Multidimensional Scalling Analysis of the World Economy During the Period 1972-2012

J. A. Tenreiro Machado [1], Maria Eugénia Mata [2]

[1] Coordinator Professor, Department of Electrical Engineering, Institute of Engineering, Polytechnic of Porto, Porto, Portugal, jtm@isep.ipp.pt

[2] Associate Professor, Nova SBE, Universidade Nova de Lisboa, Faculdade de Economia, Campus de Campolide, Lisbon, Portugal, memata@fe.unl.pt

Abstract: The last 40 years of the world economy are analyzed by means of computer visualization methods. Multidimensional scaling and the hierarchical clustering tree techniques are used. The current Western downturn in favor of Asian partners may still be reversed in the coming decades

Keywords: Economic development; Multi-dimensional scaling; global hegemony

1 Introduction

The Western Golden Age of economic growth that lasted from the end of the Second World War to the early 1970s consecrated the global hegemony of the United States [21]. Following two victories in the two World Wars without suffering any destruction within the nation's territory, the United States' role as a model of modern technologies and venture capital corporations became quite clear on the international scene [23]. Superior management abilities drove capitalistic professionalism, technological superiority, considerable growth rates, and a global upper hand [29, 26].

The macroeconomic environment of the early 1970s was not considered to be especially favorable to investment, according to venture capital experts [22]. A less careful approach may consider the 1974 oil price increase as the historical turning-point for a long-run downward trend. In spite of the sizable amount of America's domestic oil production, challenges to the industry in the non-Arab world may be cited. However, the OPEC cartelization undertaken to increase oil prices was justified as a defense against the dollar depreciation, which eroded the value of oil exports from the Arab producers. The dollar depreciation was the direct consequence of the end of the Bretton Woods monetary system, bringing to an end the 1944 commitment to fixed exchange rates [18]. The extraordinary

military expenditures associated with the Vietnam War had led to a dollar glut and the need for the inconvertibility that was consecrated in the Bretton Woods system.

From then until today Europe and America have faced several recoveries and recessions, along with other important global changes that occurred in the 1980s, owing to the breakup of the Soviet Union and the end of certain communist regimes around the world, including China [14]. The expansion of corporate enterprises into those new areas brought positive capitalism and business opportunities [17]. The unification of Germany fueled hopes for a stronger European Union. A common currency, the euro, intensified European cohesion at the turn of the millennium, which was also a hopeful aspect for Western economic growth in the 1990s.

The new millennium ushered in fierce international rivalry among the largest international partners [13]. Capitalist Russia tends to control neighboring areas, threatening European tranquility. India became a talented front-runner in IT and service sectors, China welcomed large global corporations and began a rampant industrialization that floods all world markets with inexpensive convenience goods. Other Asian partners became small tigers with modern and fashionable competitive economies (Taiwan, South Korea, Indonesia, and Singapore) [7]. They all have been able to cultivate the ranks of engineers, statisticians, computer experts, chemists, and other qualified professionals that are required in modern economies and societies [27]. Thanks to generous educational programs and government policies, Asia is favoring tertiary enrollment. Nothing comparable is occuring in Africa [2].

In the 2007 crisis, the Nobel laureate Robert Fogel made the forecast that in 2040 the existence of democracy will not depend on the usual world democratic bastions, the EU, the U.S., and Japan, but will depend rather on the new global hegemony of Asian partners [15]. The aim of this paper is not to discuss democracy, an issue that is better left to political science and political economy, but to observe the convergence process, using Multi-Dimensional Scaling (MDS) methodologies, to chart the path of these partners throughout the last four decades of Western decline (1972-2012).

Four decisive World Bank social-economic indicators were selected to examine the comparative evolution of a dozen countries' processes, and describe them until 2012 (as no data are yet available for 2013). Data and methodology are discussed in Section 2. We divide the period into two twenty-year groups, and also look at the whole forty-year process. Section 3 discusses the results obtained on economic convergence, and includes many more countries [25, 20]. Section 4 summarizes the main conclusions.

By increasing "the ultimate potential of the economy", convergence may lead to global prosperity (Current World Bank calculations say that "the number of people living below the $1.25 a day line plunges from 1.2bn people in the

developing world to fewer than 600m") [16]. However, it may also bring "potential warfare economic strategies", because war may compensate victors with greater territory and more resources, thereby increasing their production possibilities frontiers [23].

Our conclusions may not be altogether surprising, but they demonstrate that the world economic system is following a path of convergence on which the Asian partners still have a way to run, while Russia is looking very much more like the most developed countries.

2 Methodology and Data

Economic growth means new and better conditions of life at per capita levels, diversified consumption that is reflected in foreign trade openness (increasing imports and exports), as well as increasing population longevity, translated into higher life expectancy, while youth education becomes a driver for equal social opportunities. As education endogenously provides the human capital that is required in more and more sophisticated technological systems, this is a decisive long-run indicator for the ultimate potential of a national economy (for life expectancy see [1]). A stock variable such as schooling years for citizens above 25 (or 15) years of age may be the most representative, as it is more revealing about labor markets and production efficiency, but flow variables on education are longer and more complete.

