Comparing Price Levels across Countries using Minimum-Spanning Trees

Keir G. Armstrong

Carleton University, Ottawa, ON K1S 5B6, Canada

karmstro@ccs.carleton.ca

August 31, 2000

This paper points up two fundamental problems with Robert J. Hill's approach to the theory of multilateral international comparisons. One problem is empirical in nature: minimum-spanning trees in an international comparisons context are inherently unstable both over time and over small variations in the underlying data set for a given year. The second problem is that the theoretical underpinnings of the MST method are extremely weak. Hill's attempt to stretch the accepted theory of bilateral comparisons into a theory of multilateral ones by restricting the admissible binary comparisons in a particular manner is both unjustified (from a theoretical economic perspective) and undesirable (from an axiomatic perspective).

Key Words: index-number theory; purchasing power parities; multilateral international comparisons.

JEL Classification Numbers: C31, C43, C81, E31, O57.

Comment

In a recent paper, Robert J. Hill (1999a) introduces a new method for making multilateral international comparisons of relative purchasing power that is based on the minimum-spanning tree (MST) of a weighted connected graph with vertices corresponding to a bloc of countries and weights given by the (bilateral) Paasche-Laspeyres spreads among these countries. There are two fundamental problems with this approach. First, MSTs in an international comparisons context are inherently unstable, both over time and over small variations in the underlying data set for a given year. Second, the theoretical underpinnings of the MST method are extremely weak. Each of these problems is dealt with, in turn, below.

Hill (1999a) assumes a bloc of countries indexed by $k = 1, \ldots, K$ with country-specific prices (p_{ki}) and quantities (q_{ki}) corresponding to a common set of well-defined types of goods and services indexed by $i = 1, \ldots, N$. The purchasing power parity (PPP) between any pair of countries in the bloc is measured as the product of the Fisher price indexes along a pre-determined path connecting these countries (via zero, one or more of the other countries). The pre-determined paths within the bloc correspond to the minimum total "distance" among the constituent countries subject to the constraint that there is a unique connection between each pair. More precisely, the MST PPP index for country k relative to country $j \in \{1, \ldots, K\}$ is defined as the chain of Fisher price indexes across the minimum spanning tree T of the weighted connected graph G of order K with vertices $\{1, \ldots, K\}$ and weights (w_{jk}) given as $w_{jk} := \left|\ln P^L_{jk} - \ln P^P_{jk}\right|$; *i.e.*,

$$P^{MST}_{jk} := P^F_{jh} P^F_{h\ell} \cdots P^F_{tk} ,$$

where $jh, h\ell, \ldots, tk$ are edges in T, $h, \ell, \ldots, t \in \{1, \ldots, K\}$, $P^F_{jk} := \left(P^L_{jk} P^P_{jk}\right)^{\frac{1}{2}}$, $P^L_{jk} :=$

0

$\sum_i p_{ki}q_{ji} \,/\, \sum_i p_{ji}q_{ji}$ and $P^P_{jk} := \sum_i p_{ki}q_{ki} \,/\, \sum_i p_{ji}q_{ki}$.

Hill's (1999a) empirical illustration of his MST method is based on the 1980 and 1985 cross-sectional data sets produced by the United Nations International Comparison Project (ICP) and supplied to him by Alan Heston (Hill, 1999a, fn *). These data sets are now freely available at the Center for International Comparisons world-wide web site (pwt.econ.upenn.edu) wherefrom they were downloaded for the purposes of the present paper. Adhering to Hill's (1999a, p. 139) methodology for dealing with missing prices and quantities, I was able to replicate his 30-country MST for 1980 up to the edge connecting India. My 30-country MST for 1985, however, differs from his with respect to eight of the twenty-nine edges. Since I was able to precisely replicate Hill's (1999b, Fig. 2.4) OECD-24 MST for 1990 as well as his associated PPPs (Hill, 1999b, Table 2.3) using the same data set[1] from the same source,[2] it is undoubtedly the case that the differences between our 1980 and 1985 MSTs are due to small differences between our data sets rather than to some procedural error on either part.

