Spatial Chaining Methods for International Comparisons of Prices and Real Expenditures

D.S. Prasada Rao
The University of Queensland

Jointly with
Robert Hill, Sriram Shankar and Reza Hajargasht

PPPs from ICP 2011

Country	Exch. Rate US$	PPP	PLI% (World=100)
P.R. China	6.461	3.506	70.0
Hong Kong	7.784	5.462	90.5
India	46.67	15.109	41.7
Australia	0.969	1.511	201
Japan	79.809	107.454	173.6
Luxembourg	0.719	0.906	162.4
Ethiopia	16.899	4.919	37.5

Source: World Bank, 2014, Results from ICP 2011.

Real and Nominal per capita GDP
(in US dollars)

Country	Real GDP 2005	Real GDP 2011	Nominal GDP 2005	Nominal GDP 2011
P.R. China	4,091	13,495	1,721	7,321
Hong Kong	36,680	50,129	26,094	35,173
India	2,126	4,735	707	1,533
Australia	32,798	42,000	37,056	65,464
Japan	30,290	34,262	35,604	46,131
Luxembourg	70,014	88,670	80,315	115,689
Ethiopia	591	1,214	154	353
World GDP	54,975 (bill)	90,646 (bill)	44308 (bill)	70,294 (bill)

Source: World Bank, Results from ICP 2005 and 2011.

Objectives

- **Refocus on spatial chaining methods for international comparisons.**
  - **Minimum Spanning Tree (MST) linked comparisons**
  - **Shortest path (SP) chained comparisons**
- **Explore links between spatial chaining and the methods currently in use:**
  - **Equivalence of Weighted GEKS and MST comparisons**
  - **MST as a limiting case of Weighted GEKS**
  - **MST linked comparisons and CPD based comparisons**
- **Choice of a similarity measure**
  - **Laspeyres-Paasche spread**
  - **Weighted relative price dissimilarity (WRPD) measure**
  - **Allen-Diewert measure**

Objectives - continued

- To improve upon the method of minimum spanning trees for determining the links
  - Spanning trees are generally unstable
  - Links obtained are not necessarily intuitive
  - The Hill method does not necessarily give the best possible binary comparisons
- In this paper we introduce the notion of shortest path comparisons between pairs of countries
  - Implement the new concept using different measures of reliability
  - Examine the differences in the links between MST and Shortest path (SP) methods
- We establish a link between weighted GEKS and MST and SP methods of linking
  - We establish algebraic equivalence between MST comparisons and weighted GEKS

GEKS

- **The International Comparison Program makes use of Gini-Elteto-Koves-Szulc (GEKS) method for purpose of aggregating price data and making international comparisons.**
- **The GEKS formula is built on the basis of binary Fisher index numbers using the following formula.**

$$P_{jk}^{GEKS} = \prod_{\ell=1}^{M} \left[P_{j\ell}^{F} \cdot P_{\ell k}^{F} \right]^{1/M}$$

- **GEKS is obtained by solving the following minimization problem:**

$$\min \sum_{j=1}^{M} \sum_{k=1}^{M} \left[\ln P_{jk}^{GEKS} - \ln P_{jk}^{F} \right]^{2}$$

$$subject\ to\ P_{jk}^{GEKS} = P_{jl}^{GEKS} \cdot P_{lk}^{GEKS}$$

Transitivity

Weighted GEKS

- **GEKS is based on the premise that a direct binary comparison is the best way to compare two countries.**
- **GEKS provides transitive comparisons that are the closest to the binary comparisons**
- **Given that ICP covers the whole world - comparisons are sometimes made between countries which are quite dissimilar.**
  - **ICP includes comparisons between USA and Mozambique, and Germany and Laos**
  - **Comparisons between dissimilar countries are intrinsically less reliable and should be given less weight in GEKS.**
- **Weighted GEKS extends the GEKS approach to accommodate dissimilar comparisons. This is given by minimising**

$$\min \sum_{j=1}^{M} \sum_{k=1}^{M} w_{jk} \left[\ln P_{jk}^{GEKS} - \ln P_{jk}^{F} \right]^2$$

Choosing weights

The following properties are expected of the weights:

$$1.\ w_{ii} = 0 \qquad 2.\ w_{jk} = w_{kj} \qquad 3.\ 0 \le w_{jk} \le 1$$

$$4.\ w_{jk} = 1 \quad p_{ki} = \lambda p_{ji}\ \forall i \qquad 5.\ If\ p_{ki} \neq \lambda p_{ji}\ then\ w_{jk} < 1$$

We construct weights using three different measures of similarity:

Laspeyres-Paasche Spread $D^1_{jk} = \left| \ln \left(\frac{P^L_{jk}}{P^F_{jk}} \right) \right|$ $\qquad$ *weights* $= w_{jk} = \frac{1}{1 + d_{jk}}$

Diewert (2009) WPRD $D^2_{jk} = \sum_{n=1}^{N} \left[\left(\frac{s_{j,n} + s_{k,n}}{2} \right) \left(\frac{p_{k,n}}{P^F_{jk} \times p_{j,n}} + \frac{P^F_{jk} \times p_{j,n}}{p_{k,n}} - 2 \right) \right]$