Data were collected from the World Bank national development indicators covering the period from 1972 to the present (2012), sourced in [3]. The series are homogeneous, and the tertiary education level is the variable that most relates human capital with technological achievement throughout the period analyzed. To discuss the similarities among the 12 countries, the potential economic growth, and global hegemony of the main world powers, the selected indicators are GDP per capita, the weight of international trade in GDP, life expectancy at birth, and tertiary enrollment. Moreover, the human development index concept, based on indicators for human development, also recommends these same indicators [24].

The 12 selected countries (representing all continents with the exception of Africa) are {BRA, CHN, DEU, FRA, GBR, HUN, IND, ITA, JPN, PRT, RUS, USA} ≡ {Brazil, China, Germany, France, UK, Hungary, India, Italy, Japan, Portugal, Russian Federation, US}.

Here is the complete description of the four indicators:

- GDP per capita, comes from NY.GNP.PCAP.KD. It is the GNI per capita (constant 2000 USD$). The preference for per capita values has to do with the need to observe countries with very different economic

dimensions. Partners such as China, Russia, and India are much larger economies in which larger populations contribute to the GDP. The per capita indicator makes comparisons more plausible, in spite of the many differences that international comparisons always involve (e.g. climate, culture, natural resources, politics). GNI per capita is gross national income divided by midyear population. GNI (formerly GNP) is the sum of value added by all resident producers plus any product taxes (less subsidies) not included in the valuation of output plus net receipts of primary income (compensation of employees and property income) from abroad. Data are in constant 2000 U.S. dollars.

- Annual exports of goods and services (percent of GDP), from NE.EXP.GNFS.ZS. Exports of goods and services represent the value of all goods and other market services provided to the rest of the world. They include the value of merchandise, freight, insurance, transport, travel, royalties, license fees, and other services, such as communication, construction, financial, information, business, personal, and government services. They exclude compensation of employees and investment income (formerly called factor services) and transfer payments. It is a weighted average. Data are expressed as a percentage of GDP.
- Life expectancy comes from SP.DYN.LE00.IN. Life expectancy at birth indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.
- School enrollment, tertiary (gross percent) comes from SE.TER.ENRR. Gross enrollment ratio is the ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to the level of education shown. Tertiary education, whether or not to an advanced research qualification, normally requires, as a minimum condition of admission, the successful completion of education at the secondary level. Data are expressed as a percentage [4].

The time series for tertiary enrollment is the least dense, but missing observations occur in all available education time series. Barro & Lee's (2012) database for 1950-2010 is an excellent source for education, but the information is only given for every five years [12]. The UNESCO education database has yearly information, but covers only 1970-1997, with missing points.

MDS is a statistical, computational visualization methodology that examines similarity between objects through distance estimations, by means of graphical relatedness representation [10]. Similarity means the "likeness" between the objects in a mathematical degree [19].

Any MDS representation of n objects departs from estimating a $n \times n$ symmetrical matrix **R** for cross correlation measures between all objects. In the MDS map

objects are points and they represent the data vectors that describe the selected variables of the exercise. If the correlation between two objects is one, this means that the distance between them is zero. A zero correlation means an infinite distance, and the corresponding MDS points are graphically very far apart. If two points almost coincide there is a large correlation, while if they are located far away in the MDS plot, the correlation between the corresponding data vectors is small [8].

In order to estimate the coordinates of the points in an *m*-dimensional plot, the MDS numerically optimizes the estimation of the distance between all pairs of objects, according to matrix **R** of correlations [9]. The use of $m = 2$ or $m = 3$ is very frequent, because it represents a direct graphical visualization of the MDS map in 2 or 3 dimensions, respectively.

In this paper the country economies are objects characterized by 4 variables performing during 40 years. The main diagonal of the **R** symmetric matrix $\mathbf{R} = [r_{ij}]$ is composed of ones/zeros for correlation/distances, while the other elements of the matrix must be described as $0 \le r_{ij} \le 1$, $r_{ij} \ge 0$, $i,j = 1,..., n$. MDS maps are not sensitive to translations or rotations because the method uses relative measurements. Axes have only the meaning and units (if any) of the measuring index. Objects are rearranged so that the map obtained can best approximate the similarities that exist. For example, study [28] applies it to genomic datasets. The accuracy of the MDS solution is measured using the raw stress. The smaller its value, the more accurate is the fit. If stress is plotted versus the m dimensions of the MDS map, a monotonic decreasing chart is obtained. Users may choose the best dimension as a trade-off between lower stress levels and the dimensions for the map representation.

Some literature refers to MDS as a statistical tool, while other authors mention it as a computer representation scheme. The important thing is that it can provide a means of visualizing items without *a priori* restrictions or any additional constraints. Of course different indices produce different maps, because MDS maps reflect the measures for similarity. The great value of MDS plots is the direct visualization of results they provide.

