Space-Time (In)Consistency
in the National Accounts:
Causes and Cures

Nicholas Oulton

Centre For Macroeconomics,

London School of Economics

October 2015
(revision of April 2015 version)

Address
Centre For Macroeconomics
London School of Economics
Houghton Street
LONDON
WC2A 2AE
United Kingdom

Email
n.oulton@lse.ac.uk

Abstract

Space-time consistency in the national accounts means that the relative size or standard of living of a country today is the same whether we measure it by an earlier PPP extrapolated to the present using relative inflation rates from the national accounts or by the current PPP. Empirically, space-time *inconsistency* is extensive. Theoretically, space-time *consistency* prevails if the consumer's utility function (or the revenue (GDP) function) is homothetic and if Divisia price indices are used to deflate nominal consumption (or GDP), both over time and across countries. Hence the inconsistency we observe is due to either (a) non-homotheticity in consumption (or production); (b) approximation error when discrete chain indices are used instead of continuous Divisia indices; or (c) errors in domestic price indices and PPPs. Based on detailed data from the 1980 and 2005 International Comparisons Program and the Penn World Table, I conclude that errors in price indices and PPPs are the major cause of inconsistency.

JEL codes: C43; E01; O47; I31; E23
Key words: PPP; Divisia; Konüs; price index; path-dependence; consistency

1. Introduction[1]

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 national accounts.

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*., 2015a). 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 version 8. 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) and (2015a) 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.[2] 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.[3] 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 (or GDP function) is homothetic, I show that in principle the national accounts are 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 of 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 the first two possible causes of space-time inconsistency. To address the non-homotheticity issue, I estimate PPPs for household consumption in countries which participated in both the 1980 and 2005 rounds of the World Bank's ICP; these PPPs are adjusted for the growth in real income between 1980 and 2005. I then compare a measure of inconsistency over 1980-2005 using these adjusted PPPs with the same measure calculated using actual PPPs. To address

---

[2] See Balk (2009) for a review of different approaches to international comparisons.
[3] My approach is therefore in the same spirit as that of Neary (2004) and Feenstra, Ma and Rao (2009).

approximation error, I compare inconsistency calculated using two widely-employed index number formulas: chained Fisher and chained Törnqvist. The likely importance of data errors can then be assessed as the residual.

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 52 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 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 national accounts are space-time consistent (Proposition 2). These theoretical results are then used in Section 5 to estimate how much of the measured inconsistency in the national accounts between 1980 and 2005 is due to non-homotheticity (non-constant returns) and how much to chain index approximation error; the importance of data errors is then assessed as the residual. Section 6 concludes.

2. The size of the inconsistency problem

2.1 Inconsistency in theory

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:

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

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.

In the second, indirect way (which might be needed if a PPP is only available at time r but not at time t) 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:

$$\left[Z_t\left(A,B\right) \right]_{indirecct} = \frac{\left(Y_t^B / Y_r^B \right) V_r^B}{\left(Y_t^A / Y_r^A \right) V_r^A} \left(\frac{1}{PPP_r\left(A,B\right)} \right)$$

$$= \left(\frac{V_t^B}{V_t^A} \right) \left(\frac{P_r^B / P_t^B}{P_r^A / P_t^A} \right) \left(\frac{1}{PPP_r\left(A,B\right)} \right)$$

(2)

The second line follows by applying the definition of nominal GDP per capita.

Consistency across space and time requires that the direct and indirect measures should yield the same answer. After equating the right hand sides of equations (1) and (2), the criterion for consistency is seen to be:

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

(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 national accounts*.[4] 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

---

[4] 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.

and these will lead to errors in measuring living standards. But such errors will not lead to inconsistency in measurement.

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 in practice, i.e. that condition (3) will be satisfied. 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? The aim of this paper is to clarify the reasons why inconsistency might occur even in the absence of data errors. But first I consider how large the space-time inconsistency problem is in practice. I look first at OECD and Eurostat data and next at the Penn World Table.

2.2 Inconsistency in practice: OECD data

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.[5] (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

---

[5] The OECD's PPPs for GDP are constructed initially as bilateral Fisher indices to which the EKS procedure is then applied. They therefore differ from the PPPs which appear in the latest Penn World Table (version 8): see below.

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 (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:

$$I = \frac{PPP_r(A,B)\left(\dfrac{P_t^B / P_r^B}{P_t^A / P_r^A}\right)}{PPP_t(A,B)} \tag{4}$$

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 percent higher than the actual one. So relative to the United States these countries are about 10 percent 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 percent and 56 percent 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.

2.2 Inconsistency in practice: data from the Penn World Table

Version 8 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 (at the time of writing the results of the 2011 round have not yet been incorporated into the PWT). Participation in the ICP has been patchy. For example, 146 countries participated in the 2005 round. 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*. (2015a) argue that the output-side measure is more appropriate for measuring output and the expenditure-side one for 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 multilateral Fisher indices. That is, bilateral Fisher indices are converted to multilateral indices using the EKS (or GEKS) procedure: see the Appendix. 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

---

[6] Version 8 comes in two flavours, the earlier 8.0 and the later 8.1; the latter became available after much of the empirical work in this paper was completed but is also used below. For present purposes the main difference between versions 8.1 and 8.0 lies in the PPPs. In version 8.1 the 2005 PPPs at the Basic Heading level have been adjusted to bring them more into line with the methodology of the 2011 ICP. This mainly affects countries in the Asia-Pacific region and in Africa, lowering their price level and raising their standard of living in 2005: see Feenstra *et al*. (2015b). The PWT is freely available at www.ggdc.net/pwt. A full account of Version 8.0 of the PWT is in Feenstra *et al*. (2013) and of Version 8.1 is in Feenstra *et al*. (2015a).

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 1. 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 percent p.a. in annual 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 $+0.69$ percent p.a. Overall the mean difference between the two measures of the growth rate in these 53 countries was a modest 0.13 percent p.a. But the variation around the mean was quite large: a standard deviation of 1.14 percent p.a.[7]

For some countries the inconsistency is startlingly large. For example in Argentina GDP per head was only 8 percent 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 percent p.a. in Tanzania according to the national accounts but it *fell* at 1.12 percent 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.

---

[7] In Chart 1 the PPP-based measure of GDP is the PWT variable *rgdpo* and the national-accounts-based measure is the PWT variable *rgdpna* (both measured in millions of 2005 US dollars). So the inconsistency index is:
$$[rgdpo(2005)/rgdpo(1980)]/[rgdpna(2005)/rgdpna(1980)]$$
i.e. both measures of GDP per head are expressed as relative to their respective 1980 levels (note that population cancels out). By construction in PWT 8.0 *rgdpna* = *rgdpo* in 2005. So the inconsistency index reduces to $[rgdpna(1980)/rgdpo(1980)]$. The inconsistency index therefore measures the ratio of total growth between 1980 and 2005 on the PPP-based measure to total growth on the national-accounts-based measure.

