconsistency is due entirely to the price indices. There may be errors in population estimates or in nominal GDP but these will not lead to inconsistency.

2.1 Inconsistency in practice: OECD data

How large is space-time inconsistency in practice? The OECD and Eurostat have been publishing PPPs for their own member countries and for a few selected outsiders for a number of years now.⁴ (The most fully available of the outsiders is the Russian Federation). For the EU the first available true comparison is for 1995 while for non-EU OECD members and for Russia it is 1999; earlier years are estimates. For the EU the latest year available at the time of writing is 2012; for the other countries it is 2011. This group of 35 countries has the advantage (by contrast with the much larger group available from the World Bank's ICP), that the methodology for constructing PPPs has not changed all that much over time and because the quality of their national statistics is mostly quite high. To assess space-time consistency we also need GDP deflators which can be estimated by dividing nominal GDP

³ The OECD uses the right hand side of (3) to estimate PPPs for years when no direct evidence exists; see the OECD's FAQs on PPPs (www.oecd.org/std/ppp/faq). The Penn World Table also uses this formula for interpolating PPPs for missing years between ICP benchmarks and for extrapolating PPPs before the earliest or beyond the latest ICP benchmark (currently 2005).

⁴ The OECD's PPPs are constructed initially as Fisher indices to which the EKS procedure is then applied. They therefore differ from the PPPs which appear in the latest Penn World Table (PWT8.0): see below.

(expenditure measure) by real GDP. (All these data were downloaded from the OECD website on 26 March 2014.) For this group of countries I have projected forward the 1995 (or 1999) PPPs to 2011, the latest available year for this group of countries though 2012 is also available for EU countries. I then calculate the inconsistency index which is the ratio of the projected PPP, the right hand side of equation (3), to the actual PPP, the left hand side. Table 1 shows the results which are for both 2011 and 2012 and two alternative base years, 1995 and 1999. Use of 1995 as a base or extending the projection to 2012 biases the results against finding inconsistency because then some of the PPPs are estimated from the very relationship that we are testing, equation (3). However, the results do not differ much between the different columns of Table 1.

In a perfectly consistent system the ratio of the projected to the actual PPP would always equal one but this is far from the case here. On average the projected PPP is about 10% higher than the actual one. So relative to the United States these countries are about 10% poorer using projected PPPs than they would appear using actual PPPs. Given that many of these countries are quite similar in living standards and that the time period is a maximum of 18 years, the average inconsistency ratio seems high. Also the average conceals a lot of individual variation. There are countries like Japan where the ratio is virtually 1 while for Norway and Russia the gap is huge: 60% and 56% respectively. For 28 of the 35 countries the inconsistency index exceeds one but for 7 countries it is less than one. Amongst the latter Israel stands out with a ratio in 2012 of 0.84.

A simple empirical model can account for some of the variation in the inconsistency index across these countries. It seems likely that if inflation is rapid or if an economy is growing fast (leading to rapid structural change) then the quality of the underlying statistics may suffer. After some experimentation I find that the following model fits reasonably well:

$$\ln(\text{Inconsistency index}) = 4.67*(\text{Inflation}) - 13.52*(\text{Inflation}^2) - 4.38*(\text{GDP growth})$$

(3.11) (2.27) (2.18)

+ residual

Absolute values of *t*-ratios in parentheses; *N* = 34 (US omitted); *R*² = 0.3458.

Here the inconsistency index is for 2012 using 1995 PPPs as the base. Inflation and GDP growth are measured by the log difference, 2012 over 1995, of the GDP deflator (defined as nominal divided by real GDP) and of real GDP respectively. Nominal and real GDP are taken from the OECD's national accounts data.⁵ A non-EU dummy was tried but was not significant and neither was the initial level of real GDP per head. Interestingly the largest outlier is Norway (see the added variable plot of Chart 1).

2.2 Inconsistency in practice: data from the Penn World Table

The recently published version 8.0 of the Penn World Table (PWT) contains PPPs and national accounts data for 167 countries.⁶ The PPPs in the PWT are based on successive rounds of the World Bank's International Comparison Program (ICP) which took place in 1970, 1975, 1980, 1985, 1996, 2005 and 2011 (the results of the 2011 round have not yet been fully published and so have not yet been incorporated into the PWT). Participation in the ICP has been patchy. For example, 146 countries participated in the 2005 round (the latest to be published at the time of writing). But if we count countries which participated in both the 1980 and the 2005 rounds that number falls to 53. And only 16 countries participated in both the 1970 and the 2005 rounds. So I concentrate here on measuring inconsistency over 1980-2005.

Before presenting the numbers it should be noted that the PPPs which appear in the PWT are not the same as the ones published by the OECD (even where the countries overlap) for several reasons:

1. The OECD presents just one overall PPP for each country. The PWT presents two overall PPPs for each country, which they refer to as output-side and expenditure-side. In the output-based measure, exports are deflated by an export price index and imports by an import price index; in the expenditure-side PPP both exports and imports are deflated by the PPP for final expenditure (consumption plus investment plus government expenditure). Feenstra *et al.* (2013) argue that the output-side measure is more appropriate for measuring output and the expenditure-side one for

⁵ http://stats.oecd.org/Index.aspx?datasetcode=SNA_TABLE1, accessed 26 March 2014.

⁶ The PWT is freely available at www.ggdcd.net/pwt. A full account of Version 8.0 of the PWT is in Feenstra *et al.* (2013).

measuring welfare. Here I use the output-side PPP as I want to compare this with results based on the GDP deflator from the national accounts.

2. The OECD's PPPs are Fisher indices. That is, bilateral Fisher indices are converted to multilateral indices using the EKS procedure (see below). The overall PPPs in the PWT are derived in a two-step procedure. First, PPPs for each of consumption, investment and government expenditure are derived as multilateral Fisher indices. Second, these three PPPs are aggregated using the Geary-Khamis procedure. That is, in the final result each PPP is the ratio of the value of output measured in domestic prices to the value of output measured in international prices; the international price of each product (Basic Heading) is a quantity-weighted average of the domestic price in every country.

It should also be noted that the number of products (Basic Headings) out of which overall PPPs are constructed has varied between successive rounds of the ICP. In 1980 there were 151 Basic Headings and in 2005 129.

53 countries participated in both the 1980 and the 2005 ICP and results for these countries appear in Table 2 and Chart 2. In 21 out of these 53 the PPP-based measure of GDP per head grew more slowly than the national accounts measure over 1980-2005; the average inconsistency in these countries was -0.73% p.a. in growth rate terms. In 32 countries the reverse was the case with the PPP-based measure growing more rapidly than the national accounts one; the average discrepancy was $+0.69\%$ p.a. Overall the mean difference between the two measures of the growth rate in these 53 countries was a modest 0.13% p.a. But the variation around the mean was quite large: a standard deviation of 1.14% p.a.

For some countries the inconsistency is startlingly large. For example in Argentina GDP per head was only 8% higher in 2005 than in 1980 according to the Argentinian national accounts. But according to the PPP-based measure it more than tripled over this 25 year period. At the other end of the scale GDP per head grew at 1.15% p.a. in Tanzania according to the national accounts but it *fell* at 1.12% p.a. using PPPs.

The degree of inconsistency seems large enough to cause concern. But what does theory have to say about this? Specifically, how great a degree of inconsistency is consistent with

economic theory? The next section will attempt to answer this question by a consideration of Divisia index numbers.

3. The inconsistency issue: insights from Divisia index numbers

3.1 Divisia index numbers: theory and derivation

The Divisia approach to index numbers (Divisia 1925-1926; Richter ; Hulten 1973; Balk 2005) sets up an ideal standard to which real life index numbers can never attain. It is nevertheless highly useful in clarifying conceptual issues. And Divisia index numbers can be approximated by chain indices so there is hope that their theoretical properties will carry over in practice to real world index numbers.

Consider an aggregate like nominal GDP per capita, $\sum_{i=1}^N p_i(t) y_i(t)$, where p_i is the price and y_i is the quantity per capita of the i th component of GDP ($i = 1, \dots, N$). We would like to separate this into an aggregate price P (“the price level”) and an aggregate quantity Y (“real GDP per capita”). So we write:

$$P^D(t, r) Y^D(t; r) := \frac{\sum_{i=1}^N p_i(t) y_i(t)}{\sum_{i=1}^N p_i(r) y_i(r)} \quad (4)$$

Here superscript D commemorates Divisia. The price and quantity indices are considered to be functions of time (t); the reference period r is also included as an argument of the function since the value of the index depends on it. The right hand side is the value index: nominal GDP per capita at time t relative to nominal GDP per capita in the reference period r (the value index equals 1 in the reference period, i.e. when $t = r$). The left hand side is an implicit definition of the Divisia price and quantity indices P^D and Y^D . So as this is a definition, we can take the total derivative of both sides with respect to time (treating all variables as continuous and differentiable) and the equality of the left and right hand sides will continue to hold:

$$\frac{d \ln P^D(t, r)}{dt} + \frac{d \ln Y^D(t, r)}{dt} = \sum_{i=1}^N w_i(t) \frac{d \ln p_i(t)}{dt} + \sum_{i=1}^N w_i(t) \frac{d \ln y_i(t)}{dt}$$

where $w_i(t) := \frac{p_i(t) y_i(t)}{\sum_{i=1}^N p_i(t) y_i(t)}$, the expenditure share of the i th component of GDP.