Time dynamics analysis may require the explicit division of the total time period into several sub-periods of width *h* to be considered by MDS as independent objects. For a total time period *T*, $p = T/h$ samples are produced and the number of MDS points increases proportionally in the plots. The time samples to be adopted are a compromise between the possibility of understanding fast dynamics (with small values of h, meaning many sub-periods), and the advantage of plotting a limited number of MDS points (which requires large values of *h*). As time series cover a period $T = 40$ years (1972-2012), two cases are developed, namely a comparison based on the whole period of time, that is $h = 40$ ($p = 1$) and a division into two sub-periods of 20 years each $h = 20$ ($p = 2$) for 1972-1992 and 1992-2012. We obtain $n = n_c \times p$ objects to be analyzed in the MDS of h years length

each. Therefore, the two cases consist of $p = 2$, $h = 20$, $n = 40$, and $p = 1$, $h = 20$, $n = 20$. Points labeled as "USA1" mean USA during the first twenty-year period (1972-1992), and JPN2 means Japan during the second twenty-year period (1992-2012). Of course $l = 4$ economic variables {GDP per capita, openness, life expectancy, tertiary education enrollment} are adopted for each country economy, having identical weights.

The data set of the 4 economic variables had missing values. Some countries had some years without data values, which had to be estimated by means of linear interpolation between adjacent years (in particular the Russian Federation, for all variables). Germany had many missing values for the tertiary education enrollment. An iterative procedure was adopted to avoid anomalous values: (*i*) all variables' time series were plotted, for the whole set of countries (*ii*) a non-linear individual trend line was estimated for each country, (*iii*) the values emerging from a given trend line were compared with the succeeding values (in order to test the consistency with the real values) and with the values of the remaining countries (to test the relative positions), and (*iv*) the trend line was accepted or replaced by another one, depending on whether the visual comparison did or did not confirm. This was a time-consuming procedure, but it prevented the acceptance of misleading values. The practical implementation of this careful approach proved to be faster than expected because of the smooth path of the variables, which tend to evolve similarly. Different trend lines have produced values relatively close and, therefore, the estimation error is not significant.

Two methods can be used to construct the matrix $\mathbf{R} = [r_{ij}]$, the cosine correlation and the Euclidean distance, defined as:

$$r_{ij} = \frac{\sum_{t=1}^{h}\sum_{k=1}^{l} x_i(k,t)x_j(k,t)}{\sqrt{\sum_{t=1}^{h}\sum_{k=1}^{l} x_i^2(k,t)\cdot\sum_{t=1}^{h}\sum_{k=1}^{l} x_j^2(k,t)}}, \quad i,j = 1,\cdots,n \tag{1}$$

$$r_{ij} = \sum_{t=1}^{h}\sum_{k=1}^{l}\left[x_i(k,t) - x_j(k,t)\right]^2, \quad i,j = 1,\cdots,n \tag{2}$$

where x_i and x_j denote economic variables for the i-th and j-th objects, t and k are two dummy indices for time and type of economic variable, h is the sampling period, and n is the total number of objects. Equation (1), which measures the angle between two vectors, is often called the cosine coefficient because it denotes an angular metric, and is a normalized inner product [11]. Equation (2) describes the Euclidean distance over the time period. Note that expression (1) is not sensitive to the amplitude of the vectors, while expression (2) captures differences between vectors, on both amplitude and direction.

The two-decade approach for describing the relative similarities among countries seems to be successful. The use of shorter periods would allow for more time detail, but plots would contain many more points, and reading would become increasingly difficult. The twenty-year visualization also may have a special economic meaning in capturing any Kondratief business-cycle influence for the behavior of the countries' economies. Expressions (1) and (2) have an implicit embedded description of the time evolution. As a result, the next plots of cases a) and b) differ according to equations (1) and (2), because they reflect time dynamics more or less explicitly.