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 1966; Hulten 1973; Balk 2005) sets up an ideal standard to which real life index numbers can never fully 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 or consumption 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 the aggregate ($i = 1,...,N$). Let $P^D(r,t)$ be the Divisia price index at time t for this set of prices (here superscript D commemorates Divisia) and let the index take the value 1 in the reference period r; see the Appendix for a fuller characterisation of Divisia indices. The growth rate of the Divisia price index is defined as:

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

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

The level of the index at time t is found by integration:

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

Divisia indices have two desirable properties. First, by definition the value index is the product of the price and quantity indices. Second, the indices are consistent in aggregation. 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 index is given by the chained Törnqvist price index (P_t^{ChT}) whose discrete growth rate is (compare equation (5)):

$$\Delta \ln P_t^{ChT} = \sum_{i=1}^{N} \left(\frac{w_{it} + w_{it-1}}{2} \right) \Delta \ln p_{it} \qquad (7)$$

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 $P^D(r,t) = P^D(r,r)$. 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 (the utility function or the revenue (GDP) function[8]) be homothetic. 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. 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).[9]

Above I called the possibility of path-dependence a "drawback" of Divisia index numbers. In reality, and using the language of software, it should be thought of as a feature, not a bug, since it alerts us to the possibility of non-homotheticity in the utility or GDP function.

---

[8] The revenue or GDP function is the analogue in production theory of the expenditure function in consumer theory: see Woodland (1982) (though his term for it is the GNP function). Under competitive assumptions producers who take prices as given behave as if they are maximising the value of output (GDP) subject to the various production functions and the stocks of resources.

[9] Mathematically a Divisia index is a line integral. Line integrals and the concept of path-independence are covered in undergraduate-level textbooks aimed at students of physics and engineering, e.g. Salas, Hill and Etgen (2007). Oddly, these topics are omitted even from quite advanced textbooks aimed at economics students.

3.2 The Divisia approach applied to cross-country comparisons

So far we have interpreted the variable τ in the solutions for the Divisia price index, equation (6), 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. It may be objected that there are only about 200 countries in the world and moreover they differ radically in numerous other ways quite apart from prices and expenditure shares. In response three points can be made:

1. The assumption that prices and expenditure shares vary continuously over time is clearly an abstraction. Even if physical time is continuous, economic decisions are not made (and are certainly not observed) continuously. So the assumption of continuity is made for convenience in the hope that it is approximately true.

2. The assumption of a continuum of countries is really saying something about prices and expenditure shares, namely that they can vary continuously. So we can conceive of an (imaginary) country, identical in every way to an observed one except that its prices and expenditure shares differ by epsilon from those of the observed one. Of course, the assumption that prices are continuous is certainly not literally true since there is a minimum size of the monetary unit, e.g. a cent or a penny. But this is not thought of as a great difficulty in micro theory generally so should not worry us unduly in the present context.

3. It is true that countries differ radically on many factors such as climate, demography, culture and religion that may be relevant for economic behaviour. These factors change only slowly if at all within a given country so for time series comparisons they can usually be ignored. For cross-country comparisons they must be taken into account. But this is a different issue from whether it is reasonable to assume a continuum of countries. The way to deal with it is to allow explicitly for such factors in modelling economic behaviour: more on this below in Section 5.

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

$$\ln P^D(r,t) = \sum_{i=1}^{N} \left[\int_r^t w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right] \qquad (8)$$

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

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 (for most countries) by applying the GEKS procedure. So it is reassuring to note that Divisia indices also satisfy transitivity:

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

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

for any three countries r, s and t.

Proof

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

$$= \sum_{i=1}^{N} \left[\int_r^s w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right] + \sum_{i=1}^{N} \left[\int_s^t w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right]$$

$$= \ln P^D(r,s) + \ln P^D(s,t)$$

$\square$

The relationship between chain indices which approximate Divisia indices and the GEKS is further explored in the Appendix.

We can now 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.[10]

Proof   We can write the Divisia price index (8) in non-parametric form by eliminating the variable indexing countries or time (τ):

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

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

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

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 climate) 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.

Now take logs in the condition for space-time consistency, equation (3):

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

Translate this condition into Divisia price indices:

$$\ln P_t^D(A,B) = \ln P_r^D(A,B) + \ln P_B^D(r,t) - \ln P_A^D(r,t) \tag{9}$$

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. In line integral terms equation (9) can be written as:

$$\int_G \mathbf{w}(\mathbf{p},\mathbf{X}) \cdot d \ln \mathbf{p} = \int_H \mathbf{w}(\mathbf{p},\mathbf{X}) \cdot d \ln \mathbf{p} + \int_I \mathbf{w}(\mathbf{p},\mathbf{X}) \cdot d \ln \mathbf{p} - \int_J \mathbf{w}(\mathbf{p},\mathbf{X}) \cdot d \ln \mathbf{p} \tag{10}$$

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

---

[10]   Proposition 2 was stated and proved in Oulton (2015).

Path	Type	Endpoints
G	Cross-country, time t	$\ln \mathbf{p}(A,t), \ln \mathbf{p}(B,t)$
H	Cross-country, time r	$\ln \mathbf{p}(A,r), \ln \mathbf{p}(B,r)$
I	Time series, country B	$\ln \mathbf{p}(B,r), \ln \mathbf{p}(B,t)$
J	Time series, country A	$\ln \mathbf{p}(A,r), \ln \mathbf{p}(A,t)$

So taking into account the minus sign we see that the right hand side of (10) 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 (10) 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 (e.g.) income or climate, the latter do not change on this particular subset of paths.

Remark 2.3   The condition for path-independence is that expenditure shares should depend only on prices and not on any other factors, at least over the paths under consideration. Path-dependence occurs if the utility function is not homothetic and if real income varies over the paths under consideration. Path-dependence also occurs if expenditure shares depend on other factors as well such as climate, demography and culture and if these vary across countries (or over time though this is less likely to be important empirically).

   Proposition 2 shows that there are three reasons why the national accounts 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 underlying data for prices and quantities.[11]

3.3 A discrete analogue to Proposition 2

There is a discrete analogue to Proposition 2 but it is more restrictive in scope. At this point we need to bring in the Konüs price index as the theoretically preferable way to measure changes in the cost of living (Konüs 1939). A similar concept, derived from the GDP function, is the right one when we want to measure aggregate output.[12]

Let the consumer's expenditure function be

$$x = E(\mathbf{p}, \mathbf{h}, u), \quad \partial x / \partial u > 0 \tag{11}$$

This shows the minimum expenditure x needed to reach utility level u when $\mathbf{p} = (p_1\ p_2 \dots p_N)$ is the Nx1 price vector faced by the consumer ($x = \sum_i p_i y_i$ where the y_i are the quantities purchased) and where $\mathbf{h}$ is a vector of background variables such as climate, demography and culture which influence utility. 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, in both cases with the background factors fixed at the same level $\mathbf{h}(b)$:

$$P^K(t, r, b) = \frac{E(\mathbf{p}(t), \mathbf{h}(b), u(b))}{E(\mathbf{p}(r), \mathbf{h}(b), u(b))} \tag{12}$$

The Konüs price index is in general not unique but depends on the base utility level and background factors (Konüs 1939; Samuelson and Swamy 1974; Deaton and Muellbauer, chapter 7, 1980b).