Allen – Diewert measure $D^3_{jk} = \sum_{n=1}^{N} \left\{ \left(\frac{s_{j,n} + s_{k,n}}{2} \right) \left[\left(\frac{p_{k,n}}{P^F_{jk} \times p_{j,n}} - 1 \right)^2 + \left(\frac{P^F_{jk} \times p_{j,n}}{p_{k,n}} - 1 \right)^2 \right] \right\}$

Spatial Chaining

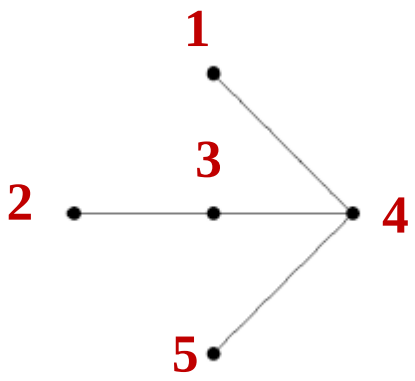
- **For temporal comparisons, we have a natural order to chain comparisons**

  2010 → 2011 → 2012 → 2013 → 2014 → 2015

- **Spatial chaining is where countries or regions are compared with other countries using chained links**
  - **In spatial comparisons, there is no natural ordering**
- **How does one order the countries to determine the chains?**
- **Question then is whether it is possible to devise a method of finding spatial chains to making comparisons between countries.**
- **Hill (1999, 2001, 2004, 2009) advocated the use of spatial linking based on minimum spanning trees.**
  - **Spanning tree is a concept used in *Graph Theory***
  - **Spanning tree provides an order which countries can be linked.**

Price comparisons using a Spanning tree

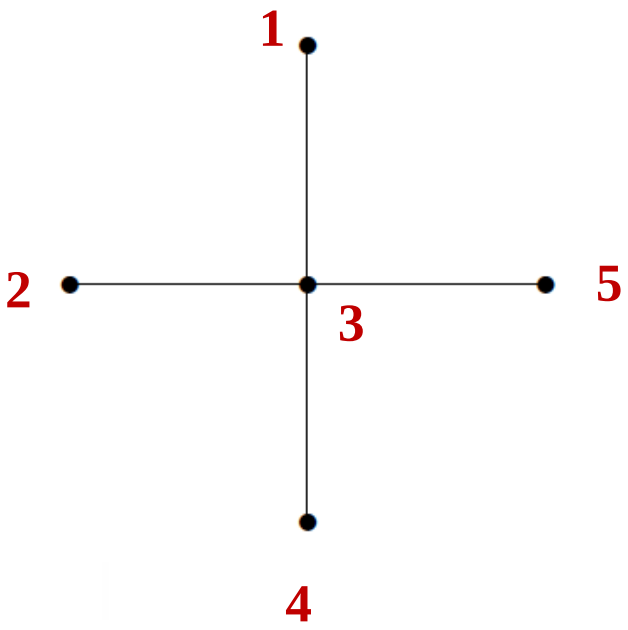
- **First we choose a binary index number that satisfies time/country reversal test – e.g., Fisher and Tornqvist.**
- **A spanning tree is a connected graph where there is an unique path between any pair of countries.**
- **Suppose we wish to use the following spanning tree for a set five countries.**

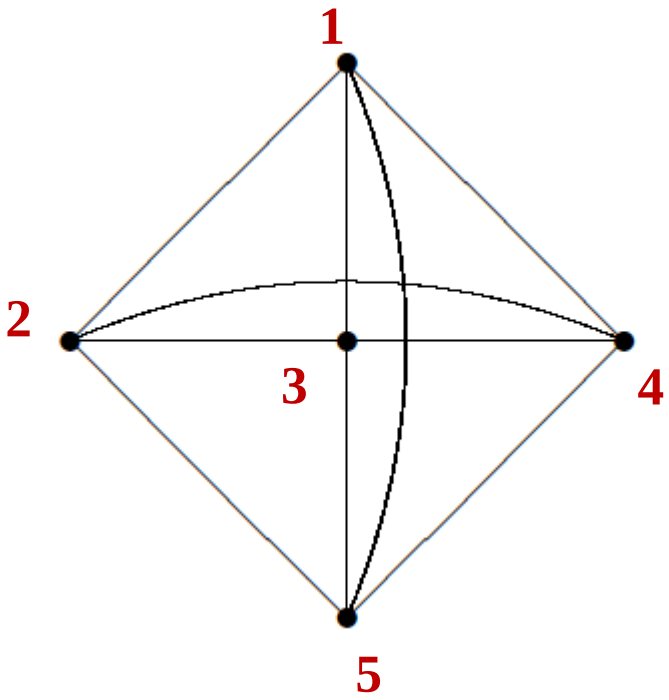


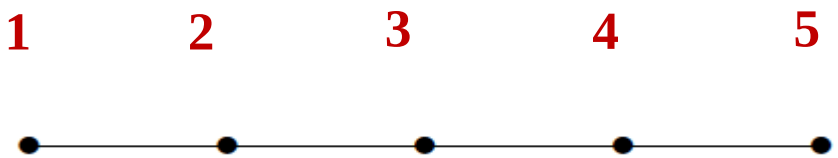
- **The comparisons between countries are made using the chains shown in the spanning tree.**

$$P_{12}^{ST}(F) = P_{14}^{F} \cdot P_{43}^{F} \cdot P_{32}^{F} \qquad P_{15}^{ST}(F) = P_{14}^{F} \cdot P_{45}^{F} \qquad P_{35}^{ST}(F) = P_{34}^{F} \cdot P_{45}^{F}$$