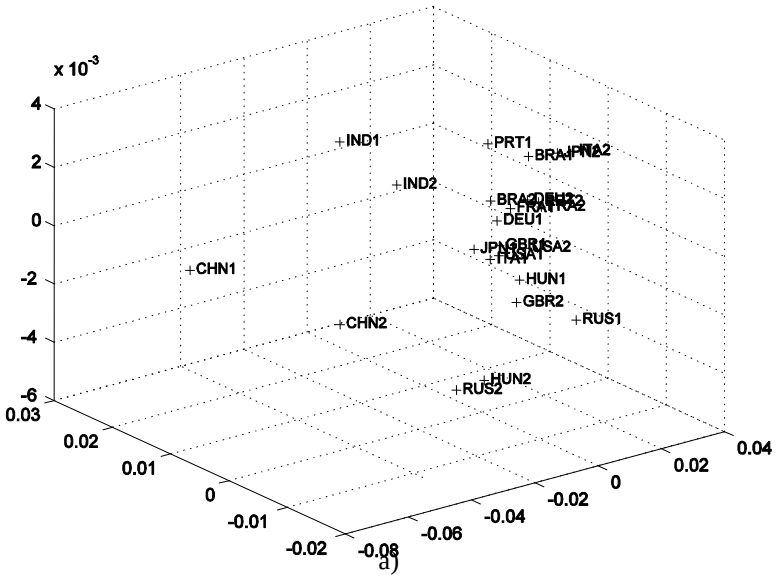
3 Discussing the Estimations

Looking at the cosine correlation maps (Fig. 1a), one can follow countries' evolution along distance vectors, revealing the path of the comparison among the 12 sampled countries during the two twenty-year periods (1972-2012). China, India, and Russia occupy different positions according to the four socio-economic indicators selected, in a three-dimensional view. In spite of their common policies of promoting high economic growth rates, they still remain far from the most developed countries of the world economic system, which are so similar that a real cloud is formed by the Western World set of European countries, and the USA [5].

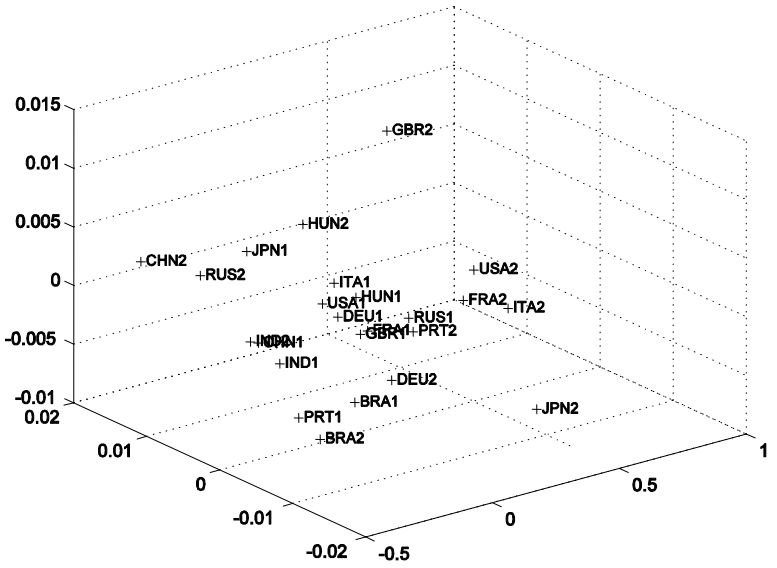
The Euclidian-distance plot (Fig. 1b) shows how China and Russia could perform well in becoming more similar (closer) to Japan and the most developed partners. India also converged, in spite of the social turmoil resulting from the prevailing caste system, "which divides the population into a hereditary hierarchy that determines economic and social opportunities" [15]. This handicap has relaxed somewhat as a result of government policies: "The government has sought to offset the discrimination against lower castes with educational subsidies" [15]. Nevertheless, the caste system remains rigid in rural areas.

The MDS method also allows for a projection in two dimensions (Fig. 2), which confirms that Asian partners still have a long way to go in to become "similar" to the Western partners. Russia and Brazil are the two partners that moved to positions closer to the large cluster of developed countries in the last 20 years. Brazil shows the success of the government policies implemented, as well as the discovery of oil in the Brazilian offshore Atlantic. Russia also followed a move from its 2002 position until now (2012 data), reflecting its openness, as well as oil and gas exports. How to qualify this move?

Using the Euclidean distance instead of the cosine correlation pattern for distance measuring, the same messages can be seen in this different representation. In becoming more similar to the large cluster of developed countries, Russia also converged to a position closer to the Japanese (and Chinese) profile.



a)



b)

Figure 1

Three-dimensional MDS representation with $m = 20$ for the 12 countries and two 20-year periods, based on the 4 selected variables). Using: a): Cosine correlation (1), b) Euclidean distance (2)