---

[11]   Feenstra *et al*. (2015a, Appendix D) argue that updating PPPs by relative inflation rates is likely to understate true PPPs (or to overstate true PPPs when a current PPP is extrapolated backwards) due to the Balassa-Samuelson effect: the tendency for the relative prices of non-traded goods to rise as a country gets richer. However that may be, Proposition 2 shows that the root of the matter is not Balassa-Samuelson but non-homotheticity. Our empirical results above for GDP and below for consumption show that inconsistency is two-sided: it can lead just as often to an updated PPP exceeding a later one as to the opposite effect.

[12]   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.

Now consider the more restrictive case where consumption behaviour is described by a homothetic expenditure function:

$$x = c(\mathbf{p}) f(u, \mathbf{h})$$

where $f(u, \mathbf{h})$ 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 $f(u, \mathbf{h})$ units of utility. This 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, equation (12), 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 u_0 and of the background factors $\mathbf{h}_0$, is defined as:

$$P^K(1,2) = \frac{c(\mathbf{p}^2) f(u_0, \mathbf{h}_0)}{c(\mathbf{p}^1) f(u_0, \mathbf{h}_0)} = \frac{c(\mathbf{p}^2)}{c(\mathbf{p}^1)} \tag{13}$$

In other words the price index is independent of the utility level and of background factors, 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

$$\ln PPP_t(A, B) = \ln c(\mathbf{p}_{Bt}) - \ln c(\mathbf{p}_{At}) \tag{14}$$

Recall that we are assuming that economic behaviour is completely captured by the unit cost function $c(\mathbf{p})$ and that there exists a price index which is exact for this unit cost function. The right hand side of equation (14) is then the directly calculated PPP. The projected PPP, calculated from the right hand side of equation (3), is (in logs)

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

where the last line follows from (14). So equation (3) is satisfied and the system is space-time consistent. ☐

For this proposition to be of any use we need to find a cost function which can be calculated in practice, i.e. we need a price index which is exact for the unit cost function which describes behaviour. Such price indices certainly exist for some unit cost functions. Suppose that the unit cost function is a quadratic mean of order r, as defined by Diewert (1976):

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

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 we take the limit as $r \to 0$ 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* parameter values (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 assuming that the unit cost function is a quadratic mean of order r is much stronger than anything required for Proposition 2.[13]

---

[13]   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.

4. The relationship between Divisia and Konüs index numbers

4.1 Konüs index numbers

The Divisia price index is closely related to the true cost-of-living or Konüs price index. To show the relationship between the two, totally differentiate the definition of the Konüs price index, equation (12), with respect to time and apply Shephard's Lemma to get:

$$\frac{d \ln P^K(t,r,b)}{dt} = \sum_{i=1}^{i=N} w_i(t,b) \frac{d \ln p_i(t)}{dt} \tag{15}$$

(Balk 2005; Oulton 2008 and 2012c). Here the $w_i(t,b)$ 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 and background factors also at levels found in period (country) b, $\mathbf{h}(b)$:

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

The Konüs has the same form as the Divisia except that it uses *compensated* budget shares, $w_i(t,b)$, not actual shares: compare equation (5). The actual, observed budget shares in period (country) t can be written as $w_i(t,t)$.

By definition, utility and background factors are held constant in a Konüs price index. So Konüs price indices are path-independent and therefore exhibit space-time consistency since only prices influence compensated budget shares.[14] Hence, adapting equation (9), the condition for space-time consistency is satisfied for Konüs price indices:

$$\ln P_t^K(A,B) = \ln P_r^K(A,B) + \ln P_B^K(r,t) - \ln P_A^K(r,t) \tag{16}$$

where both utility and background factors are held constant at some base level (the same for all countries and time periods).

4.2 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

---

[14]   Oulton (2012b) and (2012c) finds that background factors explain even more of the cross-country variation in budget shares in 2005 than do income and prices.

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):

$$\ln x = \ln A(\mathbf{p}) + \frac{B(\mathbf{p})\ln u}{1 - \lambda(\mathbf{p})\ln u} + D(\mathbf{p})\ln h \qquad (17)$$

The last term, $D(\mathbf{p})\ln h$, is not in Banks *et al.* (1997), but is included to allow for background factors, in a parsimonious way. For simplicity of exposition I allow for just a single factor, representing say climate ("heat"), but it is straightforward to extend this to any number of factors. In equation (17) $A(\mathbf{p})$ is a function homogeneous of degree 1 and $B(\mathbf{p})$, $D(\mathbf{p})$ and $\lambda(\mathbf{p})$ are functions homogeneous of degree zero.[15] Banks *et al.* (1997) suggested the following specification for $B(\mathbf{p})$ and $\lambda(\mathbf{p})$:

$$B(\mathbf{p}) = \prod_{k=1}^{k=N} p_k^{\beta_k}, \quad \sum_{k=1}^{k=N} \beta_k = 0 \qquad (18)$$

$$\lambda(\mathbf{p}) = \sum_{k=1}^{k=N} \lambda_k \ln p_k, \quad \sum_{k=1}^{k=N} \lambda_k = 0 \qquad (19)$$

And in the same spirit we can add:

$$D(\mathbf{p}) = \prod_{k=1}^{k=N} \delta_k \ln p_k, \quad \sum_{k=1}^{k=N} \delta_k = 0 \qquad (20)$$

The expenditure shares in this demand system are derived by applying Shephard's Lemma:

$$w_i = \frac{\ln A(\mathbf{p})}{\partial \ln p_i} + \beta_i[\ln x - \ln A(\mathbf{p}) - D(\mathbf{p})\ln h] + \frac{\lambda_i}{B(\mathbf{p})}[\ln x - \ln A(\mathbf{p}) - D(\mathbf{p})\ln h]^2 + \delta_i \ln h$$

The relationship between actual and compensated shares is then given by the next Proposition:

Proposition 4   For the Generalised PIGLOG demand system with background factor(s) included, equations (17)-(20), the relationship between compensated and actual budget shares is:

$$w_i(t,b) = w_i(t,t) - \beta_i z(t,b) - \frac{\lambda_i}{B(\mathbf{p}(t))}[z(t,b)]^2 - \delta_i \ln[h(t)/h(b)], \quad i = 1,...,N \qquad (21)$$

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

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

---

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

and $x(t,v)$ is the minimum expenditure necessary to achieve utility level $u(v)$ with background factor $h(v)$ at the prices of time t; $x(t,t)(x(b,b))$ is actual expenditure in country or period t (country or period b). Recall that the compensated shares are the shares we would observe if prices were those of time (country) t but consumers' utility were at the base period (country) level $u(b)$ and the background factor at the base level $h(b)$.

Proof   The proof, which draws on Oulton (2008) and (2012b, Appendix A.1), is in the Appendix.

In summary, to estimate the Konüs price index we need to estimate the compensated shares. The latter can be calculated from the actual shares given knowledge of the β_i, λ_i and δ_i parameters. These parameters can be estimated by fitting the generalised PIGLOG demand system econometrically. This method is quite feasible since it only requires us to estimate the parameters measuring the consumer's response to income changes or background factors, not the more numerous parameters measuring response to price changes. So the price variables can be replaced by their principal components in order to conserve degrees of freedom. Then the Konüs price index can be calculated by an iterative process using equations (15) and (21); see Oulton (2008) and (2012a) for detailed discussion. In the next section I use these results to estimate the proportion of space-time inconsistency attributable to non-homotheticity (as opposed to other causes).