The reasonableness of this conclusion is underscored by a close examination of the procedure used to generate T, *viz.* Kruskal's algorithm. To begin with, an edge in G of minimum weight is chosen. Successive edges with progressively higher weights are examined and chosen if and only if doing so induces an acyclic subgraph of G. The algorithm stops once $K-1$ edges have been chosen. It is the acyclicity restriction that makes it likely that a given MST will be sensitive to small perturbations in the associated weights. Suppose, for example, that Kruskal's algorithm is at a point in its edge-selection process where countries h and ℓ have been chosen but country j has not. Suppose further that w_{jh} is the lowest unexamined weight in G and $w_{j\ell}$ is

[1] *Viz.* the electronic spreadsheets "par90.wk1" and "nom90.wk1" comprising the basic-heading prices and expenditures for 1990 produced by the Eurostat-OECD PPP Programme.

[2] *Viz.* the OECD Statistics Directorate.

among the next few lowest. The smaller is the difference $w_{j\ell} - w_{jh}$, the smaller is the increase in w_{jh} required to compel the algorithm to choose edge $j\ell$ rather than edge jh. The greater the number of edges that end up sharing the h vertex when $w_{j\ell} > w_{jh}$, the greater the impact of such an increase on the structure of T.

A more dramatic illustration of MST sensitivity with respect to the weights is provided by the addition of new countries. After noting that "[a] total of 42 countries were present in both the 1980 and 1985 ICP data sets," Hill (1999a, fn 13) goes on to say that "five smaller European countries and seven smaller African countries were removed from the sample to make the regional groupings of countries more balanced."[3] Among the deleted European countries are the Netherlands and Belgium with populations (in the early 1980s) of about 14 million and 10 million, respectively. Among the deleted African countries are Ethiopia and Madagascar with respective populations of about 44 million and 9 million. Since Hill did not delete six European countries with populations that were roughly the same or smaller than those of the Netherlands and Belgium, and did not delete three African countries with populations that were smaller than Madagascar's, one is forced to conclude that by "smaller" he means "relatively unimportant" in some non-demographic sense.

I tested the implied proposition that the Netherlands, Belgium, Ethiopia and Madagascar are relatively unimportant to the structure of T by including these countries in G and then applying Kruskal's algorithm. The resulting 34-country MSTs for 1980 and 1985 are depicted, respectively, in Figures 1(a) and 1(b). Each solid or dotted line therein represents an admissible binary link among the countries in G (*i.e.*, an edge in T). Dashed lines are edges in the corresponding

[3] It is interesting that, in the very next paragraph, Hill talks about "the disadvantage of throwing away useful information" and describes the steps he took to "avoid throwing away data" from the remainder of the sample.

32-country MST that are no longer admissible, dotted lines are new edges, and solid lines are edges that are invariant to the added countries (labelled in italics). While Ethiopia (ETH) and Madagascar (MDG) appear to be relatively unimportant to the structure of the 1980 MST, the broken link between Côte d'Ivoire (CIV) and Zimbabwe (ZWE) in Figure 1(b) suggests that this may not be the case for the 1985 MST. Similarly, the Netherlands (NLD) appears to be relatively unimportant to the structure of the 1985 MST but not to that of the 1980 MST. The presence, in both figures, of many dashed and dotted lines in the immediate vicinity of Belgium (BEL), however, strongly suggests that this country is very important to the structure of both MSTs.

Further evidence of the instability of the 1980 and 1985 MSTs with respect to the addition of new countries is provided by Figures 1(c) and 1(d). These figures show that the result of reinstating the remaining eight countries thrown out by Hill is a fragmentation of the 34-country MST into several clusters of solid-line -linked countries.

A comparison of Figures 1(c) and 1(d) reveals that there are very few features of the full 42-country MST which are robust: The African countries are grouped together with Tanzania (TZA) as the portal to the rest of the world, and the "Europe Group 2" countries[4] are closely linked.[5] None of the other features that Hill (1999a, p. 139) characterizes as "common to both spanning trees" holds up to the addition of the twelve omitted countries. Hence, the MST cannot be said to be stable over time.

The second fundamental problem with the MST method is that it is not justifiable in terms of economic theory. As demonstrated below, the justification that Hill (1999a, pp. 137–38)

[4] *Viz.* Austria (AUT), Yugoslavia (YUG), Hungary (HUN) and Poland (POL).