It is natural to identify the first term on the right hand side with the price index and the second with the quantity index:

$$\frac{d \ln P^D(t, r)}{dt} := \sum_{i=1}^N w_i(t) \frac{d \ln p_i(t)}{dt} \quad (5)$$

$$\frac{d \ln Y^D(t, r)}{dt} := \sum_{i=1}^N w_i(t) \frac{d \ln y_i(t)}{dt}$$

We can now recover the levels of the price and quantity indices at time t by integration:

$$\begin{aligned} \ln P^D(t, r) &= \sum_{i=1}^N \left[\int_r^t w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right] + \ln P^D(r, r) \\ &= \sum_{i=1}^N \left[\int_r^t w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right] \end{aligned} \quad (6)$$

and

$$\begin{aligned} \ln Y^D(t, r) &= \sum_{i=1}^N \left[\int_r^t w_i(\tau) \frac{d \ln y_i(\tau)}{d\tau} d\tau \right] + \ln Y^D(r, r) \\ &= \sum_{i=1}^N \left[\int_r^t w_i(\tau) \frac{d \ln y_i(\tau)}{d\tau} d\tau \right] \end{aligned} \quad (7)$$

Here I have normalised the price and quantity indices to be 1 in the reference period:

$$P^D(r, r) = Y^D(r, r) = 1 \quad (8)$$

This normalisation is consistent with the definition of equation (4) when $t = r$. This completes the characterisation of the Divisia price and quantity indices.

Divisia indices have many desirable properties. First, by definition the value index is the product of the price and quantity indices. Second, it is easy to see that the indices are consistent in aggregation.⁷ But they also suffer from two drawbacks. First, they are defined in continuous time so in practice they cannot be calculated exactly. However they can be approximated by chain indices. For example, a discrete approximation to the Divisia price

⁷ We can construct the Divisia price or quantity index from the ground up, using the most detailed components. Or we can proceed in two stages: first construct Divisia indices for sub-aggregates (e.g. consumption goods, investment goods, etc) and then construct a Divisia index of the sub-aggregate Divisia indices. Consistency in aggregation means that the two approaches yield the same answer at the aggregate level.

index is given by the chained Törnqvist price index (P_t^{ChT}) whose discrete growth rate is (compare equation (5)):

$$\Delta \ln P_t^{ChT} = \sum_{i=1}^N \left(\frac{w_{it} + w_{it-1}}{2} \right) \Delta \ln p_{it} \quad (9)$$

The second drawback is path dependence. This means that a Divisia index may fail the circularity test. Suppose that prices and quantities vary over some path between reference period r and end period t but in such a way that prices and quantities return to their original, period- r values in period t . Then we want the price and quantity indices to be unchanged in period t , i.e. we want $P^D(t; r) = P^D(r; r)$ and $Y^D(t; r) = Y^D(r; r)$. But this is not guaranteed to be the case. In general the value of a Divisia index at the terminal point of a path depends on the path as well as on the values of the prices and quantities at the initial and terminal points.

As is well known (Hulten 1973; Balk 2005), there is one condition which is necessary and sufficient for a Divisia index not to be path-dependent, i.e. for it to be *path-independent*. This is that the aggregate in question exhibit homotheticity (in the case of utility functions) or constant returns to scale in the case of revenue (GDP) functions. If we take prices to be exogenous to individual producers or consumers, this boils down to the requirement that the expenditure (output) shares must depend only on the prices and not on some other factor which may change over the path such as scale or income. Mathematically a Divisia index is a line integral. Path-independence requires that it be possible to find a “potential function” such that the total differential of the potential function depends only on prices. A potential function $\Phi(\ln \mathbf{p})$ must be such that

$$d\Phi = \sum_{i=1}^{i=N} w_i(t) d \ln p_i(t)$$

The point is that this must be a total differential, i.e. Φ must depend only on the prices and nothing else (or at least nothing which changes over any path under consideration).⁸ Then

$$\Phi(t) - \Phi(r) = \sum_{i=1}^{i=N} \int_{\ln p_i(r)}^{\ln p_i(t)} w_i(\tau) d \ln p_i(\tau)$$

The economic interpretation is that

$$\Phi(t) - \Phi(r) = \ln P^D(t, r)$$

⁸ Line integrals and the concept of path-independence are covered in undergraduate-level textbooks aimed at students of physics and engineering. Oddly, these topics are omitted even from quite advanced textbooks aimed at economics students.

3.2 The relationship between Divisia and Konüs index numbers

The Divisia price index is closely related to the Konüs price index. The Konüs price index (Konüs 1939) is the appropriate one in theory when we are trying to measure welfare. A similar concept is the right one when we want to measure aggregate output.⁹

To show the relationship between the Divisia and Konüs, let the consumer's expenditure function be

$$x = E(\mathbf{p}, u), \quad \partial x / \partial u > 0$$

This shows the minimum expenditure x needed to reach utility level u when $\mathbf{p} = (p_1 p_2 \dots p_N)$ is the $N \times 1$ price vector faced by the consumer; $x = \sum_i p_i y_i$ where the y_i are the quantities purchased. The Konüs price index is defined as the ratio of the cost of reaching the utility level of the base period (or base country) b at the prices of period (country) t to the cost of reaching the same utility level at the prices of the reference period (country) r :

$$P^K(t; r, b) = \frac{E(\mathbf{p}(t), u(b))}{E(\mathbf{p}(r), u(b))} \quad (10)$$

Applying Shephard's Lemma it is straightforward to show that the growth rate of the Konüs price index is given by

$$\frac{d \ln P^K(t, r, b)}{dt} = \sum_{i=1}^{i=N} w_i^K(t, b) \frac{d \ln p_i(t)}{dt} \quad (11)$$

(Balk 2005; Oulton 2008 and 2012c). Here the $w_i^K(t, b)$ are the *compensated* budget shares in period (country) t , the shares which would be observed if the consumer faced the prices of period (country) t while utility was held at the level of period (country) b :

$$w_i^K(t, b) = \frac{\partial \ln E(\mathbf{p}(t), u(b))}{\partial \ln p_i(t)}$$

The Konüs has the same form as the Divisia except that it uses *compensated* budget shares, not actual shares: compare equation (5). (The actual budget shares in period (country) t can be written as: $w_i(t) = w_i^K(t, t)$.) The Konüs price index is not unique but depends on the

⁹ Using an axiomatic approach, van Veelen (2002) has proved an impossibility theorem which purports to rule out an economically acceptable solution to the problem of measuring the standard of living, both internationally and intertemporally. However, his 4th and final axiom, "Independence of irrelevant countries" (or irrelevant time periods), would rule out the use of chain indices. On the economic approach the latter are essential to derive good approximations to Divisia or Konüs indices.

reference utility level. But for any given reference utility level, the Konüs price index is path-independent since by definition utility is being held constant. As is well known, the Konüs index is independent of the utility level and depends only on the prices if and only if demand is homothetic, i.e. if all income elasticities are equal to one (Konüs 1939; Samuelson and Swamy 1974; Deaton and Muellbauer, chapter 7, 1980b).