Price comparisons using a Spanning tree







Weighted GEKS and MST Price comparisons

We prove the following two theorems:

Theorem 1: Consider a spanning tree that connects all the countries. Let w_{jk} represent weights such that $w_{jk} = 1$ if country *j* is directly connected to country *k* and zero otherwise. Then price comparisons based on the MST are identical to the indexes obtained using weighted GEKS method with weights w_{jk} implied by the MST - can be proved using induction.

Theorem 2: Consider the following system of generalized weights

$$\bar{w}_{jk}^{a} = \frac{(\bar{w}_{jk}^{a})^{x}}{\max_{j,k=1}^{J}[(\bar{w}_{jk}^{a})^{x}]}, \qquad x \geq 0,$$

In the limit as *x* tends to infinity, the weighted-GEKS method converges to the minimum spanning tree method

Spatial chaining and CPD

- **When it comes to spatial chaining the following question is often raised:**

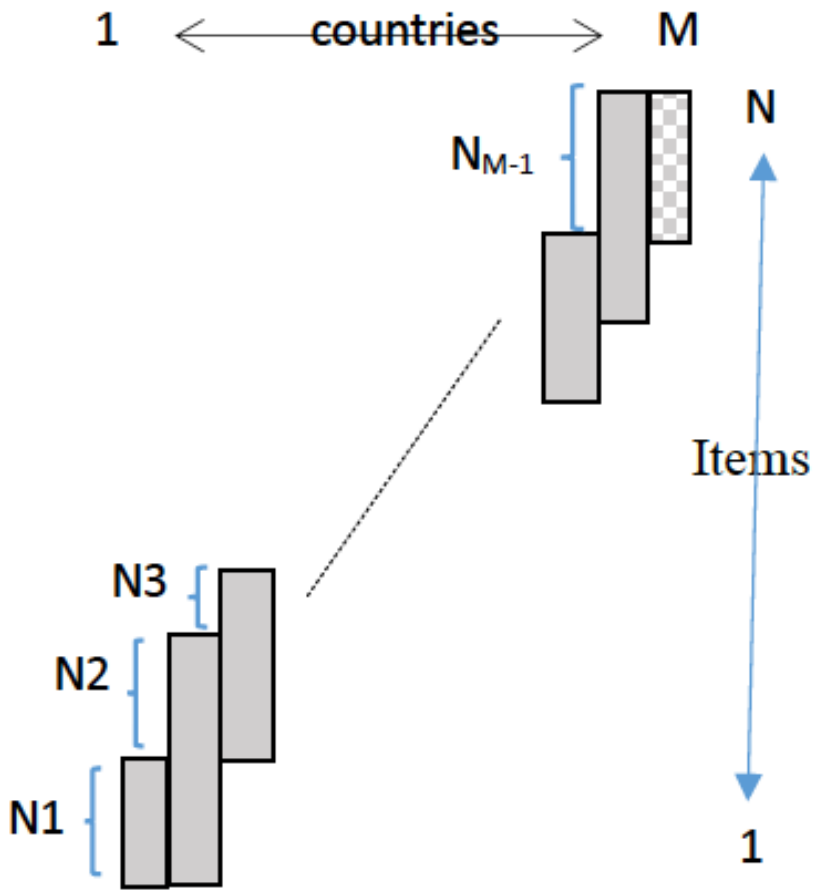
  Is it meaningful to obtain spatially chained comparisons between two countries that have no commodities that are commonly consumed?

The answer to this is that it is not meaningful to use spatial chaining.

- **We prove the following theorem which establishes that comparisons based on spatial chaining are identical to those obtained using the Country-Product-Dummy (CPD) method which is the currently accepted method.**

Spatial chaining and CPD

We consider the following scenario:



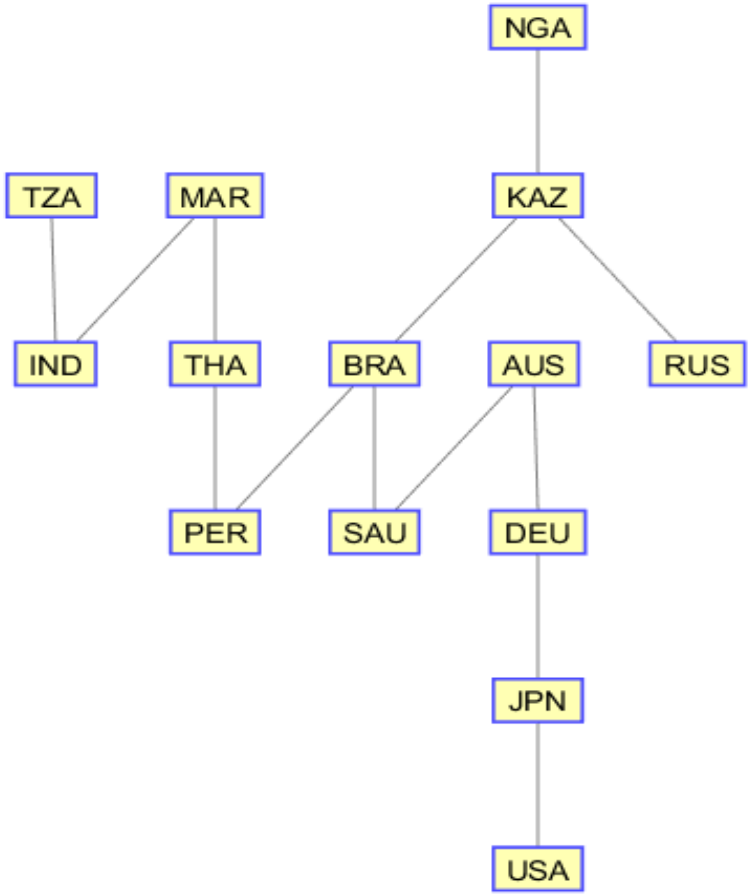
Theorem: The PPPs computed for this data matrix using CPD method and spatial chaining are identical.

Proof uses the structure of data and the algebraic derivation of PPPs using the CPD method

Which spanning trees to use?

- For a given set of *M* countries, there can be M^{M-2} number of spanning trees that can be used. For example, if there are five countries, there can be 125 different spanning trees.
- Which spanning tree should we choose?
  - Hill (1999) and subsequent work advocates the use of *minimum spanning tree* (MST) for price comparisons.
- To identify the minimum spanning tree, we need to associate weights to each binary comparison. This is like a measure of cost associated with the comparison.
- In rest of this work, we make use of the three distance measures described before – LPS; WPRD and Allen-Diewert measures.
  - The minimum spanning tree is identified using Kruskal's algorithm.

Minimum Spanning tree - example



Shortest Path (SP) Approach

- **Main starting point is that MST may actually make some comparison worse than the original binaries.**
- **The shortest path between a pair of countries *j* and *k* is defined here as the path with the minimum sum of weights**
  - **In principle, the SP approach identifies the best possible comparison between any pair of countries.**
- **The basic approach is to choose the shortest path among paths of length 1, 2,…,M-1.**
- **What distance metric do we use?**

Which distance metric do we choose?

- **Choice of distance metric for computing shortest paths is not equivalent to the choice of distance metric for spanning tree.**
  - **In the case of minimum spanning tree all that matters is the ordinal ranking of edges.**
- **In the case of shortest paths, the metric has to be economically meaningful to sum the distance metric along a chain path**
  - **We provide two theoretical results that narrow our choice to the use of LPS and the WPRD metrics.**

$$E[\ln(P^L_{jk}/P^P_{jk})] + E[\ln(P^L_{kl}/P^P_{kl})] < E[\ln(P^L_{jl}/P^P_{jl}) \Rightarrow var(\ln P_{jk} + \ln P_{kl}) < var(\ln P_{jl})$$

$$E[\ln(P^L_{jk}/P^P_{jk})] + E[\ln(P^L_{kl}/P^P_{kl})] < E[\ln(P^L_{jl}/P^P_{jl})]$$

$$\Leftrightarrow E(D^{WRPD}_{jk}) + E(D^{WRPD}_{kl}) + \sigma_{jl} - (\sigma_{jk} + \sigma_{kl}) < E(D^{WRPD}_{jl})$$

Shortest Path (SP) Approach

- **If the MD path between two countries *j* and *k* is defined by countries with labels {$i_1, i_2, \ldots, i_P$ }, then**

$$MD_{jk}(Fisher) = F_{j,i_1} \cdot \left[\prod_{l=1}^{P-1} F_{i_l, i_{l+1}} \right] . F_{i_P, k}$$

- **Properties:**

1. $d_{SP}(x_j, x_k) \leq d_{MST}(x_j, x_k)$ *for all* j *and* k

2. $d_{SP}(x_j, x_k) \leq d(x_j, x_k) \forall j, k$

3. $d_{SP}(x_j, x_k)$ **is a proper distance metric**

4. **The SP chained index is not transitive – by construction. So we can use GEKS on the SP index.**

5. **The Shortest Paths are identified using Dijkstra algorithm – this identifies minimum paths for all countries starting from a given source country.**

SP Approach – Some analytical Results

- **Shortest paths from a given country to all the other countries combined together form a spanning tree.**
- **This means we can consider SP spanning tree (SPST) for each country**
- **SPST from each origin country can be different.**
- **Shortest path based binary comparisons are <u>not transitive</u>**
- **Since the shortest path comparisons provide the best binary comparisons, we can use GEKS on the matrix of shortest path binary comparisons – this is referred to as SP GEKS.**