[5] The robustness of these clusters may be an artifact of the many regionally-based imputations the ICP had to make for the constituent countries (Austria excepted) due to the low quality of the underlying raw data. See Summers and Heston (1988) for a discussion of this issue.

appears to provide does not stand up to scrutiny.

Under the economic approach to index-number theory, the "true" PPP for country k relative to country j is given by the cost-of-living index. If the base utility level is chosen to be that of the reference country j, then the associated cost-of-living index P_{jk}^{j} is bounded from above by the corresponding Laspeyres price index P_{jk}^{L} and from below by the minimum price relative p_{ki}/p_{ji}, $i \in \{1, \ldots, N\}$. Equivalently, if the base utility level is chosen to be that of the comparison country k, then the associated cost-of-living index P_{jk}^{k} is bounded from below by the corresponding Paasche price index P_{jk}^{P} and from above by the maximum price relative p_{ki}/p_{ji}, $i \in \{1, \ldots, N\}$.[6] A sufficient condition for the cost-of-living index to be bounded both from above and from below by the corresponding Laspeyres and Paasche price indexes, respectively, is that the cost-of-living index be independent of the base utility level or, equivalently, that the underlying preference ordering be homothetic to the origin.[7] Since this is the only known condition for such bounds to exist, it must be the case that Hill (1999a) is implicitly assuming that preferences are identical and homothetic across countries.

My interpretation of Hill's theoretical justification for his MST method is as follows. Let P_{jk}^{*} denote the "true" PPP for country k relative to country j. Since preferences are identical and homothetic across countries, P_{jk}^{*} is a (unknown) base-independent cost-of-living index that is bounded by P_{jk}^{P} and P_{jk}^{L}. Since P_{jk}^{*} could lie anywhere in this range, the best bilateral estimate of its value is the Fisher index P_{jk}^{F}. Since the Fisher index is not transitive with respect to the reference and comparison prices, P_{jk}^{F} is not necessarily equal to $P_{j\ell}^{F} P_{\ell k}^{F}$ for any $\ell \neq j, k$ and $\left(P_{jk}^{F}\right)$

[6] The equivalence of these statements follows from the fact that $P_{jk}^{k} = 1/P_{kj}^{k}$, $P_{jk}^{P} = 1/P_{kj}^{L}$ and $\max_i \{p_{ki}/p_{ji}\} = 1/\min_i \{p_{ji}/p_{ki}\}$.

[7] See Pollak (1989, pp. 18–20).

is not an appropriate basis for a set of multilateral PPP indexes.[8] How do we cope with this? Select $K-1$ of the K^2 indexes in $\left(P_{jk}^F\right)$ such that they admit exactly one way to compare every pair of countries, either directly or indirectly using the chain principle, and such that the sum of the associated Paasche-Laspeyres spreads is a minimum. The idea here is to admit only those Fisher indexes for which the known bounds are tightest, subject to the constraint that the resulting subset is transitive. The theoretical justification for this procedure is loosely based on Diewert's (1983, pp. 186–87) argument for chain indexes in the intertemporal context:

> Theorems 4 and 10 above showed that is very useful to have the Paasche and Laspeyres indexes close to each other since this will lead to a very close approximation to the ["true"] cost-of-living index .... This implies that we should use the chain principle for constructing indexes rather than the fixed base principle.

An important caveat to this argument is provided by Szulc (1983, p. 555):

> [A] chain of Laspeyres fixed-basket indices can be higher than the corresponding direct Laspeyres price index and a chain of Paasche fixed-basket indices can be lower than the corresponding direct Paasche price index. ... [Chaining], therefore, should be avoided in a situation when such undesirable results tend to materialize.

The obvious question at this juncture is: does Szulc's caveat apply in the present context? To find out, I calculated the chained Paasche and Laspeyres price indexes relative to the United States across each of the two 42-country MSTs and then compared them with their direct counterparts. For both 1980 and 1985, more than three-quarters of the relevant forty-one countries exhibit the "undesirable results" described by Szulc. Consequently, the P_{jk}^{MST}s are very likely to be biased away from the P_{jk}^*s they purport to measure.