3.3 Compensated versus actual budget shares

Compensated budget shares are of course unobserved (except in the base period (country) when they equal actual shares). So the Konüs price index might seem of little practical use. But the relationship between compensated and actual budget shares can be made explicit by assuming a specific demand system such as the generalised PIGLOG or Quadratic Almost Ideal Demand System (QAIDS) of Banks et al. (1997) (see also Deaton and Muellbauer 1980a and 1980b, chapter 3):

$$\ln x = \ln A(\mathbf{p}) + \frac{B(\mathbf{p}) \ln u}{1 - \lambda(\mathbf{p}) \ln u}$$

Here $A(\mathbf{p})$ is any function homogeneous of degree 1 and $B(\mathbf{p})$ and $\lambda(\mathbf{p})$ are functions that are homogeneous of degree zero.¹⁰ Banks *et al.* (1997) suggested the following specification for $B(\mathbf{p})$ and $\lambda(\mathbf{p})$:

$$B(\mathbf{p}) = \prod_{k=1}^{k=N} p_k^{\beta_k}, \quad \sum_{k=1}^{k=N} \beta_k = 0 \quad (12)$$

and

$$\lambda(\mathbf{p}) = \sum_{k=1}^{k=N} \lambda_k \ln p_k, \quad \sum_{k=1}^{k=N} \lambda_k = 0 \quad (13)$$

Adopting this specification the relationship between compensated and actual budget shares is given by (see Oulton 2008 and 2012c):

$$w_i^K(t, b) = w_i(t) - \beta_i z(t, b) - \frac{\lambda_i}{B(\mathbf{p})} [z(t, b)]^2, \quad i = 1, \dots, N \quad (14)$$

Here $z(t, b)$ is the log of real income in period (country) t relative to real income in the base period (country) b :

$$z(t, b) = \ln \left[\frac{x(t) / x(b)}{P^K(t, b, b)} \right]$$

¹⁰ The function $A(\mathbf{p})$ is usually assumed to have the Almost Ideal form but this assumption is not necessary here.

So provided that the β_i and λ_i parameters can be estimated the Konüs price index can be calculated from observed budget shares and prices by an iterative process using equations (11) and (14). We will use this result below in estimating the proportion of inconsistency attributable to non-homotheticity (as opposed to other causes).

4. The Divisia approach applied to cross-country comparisons

So far we have interpreted the variable τ in the solutions for the Divisia price and quantity indices, equations (6) and (7), as time. But we could also interpret it as indexing countries. Just as time was assumed to be continuous so we can now think of a continuum of countries. This may seem rather artificial but it is common elsewhere in economics, e.g. a consumer, firm or product is often assumed to be an infinitesimal part of a continuum.¹¹

The PPPs for individual products in the ICP (some 120 of them in 2005) are measured as units of local currency per US dollar, i.e. they are like an exchange rate. Conceptually, they are the price of some product, say rice, in for example the UK, measured in UK pounds sterling, relative to the price of the same product in the US, measured in US dollars. They can also be thought of as the price in the UK in pounds sterling of a certain quantity of rice, namely the quantity which costs \$1 in the US in the comparison period. So we can apply Divisia index number theory and measure overall PPPs as

$$\ln P^D(t, r) = \sum_{i=1}^N \left[\int_r^t w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right] \quad (15)$$

(setting $\ln P^D(r, r) = 0$). Here τ indexes countries with the reference country r typically taken as the US.

¹¹ For Divisia index number purposes we require prices to vary smoothly as we move between countries. Prices in different countries are of course expressed in different currencies. This could lead to large jumps as we move between countries even if relative prices in the countries are similar. This difficulty can be overcome by normalising prices in some way. For example in each country each price could be divided by a weighted average of all prices where the weights represent (say) global shares in GDP. Alternatively each price (PPP) could be converted into US dollars using the official exchange rate. Both normalisations in effect change the monetary unit in which prices are measured. They therefore require equivalent adjustments to nominal (but not real) GDP.

A useful and desirable property of a cross-country price index is that it be transitive. Transitivity of the price index entails transitivity of cross-country comparisons of output and living standards. The World Bank's official PPPs are required to be transitive and this is enforced by applying the GEKS procedure (for most countries). So it is reassuring to note that Divisia indices satisfy transitivity (see section 6 for more on transitivity):

Proposition 1 The Divisia PPP (price) index is transitive:

$$P^D(t, r) = P^D(s, r)P^D(t, s).$$

for any three countries r , s and t .

Proof

$$\begin{aligned} \ln P^D(t, r) &= \sum_{i=1}^N \left[\int_r^t w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right] \\ &= \sum_{i=1}^N \left[\int_r^s w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right] + \sum_{i=1}^N \left[\int_s^t w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right] \\ &= \ln P^D(s, r) + \ln P^D(t, s) \end{aligned}$$

□

(See below, section 5, for more on transitivity).

We are now in a position to state the main theoretical result of this paper:

Proposition 2 If the Divisia price index is path-independent, then in the absence of data errors comparisons made using Divisia price indices are consistent across space and time. That is, the direct and indirect methods of equations (1) and (2) yield the same answer and the space-time consistency condition, equation (3), is satisfied.

Proof Take logs in the condition for space-time consistency, equation (3):

$$\ln PPP_t^{B/A} = \ln PPP_r^{B/A} + \ln(P_t^B / P_r^B) - \ln(P_t^A / P_r^A)$$

Translate this condition into Divisia price indices:

$$\ln P_t^D(B, A) = \ln P_r^D(B, A) + P_B^D(t, r) - P_A^D(t, r) \quad (16)$$

Here I have added a time subscript (r or t) to indicate the date to which a cross-country index applies and a country subscript (A or B) to indicate the country to which a time-series index

applies. We can write the Divisia price index (15) in non-parametric form by eliminating the variable indexing countries or time (τ):

$$\ln P^D(t, r) = \sum_{i=1}^N \left[\int_{\ln p_i(r)}^{\ln p_i(t)} w_i d \ln p_i \right]$$

We can then re-write this as a line integral using vector form and dot product notation:

$$\ln P^D(t, r) = \int_G \mathbf{w}(\mathbf{p}, \mathbf{X}) \cdot d \ln \mathbf{p}$$

where $\ln \mathbf{p}$ is the vector of log prices and $\mathbf{w}(\mathbf{p}, \mathbf{X})$ is the vector of expenditure shares; the latter are shown as functions of prices and possibly other variables (e.g. income or scale) represented by the vector \mathbf{X} . The integral is taken over a path G which commences with the share and price vectors \mathbf{w} and $\ln \mathbf{p}$ having the values of country or period r and finishes with these same vectors having the values of country or period t .

In line integral terms equation (16) can be written as:

$$\int_G \mathbf{w}(\mathbf{p}, \mathbf{X}) \cdot d \ln \mathbf{p} = \int_H \mathbf{w}(\mathbf{p}, \mathbf{X}) \cdot d \ln \mathbf{p} + \int_I \mathbf{w}(\mathbf{p}, \mathbf{X}) \cdot d \ln \mathbf{p} - \int_J \mathbf{w}(\mathbf{p}, \mathbf{X}) \cdot d \ln \mathbf{p} \quad (17)$$

Let $\ln \mathbf{p}(X, s)$ be the log price vector for country X at time s . Then the paths have the following endpoints:

$$\begin{aligned} G \text{ (cross-country, time } t) &: \ln \mathbf{p}(A, t), \ln \mathbf{p}(B, t) \\ H \text{ (cross-country, time } r) &: \ln \mathbf{p}(A, r), \ln \mathbf{p}(B, r) \\ I \text{ (time series, country } B) &: \ln \mathbf{p}(B, r), \ln \mathbf{p}(B, t) \\ J \text{ (time series, country } A) &: \ln \mathbf{p}(A, r), \ln \mathbf{p}(A, t) \end{aligned}$$

So taking into account the minus sign we see that the right hand side of (17) describes a path (JHI) whose endpoints are $\ln \mathbf{p}(A, t)$ and $\ln \mathbf{p}(B, t)$, the same as those of path G on the left hand side. Hence by path-independence the two sides of (17) are equal. In other words space-time consistency holds for Divisia price indices when the indices are path-independent.

Figure 1 illustrates for the two-good case □

Remark 2.1 In Figure 1 the overall path has sharp corners when it switches from time series to cross section variation. This doesn't affect the argument as the main theorems on line integrals continue go through. All that is required is that the overall path be continuous, but not necessarily differentiable, at every point.

Remark 2.2 Path-independence is a sufficient condition for space-time consistency. Path-independence is also necessary for space-time consistency if all possible paths are under

consideration. It is not necessary for a restricted subset of paths such that while quantities depend in principle on scale or income, the latter do not change on this particular subset of paths.

Corollary 2.1 By definition, utility (or real output) is held constant in a Konüs price index. So Konüs price indices are path-independent and so exhibit space-time consistency.

Proposition 2 shows that there are three reasons why the SNA might not exhibit space-time consistency in practice:

1. The Divisia price indices may not be path-independent
2. Even if they are path-independent they have to be approximated, e.g. by chain indices, and this leads to error (approximation error).
3. There are errors in the data.