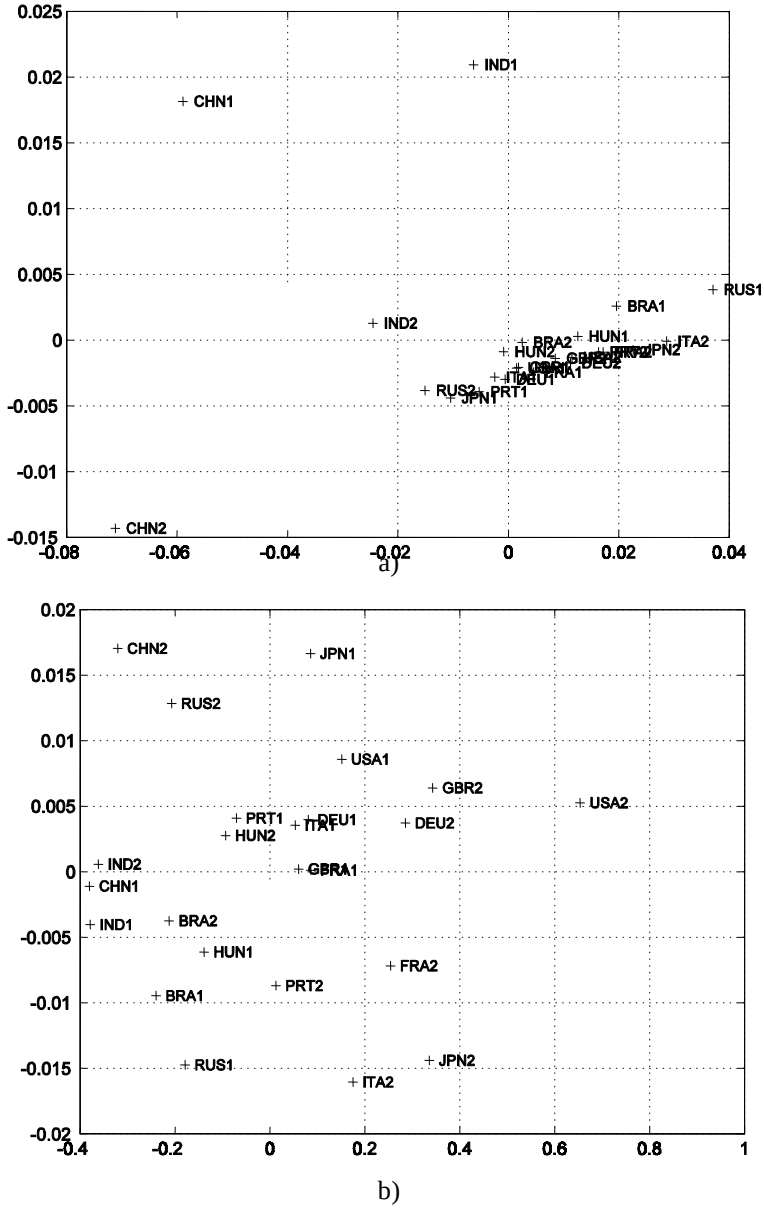
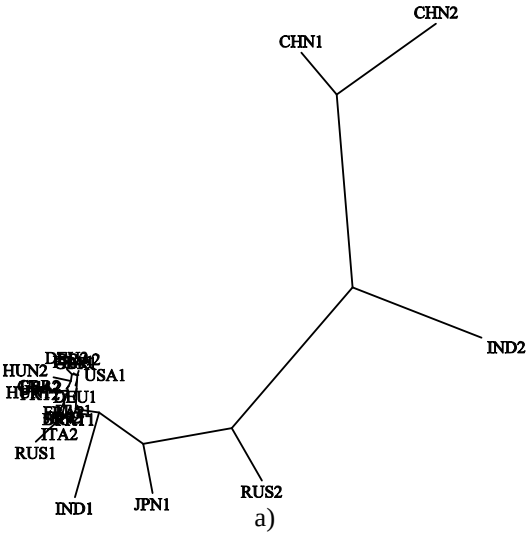
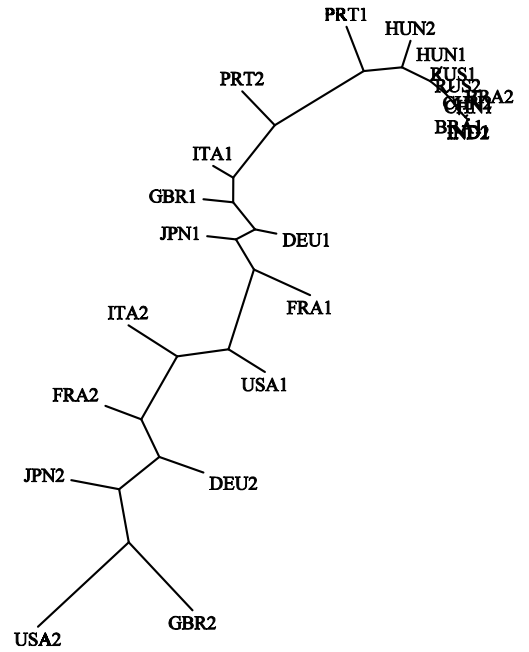


Figure 2

Two-dimensional MDS representation with $m = 20$ for the 12 countries and two 20-year periods, based on the selected variables). Using a) Cosine-correlation (1), b) Euclidean distance (2)



a)



b)

Figure 3

Hierarchical clustering tree representation with $m = 20$ for the 12 countries and two 20-year periods, based on the selected variables). Using a) Cosine correlation (1), b) Euclidian distance (2)

The robustness of the MDS representations and related conclusions were always checked through Sheppard tests and stress tests, which indicated good results.

Alternative ways of plotting the countries' relative positions are given now, using hierarchical clustering tree representations. Hierarchical clustering is a common statistical technique for data analysis. The goal is to build a hierarchy of clusters, in such a way that objects in the same cluster are, in some sense, similar to each other. Clusters are combined or, alternatively, split, based on a measure of their dissimilarity. This is achieved by adopting an appropriate metric, quantifying the distance between pairs of objects, and a linkage criterion, defining the dissimilarity between clusters. They are, perhaps, the most suggestive representations for visualizing results.