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

5.1 Inconsistency due to path-dependence

The previous section showed that there is a practical method for estimating compensated shares and therefore Konüs price indices. So we are now in a position in principle to test the extent to which the condition for space-time inconsistency, equation (16), is satisfied for Konüs price indices. In other words, if we have estimates of each of the Konüs indices in equation (16) we can see how far the left hand side differs from the right hand side. But there are some further practical difficulties to resolve before we can carry out this test. We do not have Konüs price indices for each country's own inflation rate, i.e. the intertemporal price

indices. We only have price indices based on each country's national accounts. Insofar as background factors are concerned this may not matter since these change little if at all over time (at least over a 25 year interval such as 1980-2005 with which we will be concerned). Income changes over this period are more of a worry. But cross-country variation in real income which we will control for is much greater than intertemporal variation. Also the difference between a Konüs and an ordinary price index depends on how large are the changes in relative prices, not on pure inflation. Hopefully in most countries changes in relative prices over a period of 25 years are fairly small.[16]

As we have just seen a Konüs price index depends on the reference utility level and background factors 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).[17] If background factors matter too, then by analogy a chained, superlative index is likely to be approximately equal to a true price index with both background factors and utility level at the midpoint of the sample. 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 and background factors across the countries 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. As far as income is concerned this is very likely to be higher though background factors may be unchanged.

---

[16]   I found that a conventional, chained Laspeyres, index of the cost of living for the UK, the Retail Prices Index, grew on average at 6.21 per cent per year over 1974-2004, while a Konüs index based on 1974, the year when the standard of living was at its lowest over this period, grew at 6.32 per cent per year, i.e. more rapidly by 0.11 per cent per year (Oulton 2008). So over this 30-year period the cost of living rose cumulatively by an additional 3.4 per cent on the Konüs measure. The method of estimating the Konüs cost-of-living index over time was the same as the one employed in the present paper for PPPs.

[17]   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.

To assess the size of any inconsistency due to path-dependence (non-homotheticity) 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 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. 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).[18] These regressions also included 30 background variables covering climate, demography, religion, political history, and culture (for details see again Oulton 2012b and 2012c). 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 consumption per head rose on average by 53 percent between these two years (source: national accounts data from PWT8.0). I therefore set each country's real consumption per head in 2005 at 65 percent (=100/1.53) of its actual value. For each country, I calculate compensated budget shares for the 100 products within household consumption in 2005 with these new, lower income levels.[19] I assume that in each country the levels of the background variables were the same in 1980 as in 2005. Then I calculate PPPs for consumption for each of the 141 countries assuming real consumption per head is only 65 percent of its actual 2005 value. Both the new and the original 2005 PPPs are calculated firstly as a chained 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 same country order results from other common ways to measure aggregate PPPs, such as the World Bank's own PPPs for household consumption). I then back project these new 2005 PPPs to 1980 using each country's own price index and that of the US for consumption.

So finally, I compare the inconsistency index of equation (4) calculated using actual 2005 PPPs with the index calculated using compensated 2005 PPPs (in this case, $r = 2005$ and

---

[18]   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 expenditures at the Basic Heading level which the World Bank kindly made available to me.

[19]   The compensated shares were calculated using both a linear PIGLOG system (where all the λ_i in equation (21) are zero) and a quadratic PIGLOG.

$t = 1980$). If the index using compensated PPPs has a mean closer to 1 and a lower standard deviation than that of the index which uses actual PPPs then we may conclude that non-homotheticity has some role in accounting for inconsistency.

In more detail, the right hand side of equation (16), which is the 2005 Konüs PPP for country B calculated assuming average 1980 real income levels and back-projected to 1980, is calculated as

$$\ln PPP_{2005}^{K}(A,B) + \ln P_{C}^{B}(2005,1980) - \ln P_{C}^{A}(2005,1980)$$

where P_{C}^{A}, P_{C}^{B} are the national accounts deflators for consumption in the two countries and $PPP_{2005}^{K}(A,B)$ is the Konüs PPP for country B relative to A in 2005, with real consumption per head held at the mean level of 1980. The Konüs PPP is estimated as the actual PPP for that year ($PPP_{2005}^{PWT}(A,B)$), taken from the Penn World Table[20], multiplied by an adjustment factor:

$$PPP_{2005}^{K}(A,B) = PPP_{2005}^{PWT}(A,B)\left(\frac{P_{2005}^{WB,K}(A,B)}{P_{2005}^{WB}(A,B)}\right) \qquad (22)$$

Here $PPP_{2005}^{WB}(A,B)$ is a chain index of 2005 PPPs based on World Bank data and $PPP_{2005}^{WB,K}(A,B)$ is a Konüs PPP with real income reduced in each country by 35 percent, again based on World Bank data. Country B is in turn each of the 52 countries[21] in the sample and country A is the United States. I use Penn World Table PPPs for consistency since I use national accounts data from that source too.

The left hand side of equation (16) is taken to be the actual, 1980 PPP for consumption in country B (taken from the Penn World Table). The inconsistency index I on a compensated (Konüs) basis is then calculated as the ratio of the estimated right hand side of equation (16) to the estimated left hand side or in log terms:

$$\ln I_{Kon\ddot{u}s} = \ln PPP_{2005}^{K}(A,B) + \ln P_{C}^{B}(2005,1980) - \ln P_{C}^{A}(2005,1980) - \ln PPP_{1980}^{PWT}(A,B) \quad (23)$$

---

[20] PPPs do not appear directly in the PWT but they can be calculated from other variables which are present. Thus, using PWT notation, the PPP for household consumption in any country (in local currency units per US dollar) is given by $pl\_c \times xr / pl\_c[US]$ where pl_c is the price level for consumption, xr is the exchange rate (in local currency units per US dollar) and $pl\_c[US]$ is the price level for consumption in the US.

[21] 53 countries participated in both the 1980 and 2005 ICP. For one of these, Tanzania, I judged the 2005 expenditure data to be unreliable so a compensated PPP is not available for this country.

where the 1980 PPP, $PPP_{1980}^{PWT}(A,B)$, is taken from the Penn World Table. We want to compare this to the actual inconsistency index, i.e. the index using only actual PPPs:

$$\ln I_{actual} = \ln PPP_{2005}^{PWT}(A,B) + \ln P_C^B(2005,1980) - \ln P_C^A(2005,1980) - \ln PPP_{1980}^{PWT}(A,B) \quad (24)$$

Panel (a) of Table 3 shows means and standard deviations of the inconsistency index, both actual and compensated, as calculated by equations (22), (23) and (24). The left hand side of panel (a) shows results from PWT version 8.0; the right hand side shows results from version 8.1 which differs from 8.0 in that some PPPs have been adjusted to give what the authors of PWT regard as more reliable results (Feenstra *et al*. 2015b). Using version 8.1 both the mean and the standard deviation of the inconsistency index are actually somewhat higher for these countries than using version 8.0.