4.1 A discrete analogue to Proposition 2

There is a discrete analogue to Proposition 2 but it is more restrictive in scope. Suppose that economic behaviour is described by a homothetic expenditure function:

$$E = c(\mathbf{p})f(u)$$

where $f(u)$ is a monotonically increasing function of utility u , and $c(\mathbf{p})$ is a homogeneous function of degree 1 in the price vector \mathbf{p} . $c(\mathbf{p})$ can be thought of as the unit cost function, the cost of buying each of the $f(u)$ units of utility. The expenditure function is assumed to be the same for all countries and time periods under consideration. Then there is a discrete analogue to Proposition 2 (the space-time consistency of Divisia indices):

Proposition 3 There is space-time consistency when the price index which is exact for the unit cost function $c(\mathbf{p})$ is the one employed to measure real incomes.

Proof Applying the definition of the Konüs price index to the homothetic case, the Konüs price index relating prices in situation 2 (\mathbf{p}^2) to prices in situation 1 (\mathbf{p}^1), for some pre-specified level of utility \bar{u} , is defined as:

$$P_{12}^K = \frac{c(\mathbf{p}^2)f(\bar{u})}{c(\mathbf{p}^1)f(\bar{u})} = \frac{c(\mathbf{p}^2)}{c(\mathbf{p}^1)} \quad (18)$$

In other words the price index is independent of the utility level, a consequence of homotheticity.

To distinguish between different time periods and countries, add two subscripts to the price vector. Thus $c(\mathbf{p}_{Js})$ is the value of the unit cost function for country J in period s . Then the PPP for country B relative to country A at time t is

$$\ln PPP_t^{B/A} = \ln c(\mathbf{p}_{Bt}) - \ln c(\mathbf{p}_{At})$$

This is the directly calculated PPP. The projected PPP, calculated from the right hand side of equation (3), is

$$\begin{aligned} & [\ln c(\mathbf{p}_{Br}) - \ln c(\mathbf{p}_{Ar})] + [\ln c(\mathbf{p}_{Bt}) - \ln c(\mathbf{p}_{Br})] - [\ln c(\mathbf{p}_{At}) - \ln c(\mathbf{p}_{Ar})] \\ &= \ln c(\mathbf{p}_{Bt}) - \ln c(\mathbf{p}_{At}) \\ &= \ln PPP_t^{B/A} \end{aligned}$$

So equation (3) is satisfied and the system is space-time consistent, at least if the unit cost function is known so that the Konüs price index can be calculated. \square

For this proposition to be of any use we need to find a cost function which can be calculated in practice. Suppose that the unit cost function is a quadratic mean of order r , as defined by Diewert (1976):

$$c(\mathbf{p}) = \left[\sum_{i=1}^N \sum_{j=1}^N b_{ij} p_i^{r/2} p_j^{r/2} \right]^{1/r}, \quad b_{ij} = b_{ji}, \quad \forall i \neq j, r > 0$$

Diewert (1976) showed that these functional forms are flexible in the sense that they provide a second order approximation locally to any function acceptable to economic theory. He further showed that corresponding to any such flexible functional form there exists a Konüs price index given by:

$$P_{12}^K = \left[\frac{\sum_{i=1}^N (p_i^2 / p_i^1)^{r/2} (p_i^1 y_i^1 / \mathbf{p}^1 \cdot \mathbf{y}^1)}{\sum_{k=1}^N (p_k^1 / p_k^2)^{r/2} (p_k^2 y_k^2 / \mathbf{p}^2 \cdot \mathbf{y}^2)} \right]^{1/r}$$

(Here \mathbf{y} is the quantity vector and superscripts 1 and 2 denote the two different price-quantity situations being compared). This means that the price index which is exact for the quadratic mean of order r will satisfy space-time consistency. For example if $r = 0$ the Törnqvist is the correct index or if $r = 2$ the Fisher is correct.

Note that Proposition 3, the discrete analogue to Proposition 2, is more restrictive in that the *same* function with the *same* parameters (the b_{ij} and r) is assumed to apply to all countries and all periods. Also the quadratic mean of order r is only a good approximation locally: over realistic distances it may not be so good (Hill 2006). So the assumption that the unit cost function is a quadratic mean of order r is much stronger than anything required for Proposition 2.¹²

5. Chain indices compared to the GEKS multilateral price index (PPP)

A chain index is the discrete counterpart to a Divisia index. In the national accounts chain indices are now widely employed. But in cross-country comparisons of prices typically different methods are used. For PPPs the multilateral GEKS index, based on either bilateral Fisher or bilateral Törnqvist Indices, has found wide acceptance. The purpose of this section is to compare chain indices with the GEKS.

Let p_{ij} be the price level in country j relative to the price level in country i , i.e. the bilateral PPP between i and j calculated using the expenditure shares of just these two countries. These bilateral price indices are assumed to be symmetric so that $p_{ij} = 1/p_{ji}$ and $p_{ii} = 1$; examples of price indices with these properties are the bilateral Fisher and the bilateral Törnqvist. Then with C ($C \geq 3$) countries the GEKS multilateral price index (PPP) for country j relative to country i is defined as

$$P_{ij}^{GEKS} := \left[\prod_{k=1}^{k=C} P_{ik} P_{kj} \right]^{1/C}$$

Taking logs,

$$\ln P_{ij}^{GEKS} = \frac{1}{C} \sum_{k=1}^C (\ln p_{ik} + \ln p_{kj}) = \frac{1}{C} \left[\sum_{k=1}^C \ln p_{ik} + \sum_{k=1}^C \ln p_{kj} \right]$$

Now define $\mathbf{A} = \ln \mathbf{p} \cdot \mathbf{1}$ and $\mathbf{B} = \ln \mathbf{p}' \cdot \mathbf{1}$. Here $\mathbf{1}$ is a $C \times 1$ column vector of ones and $\ln \mathbf{p}$ is a $C \times C$ matrix of the log of the bilateral price indices: $\ln \mathbf{p} = [\ln p_{ij}]$. So $\ln \mathbf{p} \cdot \mathbf{1}$ is a $C \times 1$ column

¹² It may be that there are other flexible functional forms with superior properties for longer distance comparisons and for which there is an exact price index. But I am not aware that any such have yet been discovered.

vector of the row sums of $\ln \mathbf{p}$, with the i th element being the i th row sum of $\ln \mathbf{p}$, etc, and $\ln \mathbf{p}' \cdot \mathbf{1}$ is a $C \times 1$ column vector of the column sums of $\ln \mathbf{p}$, with the j th element being the j th column sum of $\ln \mathbf{p}$. Then

$$\ln P_{ij}^{GEKS} = \frac{1}{C} [\mathbf{A}_i + \mathbf{B}_j]$$

Note that $\mathbf{A}_i + \mathbf{B}_i = 0$ for all i from the definitions of \mathbf{A} and \mathbf{B} (since $\ln p_{ij} = -\ln p_{ji}$). Then we have immediately that

$$\ln P_{ij}^{GEKS} = -\ln P_{ji}^{GEKS}$$

and

$$\ln P_{ii}^{GEKS} = \frac{1}{C} [\mathbf{A}_i + \mathbf{B}_i] = 0$$

Also the GEKS index numbers are transitive though the bilateral price indices are not. For any countries i, j and k :

$$\ln P_{ik}^{GEKS} + \ln P_{kj}^{GEKS} = \frac{1}{C} [\mathbf{A}_i + \mathbf{B}_k + \mathbf{A}_k + \mathbf{B}_j] = \frac{1}{C} [\mathbf{A}_i + \mathbf{B}_j] = \ln P_{ij}^{GEKS}$$

The matrix representation is useful for comparing the GEKS with chain indices. For $C > 3$ the GEKS makes no use of some links in the chain between any pair of countries. For example with $C = 4$ the matrix of bilateral price indices is

$$\mathbf{p} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & \mathbf{p}_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix} = \begin{bmatrix} 1 & p_{12} & p_{13} & p_{14} \\ 1/p_{12} & 1 & \mathbf{p}_{23} & p_{24} \\ 1/p_{13} & 1/p_{23} & 1 & p_{34} \\ 1/p_{14} & 1/p_{24} & 1/p_{34} & 1 \end{bmatrix}$$

Then the GEKS index for country 4 relative to country 1, P_{14}^{GEKS} , is calculated from the prices in the first row and fourth column of the matrix. The corresponding chain index is calculated from the prices on the diagonal above the principal diagonal (p_{12}, p_{23}, p_{34}). So each index makes use of some information which the other ignores. The chain index considers only adjacent price indices while the GEKS ignores some adjacent price indices. For example, in calculating P_{14}^{GEKS} no use is made of p_{23} (in bold above) while in calculating the chain index P_{14}^{Ch} no use is made of p_{13}, p_{14} or p_{24} . If the ordering of the countries is determined by some relevant measure of economic “closeness” then the chain index would seem to be superior, at least if we adopt the economic viewpoint on index numbers. If the gap between countries 1 and 4 is “large” then we are right to ignore price indices like p_{14} . The reason is that a superlative index number like the Fisher or the Törnqvist is only guaranteed to be a good

(second order) approximation locally; these indices are exact for some flexible functional form but the “flexibility” is a local property (Diewert 1976; Hill 2006). So the chain index may give a better approximation than the GEKS since the former allows the parameters of the flexible functional form to be different at each link in the chain. To put it more colourfully, in comparing the United States with Canada we probably don’t want to take a detour through the Democratic Republic of Congo (DRC). But in comparing the latter with the United States we may well want to go via intermediate countries like Nigeria, Tunisia and Brazil, indeed as many countries as possible given the enormous gap between the US and the DRC.