Empirical Results

Data used: ICP 2011 data for Household Consumption using 110 categories

Results:

- **We have results for the full set of 177 countries but it is difficult to present and discuss graphs**
- **We present graphs with results compiled for a selected sub-group of thirteen countries**
- **Countries selected are: Australia; Brazil; Germany; India; Japan; Morocco; Nigeria; Peru; Russia; Saudi Arabia; Thailand; Tanzania; and USA**

Empirical Results

We construct the following set of comparisons:

MST (LPS)

MST (WRPD)

Shortest path GEKS (LPS with L>P)

Shortest path GEKS (WRPD)

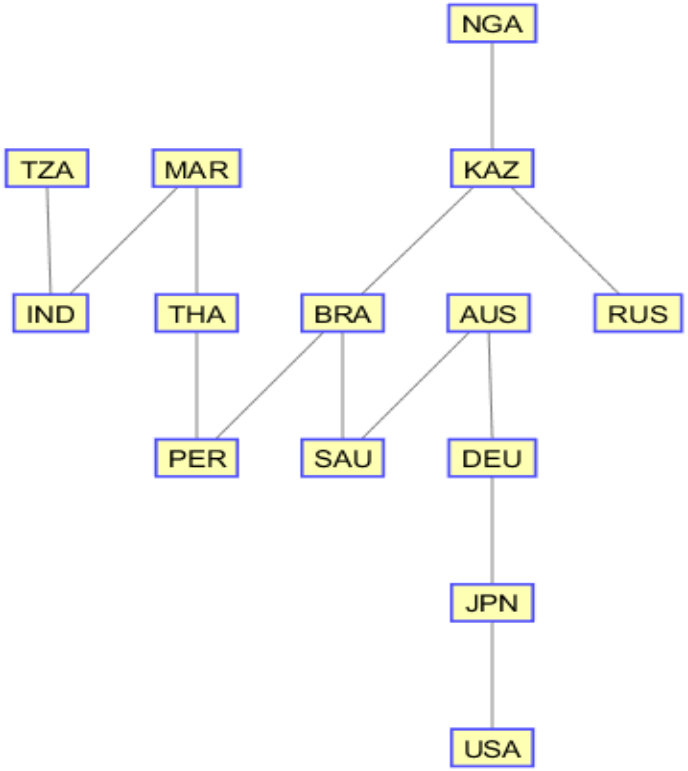
Weighted GEKS (with weights of 1/(1+LPS))

Weighted GEKS (with weights of 1/(1+WRPD))

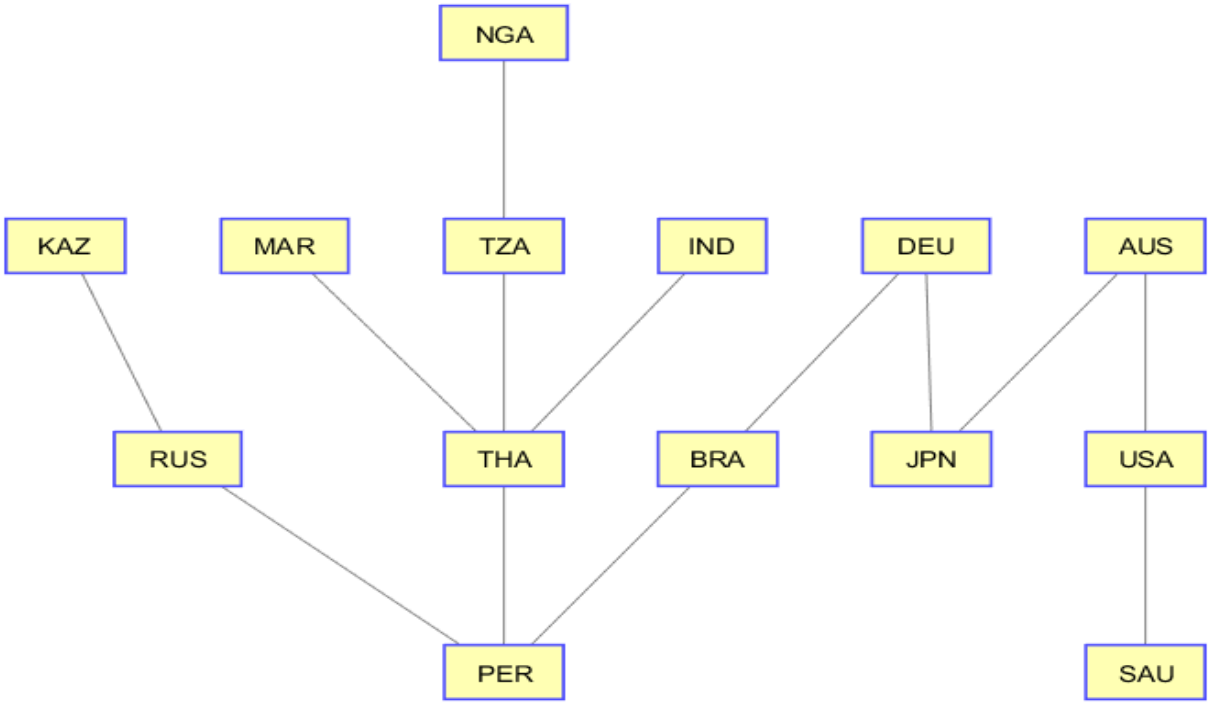
Weighted GEKS (on matrix of ones and zeros derived from union of SPSTs – LPS with L>P)