Two different models are employed for estimating compensated (Konüs) PPPs in 2005: the linear and the quadratic PIGLOG. Two different index number formulas are used, the chained Törnqvist and the chained Fisher. Using the linear PIGLOG model to generate compensated PPPs has very little effect on the inconsistency index: the mean and standard deviation are virtually the same whether we use version 8.0 or version 8.1. On the other hand using the quadratic PIGLOG model to generate compensated PPPs does have some effect on the distribution: the mean is reduced by about 6 per cent. The effect on the standard deviation is less clear cut. There is virtually no effect using version 8.0 but with version 8.1 the standard deviation is reduced by about 12%. None of these results is much affected by the choice of index number formula, Törnqvist or Fisher.

This conclusion is reinforced by the results in panel (b) of Table 3. Here no use is made of the Penn World Tables' PPPs for consumption and instead results are based entirely on my own estimates (using Basic Heading level data from the 1980 and 2005 rounds of the World Bank's ICP). The Konüs and the actual inconsistency indices are now calculated as follows:

$$\ln I_{Kon\ddot{u}s} = \ln PPP_{2005}^{K}(A,B) + \ln P_C^B(2005,1980) - \ln P_C^A(2005,1980) - \ln PPP_{1980}^{WB}(A,B) \quad (25)$$

$$\ln I_{actual} = \ln PPP_{2005}^{WB}(A,B) + \ln P_C^B(2005,1980) - \ln P_C^A(2005,1980) - \ln PPP_{1980}^{WB}(A,B) \quad (26)$$

Here the PPPs are calculated as either Fisher or Törnqvist chained indices and the Konüs PPPs use compensated shares, with 2005 real consumption per head in each country reduced by the average growth since 1980, i.e. by a factor of 0.65.

Looking at panel (b) of Table 3, we see that the linear PIGLOG does little to reduce either the mean or the standard deviation of the inconsistency index. The quadratic PIGLOG has some effect on the mean, reducing it by about 10 per cent, but little effect on the standard

deviation. The index number formula now has some effect: the mean is about 15% lower for the Fisher whether we look at the actual or the compensated index. [22]

This suggests that taking account of real income changes over 1980-2015 does little to reduce inconsistency. [23] If non-homotheticity were an important factor in accounting for inconsistency, then we would expect the mean of the compensated index to move closer to one and the standard deviation to be reduced by comparison to the actual index. But this is the case only with my own estimates of the consumption PPP and only to a limited extent. So there is not much evidence here that non-homotheticity can account for a significant proportion of inconsistency.

5.2 Inconsistency due to chain index approximation

Table 3 also sheds light on the likely importance of chain index approximation error. As just noted the index number formula makes virtually no difference to the compensated inconsistency index when PWT PPPs are used: see lines 2-4 of panel (a). Results using my own PPPs for household consumption are in panel (b) of Table 3. Lines 6 and 7 compare two different measures of the actual index: the chained Törnqvist (line 6) and the chained Fisher (line 7), both applied to underlying World Bank data at the Basic Heading level in 1980 and 2005. See Chart 2 for these two distributions of the actual index compared with the one based entirely on PWT data (as in line 1). Under the chained Törnqvist inconsistency is quite a bit higher than under the chained Fisher though the latter is close to the Penn World Table figure (line1). The main reason for the difference between Törnqvist and Fisher is that while the estimated PPPs for 1980 are very similar, those for 2005 diverge more with the Fisher being on average about a third higher (though the two indices are still highly correlated: $r = 0.992$). The Fisher and Törnqvist index numbers are the ones most commonly employed in practice. So the fact that inconsistency is significant even with the Fisher suggests that chain index approximation error is not the root cause of the problem.

---

[22]   One reason that they differ at all is that the PWT uses the GEKS formula applied to bilateral Fisher indices while my estimates use chained indices.

[23]   In estimating Konüs PPPs for 2005 I found that allowing for income differences makes a considerable difference (Oulton 2012c). But there the income differences are huge, with factors of tens or even hundreds being common. Here the income difference is a factor of only 1.5.

6. Conclusions

We have seen that space-time inconsistency is economically significant both in 35 rich countries over 1995-2011 and in a wider group of 52 rich and poor countries which participated in both the 1980 and the 2005 rounds of the World Bank's ICP. We have found that path-dependence (non-homotheticity) can probably account for only a small proportion of observed space-time inconsistency in the national accounts. And a similar finding applies to chain index approximation error. 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 international comparisons 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 by Rao *et al*. (2010) and by Hill and Melser (2014).

Finally, though the main focus of this paper has been on the measurement of consumption (welfare), many of the same considerations apply to output (GDP). Space-time inconsistency is larger for GDP than for consumption: compare the standard deviation of the index for GDP in Table 2 with that for consumption in Table 3. If the pattern of output is affected by background factors like human and physical capital stocks or resource endowments or if there are non-constant returns to scale then non-homotheticity will be an issue for output measurement too. But the results for consumption presented here suggest that allowing for these factors will not reduce inconsistency in GDP by very much.

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

Country	Code	Projection based on 1995 PPPs		Projection based on 1999 PPPs	
		2011	*2012*	*2011*	*2012*
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

Mean	1.0934	1.1009	1.0907	1.0981
S.D.	0.1484	0.1528	0.1388	0.1424
Minimum	0.8350	0.8407	0.8350	0.8407
Maximum	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 national accounts: Penn World Table data for 53 countries, 1980-2005
(in ascending order of the inconsistency index, column (3))

	GDP per head in 2005, 1980=100			Growth of GDP per head, percent p.a.		
	National Accounts	PWT	Inconsist-ency index (column 2 ÷ column 1)	National Accounts	PWT	Difference, column 5 *minus* column 4
Country	(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
Mean	159.9	168.8	1.069	1.54	1.66	0.13
S.D.	75.0	87.8	0.323	1.62	1.84	1.04

Source Penn World Table 8.0 and own calculations (http://www.rug.nl/research/ggdc/data/pwt/pwt-8.0). 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*). By construction in PWT, *rgdpo* = *rgdpna* in 2005. Both *rgdpo* and *rgdpna* are measured in 2005 US dollars in all years.

Table 3
Compensated and actual inconsistency indices for household consumption, 1980-2005
(52 countries which participated in both the 1980 and 2005 ICP)

(a) Penn World Table PPPs for household consumption	Version 8.0		Version 8.1	
	Mean	*S.D.*	*Mean*	*S.D.*
1. Actual (I_{Actual})	1.062	0.277	1.166	0.324
Compensated (adjusted for growth of real income per head 1980-2015)				
Linear PIGLOG $(I_{Konüs})$				
2. Chained Törnqvist	1.071	0.282	1.177	0.333
3. Chained Fisher	1.062	0.279	1.167	0.326
Quadratic PIGLOG $(I_{Konüs})$				
4. Chained Törnqvist	0.990	0.275	1.079	0.284
5. Chained Fisher	0.991	0.283	1.080	0.290

(b) My PPPs for household consumption	*Mean*	*S.D.*
6. Actual chained Törnqvist (I_{Actual})	1.253	0.370
7. Actual chained Fisher (I_{Actual})	1.076	0.285
Compensated (adjusted for growth of real income per head 1980-2015)		
Linear PIGLOG $(I_{Konüs})$		
8. Chained Törnqvist	1.266	0.380
9. Chained Fisher	1.076	0.287
Quadratic PIGLOG $(I_{Konüs})$		
10. Chained Törnqvist	1.159	0.323
11. Chained Fisher	1.005	0.289