The chain index between countries i and $i+n$ is also transitive in the following sense:

$$\begin{aligned} P_{i,i+n}^{Ch} &= \prod_{j=0}^{j=n-1} p_{i+j,i+j+1} = \prod_{j=0}^{j=m-1} p_{i+j,i+j+1} \prod_{j=m}^{j=n-1} p_{i+j,i+j+1} \\ &= P_{i,i+m}^{Ch} P_{i+m,i+n}^{Ch} \end{aligned}$$

However it is not in general invariant to the ordering of the countries. Suppose there are four countries numbered 1-4 and we want to calculate a chain index between countries 1 and 4.

Let the initial ordering be 1,2,3,4. Then the chain index is

$$P_{14}^{Ch} = P_{12} P_{23} P_{34}$$

Now change the ordering to 1,3,2,4. The chain index becomes

$$\tilde{P}_{14}^{Ch} = P_{13} P_{32} P_{24}$$

Suppose that the chain index is Törnqvist. Then it is easy to check that in general $P_{14}^{ChT} \neq \tilde{P}_{14}^{ChT}$, at least not without further restrictions on behaviour. The GEKS index on the other hand is invariant to the ordering. In the 4-country case, changing the ordering to 1,3,2,4 is equivalent to interchanging the second and third rows and the second and third columns of the \mathbf{p} matrix above. This leaves row and column sums unchanged. To illustrate, let matrix $\tilde{\mathbf{p}}$ be the same as \mathbf{p} except that rows 2 and 3 and columns 2 and 3 are interchanged:

$$\tilde{\mathbf{p}} = \begin{bmatrix} p_{11} & p_{13} & p_{12} & p_{14} \\ p_{31} & p_{33} & p_{32} & p_{34} \\ p_{21} & p_{23} & p_{22} & p_{24} \\ p_{41} & p_{43} & p_{42} & p_{44} \end{bmatrix}$$

Then for example

$$\ln(\tilde{P}_{14}^{GEKS}) = \sum_{j=1}^4 \ln(p_{1j}) + \sum_{i=1}^4 \ln(p_{i4}) = \ln(P_{14}^{GEKS})$$

But the corresponding chain indices \tilde{P}_{14}^{Ch} and P_{14}^{Ch} are not in general equal since

$$\tilde{P}_{14}^{Ch} = P_{13}P_{32}P_{24} \neq P_{12}P_{23}P_{34} = P_{14}^{Ch} \text{ in general.}$$

If we are completely agnostic about the correct ordering of the countries, then the GEKS, which is invariant to ordering, is superior to a chain index. But if we can develop criteria for ordering the countries on the basis of an economically relevant notion of “closeness” then a chain index is to be preferred. The minimum spanning tree method of Hill (1999) and (2009) is a systematic way of doing this.

6. How much of measured inconsistency is due to path-dependence or chain index approximation?

6.1 Inconsistency due to path-dependence

As we have just seen a Konüs price index depends on the reference utility level chosen. If so, how are we to interpret real world price indices or PPPs? The answer is that a chained, superlative index is likely to be approximately equal to a true price index with reference utility level at the midpoint of the sample (Diewert, 1976 and 1981; Feenstra and Reinsdorf, 2000; Balk 2004).¹³ Consider two sets of PPPs from two rounds of the ICP. In the first round, the PPPs in effect take the viewpoint of the average income level across the countries

¹³ Suppose a utility function exists which rationalises the data but may be non-homothetic. Diewert (1981) showed that there exists a utility level which is intermediate between the levels at the endpoints of the interval under study such that a Konüs price index over this interval, with utility fixed at the intermediate level, is bounded below by the Paasche and above by the Laspeyres. Balk (2004) showed that when the growth of prices is piecewise log linear a chained Fisher price index approximates a Konüs price index over an interval when the reference utility level is fixed at that of some intermediate point in the interval. More precise results are available for specific functional forms. Diewert (1976) showed that a Törnqvist price index is exact for a non-homothetic translog cost function when the reference utility level is the geometric mean of the utility levels at the endpoints; see also Diewert (2009) for extensions. For the AIDS, Feenstra and Reinsdorf (2000) showed that, if prices are growing at constant rates, the Divisia index between two time periods equals the Konüs price index when the reference utility level is a weighted average of utility levels along the path.

included in the comparison at that point in time; in the next round the viewpoint is different, the average across the countries in the later period which is likely to be higher.

To assess the size of any inconsistency due to path-dependence we can calculate what the PPPs would have been in the later round if the reference income level was the same as in the earlier one. In Oulton (2012b and 2012c) I estimated Konüs price indices (true PPPs) for 141 countries (out of 146) in the 2005 round of the ICP. The estimated PPPs are for household consumption with the US equal to 1 and were calculated from budget shares and Basic Heading PPPs for 100 products within household consumption; this is the product level at which the World Bank's own estimates of aggregate PPPs are constructed (for details see Oulton 2012b and 2012c). This involved estimating the income response parameters in both a Linear and a Quadratic PIGLOG system for the 100 products (one or two parameters per product).¹⁴ Using these results I can adjust the 2005 PPPs so that they reflect any desired income level. In the sample of 53 countries which participated in both the 1980 and 2005 rounds of the ICP real GDP per head rose by between 60 and 70% between these two years, depending on the measure (national accounts or using the PWT's PPPs). (123 countries participated in the 2005 ICP and also have GDP and population data for 1980 in the PWT. The rise in average real income per head in this group was very similar: 69% on the national accounts-based measure.)

In this experiment, I assume no growth on average in real income between 1980 and 2005. Assuming an average increase of 70% between 1980 and 2005, I reduce each country's real income per head in 2005 by 41%. For each country, I calculate compensated budget shares for the 100 products within household consumption in 2005 with these new, lower income levels.¹⁵ Then I calculate PPPs for consumption for each of the 141 countries assuming real income is 41% lower. Both the new and the original PPP are calculated firstly as a chained

¹⁴ Five of the original 146 countries in the 2005 ICP were dropped since their data on household expenditure seemed suspect. My study employed the unpublished and confidential data on prices and quantities at the Basic Heading level which the World Bank kindly made available to me.

¹⁵ The compensated shares were calculated using a linear AIDS, i.e. all the λ_i in equation (14) are zero. Where necessary, any estimated compensated shares that were initially negative were set to zero and all estimated compensated shares were further constrained to sum to 1. I also tried using a quadratic AIDS but the initial estimates of the compensated shares were frequently large and negative, so these results are not shown. This difficulty is inherent in the PIGLOG framework since nothing constrains the estimated shares to be positive fractions.

Törnqvist index and secondly as a chained Fisher. To construct the chain indices, the countries were ordered in accordance with a conventional GEKS-Fisher index of real consumption. (The order is highly insensitive to the formula used for the aggregate PPP). Finally, I take the ratio of the new PPP to the original one, the latter calculated using actual budget shares. This ratio of the new and the original PPP gives a measure of the size of the inconsistency due to non-homotheticity in household demand. These results are in Table 3.

Unfortunately the two ratios, Törnqvist and Fisher, send a rather different message. Using the Törnqvist ratio the new PPP is on average 60% higher than the original one and the dispersion around the mean is high: the standard deviation is 0.51. By contrast the mean of the inconsistency ratio is 1.07 with standard deviation of 0.32 (Table 2). However the Fisher ratio has mean 1.11 with standard deviation of 0.16, comparable to the inconsistency index.

6.2 Inconsistency due to chain index approximation

To test for the potential size of the error due to chain index approximation, I have calculated chained Fisher and chained Törnqvist PPPs for the same 141 countries in the 2005 ICP, using actual budget shares. As before, the countries were ordered in accordance with a conventional GEKS-Fisher index of real consumption. I consider the ratio of the chained Fisher to the chained Törnqvist PPP as an indicator of chain index approximation error. In time series for individual countries this ratio is often close to one. Summary statistics of this ratio are displayed in Table 4.