Given the way that the chains relevant to the calculation of Hill's PPPs are constructed, one should not be surprised by this result. The imposed transitivity requirement impedes the effective

[8] Note that P_{jk}^* is transitive with respect to the reference and comparison prices because it is a cost-of-living index.

use of the interpretation of the data set as K observations on a single global consumer afforded by the assumption of identical homothetic preferences across countries. Ideally, this information would be used to systematically narrow the Paasche-Laspeyres spreads as much as possible. Once the bounds on the P^*_{jk}s are as tight as the data set allows, a set of multilateral indexes respecting them could be constructed.

The approach just described is precisely the one followed by Dowrick and Quiggin (1997). The particular method they develop has the desirable feature that it converges to the tightest possible bounds if and only if the data set admits a homothetic representation. Thus, the method itself is capable of invalidating the critical assumption on which it is based. When applied to the two 42-country ICP data sets employed above, the Dowrick-Quiggin method generates PPPs which are substantially different from those of the MST method for 1980,[9] and fails to converge for 1985. The former result illustrates the expected bias in the MST method, and the latter reveals that its application may be wholly inappropriate.

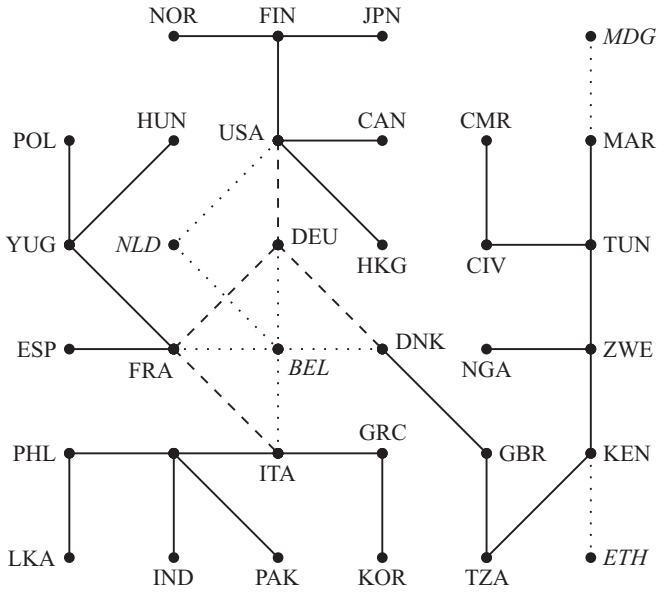
References

1. Diewert, W.E., "The Theory of the Cost-of-Living Index and the Measurement of Welfare Change," in W.E. Diewert and C. Montmarquette (eds), *Price Level Measurement* (Minister of Supply and Services Canada, 1983), 163–233.
2. Dowrick, S. and J. Quiggin, "True Measures of GDP and Convergence," *American Economic Review* 87 (1997), 41–64.
3. Hill, R.J., "Comparing Price Levels across Countries using Minimum-Spanning Trees," *Review of Economics and Statistics* 81 (1999a), 135–42.
4. Hill, R.J., "International Comparisons using Spanning Trees," in A. Heston and R.E. Lipsey (eds), *International and Interarea Comparisons of Income, Output and Prices* (University of Chicago Press, 1999b), 109–120.
5. Pollak, R.A., *The Theory of the Cost-of-Living Index* (Oxford University Press, 1989).

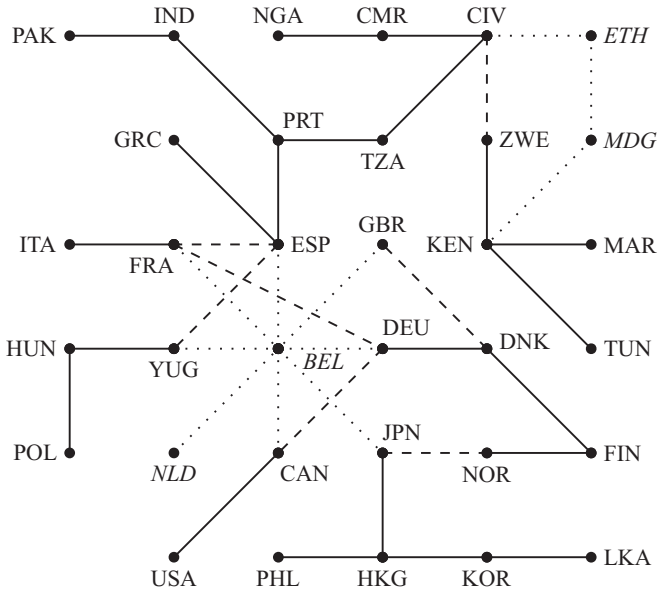
[9] The average absolute difference between the corresponding PPPs is about 3.7 percent.

