Provisional draft

Prepared for presentation at European Population Conference, Budapest, 25-28 June 2014

Demographic Trends in Developing Countries: Convergence or Divergence Processes?

Isabella Corazziari (*), Giuseppe Gabrielli (**), Anna Paterno (***), Silvana Salvini (****)

(*) ISTAT (**) University of Naples Federico II (***) University of Bari (****) University of Florence

1. Introduction and aim of the study

Many developing countries (DCs) are characterized in recent years by both a rise in life expectancy and a further fall in fertility. The trends of demographic behaviours have changed together with many modifications of social aspects (Salvini, 2004). Based on these empirical findings, various theories of social change shared the assumption that societies would converge toward a condition of similarity (Coughlin, 2000; La Croix, Mason and Shigeyuki, 2002).

The convergence theory argued that, as countries achieve similar levels of economic development, they become more similar in terms of these and other aspects of social life. This theory hypothesizes link between economic development and concomitant changes in social organization. Nevertheless, these processes may coexist with a large heterogeneity in the living conditions and huge economic gaps among countries.

The convergence hypothesis has been joined with that of modernization. In particular, in most of population studies, the concept of convergence is linked to the demographic transition theory (Chesnais, 1986). Observing the experience of developed countries, the essence of this theory expects that fertility and mortality rates covary over time in a predictable and uniform manner. In this approach, the transition from a high fertility and mortality situation to another one characterized by low vital rates may be seen as an example of convergence, that describes a world going towards a new “demographic equilibrium” (Wilson, 2001).

The aim of our research is to analyse the trends of specific demographic parameters regarding mortality and fertility, jointly with some socio-economic characteristics (living condition, socio-sanitary situation) of more than 100 DCs, to assess if convergence patterns in demographic behaviours prevail or if marked differences persist. As the paths of mortality and fertility in fact differ deeply over space and time, we need a specific statistical multi-way analysis technique that consider the time series dimension.

Since 70s’ a growing interest has been given to multiway data, classified according to more than two dimensions (the classic units x variables), and many methods have been developed (Coppi, Bolasco 1989, Coppi 1994). Few methods consider a specific statistical treatment of the third dimensions (be it time or space or another criteria of classification), usually considered symmetrically with respect to the other two. Also for the methods that treat asymmetrically the third dimension, the ordering feature of time is not considered, with very few exceptions (Corazziari

1999). The Dynamic Factor Analysis has been proposed and developed in ‘70s (Coppi, Zannella 1979) for multiway data of the type unit x variables x time, considering explicitly the third ordered dimension representing time. The method is based on the joint application of a factorial analysis and regression over time to centers of specific dimensions. The method allows a descriptive and explorative analysis of data.

We apply Dynamic Factor Analysis (DFA) and Cluster Analysis of trajectories in order to evaluate at macro-level the main demographic trends of DCs in the 1995-2010 period. Results let us reconsider the processes of convergence and enlighten the heterogeneity among clusters. In paragraph 2 we make an overview of the theoretical and empirical literature on convergence and demographic transition. In paragraph 3 we describe data and methods used to analyze the trend of DCs to the aim of detecting the process of convergence/divergence in act. Paragraph 4 includes the results of our research observing separately mortality and fertility, and lastly paragraph 5 contains discussion of obtained results taking into account the initial considerations on convergence.

2. Convergence and demographic transition: theoretical and empirical literature overview

2.1. Theoretical background

In the past two hundred years, many social scholars hypothesized that the differences among societies would decrease over time (Inkeles, 1999) and that demographic behaviors would converge in the future (Wilson, 2011). In the 1950s and 1960s, the hypothesis of convergence has been associated primarily with the theory of demographic transition driven by modernization, which have generally assumed that developing countries would follow a path of economic and social progress similar to that of developed ones. The term modernization refers to the set of processes of change on a large scale that involves a particular society, profoundly transforming its structures and patterns of social organization. This concept refers more specifically to the trend of a society affected by these processes to acquire the economic, political, social and cultural characteristics typical of modernity, which therefore reflect aspects such as individualism and rationalism. The modernization is also closely related to the concept of economic development, while the social dimension of modernization manifests itself in phenomena related to demographic change such as urbanization and extensive migration processes.

Other large-scale social transformations are included in modernization, such as population passages from traditional societies characterized by high levels of mortality and fertility (a demographic situation called “Ancien Régime”) to modern demographic regimes where vital rates are low. This passage is defined “demographic transition”, essentially related to the sociological concept of convergence, implying also the transformation of woman status, the so-called “female empowerment”.

More recently, the study of “post-industrial” society and the debate over “post-modernist” aspects of contemporary society also reflect the idea that there is a tendency for broadly similar conditions or attributes to emerge among a range of otherwise distinct and dissimilar societies (Salvini, 2004). As well as population dynamics are concerned, Chesnais (1986; 1997) and Oeppen (1999) were among the first to deal with convergence explicitly, while Heuveline (1999) considered the consequences of convergence on a regional and global scale. Similarly, the United Nations (http://esa.un.org/wpp/index.htm) base their projections on the assumption of convergence, anticipating a homogeneous world in which almost all the demographic variability has disappeared (Wilson 2001).

From a methodological point of view, demographers draw on theoretical, methodological, and empirical literature, developed by economists too, where convergence lies at the heart of modern growth theory (Barro and Sala-i-Martin, 1992;1995) and they applied a variety of statistical methods to test for convergence within and between countries.

Two of the most common methods are referred to as beta-convergence and sigma-convergence. Applied to demographic behaviours, beta-convergence occurs when countries that are laggards in the demographic transition show more movement toward convergence than those further along in that process, while sigma-convergence describes the overall spread of the observed distribution and refers to a reduction of disparities among countries in time (Sala-i-Martin 1996; Neumayer 2004). Referred to fertility and/or mortality, sigma-convergence implies that if the repeated cross-sectional standard deviation increases faster relative to the mean of the observed phenomenon, countries are diverging, and if the standard deviation declines faster relative to the mean, countries are converging.

A third method, using measures of inequality, estimates the spread of a distribution (Dorius, 2008; Firebaugh, 2003). With reference to fertility, inequality refers to the relative differences across countries in national Total Fertility Rate (TFR) estimates, basing on the idea that an absolute decline in fertility gaps is neither a necessary nor a sufficient condition for declining inequality in fertility. Dorius (2008) estimates change in the level of inequality using three population-weighted measures of inequality: the Gini coefficient, the mean log deviation (MLD), and the Theil index. Each of these measures tells us something slightly different about fertility inequality and, taken together, they help us to identify the source of change in the fertility distribution.