In the tree of Fig. 3 the United States and the European partners in general remain in top positions on the world scene. The Asian dissimilarity is confirmed, and Fogel's forecasts may have been too pessimistic.

For a better sum-up of the 40-year process (1972-2012), the following 3-dimensional MDS maps (Fig. 4) help to visualize the distances that separate the partners in the sample, in considering the four selected socio-economic indicators.

The use of a 2-dimensional plot (Fig. 5), or hierarchical clustering tree representations (Fig. 6), may be also useful.

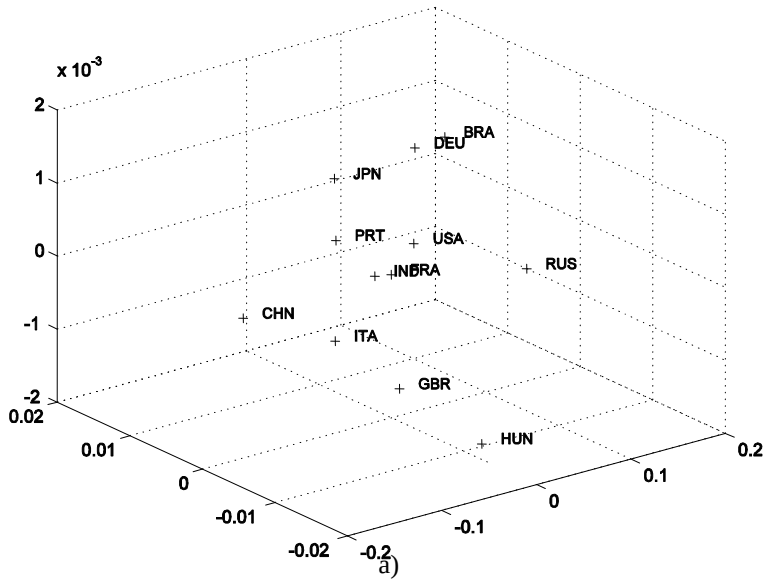
The most notable similarities occur among the Europeans, the two American countries (USA and Brazil), and Russia. The Asian partners will need some time to reach the current degree of similarity seen in the most developed world. Using Euclidean distances the proximity of Russia to small European partners stands out. If some worries exist in terms of possible military aggression, invasion, or hegemony, these relative positions tell us something about the current peaceful global balance.

The same robustness tests were carried out to assure confidence and trust in the results obtained. All tests give credibility and robustness to the plots and their respective conclusions.

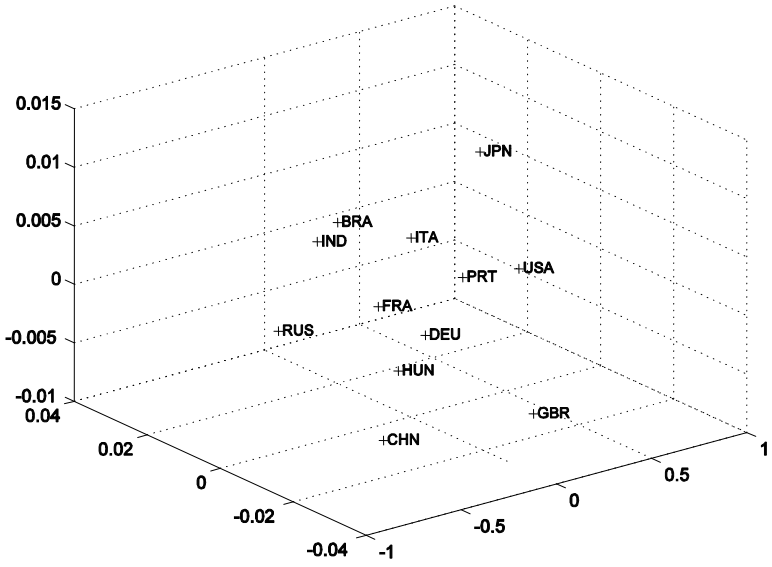
Conclusions

There is considerable convergence in progress among the world's leading countries, and data for 12 partners including European, American, and Asian nations, leads to estimations for comparative views among them, using the MDS methodology.

For economists the problem seems to be more difficult for Europe, where unemployment undermines social stability and reinforces the stagnant local demography. As it is mainly affecting the European youth, it postpones marriage and hinders fertility and birth rates, leading to a precarious social stability. This crisis may be a cyclical problem, as in long-run MDS views trends go favorable to European countries' economic growth [6].