Sources   Penn World Tables 8.0 and 8.1 (http://www.rug.nl/research/ggdc/data/pwt/pwt-8.1), PPPs and expenditures at Basic Heading level from the World Bank's 1980 and 2005 ICP, and own calculations.
Notes   Inconsistency ratio: ratio of 2005 PPP, projected back to 1980, to 1980 PPP. The compensated ratio is the ratio of the compensated PPP for consumption in 2005, projected back to 1980 using national accounts deflators for consumption, to the actual 1980 PPP for consumption. The actual inconsistency ratio is the ratio of the actual 2005 PPP for consumption, projected back to 1980 using national accounts deflators for consumption, to the actual 1980 PPP for consumption. Compensated chained Törnqvist and chained Fisher PPPs for 2005 estimated from unpublished World Bank data from the 2005 ICP comprising PPPs and expenditures for 100 Basic Headings within household consumption; see Oulton (2012b) and (2012c) for details. Compensated PPPs for 2005 are estimated from either a linear or a quadratic PIGLOG system after setting real income in each country to 65 percent of its actual level in 2005. The compensated PPP in 2005 (Törnqvist or Fisher) is the actual PPP in 2005 (from PWT8.0) times the 2005 ratio of compensated to actual PPPs, this last ratio derived from World Bank data as just described. National accounts deflators are from PWT8.0. Rows numbered 1-5 use the PWT estimate of the 1980 and 2005 PPP for household consumption. Rows numbered 6-11 use my own estimates of the 1980 and 2005 PPP for household consumption. For 1980 the source is the underlying data from the 1980 ICP: PPPs at the Basic Heading level for 115 Basic Headings within consumption and the corresponding expenditures, also available from the PWT website (a spreadsheet named icp1980.xls). For 2005 the source is unpublished World Bank data from the 2005 ICP as described above. Sample is 52 countries rather than the 53 in Table 2 since Tanzania, though present in both the 1980 and 2005 ICP, was excluded from the estimates in Oulton (2012b) and (2012c) due to data unreliability. See text for more detail.
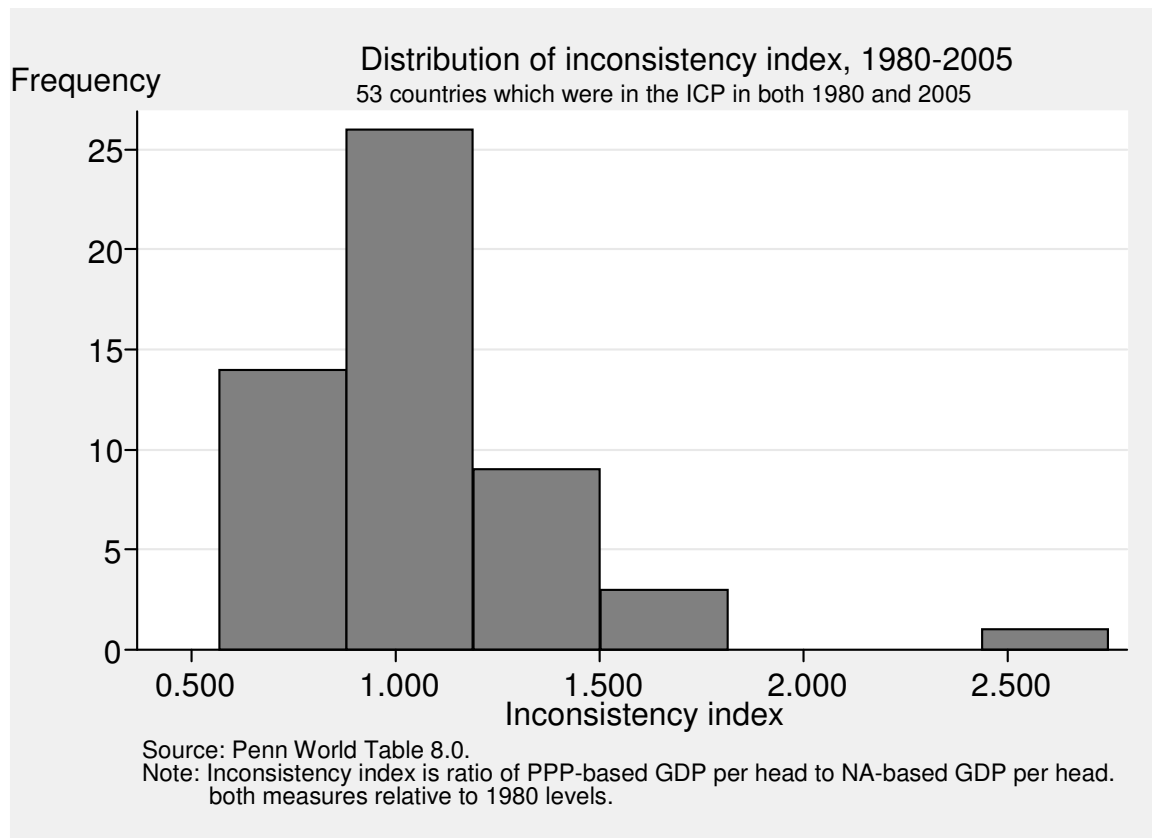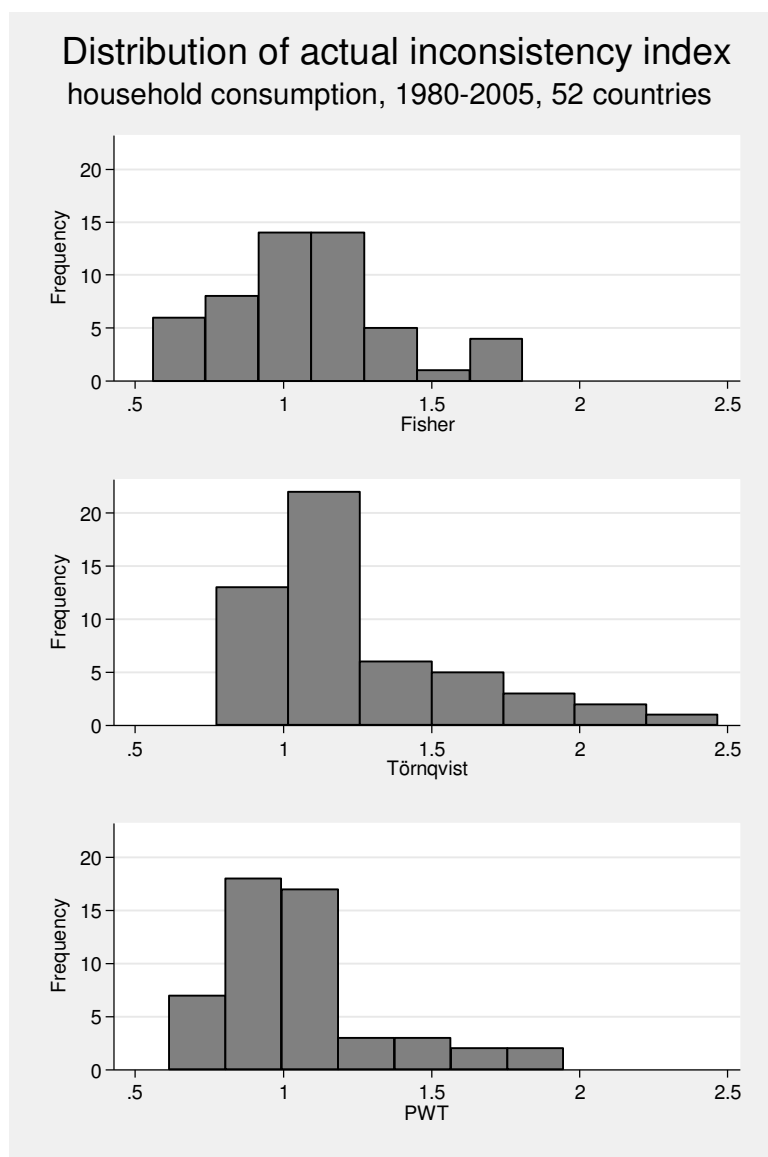
Chart 1