Lastly, we can remember the so-called “convergence clubs” approach, derived from the economic Solow model. This approach refer to groups of countries that show common trends, even if they differ from the more general patterns of convergence (Sala-i-Martin, 1996; Solow, 1956). This approach lies on the idea of conditional convergence. The equilibrium that each “club” will reach depends on the initial position and/or on other specific factors. The extended version of the Solow model (Lehmijoki and Pääkkönen, 2006) assumes that convergence should arise in demographic homogenous samples of countries, and that economic growth should be sensitive to demographic growth.

2.2. Empirical findings

The convergence hypothesis has attracted a growing interest of researchers producing interesting and often discordant results. We focus our attention on the main empirical results strictly regarding the topic of this article and produced in the most recent years, that can be summarized as follows.

At the very beginning of the new millennium, Wilson (2001) observed the trends of life expectancy at birth and of TFR in DCs. He highlighted that the second half of the twentieth century witnessed a steady increase in the share of the world’s people living under conditions of declining fertility and rising life expectancy, describing this process as “global demographic convergence”. More in detail, observing that social and demographic change had progressed more rapidly than economic development, he saw the demographic convergence as one aspect of increasing social similarity.

With reference to fertility, Casterline (2001) modeled the pace of decline in less developed countries from 1950 to 2050 and found a significant level of intercountry and intraregional variation in this process. Wilson and Pison (2004), observing the cumulative distribution of the world’s population by fertility level in 1950-2003, suggested that, despite significant change in the middle of the distribution, the overall range did not decrease.

Dorius (2008) argued that the observed variation in intercountry fertility decline for much of the last 50 years pointed to divergence, rather than convergence, and that countries began to converge only around 1995. He showed that the fall of fertility rates around the world did not necessarily mean that fertility rates were converging, defining the delayed onset of the fertility transition for many DCs as the single biggest source of divergence in the TFR. He identified a possible explanation for convergence in health, wealth, and life expectancy with the consistent link between economic and social development, and considered fertility less consistently linked to development. We will pay particular attention on this issue in the following sections.

With reference to mortality, empirical researches generally confirm the occurrence of rising and converging life expectancy levels. In fact, a number of studies demonstrated that most of the period from 1920 to 2000 was one of convergence for many countries (Becker, Philipson, and Soares 2003; Bourguignon and Morrisson 2002; Easterlin 2000; Goesling and Firebaugh 2004; Neumayer 2003, 2004; Pradhan, Sahn, and Younger 2003; Ram 2006). Nevertheless, some researchers defined convergence as “modest” at any time (Moser, Shkolnikov, and Leon 2005). McMichael et al. (2004) suggested a recurring transition process of health, observing that reversals have occurred in many countries in recent decades due in large part to declining male life expectancy in Eastern Europe and Russia and the spread of HIV-AIDS primarily in Sub-Saharan Africa. In addition, Moser, Shkolnikov and Leon (2005) confirmed that a long period of global convergence in life expectancy at birth has been replaced since the late 1980s by divergence of the already cited regions of the world (many Sub-Saharan countries and many Eastern European countries), despite the improving of global life expectancy at birth in the period 1950-2000. Moreover, they showed that the shift from global convergence to divergence has been driven by reverse in adult mortality and that divergence appeared relatively small and of limited duration compared with the earlier convergence.

Other researchers showed, in a cross-country perspective, that convergence in demography might occur conditioned on comparable socio-economic and environmental characteristics (Mishra, Ouattara and Parhi, 2011). Considering different variables, Angeli and Salvini (2009) carried out a descriptive analysis of population characteristics of countries at low and medium Human Development Index (HDI) levels, showing that some exceptions emerged in the convergence process in the mean values of parameters. The same authors highlighted both the strong link between social, economic and demographic development, and the need to examine the paths taken by different countries on the way of the globalization of behaviours.

Recently, Wilson (2011) viewed most demographic change over the past half century as falling along a “main sequence” of demographic transition and the large majority of the world’s population as engaged in a process of convergence. The principal differences between the regions of the developing world are identified in when they enter this main sequence and how rapidly they move along it. Observing the similarity of the regional trends along the main sequence, this author suggested that the health and the fertility transitions are tightly connected. In particular, he considered the fertility transition as a truly global process, with no evidence of significant reversals and only a few countries still to embark upon it. In contrast, he defined the health transition as a slower transformation characterized by a “disturbing” evidence of its fragility, with stagnation and reversals affecting hundreds of millions of people. Moreover, he showed that, when considering the progress of the health transition, the world is not a single demographic system, but it is divided by deep faults into a number of blocs, each with its own distinctive trajectory of life expectancy.

Our research fits into this theoretical and empirical framework. We will consider the recent trends of socio-demographic parameters of DCs to understand if a process of convergence is in act and if clusters of countries in terms of time trajectories of demographic behaviors may be outlined and interpreted.

3. Data and methods

The analyses are carried out on 103 countries with a population of at least 1 million in 2010, defined by UN “less” or “least” developed countries. Data in analysis come from major international sources (World Health Organization, World Bank, United Nations), and they refer to the main vital events and socio-economic indicators for the DCs. Unfortunately it was not possible to consider the years prior to 1995 for the presence of too many missing data in the considered variables.

We analyse separately fertility and mortality processes, measured through their main indicators, but on the same set of countries, so three of the initial ones (Lesotho, Sierra Leone and Panama) have been excluded from the analysis as fertility and mortality indicators were not fully available. Both fertility and mortality have been associated in the analysis with the most commonly correlated socio-economic variables (i.e. for fertility, contraceptive prevalence, gender parity index in school enrolment (GPI), HDI, etc., and, for mortality, immunization coverage among infants, access to improved sanitation, etc.) that may contribute to tracing different “patterns” in observed DCs.

To study the process of convergence/divergence of the temporal dynamics of fertility and mortality, we use a method for multi-way data based on the joint application of a factorial analysis and regression over time called Dynamic Factor Analysis (DFA). The method has been developed in 1970s by Coppi and Zannella (1979) and released in 1990s by Corazziari (1999).

DFA considers quantitative array of data classified according to the following three criteria (or modes, see Tucker 1966): statistical unit, quantitative variable and time of data collecting. This kind of data may be represented in a cubic matrix X (Law et al., 1984) whose generic element is

$$\boldsymbol{X(I, J, T)} = \{x_{ijt}\},\ i=1,\ldots I,\ j=1\ldots J,\ t=1\ldots T$$

where i is the unit index, j the variable index and t the time index, and the same units and variables are observed in each time (or occasion). Broadly speaking, this kind of methodology manages to combine, from a descriptive point of view, the Principal Component Analysis of a compromise matrix over time, and the analysis of the time dynamic of the array by linear regression models of polynomials in t of k order.