It is clear that the gap between the two indices is quite wide. The Fisher yields a higher price level (lower standard of living) than does the Törnqvist. For all 141 countries the Fisher is 26% higher on average. But the gap is narrower for richer countries: the Fisher is only 6% higher on average for the 70 countries with above median consumption per head. For these countries the mean of the ratio is 1.06 with standard deviation 0.09. The difference between richer and poorer countries is quite puzzling. We might expect the PPPs and budget shares to be less accurate in poorer countries but it is not clear why this would make the Fisher-Törnqvist ratio higher and more variable in poorer countries.

6.3 Conclusion on empirical importance of path-dependence and chain index approximation

The two ratios we have considered, the ratio of compensated PPPs to actual PPPs and the ratio of chained Fisher to chained Törnqvist PPPs, both have the potential to explain much or even all of the observed inconsistency in the SNA. And both have similar distributions to that of the inconsistency index as measured by their means and standard deviations. However the correlation between either of these measures and the inconsistency index of Table 2 is low and insignificant. And multiple regression experiments did not produce results that were any stronger. I tried regressing the inconsistency index on the two ratios, with and without consumption per capita, but none of the coefficients were significant at conventional levels and the overall fit was low. I therefore conclude that neither path-dependence nor chain index approximation error can explain much of the observed inconsistency. This conclusion is tentative since we are not quite comparing like with like. Our potential explanatory factors are based on chain indices while the inconsistency index derives from the PWT where PPPs are measured using a combination of GEKS and Geary-Khamis.

7. Conclusions

We have found that path-dependence (non-homotheticity) and chain index approximation can probably account for only a small proportion of space-time inconsistency in the SNA. That leaves data errors as the main cause of inconsistency.

For the rich countries the most likely source of error is in the PPPs rather than the national accounts. Statistical agencies in these countries devote considerable resources to measuring consumer prices, producer prices and nowadays service industry prices as well. By comparison, price collection for the PPPs is underfunded. In poorer countries the opposite may be the case. Domestic price collection programmes may be poorly funded and use out-of-date weights. Hence in these countries the more reliable prices may be the ones gathered for the ICP under some measure of international supervision (though these do of course rely also on expenditure weights at the Basic Heading level derived from the national accounts).

The conclusion must then be that space-time inconsistency can be reduced to manageable proportions only by reducing data errors. Of course the ideal though utopian solution would be for much greater resources to be expended on price collection both domestically and internationally. But in the absence of this we don't have to give up. The way forward for the Penn World Table may be an analogue of methods sometimes used to reconcile the three measures of nominal GDP, namely a weighted average of different measures with the weights determined by reliability. Such an approach has been suggested recently by Rao *et al.* (2010) and by Hill and Melsner (2014).

TABLES

Table 1
Space-time inconsistency: ratio of projected to actual PPPs in 2011 and 2012
(34 OECD and EU countries plus Russia)

<i>Country</i>	<i>Code</i>	<i>based on 1995 PPPs</i>		<i>based on 1999 PPPs</i>	
		<i>2011</i>	<i>2012</i>	<i>2011</i>	<i>2012</i>
Australia	AU	1.0464	1.0469	1.0477	1.0481
Austria	OS	1.0123	1.0175	1.0410	1.0464
Belgium	BE	1.0503	1.0602	1.0871	1.0974
Canada	CA	1.0099	1.0099	1.0058	1.0058
Chile	CE	1.1852	1.1852	1.1852	1.1852
Czech Republic	CZ	0.9642	0.9756	0.9793	0.9909
Denmark	DE	1.1485	1.1627	1.1369	1.1510
Estonia	ES	1.1644	1.1565	1.1721	1.1641
Finland	FI	1.0267	1.0385	1.0294	1.0413
France	FR	1.1054	1.1078	1.0923	1.0947
Germany	GE	1.0708	1.0708	1.0819	1.0819
Greece	GR	1.0566	1.0650	1.0704	1.0789
Hungary	HU	1.2367	1.2418	1.1933	1.1982
Iceland	IC	0.8888	0.8858	0.8973	0.8943
Ireland	IR	1.0365	1.0404	1.0865	1.0906
Israel	IS	0.8350	0.8407	0.8350	0.8407
Italy	IT	1.0892	1.1088	1.0664	1.0856
Japan	JA	0.9944	0.9940	0.9944	0.9940
Korea	KO	0.9098	0.9098	0.9098	0.9098
Luxembourg	LU	1.1976	1.2167	1.1881	1.2071
Mexico	ME	1.2050	1.2005	1.2050	1.2005
Netherlands	NE	1.1246	1.1284	1.0951	1.0987
New Zealand	NZ	1.0293	1.0293	1.0342	1.0342
Norway	NO	1.5510	1.6039	1.4764	1.5268
Poland	PO	1.0737	1.0844	1.0644	1.0751
Portugal	PT	1.1327	1.1515	1.1303	1.1491
Russian Fed.	RU	1.5603	1.5603	1.5603	1.5603
Slovak Republic	SR	1.1451	1.1388	1.1245	1.1183
Slovenia	SL	1.0050	1.0272	0.9887	1.0106
Spain	SP	1.1508	1.1650	1.1304	1.1443
Sweden	SW	0.9875	0.9943	1.0000	1.0069
Switzerland	CH	1.1492	1.1713	1.1467	1.1687
Turkey	TU	1.1606	1.1647	1.1607	1.1648
United Kingdom	UK	0.9650	0.9765	0.9568	0.9682
United States	US	1.0000	1.0000	1.0000	1.0000

Table 1, continued

<i>Mean</i>	1.0934	1.1009	1.0907	1.0981
<i>s.d.</i>	0.1484	0.1528	0.1388	0.1424
<i>Minimum</i>	0.8350	0.8407	0.8350	0.8407
<i>Maximum</i>	1.5603	1.6039	1.5603	1.5603

Note Projected PPPs calculated from right hand side of equation (3), with either 1995 or 1999 as the base year.

Source OECD (<http://stats.oecd.org>, accessed 26 March 2014). For non-EU countries 1995 PPPs are estimates based on 1999 PPPs and 2102 PPPs are estimates based on 2011 PPPs.

Table 2
Space-time inconsistency in the SNA: 53 countries, 1980-2005
(ordered by degree of inconsistency, column (3))

<i>Country</i>	GDP per head in 2005, 1980=100			Growth of GDP per head, % p.a.		
	National Accounts	PWT	Ratio, column 2 ÷ column 1	National Accounts	PWT	Difference, column 5 <i>minus</i> column 4
	(1)	(2)	(3)	(4)	(5)	(6)
Tanzania	133.4	75.7	0.567	1.15	-1.12	-2.27
Nigeria	100.3	61.2	0.610	0.01	-1.96	-1.98
Kenya	100.3	70.3	0.701	0.01	-1.41	-1.42
Colombia	135.5	95.6	0.705	1.21	-0.18	-1.40
Indonesia	233.2	169.8	0.728	3.39	2.12	-1.27
Sri Lanka	233.6	180.0	0.770	3.39	2.35	-1.04
Pakistan	173.6	133.9	0.771	2.21	1.17	-1.04
Uruguay	126.1	102.5	0.813	0.93	0.10	-0.83
Tunisia	181.7	149.8	0.825	2.39	1.62	-0.77
Morocco	151.3	125.7	0.831	1.66	0.92	-0.74
Ethiopia	103.5	87.5	0.845	0.14	-0.53	-0.67
Chile	212.1	179.5	0.846	3.01	2.34	-0.67
Senegal	114.8	98.4	0.857	0.55	-0.07	-0.62
India	258.7	224.7	0.869	3.80	3.24	-0.56
Malawi	81.1	74.8	0.922	-0.84	-1.16	-0.32
Hong Kong	255.6	236.7	0.926	3.75	3.45	-0.31
Mexico	119.3	111.0	0.930	0.71	0.42	-0.29
Venezuela	83.7	79.9	0.955	-0.71	-0.90	-0.19
France	146.0	143.0	0.979	1.51	1.43	-0.08
United States	167.8	165.7	0.987	2.07	2.02	-0.05
Philippines	107.9	107.7	0.998	0.30	0.30	-0.01
Finland	172.0	172.8	1.005	2.17	2.19	0.02
United Kingdom	181.6	182.4	1.005	2.39	2.41	0.02
Israel	158.7	160.6	1.012	1.85	1.90	0.05
Mali	160.5	164.9	1.028	1.89	2.00	0.11
Cote d'Ivoire	70.0	72.8	1.040	-1.42	-1.27	0.16
Belgium	155.6	161.9	1.040	1.77	1.93	0.16
Ecuador	118.8	124.1	1.045	0.69	0.87	0.18
Denmark	159.7	167.3	1.048	1.87	2.06	0.19
Canada	151.5	159.6	1.053	1.66	1.87	0.21
Botswana	399.1	423.9	1.062	5.54	5.78	0.24
Italy	149.6	161.3	1.078	1.61	1.91	0.30
Japan	171.3	186.6	1.089	2.15	2.49	0.34
Luxembourg	251.8	280.4	1.114	3.69	4.12	0.43
Spain	178.5	200.3	1.122	2.32	2.78	0.46
Korea, Republic of	472.2	534.4	1.132	6.21	6.70	0.49
Zambia	75.9	86.0	1.133	-1.10	-0.60	0.50