Distribution of inconsistency index, 1980-2005
53 countries which were in the ICP in both 1980 and 2005

Source: Penn World Table 8.0.
Note: Inconsistency index is ratio of PPP-based GDP per head to NA-based GDP per head.
both measures relative to 1980 levels.

Chart 2

Distribution of actual inconsistency index
household consumption, 1980-2005, 52 countries

Source　See notes to Table 3. The means and standard deviations of these distributions are in lines 1 (PWT version 8.0, 6 and 7 (labelled "Actual") of Table 3.
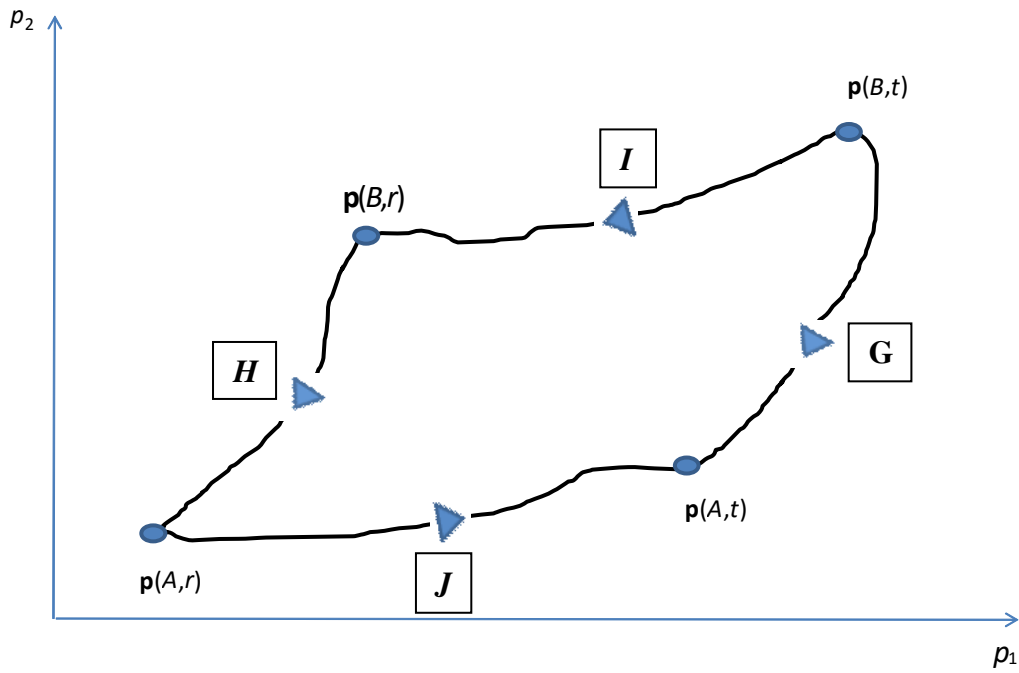
Figure 1   Price variation over time and across countries: the two-good case
Note: $\mathbf{p}(X,s)$ is the price vector for country X $(A \le X \le B)$ at time s $(r \le s \le t)$.

APPENDIX

A.1 Characterisation of Divisia 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 ($i = 1,...,N$). 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:

$$P^D(r,t)Y^D(r,t) := \frac{\sum_{i=1}^{N} p_i(t)\,y_i(t)}{\sum_{i=1}^{N} p_i(r)\,y_i(r)} \tag{A.27}$$

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 differentiable) and the equality of the left and right hand sides will continue to hold:

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

where $w_i(t) := \dfrac{p_i(t)\,y_i(t)}{\sum_{i=1}^{N} p_i(t)\,y_i(t)}$, 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:

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

$$\tag{A.28}$$

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

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

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

$$= \sum_{i=1}^{N} \left[\int_r^t w_i(\tau) \frac{d \ln p_i(\tau)}{d\tau} d\tau \right] \qquad (A.29)$$

and

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

$$= \sum_{i=1}^{N} \left[\int_r^t w_i(\tau) \frac{d \ln y_i(\tau)}{d\tau} d\tau \right] \qquad (A.30)$$

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

$$P^D(r,r) = Y^D(r,r) = 1 \qquad (A.31)$$

This normalisation is consistent with the definition of equation (A.27) 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, the indices are consistent in aggregation.[24]

A.2 Proof of Proposition 4

Proposition 4 in the main text states that, for the generalised PIGLOG demand system of equations (17)-(20) with a background factor h included, the compensated shares are given by

$$w_i(t,b) = w_i(t,t) - \beta_i z(t,b) - \frac{\lambda_i}{B(\mathbf{p}(t))} [z(t,b)]^2 - \delta_i \ln[h(t)/h(b)], \quad i = 1,...,N$$

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

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

---

[24]   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.

Proof   The proof draws on Oulton (2012a, Appendix A.1), augmented to allow explicitly for background factors. A more general proposition is proved first and then applied to the generalised PIGLOG system.

The budget shares are functions of utility, but utility is a monotonically increasing function of expenditure when prices and background factors are held constant. So the budget shares are functions of expenditure, as well as of prices and the background factor. We now need the following assumption:

Assumption   The function relating the budget shares of any product to log expenditure is *entire*: that is, it is infinitely differentiable (smooth) and its Taylor series converges to the value of the function at every point in the (economically relevant) domain, i.e. where $x > 0$. Note that polynomials, the exponential function, and the sine and cosine functions are entire. And sums, products and compositions of entire functions are also entire.