In the cubic array $\boldsymbol{X(I, J, T)}$ three sources of variation can be considered and modelled, each of which depending on the two modes of the arrays that can be considered (units x variable; variables x time; units x times). Weights for each dimension of the array can also be considered (weights for units, variables and times). The first source of variation can be attributed to the joint interaction of variables and units, a sort of structural variability or *static*, that is the undertone of the overall variability subject to time changes due to time interaction with variables and with units. The second and the third sources of variation refer to time and its relation with units and variables. In particular, the dynamic of variables over time is represented by the variability of the mean of each variable over time $(x_{.jt})$; the dynamic of units is represented by the time changes of a barycenter of each unit over the set of variables (an average of variables for each unit). As the focus of the present work is on the indicators of fertility and mortality, more relevance is given to variables and their dynamics[1], so the units’ dynamic over time is considered as differential: given the mean time changes of each variable, each unit will be observed in its net time variation, that is if it strengthens the change of the variables or it moves in other directions over time, weakening or even contrasting the overall dynamics.

[1] The DFA provides the possibility to consider both the dual and the tridual extension of each of the four models it considers, according to which dimension is considered more relevant and strategic in the analysis. The dual version focuses on the units dimension, and the tridual on time dimension.

The goal of the methodology is to linearly decompose the overall variability characterising the observed data, described by the covariance matrix of *X(I, J, T)*[2] in the three sources of variability described above, called static (a sort of mean over time), dynamic of centres ($x_{.jt}$) and units' differential dynamic (the net dynamic of single units, when the centres trends have been subtracted). It has been shown (Coppi 1979; Corazziari 1999) that the overall covariance matrix may be decomposed into the sum of three covariance matrices each of them describing one of the above sources of variation:

$$\boldsymbol{S} = {}^{*}\boldsymbol{S}_i + {}^{*}\boldsymbol{S}_t + \boldsymbol{S}_{it}$$

where S is the overall covariance matrix of the array ***X(I, J, T)***, $^{*}\boldsymbol{S}_i$ is the covariance matrix of the centres $x_{ij.}$ representing the mean structure over time, $^{*}\boldsymbol{S}_t$ is the covariance matrix of $x_{.jt}$, and $\boldsymbol{S}_{it}$ is the covariance matrix representing the differential dynamics of units, after subtracting the mean variables dynamic and the static source of variation.

The DFA consists in 4 models each of which approaching the three sources of variation with a specific strategy.

As regards time evolution of the centres $x_{.jt}$ all of the three AFD models consider a linear regression model for each variable j, where the independent variable is time. The parameters are obtained by ordinary least squares. The assumptions about residuals $e_{.jt}$ are the classic ones: $cov[e_{.jt}, e_{.j't'}] = w_j$, if j=j' e t=t', and 0 otherwise.

The variability of the centres $x_{ij.}$ is analized by factorial analysis of specific covariance matrices in each of the three AFD models. In the first model, factorial analysis is applied to the covariance matrix $\mathbf{S_t} = {}^{*}\mathbf{S_i} + \mathbf{S_{it}}$, and it can be easily shown that $\mathbf{S_t}$ is also obtained as the sum of the covariance matrix of variables in each occasion, divided the total number of times. So, by projecting the matrices $\mathbf{X_t}$ centred in each time, we obtain the factorial representation of each unit in each time, that is their dynamic, net to the overall mean time dynamic. The representation of the centres $x_{ij.}$ is obtained by projecting their matrix $^{*}\mathbf{X_i}$ centred, on the factorial plane, due to the decomposition of $\mathbf{S_t} = {}^{*}\mathbf{S_i} + \mathbf{S_{it}}$, as $^{*}\mathbf{S_i}$ is the covariance matrix of $^{*}\mathbf{X_i}$. The other models considers different matrix for the analysis of the static source of variation, and for the analysis of the differential dynamic of units (Corazziari 1997;1999).

As regards differential time evolution of the units, in the first model it is described by comparing the projection of each unit in each time, with the projection of the corresponding centre $x_{ij.}$, according to the decomposition $\mathbf{S_t} = {}^{*}\mathbf{S_i} + \mathbf{S_{it}}$. In the second and third models regression techniques are used to analyse such differential time evolution of units. The fourth model analyses the three sources of variation on the basis of a strategy similar to index numbers, considering one time matrix of data as reference where to plot the other occasions.

Indexes of the goodness of fit of each source of variation in each models are also provided. They are calculated as the ratio between the trace of the modelled covariance matrix of the specific source of variation, and the corresponding observed trace, for each of the covariance matrices described above.

In the following two applications (one for mortality and the other for fertility), the first model of the DFA has been applied. The units are the countries. For the analysis of mortality the variables are 8 yearly indicators (death probability at age 15-60 separately for males and females, under five mortality, HIV-AIDS prevalence, HDI, immunization for DPT3 - diphtheria, pertussis and tetanus -, access to improved sanitation and a macro area variable (that is the geographical region of

[2] In Corazziari (1999), it has been shown that the overall covariance matrix of X(I,J,T) is the covariance matrix of the two-way matrix X(IT,J), obtained collapsing the single matrices $X(I,J)_t$ in each time, over time.

belonging)[3]. For the analysis of fertility, 6 yearly indicators have been considered (adolescent fertility rate, TFR, contraceptive prevalence, HDI, GPI and macro area).

Times are the 4 years 1995, 2000, 2005 and 2010. Each application identifies a factor plane whose interpretation is based on the correlation coefficients between the variables and the axes of the factor plane. The trajectories of the projected countries over the plane are then analysed by a cluster analysis of trajectories (Carlier 1986), obtaining clusters of countries homogeneous as regards the level and dynamics of the considered variables, more easily interpretable in their positions over the plane.

The main feature of the cluster analysis of trajectories is the type of distance between units. When studying trajectories two type of distance between couple of units can be considered: a mean of the comparison (differences) between the two units in each occasion (*mean instantaneous distance*) and a mean of the comparison of the variations between adjacent occasions of each unit (*mean unfolding distance*). A mean of the two considered distances has been proposed, with suitable weights giving more relevance to one of the two distances. Given the above distances between units, the standard methods of cluster analysis is applied, based on the defined distance between clusters and between units and clusters. In the present work, for both the mortality and fertility analysis, the mean of the two distances proposed by Carlier (1986) has been considered, and the Ward method has been chosen among the hierarchical cluster analysis methods, confirmed by a final K-means cluster analysis based on the barycentre of the clusters of the Ward better partition.

The trajectories of the units over the factorial plane have been clustered. If the clusters' trajectory tends towards the centre of the plane, characterising the overall average dynamic of the system of data, the clusters tend to buy homogeneity (witnessing the presence of a convergence process); if they move away from the centres, the heterogeneity increases.

4. Results

4.1. Mortality and health

DFA results show first of all that the first two components of the factor analysis explain a great part of variability (I_t=94.8%) and the best represented times are the second and the third ones (respectively 2000 and 2005, with percentages equal to 95.5 and 95.1).