Table 2, continued

Netherlands	157.1	179.6	1.143	1.81	2.34	0.53
Hungary	149.4	176.4	1.180	1.61	2.27	0.66
Madagascar	68.1	81.1	1.190	-1.54	-0.84	0.70
Portugal	178.6	212.6	1.190	2.32	3.02	0.70
Cameroon	84.2	100.5	1.194	-0.69	0.02	0.71
Norway	181.7	218.0	1.200	2.39	3.12	0.73
Austria	162.2	198.7	1.225	1.94	2.75	0.81
Germany	149.0	183.2	1.230	1.60	2.42	0.83
Ireland	283.5	364.7	1.286	4.17	5.18	1.01
Greece	141.9	182.7	1.288	1.40	2.41	1.01
Paraguay	97.4	125.7	1.291	-0.11	0.91	1.02
Brazil	112.8	153.6	1.362	0.48	1.72	1.24
Bolivia	100.2	151.2	1.510	0.01	1.65	1.65
Peru	101.8	157.3	1.546	0.07	1.81	1.74
Poland	152.8	250.5	1.640	1.70	3.67	1.98
Argentina	108.5	298.2	2.747	0.33	4.37	4.04
<i>Mean</i>	159.9	168.8	1.069	1.54	1.66	0.13
<i>s.d.</i>	75.0	87.8	0.323	1.62	1.84	1.04

Source Penn World Table 8.0 and own calculations (Penn World Table 2103). The Penn World Table measure of GDP per head uses output-side PPPs and is measured by the PWT variables *rgdpo* divided by population (*pop*). The national accounts measure is *rgdpna* divided by population (*pop*).

Table 3
Ratio of PPP for consumption using compensated shares to PPP using actual shares
(compensated shares derived from linear PIGLOG demand system assuming reduction of 41 %
in real consumption per head in 2005)

<i>Index number</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>
Chained Törnqvist	1.60	0.51	0.75	2.48
Chained Fisher	1.11	0.16	0.63	1.57

Source Unpublished data from the World Bank's 2005 ICP and own calculations.

Note 141 countries in 2005. PPPs estimated from 100 Basic Headings within household consumption; see Oulton (2012b) and (2012c) for details. For chain indices, countries ordered by real consumption per head using GEKS Fisher PPP for consumption to deflate household consumption in local currency units to international dollars.

Table 4
PPPs for household consumption: ratio of chained Fisher to chained Törnqvist

	<i>N</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>
All countries	141	1.26	0.23	0.93	1.70
Less than or equal to median	71	1.46	0.12	1.17	1.70
Above median	70	1.06	0.09	0.93	1.25

Source Unpublished data from the World Bank's 2005 ICP and own calculations.

Note PPPs estimated from 100 Basic Headings within household consumption; see Oulton (2012b) and (2012c) for details. 141 countries in total of which 71 have consumption per capita less than or equal to the median and 70 are above the median. For chain indices, countries ordered by real consumption per head using GEKS Fisher PPP for consumption to deflate household consumption in local currency units to international dollars.

CHARTS AND FIGURES

Chart 1

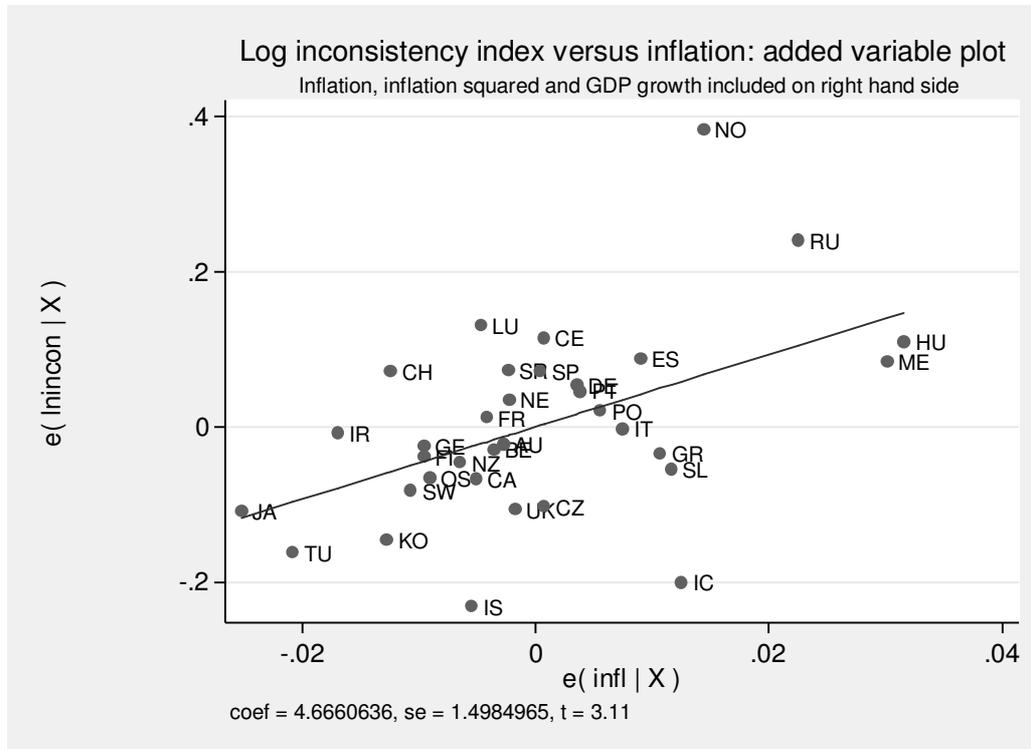


Chart 2

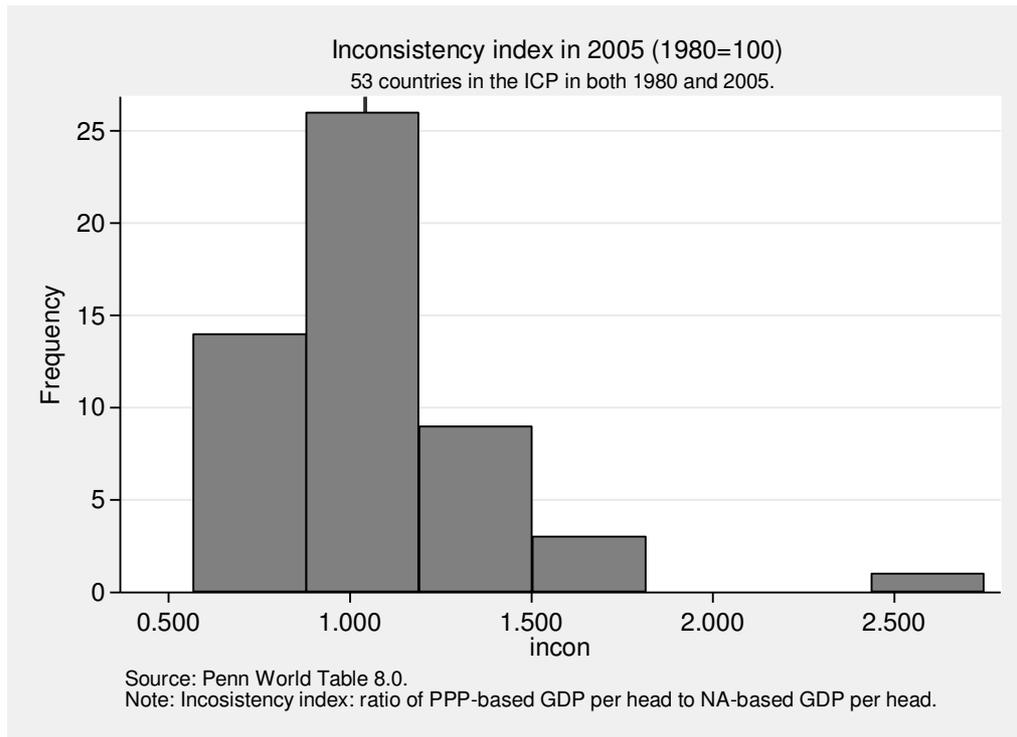
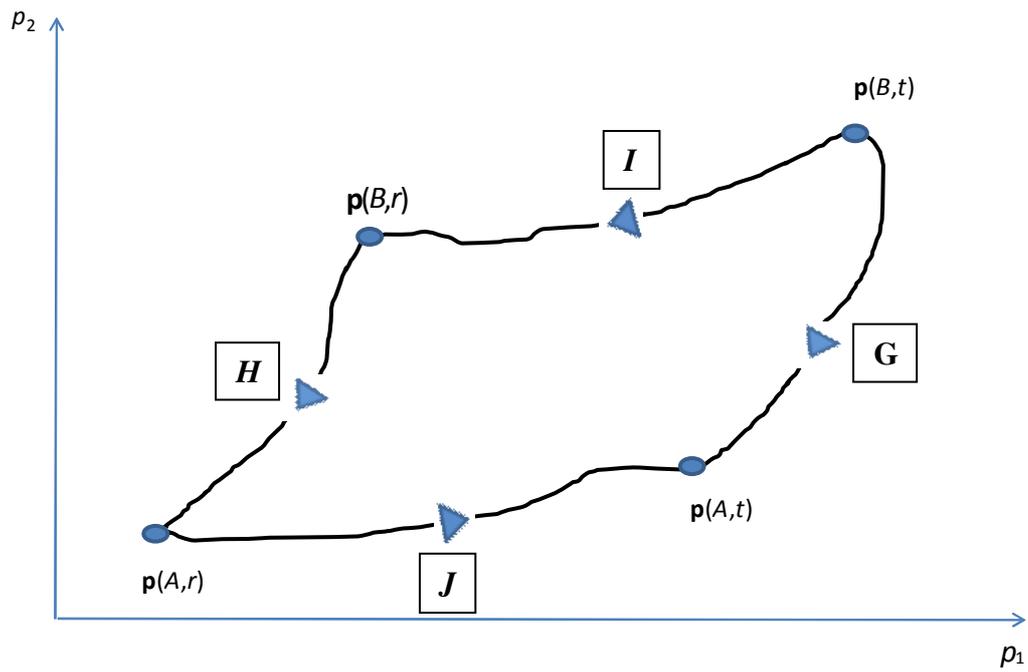