6. Summers, R. and A. Heston, “A New Set of International Comparisons of Real Product and Price Levels Estimates for 130 Countries, 1950–1985,” *Review of Income and Wealth* 34 (1988), 1–25.
7. Szulc, B.J., “Linking Price Index Numbers,” in W.E. Diewert and C. Montmarquette (eds), *Price Level Measurement* (Minister of Supply and Services Canada, 1983), 537–66.

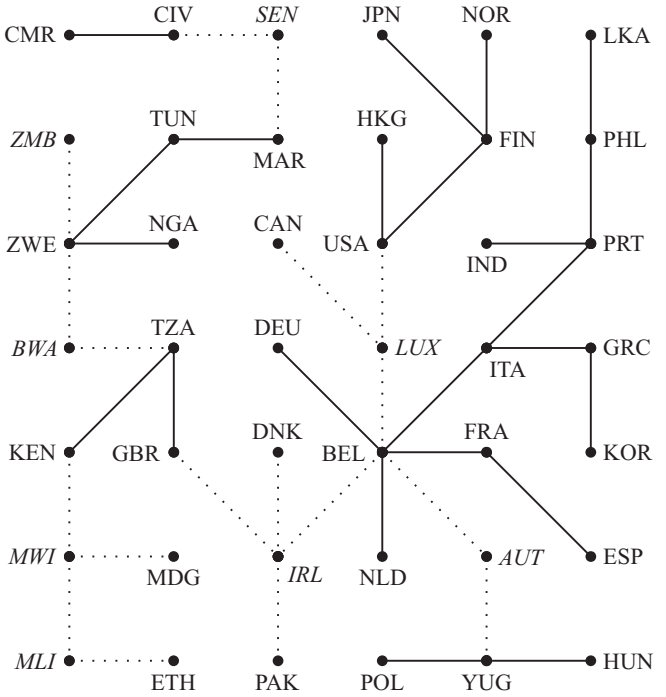
FIGURE 1.—MINIMUM-SPANNING TREES



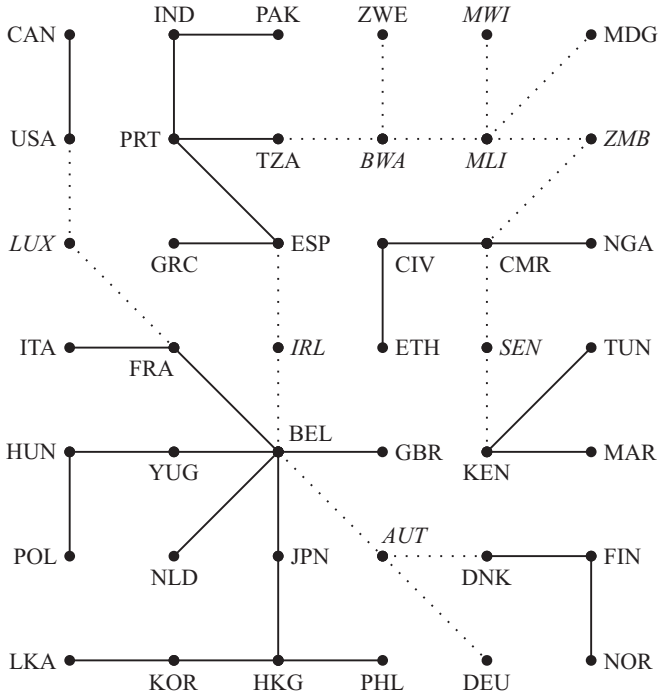
(a) 34-country MST in 1980 in relation to its 30-country counterpart



(b) 34-country MST in 1985 in relation to its 30-country counterpart



(c) 42-country MST in 1980 in relation to its 34-country counterpart



(d) 42-country MST in 1985 in relation to its 34-country counterpart