With regard to the correlation between variables and factors (table 1) we note a strong negative correlation of the first component with both the probabilities of death in adult age (separately for men and women), and the prevalence of HIV-AIDS; the correlation with under five mortality is also negative, but weaker. The second component is strongly and positively correlated with under-five mortality. The correlation is positive also with adult mortality (separately for men and women), but the values are low. Strong negative correlation of the second component is shown with HDI, access to improved sanitation and immunization for DPT3. In synthesis, the first component assumes the meaning of mortality and morbidity (increasing values means lowering mortality and morbidity rates), while the second component represents sanitary conditions and overall health status of observed countries (increasing values of the components means a worsening of such conditions).

[3] The macro-area is conceived as a proxy variable summarising other characteristics of each country not expressed by the chosen indicators, but important to limit clustering of countries too far geographically. It is conceived as a variable considering both homogeneity of countries belonging to the same macro-area, than, at a lesser extent, proximities of different macro-areas, that is macro-area with value 1 is more close to macro-area with value 2 than other macro-areas for example. Also a disjunctive set of variables each referring to a single macro-area has been considered in alternative, providing similar but more scattered results, so the macro-area unique variable has been preferred in the analysis.

Following this interpretation of our results, the positive correlation of the second component with under five mortality can be understood if we consider the latter as an indicator of the health and socio-demographic development of populations.

Table 1 - Correlation matrix between variables and the first two components.

Variables	Component 1	Component 2
Macro area	0.25	-0.34
Prob. of death 15-60 (males) – PDm	-0.79	0.44
Prob. of death 15-60 (females) – PDf	-0.86	0.47
Under five mortality (M05)	-0.43	0.88
HIV-AIDS prevalence	-0.99	-0.15
Immunization for diphtheria, pertussis and tetanus (DPT3)	0.19	-0.71
Access to improved sanitation (AIS)	0.39	-0.78
HDI	0.40	-0.83

Source: Our elaboration on data of World Health Organization, World Bank, United Nations

Before commenting factorial results the overall dynamics of centers has to be discussed, in order to better understand the cluster of units' trajectories projected on the factorial plan. The dynamic of centers ($x_{.jt}$) over time is described through time regression of a suitable order as indicated in table 2. The overall index of fitness for this type of variability is good (0.971).

Table 2 - Time regression analysis of centres of units (overall index of regression fitness *It=0.971)

Variable	R-square	Least Square Estimates					
		Constant	(std. error)	Slope coefficient	(std. error)	Second order coefficient	(std. error)
PDm	0.992	1.068	(3.2E-2)	-1.5E-2	(5.8E-3)	1.8E-2	2.9E-2
PDf	0.979	1.037	(5.2E-2)	-2.0E-2	(9.5E-3)	4.5E-2	4.8E-2
M05	0.980	1.331	(3.7E-2)	-0.133	(1.3E-2)		
HIV-AIDS	0.885	0.637	(0.136)	-6.6E-2	(2.4E-2)	0.344	(0.124)
DTP3	0.972	0.874	(1.7E-2)	5.0E-2	(6.1E-3)		
AIS	0.997	0.886	(5.2E-3)	4.6E-2	(1.9E-3)		
HDI	0.997	0.889	(4.9E-3)	4.4E-2	(1.8E-3)		

Source: Our elaboration on data of World Health Organization, World Bank, United Nations

Only for variable HIV-AIDS prevalence, and the two indicators of adult mortality a polynomial in *t* (time) of second order has been required, for all the other variables (also for variables in the analysis of fertility) a simple linear regression model in *t* has been fitted. Average (over units) indicators of mortality decrease over the considered period, especially child mortality, while the HDI, the immunization DPT3 indicator, the access to improved sanitation do increase. Also the HIV-AIDS mean prevalence decreases, mainly since 2000 (second occasion).

The projection of countries in each occasion over the factorial plane, provides their trajectories over time, differential to the centers $x_{.jt}$ dynamics described by regression. Clustering the trajectories, have lead us to choose a partition formed by 7 clusters (both if using the hierarchical Ward methods, than the K- means confirmative one, with very similar aggregations of countries and very few exceptions). The sequence of the clusters expresses the ranking according to the mean value of adult mortality levels (from the highest to the lowest) and are here shown (see table 3). A geographic characterization is quite evident for 5 of them, while 2 (the first and the second one) include countries located in different macro-areas.

Cluster 7 is the largest one, comprising 53 countries belonging to various macro-areas. This cluster is characterized mainly by the presence of low levels of adult mortality and under five mortality, as well as for a scarce presence of HIV-AIDS. It appears also the best cluster with reference to immunization coverage, access to improved sanitation and HDI. Therefore, it could be considered as including the “best” countries in terms of survival, health and socio-economic conditions. It may be considered also as a reference group when convergence dynamic is questioned. Cluster 6 includes 15 countries, also in this case belonging to various macro-areas. Its main features are levels of adult and mainly under five mortality higher than cluster 7, although lower than the other groups of countries. Cluster 5 contains 20 Sub-Saharan countries, plus Afghanistan that shows an evident delay in the health transition. This cluster presents the lowest levels of immunization and of access to sanitation. It is characterized, above all, by high levels of under five mortality and one of the lowest HDI. Cluster 4 is formed by only 5 Sub-Saharan countries. They are featured by medium-high mortality levels, and, with reference to the other variables, by an intermediate position. Cluster 3 comprises 5 Sub-Saharan countries characterized by the second highest under five mortality rates despite the presence of medium levels of immunization and access to sanitation. The only country present in cluster 2 is Zimbabwe, characterized by a very high adult mortality and high prevalence of HIV-AIDS. The same disease is widespread in Cluster 1, that is formed by only 2 countries (Botswana and Swaziland). Although the last 3 countries share many aspects of their mortality and morbidity conditions, they are divided in different clusters due to the different dynamic showed by these aspects over the time, as can be noted observing fig. 2.

Table 3 - List of analyzed countries by cluster of mortality (K-means method).