Figure 1 Price variation over time and across countries: the two-good case



References

- Balk, Bert M. (2005). "Divisia Price and Quantity Indices: 80 Years After." *Statistica Neerlandica*, 59: 119-158.
- Balk, Bert (2009). "Aggregation Methods in International Comparisons: An Evaluation". Chapter 3 of *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, edited by D.S. Prasada Rao. Cheltenham, UK: Edward Elgar.
- Banks, James, Richard Blundell and Arthur Lewbel (1997). "Quadratic Engel Curves and Consumer Demand." *The Review of Economics and Statistics*, 79(4): 527-539.
- Caves, Douglas W., Laurits R. Christensen, and W. Erwin Diewert (1982). "Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers." *Economic Journal*, 92: 73-86.
- Crawford, Ian, and J. Peter Neary (2008). "Testing for a Reference Consumer in International Comparisons of Living Standards". *American Economic Review*, 98(4): 1731-1732.
- Deaton, Angus (2010). "Price indexes, inequality, and the measurement of world poverty". *American Economic Review*, 100:1, 5-34.
- Deaton, Angus and Alan Heston (2010). "Understanding PPPs and PPP-based national accounts". *American Economic Journal: Macroeconomics 2010*, 2:4, 1-35.
- Deaton, Angus, and John Muellbauer (1980a). "An Almost Ideal Demand System." *American Economic Review*, 70: 312-336.
- Deaton, Angus, and John Muellbauer (1980b). *Economics and Consumer Behaviour*. Cambridge: Cambridge University Press.
- Deaton, Angus, and Bettina Aten (2014). "Trying to understand the PPPs in ICP2011: why are the results so different?". NBER Working Paper no. 20244.
- Diewert, W. Erwin (1976). "Exact and Superlative Index Numbers." *Journal of Econometrics*, 4: 115-46.
- Diewert, W. Erwin (1981). "The Economic Theory of Index Numbers: A Survey". In *Essays in Index Number Theory Volume 1*, edited by W. Erwin Diewert and Alice O. Nakamura, 1993. Amsterdam: North-Holland.
- Diewert, W. Erwin (2009). "Cost of Living Indexes and Exact Index Numbers". In *Quantifying Consumer Preferences*, edited by Daniel Slottje. Emerald Publishing Group.

- Divisia, François (1925-1926). "L'indice Monétaire et la Théorie de la Monnaie." *Revue d'Economie Politique*, 39(4): 842-864; 39(5): 980-1008; 39(6): 1121-1151 [1925]; 40(1): 49-81 [1926].
- Feenstra, Robert C. and Marshall B. Reinsdorf (2000). "An Exact Price Index for the Almost Ideal Demand System." *Economics Letters*, 66: 159-162.
- Feenstra, Robert C., Alan Heston, Marcel P. Timmer and Haiyan Deng (2009). "Estimating real production and expenditures across nations: a proposal for improving the Penn World Tables". *Review of Economics and Statistics*. Vol. 91, No. 1 (February), 201-212.
- Feenstra, Robert C., Hong Ma, and D.S. Prasada Rao (2009). "Consistent comparisons of real incomes across space and time". *Macroeconomic Dynamics*, 13(S2), 169-193.
- Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer (2013). "The next generation of the Penn World Table". Available at <http://www.rug.nl/research/ggdc/data/penn-world-table>.
- Hill, Robert J. (1999). "Comparing Price Levels Across Countries Using Minimum Spanning Trees". *Review of Economics and Statistics*, 81(1): 135-142.
- Hill, Robert J. (2004). "Constructing Price Indexes Across Space and Time: The Case of the European Union." *American Economic Review*, 94 (5): 1379-1410.
- Hill, Robert J. (2006). "Superlative Index Numbers: Not All of Them Are Super". *Journal of Econometrics*, 130: 25-43.
- Hill, Robert J. (2009). "Comparing per capita income levels across countries using spanning trees: robustness, prior restrictions, hybrids and hierarchies". Chapter 9 in Rao (2009).
- Hill, Robert J. and Daniel Melser (2014). "Benchmark averaging and the measurement of changes in international income inequality". Mimeo.
- Hulten, Charles R. (1973). "Divisia Index Numbers." *Econometrica*, 41: 1017-1025.
- Johnson, Simon, William Larson, Chris Papageorgiou and Arvind Subramanian (2013). "Is newer better? Penn World Table revisions and their impact on growth estimates," *Journal of Monetary Economics* 60(2): 255-274.
- Konüs, Alexander A. (1939). "The Problem of the True Index of the Cost-of-Living." *Econometrica*, 7(1): 10-29. (English translation; first published in Russian in 1924).
- Neary, J. Peter (2004). "Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income." *American Economic Review*, 94(5): 1411-1428.

- Oulton, Nicholas (2008). "Chain Indices of the Cost of Living and the Path-Dependence Problem: An Empirical Solution." *Journal of Econometrics*, 144: 1 (May): 306-324. [<http://dx.doi.org/10.1016/j.jeconom.2008.02.001>].
- Oulton, Nicholas (2012a). "How to Measure Living Standards and Productivity". Centre for Economic Performance. Discussion Paper no. 949. (<http://cep.lse.ac.uk/pubs/download/dp0949.pdf>).
- Oulton, Nicholas (2012b). "The Wealth and Poverty of Nations: True PPPs for 141 countries". Centre for Economic Performance. Discussion Paper no. 1080. (<http://cep.lse.ac.uk/pubs/download/dp1080.pdf>).
- Oulton, Nicholas (2012c). "How to measure living standards and productivity". *Review of Income and Wealth*, vol. 58, issue 3 (September), 424-456, 2012.
- Penn World Table (2013). "Technical guide to PWT8.0". Available at <http://www.rug.nl/research/ggdc/data/penn-world-table>.
- Rao, D.S. Prasada (editor) (2009). *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*. Cheltenham, UK: Edward Elgar.
- Rao, D.S. Prasada, A. Rambaldi and H. Doran (2010). "Extrapolation of Purchasing Power Parities using multiple benchmarks and auxiliary information: a new approach". *Review of Income and Wealth*, 56, Special Issue 1, S59-S98.
- Samuelson, Paul A. and Subramanian Swamy (1974). "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis." *American Economic Review*, 64: 566-593.
- van Veelen, Matthijs (2002). "An Impossibility Theorem Concerning Multilateral International Comparison of Volumes". *Econometrica* 70(1), 369-75.
- World Bank (2008). *Global Purchasing Power Parities and Real Expenditures: 2005 International Comparison Program*. Available at <http://siteresources.worldbank.org/ICPINT/Resources/icp-final.pdf>.