So consider the share w_i for the *i*-th product as a function of expenditure and the background factor. To approximate this function at the point $w_i(t,t)$, expand it in a Taylor series around the point $w_i(t,b)$: that is, hold prices constant at their levels at time *t* and vary expenditure (utility) and the background factor from their levels in the base period (period *b*), to obtain

$$w_i(t,t) = w_i(t,b) + \left(\frac{\partial w_i(\cdot,\cdot)}{\partial \ln E(\cdot,\cdot)} \right)_{\substack{\mathbf{p}=\mathbf{p}(t), \\ x=E(t,b), \\ h=h(b)}} \cdot [\ln x(t,t) - \ln x(t,b)]$$

$$+ \frac{1}{2!} \left(\frac{\partial^2 w_i(\cdot,\cdot)}{\partial \ln E(\cdot,\cdot)^2} \right)_{\substack{\mathbf{p}=\mathbf{p}(t), \\ x=E(t,b), \\ h=h(b)}} \cdot [\ln x(t,t) - \ln x(t,b)]^2 + ...$$

(A.32)

$$+ \left(\frac{\partial w_i(\cdot,\cdot)}{\partial \ln h} \right)_{\substack{\mathbf{p}=\mathbf{p}(t), \\ x=E(t,b), \\ h=h(b)}} \cdot [\ln h(t) - \ln h(b)]$$

$$+ \frac{1}{2!} \left(\frac{\partial^2 w_i(\cdot,\cdot)}{\partial \ln h^2} \right)_{\substack{\mathbf{p}=\mathbf{p}(t), \\ x=E(t,b), \\ h=h(b)}} \cdot [\ln h(t) - \ln h(b)]^2 + ...$$

(Recall that $x(t,v)$ is defined as the minimum expenditure requited to achieve $u(v)$ at the prices of time (country) *t* when the background factor is $h(v)$). Note that $\ln x(t,t) - \ln x(t,b) = \ln[x(t,t)/x(t,b]$ is the log of the ratio of the expenditure needed to

achieve the utility level of period t to the expenditure needed to achieve the level of period b, both evaluated at the prices of period t. In fact

$$\frac{x(t,t)}{x(t,b)} = \left[\frac{x(t,t)}{x(b,b)}\right] \cdot \left[\frac{x(b,b)}{x(t,b)}\right] = \frac{x(t,t)/x(b,b)}{P^K(t,b,b)} \qquad (A.33)$$

where $x(v,v)$ is actual money expenditure at time v ($v=t,b$) and we have used the definition of the Konüs price index in equation (15).

For the generalised PIGLOG system of equations (17)-(20) the expenditure shares are derived by applying Shephard's Lemma:

$$w_i = \frac{\partial \ln x}{\partial \ln p_i}$$
$$= \frac{\ln A(\mathbf{p})}{\partial \ln p_i} + \beta_i[\ln x - \ln A(\mathbf{p}) - D(\mathbf{p})\ln h] + \frac{\lambda_i}{B(\mathbf{p})}[\ln x - \ln A(\mathbf{p}) - D(\mathbf{p})\ln h]^2 + \delta_i \ln h$$

Now adopt the normalisation that $\ln(u(b)) = 0$ so that $\ln[x(t,b)/A(\mathbf{p}(t))D(\mathbf{p}(t))] = 0$ where as before $x(t,b)$ is the minimum expenditure required to reach $u(b)$ at the prices of t and background factor at the level $h(b)$. Then the derivatives in (A.32) evaluate to

$$\left(\frac{\partial w_i(\cdot,\cdot)}{\partial \ln E(\cdot,\cdot)}\right)_{\substack{\mathbf{p}=\mathbf{p}(t),\\x=E(t,b),\\h=h(b)}} = \beta_i$$

$$\left(\frac{\partial^2 w_i(\cdot,\cdot)}{\partial \ln E(\cdot,\cdot)^2}\right)_{\substack{\mathbf{p}=\mathbf{p}(t),\\x=E(t,b),\\h=h(b)}} = \frac{2\lambda_i}{B(\mathbf{p}(t))}$$

and

$$\left(\frac{\partial w_i(\cdot,\cdot)}{\partial \ln h}\right)_{\substack{\mathbf{p}=\mathbf{p}(t),\\x=E(t,b),\\h=h(b)}} = \delta_i$$

All other higher derivatives in (A.32) are zero. So substituting these values and (A.33) into (A.32) and rearranging we obtain Proposition 4. □

Note that the Taylor series is exact for the generalised (quadratic) PIGLOG.

A.3 Chain indices compared to the GEKS multilateral price index

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,

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

Now define $\mathbf{A} = \ln \mathbf{p} \cdot \mathbf{1}$ and $\mathbf{B} = \ln \mathbf{p}' \cdot \mathbf{1}$. Here $\mathbf{1}$ is a Cx1 column vector of ones and $\ln \mathbf{p}$ is a CxC matrix of the log of the bilateral price indices: $\ln \mathbf{p} = [\ln p_{ij}]$. So $\ln \mathbf{p} \cdot \mathbf{1}$ is a Cx1 column 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 Cx1 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

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

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

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

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

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

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 bilateral price indices 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}.

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

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

$$= P_{i,i+m}^{Ch} P_{i+m,i+n}^{Ch}$$

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 **p** matrix above. This leaves row and column sums unchanged. To illustrate, let matrix $\tilde{\mathbf{p}}$ be the same as **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

$$\ln \tilde{P}_{14}^{GEKS} = \left(\frac{1}{4}\right)\left[\sum_{k=1}^{4} \ln p_{1k} + \sum_{k=1}^{4} \ln p_{k4}\right] = \ln P_{14}^{GEKS}$$

But the corresponding chain indices $\tilde{P}_{14}^{Ch}$ and P_{14}^{Ch} are not in general equal since $\tilde{P}_{14}^{Ch} = P_{13}P_{32}P_{24} \neq P_{12}P_{23}P_{34} = P_{14}^{Ch}$ in general.

In the two cases just considered it could be argued that both the GEKS and the chain index make use of exactly the same underlying information, namely the prices and budget shares of each of the four countries. This is not necessarily always the case. For example, suppose we change the ordering to 1,4,2,3. Then the GEKS is unchanged but the chain index for country 4 relative to country 1 now makes use of only the prices and budget shares of these two countries: $\tilde{P}_{14}^{Ch} = P_{14}$. In general, the GEKS uses the information on prices and budget shares in all countries in order to construct each multilateral index; this information is summarised by bilateral price indices between each pair of countries. The chain index on the other hand discards some information unless the two countries being compared are the first and the last (according to the chosen ordering).

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; Cooper and McLaren 1992 and 1996). And Hill (2006) has shown that there can be large differences in practice between different superlative index numbers when the gaps between the countries or time periods being compared are large. 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 DRC 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.

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.

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.

Cooper, Russel J. and Keith R. McLaren (1992). "An Empirically Oriented Demand System With Improved Regularity Properties". *Canadian Journal of Economics* **25**(3), 652-668.

Cooper, Russel J. and K.R. McLaren (1996). "A System of Demand Equations Satisfying Effectively Global Regularity Conditions". *Review of Economics and Statistics*, 78(2), May, 1996, 359-364.

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.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer (2015a). "The next generation of the Penn World Table". *American Economic Review*, 105(10), 3150-3182.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer (2015b). "What is new in PWT 8.1?". http://www.rug.nl/research/ggdc/data/pwt/v81/what_is_new_in_pwt_81.pdf.

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.

Oulton, Nicholas (2015). "Understanding the space-time (in)consistency of the national accounts", *Economics Letters*, vol. 132, April 2015. [http://dx.doi.org/10.1016/j.econlet.2015.04.009].

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.

Richter, Marcel K. (1966). "Invariance axioms and economic indexes". *Econometrica*, 34: 739–755.

Salas, Saturnino L., Einar Hill and Garret J. Etgen (2007). *Calculus: One and Several Variables* (Tenth edition). Hoboken, NJ: John Wiley and Sons, Inc.

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.

Woodland, A. D. (1982). *International Trade and Resource Allocation.* Amsterdam: North-Holland Publishing Company.

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.