Cluster	Countries
1	Botswana, Swaziland.
2	Zimbabwe.
3	Malawi, Mozambique, Namibia, South Africa, Zambia.
4	Central African Republic (CAR), Côte d’Ivoire, Kenya, Tanzania, Uganda.
5	Afghanistan, Angola, Benin, Burkina Faso, Burundi, Cameroon, Chad, Congo, Congo Dem. Rep, Ethiopia, Gabon, Guinea, Guinea-Bissau, Haiti, Mali, Niger, Nigeria, Rwanda, Togo.
6	Bangladesh, Bolivia, Cambodia, Gambia, Ghana, India, Indonesia, Lao PDR, Madagascar, Mauritania, Mongolia, Myanmar, Nepal, Pakistan, Papua New Guinea (PNG), Senegal, Sudan, Tajikistan, Yemen.
7	Algeria, Argentina, Armenia, Azerbaijan, Belarus, Brazil, Chile, China, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, Egypt, El Salvador, Georgia, Guatemala, Honduras, Iran, Iraq, Jamaica, Jordan, Kazakhstan, Kuwait, Kyrgyzstan, Lebanon, Libya, Malaysia, Mauritius, Mexico, Moldova, Morocco, Nicaragua, Oman, Paraguay, Peru, Philippines, Qatar, Russian Federation, Saudi Arabia, Sri Lanka, Suriname, Syria, Thailand, Trinidad and Tobago, Tunisia, Turkey, Ukraine, Uruguay, Uzbekistan, Venezuela, Vietnam.

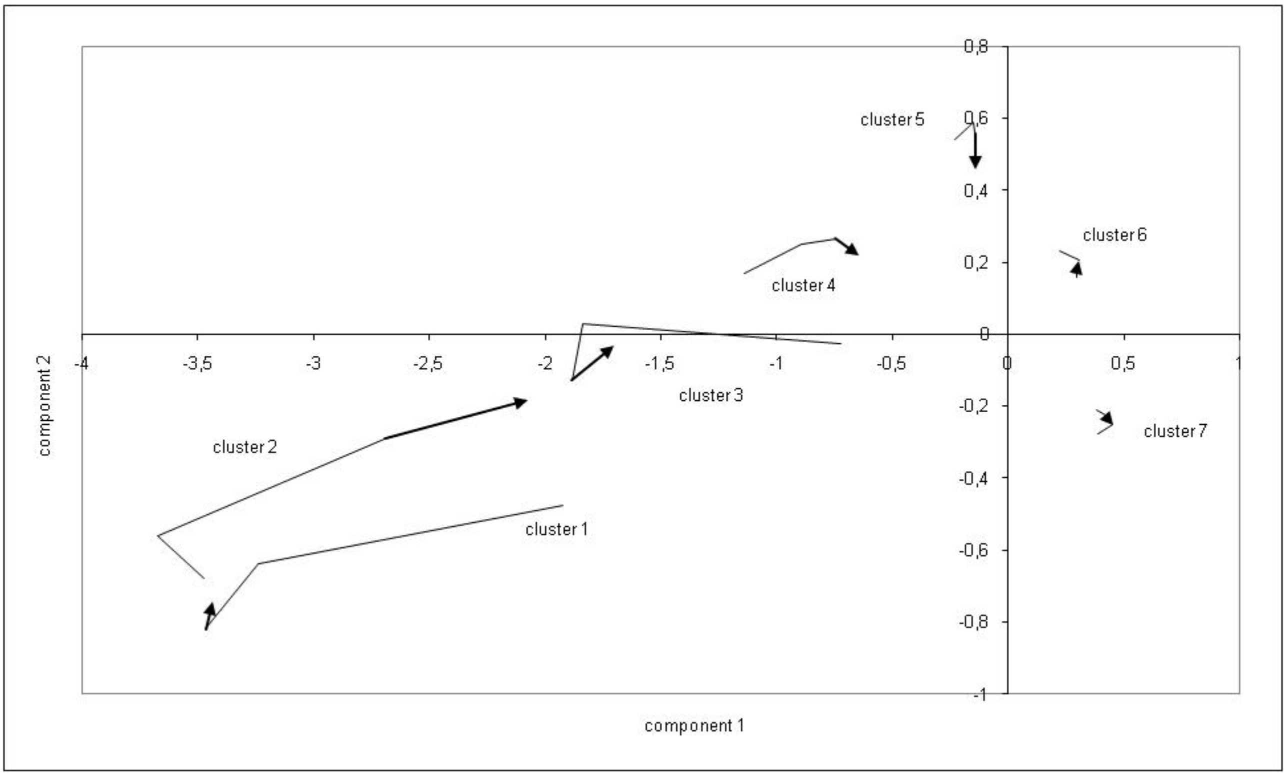
Source: Our elaboration on data of World Health Organization, World Bank, United Nations

According to the DFA method, the differential dynamic of median centers of clusters is represented in figure 1. Generally speaking, the representation of the trajectories of the clusters on the factor plane shows a slight trend toward the average situation; in fact, they show weak differential dynamics towards the center of the axes. Cluster n. 7, 6 and 5 are located in different quarters of the factorial plan, according to the different values of the active variables. However, looking at the temporal trends of each cluster, we observe that they tend to slightly converge toward the center of the axes of the factorial plan. In other words, the dynamic of countries included in such clusters (that represent almost the 90% of the global scenario included in the analysis) show a slow reduction of their differentials toward a uniformity to the centers of considered variables.

As it regards cluster 2 and cluster 4, the pattern of convergence toward the centre of the axes is much more evident. They represent those countries in which HIV-AIDS prevalence significantly decreased in the considered period: respectively from 25.0 to 15.2% in cluster 2 and from 9.3 to 5.8% in cluster 4.

Clusters 3 and 1 seem to outline a worsening of their conditions, showing a differential dynamic that moves toward left in the plane, which is toward highest levels of the aforementioned variables. This trend seems to show an opposite tendency with respect to the decreasing mean of the general mortality and of the HIV-AIDS levels, as indicated by the regression analysis of such indicators centers over units (table 2). However, both clusters show a slight but significant "reversed" path dynamic in the period 2005-2010, oriented toward the center of the axes.

Figure 1 - Differential dynamics of median centers of clusters – Health and mortality.



Note: cl_i means cluster i, t_i means time i.
Source: Our elaboration on data of World Health Organization, World Bank, United Nations

4.2. Fertility

Our results regarding fertility analysis can be synthesized as follows. The first two components of DFA explain 85.4% of the variability of the phenomenon, a little less than in the mortality analysis where, nevertheless, variables are slightly more numerous. For fertility, the times better represented are the third and the fourth (respectively 2005 and 2010, with percentages of explained variability equal to 86.5 and 85.6).

The correlation of the two components with the active variables (see table 4), indicates that the first component represents with positive values better situations in terms of human development, GPI, and contraceptive prevalence, while with negative values higher total and adolescent fertility rates.

Looking at the factorial plan, that shows graphically the correlations between factor components and variables exposed in table 4, we note the separation described above between adolescent fertility and TFR from one side, and GPI, HDI and contraceptive prevalence from the other. In synthesis, the variables measuring development are opposite to fertility behavior.

Table 4 – Correlation matrix between variables and the first two components.