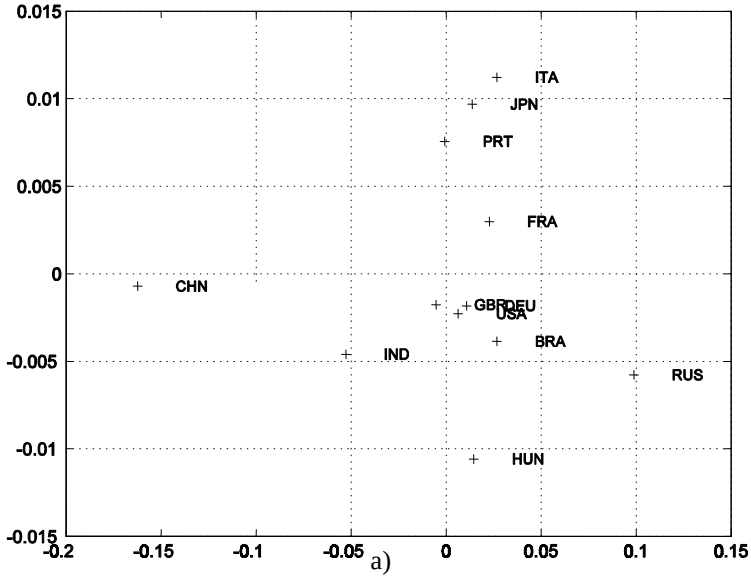
a)



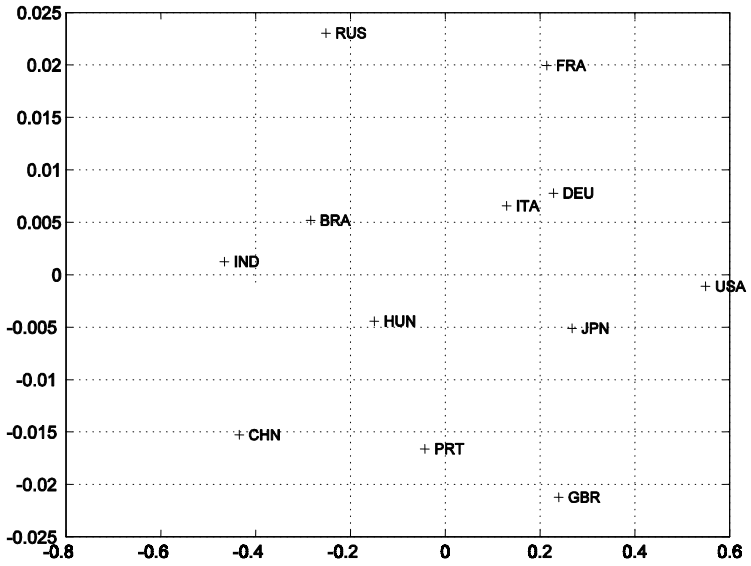
b)

Figure 4

Three-dimensional MDS map with $m = 40$ for the 12 countries and one 40-year period, based on the selected variables). Using a) Cosine correlation (1), b) Euclidean distance (2)



a)



b)

Figure 5

Two-dimensional MDS map with $m = 40$ for the 12 countries and one 40-year period, based on the selected variables). Using a) Cosine correlation (1), b) Euclidean distance (2)

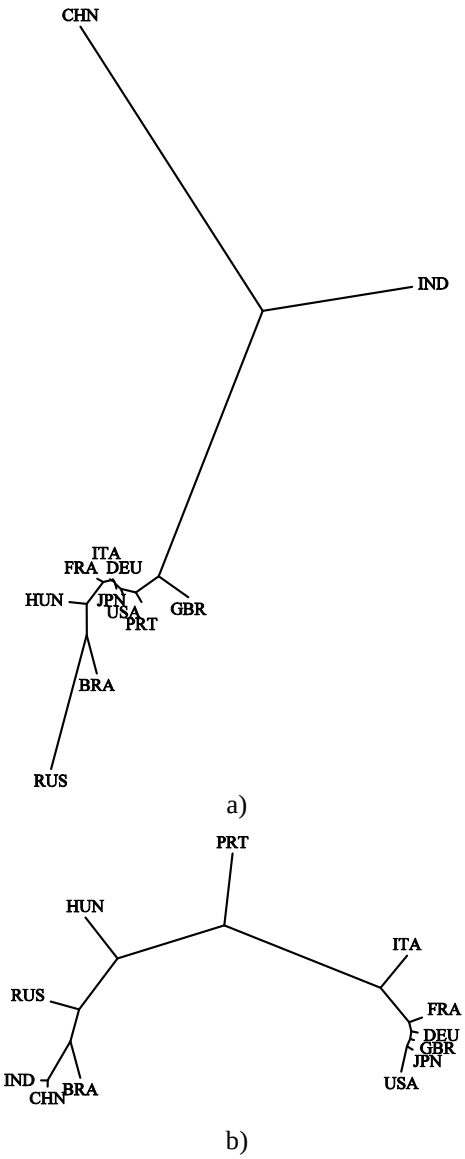


Figure 6

Hierarchical clustering tree representation with $m = 40$ for the 12 countries and one 40-year period, based on the 4 selected variables. Using a) Cosine correlation (1), b) Euclidean distance (2)

Based on data for four socio-economic indicators throughout the last 40 years, it is possible to conclude that Asian partners still need some more decades before similarities can be comparable to the those that exist among the most developed

partners. Russia, however, presents a convergence path, particularly in the last 20 years, that deserves to be emphasized. Strategists, military experts, and politicians can learn a lot from the observation of these estimations. European difficulties, even resulting from a business-cycle context, may witness a noticeably negative influence, if economic warfare takes place, because they may reduce the Milward's ultimate potential of the European partners' economies.