Weighted GEKS (on matrix of ones and zeros derived from union of SPSTs – WRPD)

MST with LPS distance measure

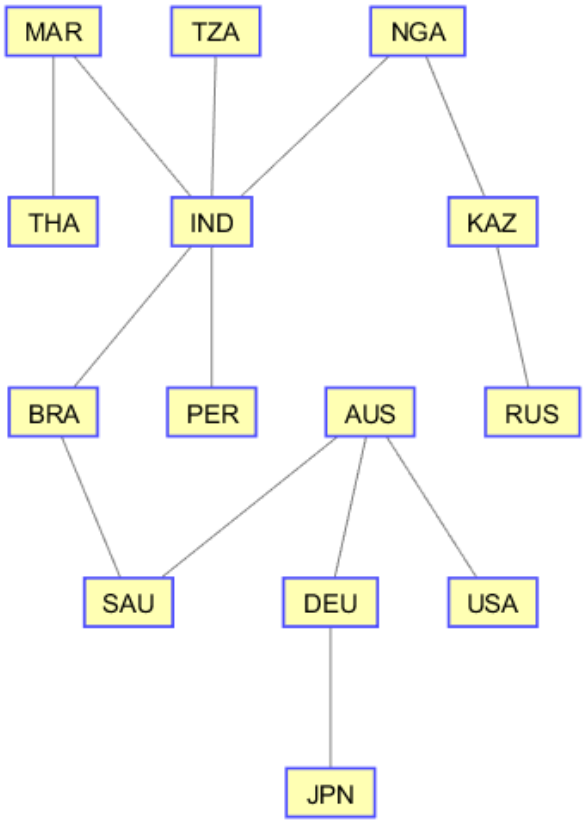


MST with weighted relative price distance measure



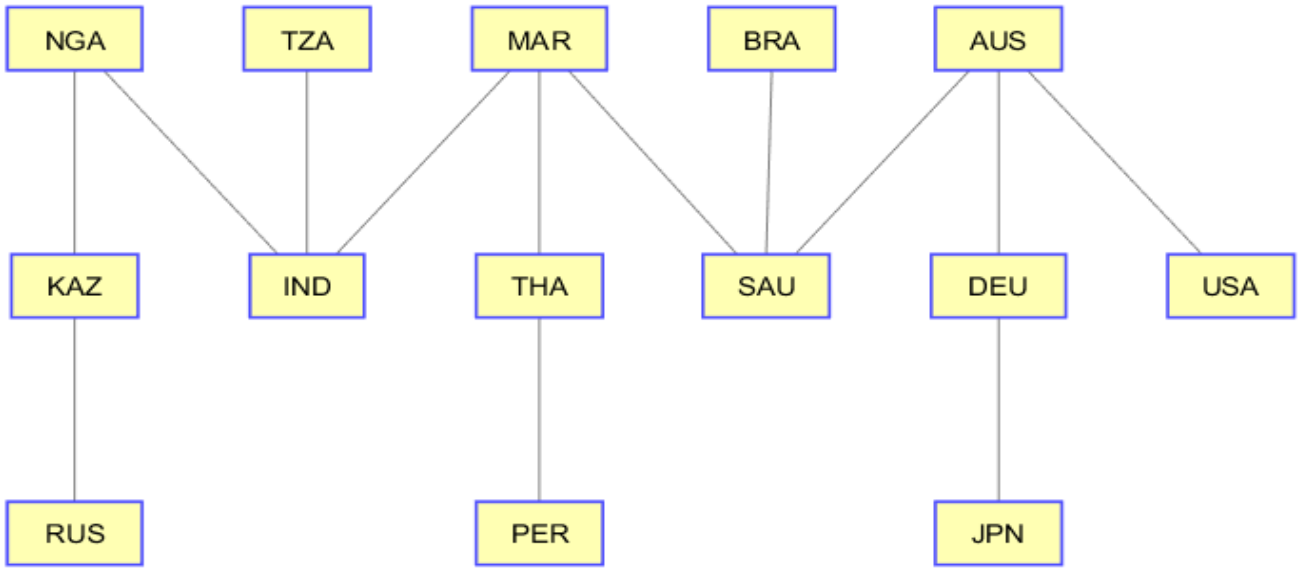
The MD Paths from Selected Countries Using LPS Measure