Variables	Component 1	Component 2
Macro area	0.28	0.75
Adolescent fertility rate	-0.89	0.41
TFR	-0.93	-0.06
HDI	0.86	0.18
Gender parity index in school enrolment (GPI)	0.74	0.24
Contraceptive prevalence	0.86	0.37

Source: Our elaboration on data of World Health Organization, World Bank, United Nations

Table 5 - Time regression analysis of centers of units (overall index of regression fitness *It=0.989)

Variable	R-square	Least Square Estimates			
		Constant	(std. error)	Slope coefficient	(std. error)
Adolescent fertility rate	0.989	1.228	1.850E-2	-9.134E-2	6.757E-3
TFR	0.984	1.229	2.292E-2	-9.171E-2	8.369E-3
HDI	0.997	0.889	4.918E-2	4.430E-2	1.796E-3
GPI	0.999	0.939	1.178E-3	2.437E-2	4.301E-4
Contraceptive prevalence	0.991	0.801	1.481E-2	7.975E-2	5.408E-3

Source: Our elaboration on data of World Health Organization, World Bank, United Nations

To better describe the trajectories of countries over the factorial plane, a look to the overall dynamic of the data is necessary. Table 5 shows the time regression parameters of the considered indicators means. All the indicators of fertility are decreasing over time, while indicators describing development are increasing on average in the considered period. According to such dynamics, the positions and trajectories of the aggregated clusters are to be described (table 6).

The cluster analysis of the countries trajectories projected on the plane have lead us, also for fertility, to choose a partition formed by 7 clusters (again both if using the hierarchical Ward method and the K-means confirmative one, with very similar aggregations of countries and very few exceptions). The number of the clusters expresses the ranking according to the mean value of TFR (from the highest to the lowest), generally describing the demographic transition stage.

Table 6 - List of analyzed countries by cluster of fertility (K-means method).

Cluster	Countries
1	Afghanistan, Angola, Chad, Congo Dem. Rep, Guinea, Mali, Niger, Uganda.
2	Benin, Burkina Faso, Cameroon, Central African Republic (CAR), Côte d'Ivoire, Guinea-Bissau, Madagascar, Malawi, Mozambique, Nigeria, Senegal, Tanzania, Yemen, Zambia, Congo, Ethiopia, Gambia.
3	Burundi, Cambodia, Ghana, Lao PDR, Mauritania, Pakistan, Rwanda, Sudan, Togo.
4	Bangladesh, Gabon, Guatemala, Haiti, India, Iraq, Kenya, Namibia, Nepal, Papua New Guinea (PNG), Swaziland, Zimbabwe.
5	Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Honduras, Jamaica, Mexico, Nicaragua, Paraguay, Peru, Uruguay, Venezuela.
6	Algeria, Armenia, Azerbaijan, Belarus, Egypt, Georgia, Kazakhstan, Kyrgyzstan, Libya, Malaysia, Moldova, Morocco, Myanmar, Russian Federation, Tunisia, Ukraine, Uzbekistan, Tajikistan, Saudi Arabia.
7	Botswana, China, Indonesia, Iran, Jordan, Kuwait, Lebanon, Mauritius, Mongolia, Oman, Philippines, Qatar, South Africa, Sri Lanka, Suriname, Syria, Thailand, Trinidad and Tobago, Turkey, Vietnam.

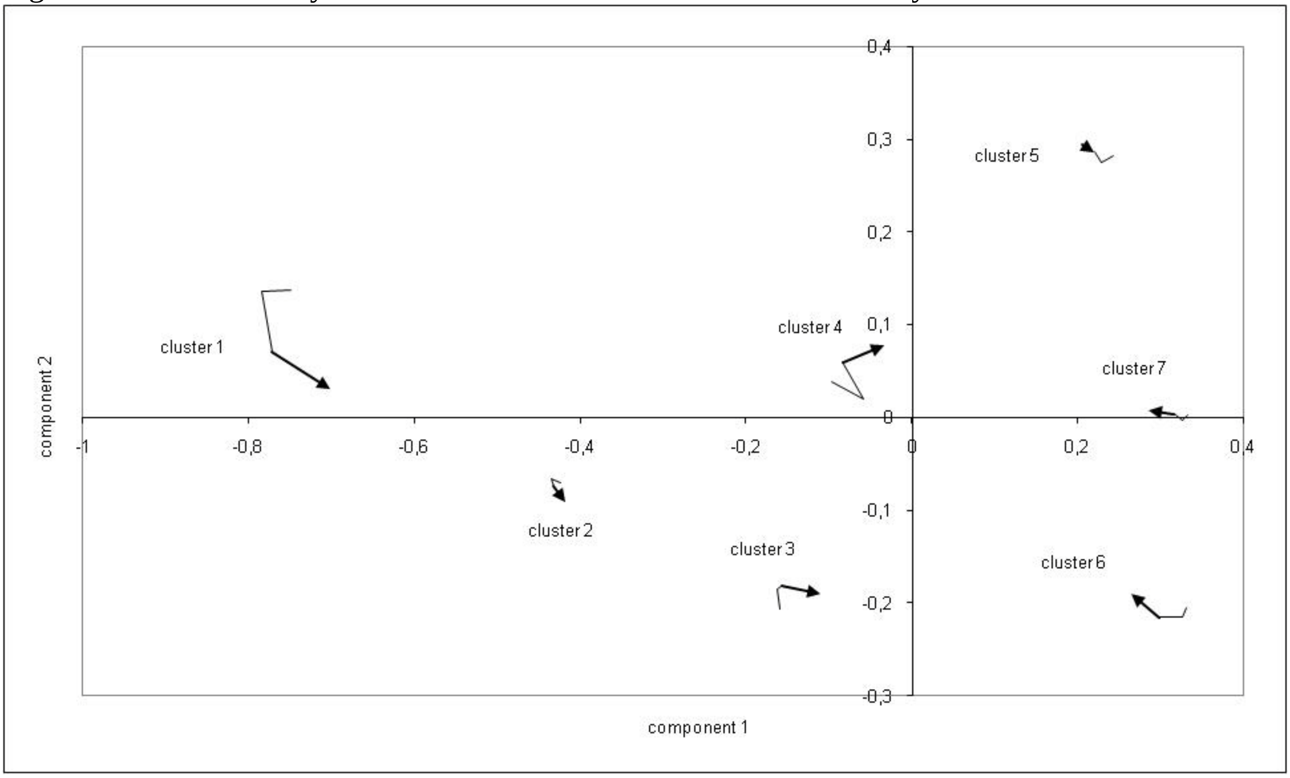
Source: Our elaboration on data of World Health Organization, World Bank, United Nations