References

[1] D. Acemoglu, and S. Johnson, Disease and Development: The Effects of Life Expectancy on Economic Growth, Journal of Political Economy, Vol. 115, n. 6, 2007, pp. 925-985

[2] D. Acemoglu, and J. A. Robinson, Why is Africa Poor, Economic History of Developing Regions, Vol. 25, n. 1, 2010, pp. 21-50

[3] World Bank, World Bank national development indicators, 2013, http://data.worldbank.org/data-catalog/world-development-indicators

[4] R. Barro, and Jong-Wha Lee, A New Data Set of Educational Attainment in the World, 1950-2010, NBER Working Paper No. 15902, April 2010, (revised 2012)

[5] J. Belak, B. Milfelner, Enterprise Culture as One of the Enterprise's Key Success Factors (Integral Management Approach): Does the Internal and External Cultural Orientation Matter?, Acta Polytechnica Hungarica, Vol. 9, No. 3, 2012, pp. 27-44

[6] P. Benedek, Compliance Management – a New Response to Legal and Business Challenges, Acta Polytechnica Hungarica, Vol. 9, No. 3, 2012, pp. 135-148

[7] M. Bordo, The Global Financial Crisis of 2007-08: Is it Unprecedented? (with John S. Landon-Lane). NBER Working Paper No. 16589, December 2010

[8] I. Borg, Modern Multidimensional Scaling-Theory and Applications, 2nd edition, Springer-Verlag, New York, 2005

[9] S-H Cha, Taxonomy of Nominal Type Histogram Distance Measures, American Conference on Applied Mathematics (MATH'08), Harvard, Massachusetts, USA, March 2008, pp. 24-26

[10] T. Cox, and M. Cox, Multidimensional Scaling, 2nd edition, Chapman & Hall/CRC, 2001

[11] E. Deza, and M. M. Deza, Dictionary of Distances, Elsevier, 2006

[12] Education database for 1950-2010, http://www.barrolee.com/data

[13] B. Eichengreen, Financial Crises and What to Do About Them, Oxford University Press, 2002

[14] B. Eichengreen, Global Imbalances and the Lessons of Bretton Woods. The MIT Press, 2006

[15] R. Fogel, Capitalism and Democracy in 2040: Forecasts and Speculations, NBER Working Paper 13184, June 2007

[16] R. Harding, US economy grows 4.6%, The World Blog, Financial Times, May, 10, 2014 http://www.ft.com/intl/world/us

[17] G. Jones, (ed.), The Making of Global Enterprise, London: Frank Cass, 1994

[18] C. Kindleberger, A Financial History of the Western World, Cambridge, Cambridge University Press, 1993

[19] J. Kruskal, Multidimensional Scaling, Newbury Park, CA, Sage Publications, Inc., 1978

[20] J. Machado, M. Mata, A multidimensional scaling perspective of Rostow's forecasts with the track-record (1960s-2011) of pioneers and latecomers, in J. Awrejcewicz, M. Kazmierczak, P. Olejnik, J. Mrozowski (editors) Dynamical Systems, Theory, Łódź, Poland, 2013, pp. 361-378

[21] A. Maddison, The World Economy, A New Millenium Perspective, Development Centre Studies, 2001

[22] M. Martorelli, Professionalizing Venture Capital, Financial History, Museum of American Finance, Vol. 109, Winter, 2014, pp. 34-37

[23] A. Milward, War, Economy, and Society, Berkeley, Los Angeles, University of California Los Angeles, 1977

[24] L. Prados, Capitalism and Human Development 1870-2007, University Carlos III, Madrid, mimeo. Accessed on 10 January 2012

[25] W. Rostow, Stages of Economic Growth, A Non-Communist Manifesto, Cambridge University Press, 1960

[26] Z. Szabo, E. Herman, Productive Entrepreneurship in the EU and Its Barriers in Transition Economies: A Cluster Analysis, Acta Polytechnica Hungarica, Vol. 11, No. 6, 2014, pp 73-94

[27] P. Tóth, Learning Strategies and Styles in Vocational Education, Acta Polytechnica Hungarica, Vol. 9, No. 3, 2012, pp. 195-216

[28] J. Tzeng, H. Lu, W. Li, Multidimensional Scaling for Large Genomic Data Sets, BMC Bioinformatics, Vol. 9, 179, 2008, pp. 1-17

[29] H. Van der Wee, Prosperity and Upheaval, World Economy, 1945-1980, California, University of California Press, Penguin Books, 1986