India with all other countries



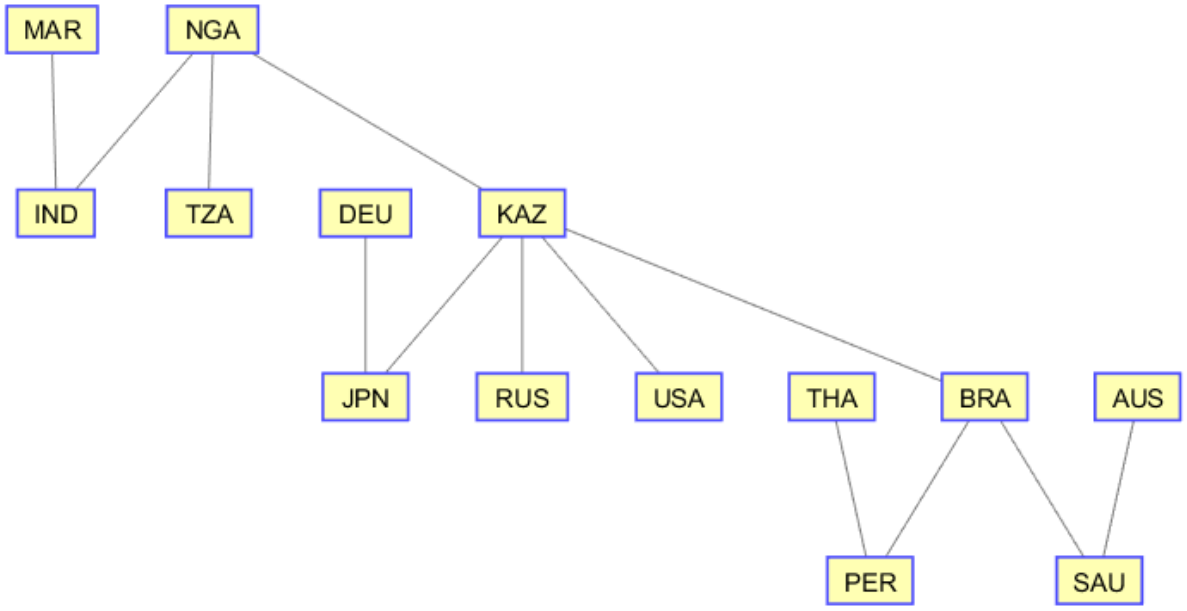
The MD Paths from Selected Countries Using LPS Measure