The chosen partition shows well separated clusters of comparable magnitude (in particular for the clusters 2, 5, 6 and 7, respectively compounded by 17, 18, 19 and 20 countries). The geographic characterization is quite marked for the cluster 1 and 2 that include Sub-Saharan countries, with the only exception of Afghanistan and Yemen. These countries are characterized by a strong delay in the demographic transition, that is high levels of fertility and adolescent fertility and a low value of HDI, the same aspects that distinguish Sub-Saharan Africa countries. Adolescent fertility, strongly related to women empowerment and female human capital, may be considered a proxy of development of women status and, consequently, of the degree of growth of the country. The cluster 3 includes countries belonging to Southern Asia and Sub-Saharan Africa, facing a medium stage of the path toward modernization. The cluster 4 is geographically more heterogeneous, including three Asian countries (India, Bangladesh and Nepal) but also some African countries, and Guatemala and Haiti, in Central American and Caribbean regions; it is characterized by high level of adolescent fertility. Latin-American countries, where fertility is lower, contraception higher, such as the development indicators HDI and GPI compound the cluster 5. Cluster 6 and 7 contain countries belonging to different continents that, nevertheless, present common demographic characteristics. Cluster 6 include several Asian countries that in the recent past made part of Soviet Union, and that actually are living the process of economic and demographic transition. This cluster includes also the South-Eastern shore of Mediterranean, where demographic characteristics approach those of European countries, even if the stage of development and women's status described by HDI and GPI are quite different. In cluster 7, China and Iran are an example of got fertility transition, with value of TFR that nowadays are below replacement level. This cluster includes also some countries located both in the South-Eastern side of Mediterranean Basin and in Latin America and is featured by low fertility, high contraception, HDI and GPI, representing a situation of well-established demographic transition.

In synthesis, cluster 5, 6 and 7 present a better situation in terms of index of development and use of contraception; in opposite, clusters 1 and 2 manifest a worse situation in terms of development status.

In figure 2 we have reported the trajectories of median centers of the clusters for fertility analysis. We observe the temporal dynamic of the clusters with respect to the center of the axes, representing on average the reference of the overall dynamic. The clusters 4, 6 and 7 tend to converge toward the mean situation of the whole group of countries in the period we examine, represented by the barycenter of the axes. This means that the differential dynamic of countries included in these cluster tends to uniform to the dynamic of the centers of variables that determine the factorial plan. Cluster 1 starts from the worst position and goes towards the barycentre, Cluster 2 and 3 seem do not show high variability over time.

Figure 2 - Differential dynamics of median centers of clusters – Fertility.



Note: cl_i means cluster i, t_i means time i.
Source: Our elaboration on data of World Health Organization, World Bank, United Nations

5. Discussion

In the previous pages we observe the trends of specific demographic parameters regarding fertility and mortality in DCs, together with some socio-economic variables , to evaluate if convergence in their demographic behaviour prevails or if marked differences remain.

After recalling the most relevant literature on this topic, and taking into account the fact that the paths of mortality and fertility differ deeply over space and time, we apply DFA and Cluster Analysis of trajectories in order to evaluate at macro-level the main demographic trends of DCs in the 1995-2010 period.

Examining results for mortality, we note that the largest cluster includes 53 countries and is characterized mainly by low levels of adult mortality, scarce prevalence of HIV-AIDS, high level of immunization coverage and access to improved sanitation and, finally, by high HDI. Therefore, it could be considered at the top of the classification according to development, including the "best" countries in terms of survival, health and socio-economic conditions. In this cluster, convergence in modernization characteristics is realizing. Another cluster compounded by countries belonging to different areas presents low level of mortality too, even if higher than the previous one. All the remaining clusters are more homogeneous from the geographical distribution point of view, and present higher adult and under five mortality. In particular, three clusters include Sub-Saharan countries, distinguished by high prevalence of HIV-AIDS and, consequently, by the highest level of adult mortality.

The dynamic analysis of mortality seems therefore to confirm that a large part of DCs are converging on a uniform model of health and mortality, leaving backward in particular those countries (almost all localized in Sub-Saharan Africa) characterized by high values of HIV-AIDS prevalence and – consequently – high mortality. In fact, looking at the trajectories of all the clusters, we observe that most of them tend to slightly converge toward the center of the axes of the factorial plan, showing a trend to uniformity.

The analysis of fertility indicators lead us to different considerations.

Our findings show well separated clusters of comparable magnitude. The geographic characterization is quite marked for the first and the second clusters including mostly Sub-Saharan countries that demonstrate a strong delay in the demographic transition.. The cluster that contains Latin-American countries, where fertility is lower, contraception higher, such as the development indicators HDI and GPI, is particularly homogenous and describes a situation of well-established demographic transition. We found then the cluster that includes several Asian countries (before making part of Soviet Union), and that actually are living the process of economic transition. Some values of variables are similar to Latin American experiences but mechanisms governing the path toward modernization are quite different. Finally, there is the cluster more modernized, geographically heterogeneous, including China and Iran, that have reached the last stage of fertility transition and where TFR is below replacement level.

We observe that the temporal dynamic of the clusters with respect to the center of the axes, representing on average the reference of the overall dynamic, is not particularly evident. In summary, although many countries show a process of convergence group by group, there is not a notable common situation, which they converge to, and a quite numerous group of countries remains that present different models of trajectories characterized by the belonging to specific territorial and cultural contexts.

Overall, if we compare the process of convergence in mortality and in fertility, we note a remarkable difference. Mortality analysis evidences a large cluster compounded by a half of all DCs showing that these countries are very similar in their trajectories. Such a result reveals a strong process of convergence. The other less numerous clusters show instead different dynamic characteristics, divergent from the largest cluster. As it regards fertility, we not find a large homogeneous group, but many clusters of comparable magnitude, showing different situations that do not evidence a process of convergence similar to mortality's one. Mortality and fertility do not present similar patterns therefore, and health and mortality seem to have reached a more advanced stage of convergence.

With reference to the theoretical and empirical framework recalled above, our results are mostly coherent (despite some exceptions) with those obtained by many of previous analyses (see "Empirical findings" section).

Firstly, considering mortality, our results are in line with those of researches that highlighted the occurrence of rising and converging survival levels (Wilson, 2001). With reference to Wilson's (2011) assertion that the world is not a single demographic system, but is divided by deep faults into a number of blocks, each with its own distinctive trajectory of life expectancy, we prove analogously the presence of heterogeneous clusters of countries. In this contest, our results are slightly different from those of Wilson in highlighting the existence of a large group of countries, that can be considered as a "point of arrival" when convergence dynamic is questioned. Secondly, consistently with McMichael et al. (2004), we show the presence of changing trends and reversals in some groups of countries, due in large part to the spread of HIV-AIDS primarily in Sub-Saharan Africa. Notwithstanding, we highlight, agreeing with Moser, Shkolnikov and Leon (2005), that divergence is relatively less widespread compared with the convergence.

