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

CFM-DP2015-24

Nicholas Oulton

Centre For Macroeconomics, London School of Economics and Political Science

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 issue has been raised again by the release of the results of the 2011 round of the World Bank's International Comparison Program (ICP) which are very different from what would be expected on the basis of extrapolating from the 2005 round.

PPPs are a critical building block in the widely-used Penn World Table (PWT). The Penn World Table enables one to observe the size of countries' economies over time (GDP), their standard of living (measured by consumption per head) and productivity (GDP per employee). The trouble is that the PWT has two measures of the growth rate of GDP and of consumption, one based on PPPs and the other based on each country's national accounts. And these two measures can differ substantially. An extreme case is Argentina. Based on PPPs the country is a star: its GDP per head rose by 4.37 % per year between 1980 and 2005. But based on its own national accounts GDP per head rose by only 0.33% per year.

This paper makes three contributions. First, I document the extent of space-time inconsistency in two datasets: (a) the 35 (mainly rich) countries for which the OECD produces annual estimates of

PPPs; and (b) 52 countries which participated in both the 1980 and the 2005 round of the World Bank's ICP. In both cases I find space-time inconsistency to be significant.

Second, I consider the theoretical conditions under which, in the absence of errors in the data, real consumption per head would be perfectly *consistent* across space and time. (Similar considerations apply to output measurement). I show that in principle this would be the case provided (1) that consumer demand is *homothetic,* which means that budget shares respond only to prices and not to real incomes or to background factors like culture, climate or demography; and (2) that Divisia (continuous) price indices are employed to measure PPPs and domestic prices. So if we observe space-time inconsistency in practice it must be due to one or more of the following causes:

1. Non-homotheticity of the consumer's expenditure 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 domestic price indices.

Non-homotheticity is obviously a very strong condition which is unlikely to be satisfied in practice. But how much does it contribute to the inconsistency which we observe? The third contribution of this paper is to estimate the size of the non-homotheticity effect empirically. I do this by estimating PPPs for 2005 which are adjusted to the real income levels of 1980. I then compare the degree of inconsistency observed with these adjusted PPPs with the degree observed with actual PPPs. I find that the degree of inconsistency is about the same. So non-homotheticity does not account for inconsistency. I also assess the sensitivity of inconsistency to discrete approximation error by comparing results using two widely employed index numbers, chained Fisher and chained Törnqvist. Again I find that inconsistency remains significant whichever index number is employed.