Morocco with all other countries

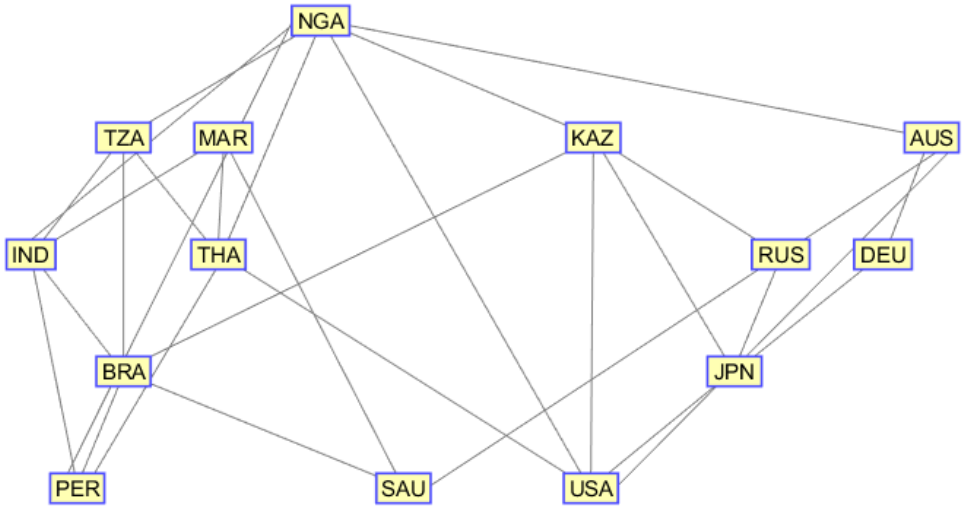


The MD Paths from Selected Countries Using LPS Measure

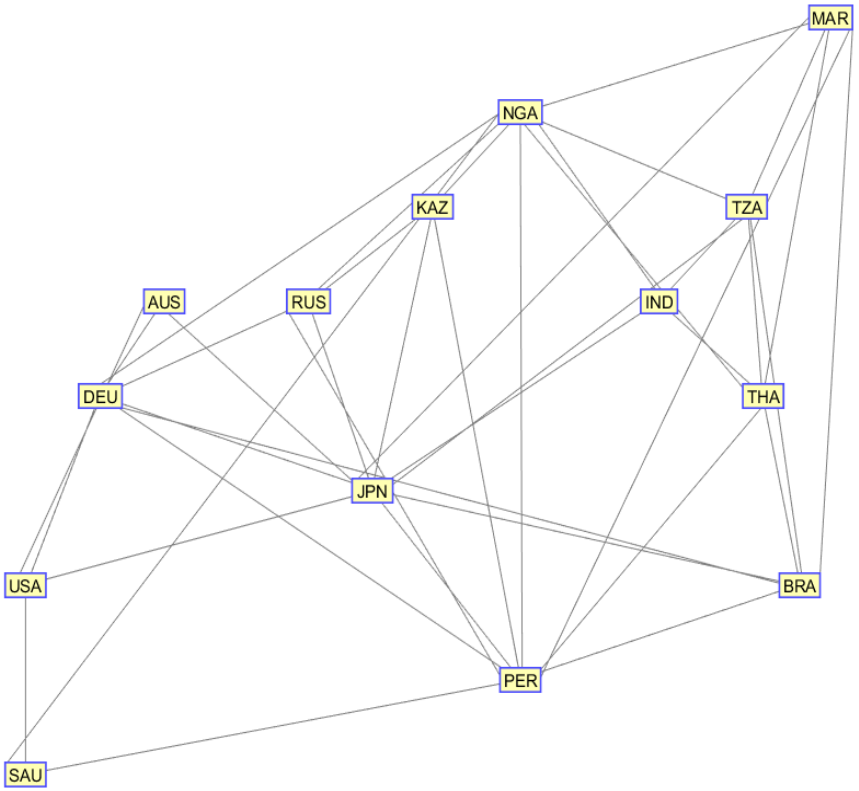
Kazhakistan with all other countries



Union of all Minimum Distance Paths - LPS



Union of all Minimum Distance Paths - WPRD



Comparisons with LPS

	Total within region comparisons	Shortest path without external countries	MST without external countries
Africa	1225	83	31
Asia_Pacific	253	17	7
CIS	36	11	5
EU_OECD	1035	57	22
Latin America	120	24	6
West Asia	55	3	2
Singleton	0	0	0

Comparisons with WRPD

	Total within region comparisons	Shortest path without external countries	MST without external countries
Africa	1225	565	43
Asia_Pacific	253	97	20
CIS	36	17	5
EU_OECD	1035	328	45
Latin America	120	76	13
West Asia	55	19	8
Singleton	0	0	0

Comparisons with LPS

Country	PPP LPS MD	PPP LPS SP GEKS	PPP LPS MST	PPP LPS MST WGEKS	PPP LPS SP WGEKS	Weighted GEKS (with weights of 1/(1+LPS))	
CHN	3.481	3.522	3.680	3.680	3.669	3.443	3.417
FJI	1.114	1.031	1.059	1.059	1.054	1.038	1.033
HKG	5.608	5.541	5.923	5.923	5.686	5.505	5.486
IND	13.590	13.430	14.370	14.370	14.490	14.753	14.632
IDN	3848.043	3558.610	3658.291	3658.291	3537.608	3523.292	3507.231
LAO	2117.799	2127.555	2340.690	2340.690	2372.286	2335.314	2322.028
MAC	5.394	5.418	5.496	5.496	5.436	4.858	4.968
MYS	1.554	1.501	1.535	1.535	1.536	1.463	1.455

Comparisons with WPRD

Country	PPP WRPD MD	PPP WRPD SP GEKS	PPP WRPD MST	PPP WRPD MST WGEKS	PPP WRPD SP WGEKS	Weighted GEKS (with weights of 1/(1+WRPD))	
CHN	3.916	3.430	2.602	2.602	3.425	3.469	3.417
FJI	1.009	0.984	0.785	0.785	1.014	1.042	1.033
HKG	5.689	5.116	3.780	3.780	5.669	5.547	5.486
IND	12.079	13.186	10.154	10.154	14.064	14.683	14.632
IDN	4260.520	3553.017	2711.796	2711.796	3502.607	3543.208	3507.231
LAO	2059.187	2178.207	1766.702	1766.702	2270.438	2325.560	2322.028
MAC	5.384	4.860	3.577	3.577	5.099	5.058	4.968
MYS	1.713	1.450	1.104	1.104	1.453	1.468	1.455

Robustness of comparisons
Various methods

We use Jack-Knife method to assess stability of comparisons from various methods. Results are reported below

	PPP LPS SP	PPP LPS SP GEKS	PPP LPS MST	PPP LPS MST WGEKS	PPP LPS SP WGEKS
CHN	0.2730	0.1650	0.7836	0.7836	0.0820
FJI	0.1898	0.0455	0.2655	0.2655	0.0278
HKG	0.2536	0.1968	1.5271	1.5271	0.1711
IND	2.0817	0.9498	5.7680	5.7680	0.4384
IDN	757.2862	184.8605	919.6091	919.6091	243.8574
LAO	281.2256	151.5773	865.4435	865.4435	89.4408
MAC	0.2417	0.1962	1.1435	1.1435	0.2782
MYS	0.0658	0.0470	0.3034	0.3034	0.0288

We are currently conducting simulation studies to assess the performance of various methods in the presence of noise in price data.

Conclusions

Spatial chaining is shown to be a promising area for research.

- **The SP approach provides better links between pairs of countries than the MST.**
- **The SP links are more stable than the MST links.**
- **We are able to provide a link between spanning tree comparisons and weighted GEKS methods.**
- **Of all the distance and similarity measures we find LPS and WPRD to be conceptually suitable for the SP approach.**
- **We are currently conducting a simulation study to assess the robustness of the SP comparisons in the presence of noise in the price and expenditure observations.**
- **Given the stability of shortest path chains between countries, it may be feasible to redesign price collection strategies that strengthen international comparisons.**

Thank you!