Observing fertility, our outcomes confirm, in line with Casterline's findings (2001) the presence of an intercountry and intraregional variation in the pace of its decline. Furthermore, our outcomes are consistent with those of Dorius (2008) and those of Wilson and Pison (2004), showing that the trends of fertility variation around the world are not necessarily converging, due to the delayed onset of this transition for many DCs.

Our findings corroborate the connection, proved by previously quoted studies (Mishra, Ouattara and Parhi, 2011; Angeli and Salvini, 2009), between economic, social, cultural and behavioural characteristics, as well as of their levels of development, and demographic convergence processes.

More in detail, a further similarity arise between our results and the analysis performed by Wilson (2001). In fact, he explained convergence in mortality with the consistent link between economic and social development, and considered fertility less consistently linked to these variables.

Furthermore, according to this contribution, the conclusion of our reflections may lead to the consideration that mortality may above all depend on the modernization path and that the convergence is largely influenced by the socio-economic evolution. On the contrary and in line with Dorius (2008), fertility convergence depends mostly on cultural variables, by modification of attitudes and tradition, which are changing more slowly with respect to economic development.

References

Angeli A., Salvini S. (2009). Demographic Heterogeneity of Developing Countries between Reality and Convergence Myth, *XLII Scientific Meeting of Italian Association for Population Studies*, Milano.

Barro, R. e Sala-i-Martin, X., 1992, Convergence. *Journal of Political Economy*, 100, n. 2 (April 1992), 223-51.

Barro, R. e Sala-i-Martin, X.,1995, *Economic growth*, New York, McGraw-Hill.

Becker G.S., Philipson T.J. and Soares R.R., 2003, The quantity and quality of life and the evolution of world inequality, Nber Working Paper Series, Working Paper 9765, http://www.nber.org/papers/w9765, National Bureau Of Economic Research.

Bourguignon, F. and Morrison, C., (2002), 'Inequality among world citizens: 1890–1992', *American Economic Review*, vol. 92, no. 4, pp. 727–744.

Carlier A. (1986). Factor Analysis of Evolution and Cluster Methods on Trajectories, Compstat .

Casterline, J., 2001. The pace of fertility transition: National patterns in the second half of the twentieth century, in Global fertility transition, R. Bulatao e J. Casterline eds., A Supplement to Volume 27, in «Population and Development Review», pp. 17-52.

Chesnais, J.C., (1986). *La transition démographique : étapes, formes, implications économiques*, Presses Universitaires de France, INED, Paris.

Chesnais, J.C., 1997, "La transition démographique: trente ans de bouleversements," in J.C. Chateland and J-C. Chesnais (eds.) La Population du Monde: Enjeux et Problemes. Paris: INED-PUF.

Coppi R., 1994. An Introduction to Multiway Data and their Applications. Computational Statistics and data analysis , 18 (4), pp. 3-13.

Coppi R., Bolasco, 1989. Multiway Data Analysis. Amsterdam, North Holland.

Coppi R., Zannella F. (1979). L'analisi fattoriale di una serie temporale multipla relativa allo stesso insieme di unità statistiche, *Atti della XXIX Riunione della SIS*.

Corazziari I. (1997). Dynamic Factor Analysis, Classification and data Analysis, Pescara 3-4 luglio 1997.

Corazziari I. (1999). *Analisi di Dati Longitudinali Multivariati*, Tesi di Dottorato XI ciclo, Napoli aprile 1999.

Coughlin, R.M., 2000, Convergence theories, in Encyclopedia of sociology, E. F.Borgatta e M.L. Borgatta (eds.), New York, Macmillan, pp. 422-431.

Dorius S.F. (2008). Global Demographic Convergence? A Reconsideration of changing Inter-country Inequality in Fertility, *Population and Development Review*, 34(3), 519-537.

Easterlin, R. A. 2000. "The worldwide standard of living since 1800," *The Journal of Economic Perspectives*, 14, pp. 7–26.

Firebaugh, G., 2003, *The New Geography of Global Income Inequality*, Cambridge and London: Harvard University Press.

Firebaugh, G., Goesling B, 2004. "Accounting for the recent decline in global income inequality," *The American Journal of Sociology*, 11, pp. 283–312.

Heuveline, P., 1999, "The global and regional impact of fertility and mortality transitions, 1950-2000", *Population and Development Review*, 25, pp. 681-702.

La Croix S. J., Mason A. and Shigeyuki A., 2002, Population and Globalization, Southeast Asian Studies, vol. 40, 3, pp. 240-267.

Lehmijoki U., Pääkkönen J. (2006). Demographic Clubs: Convergence within, divergence between, Helsinki Center of Economic Research, Discussion Papers, 136.

McMichael, A. J., McKee, M., Shkolnikov, V., Valkonen, T., (2004), Mortality trends and setbacks: global convergence or divergence? *The Lancet*, 363(9415), pp. 1155-1159.

Mishra T, Ouattara B., Parhi M. (2011). *Cross-country convergence in demography under stochastic shocks*, School of Business and Economics, Working Paper in Economics, 1 Swansea University, Singleton Park, Swansea.

Moser K., Shkolnikov V.M., Leon D.A., (2005), World mortality 1950–2000: divergence replaces convergence from the late 1980s, *Bulletin of the World Health Organization*, Vol. 83(3), pp. 202-209.

Neumayer, E., 2003, "Beyond income: Convergence in living standards, big time," *Structural Change and Economic Dynamics*, 14, pp. 275–296.

Neumayer, E., 2004. "HIV/AIDS and cross-national convergence in life expectancy," *Population and Development Review*, 30, pp. 727–742.

Oeppen, J.,1999. "The health and wealth of nations since 1820," Paper presented at 1999 Meeting of the Social Science History Association, Fort Worth, TX.

Pradhan, M., Sahn, D. E., & Younger, S. (2003). Decomposing world health inequality, *Journal of Health Economics, 22 (2), pp. 271-293.*

Ram R., 2006, "State of the 'life span revolution' between 1980 and 2000" *Journal of Development Economics, 80, pp. 518–526.*

Sala-i-Martin X. X. (1996). The classical approach to convergence analysis, *The Economic Journal*, 106, 1019-1036.

Salvini S. (2004). Demographic Convergence and Economic Divergence? Revisiting Demographic Transition in Developing Countries, *Proceedings of the XLII Scientific Meeting of the Italian Statistical Society* (SIS), Bari

Solow R. M. (1956). A Contribution to the Theory of economic Growth, *Quarterly Journal of economics*, 70, 65-94.

Wilson C. (2001). On the scale of global demographic convergence 1950-2000, *Population and Development Review*, 27(1), 155-171.

Wilson C. (2011). Understanding global demographic convergence since 1950, Population and Development Review, 37(2), 375-388.

Wilson, C. and Pison G. (2004). Half the world's population lives where fertility is below replacement level. *Population and Societies*. Oct; (405):1-4.