# Demographic Trends

in Developing Countries:

Convergence or Divergence Processes?

Maria Silvana Salvini, Giuseppe Gabrielli,
Anna Paterno, Isabella Corazziari
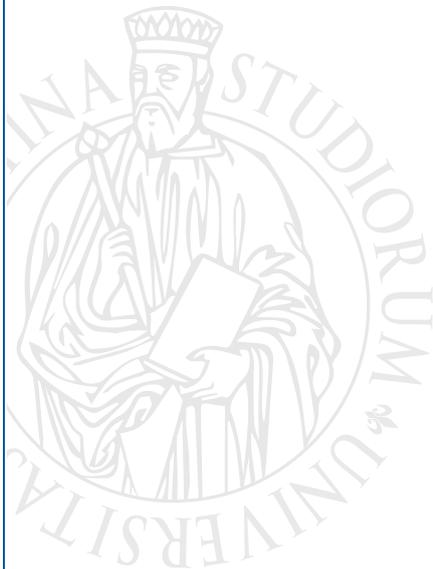
Demographic Trends in Developing Countries: Convergence or Divergence Processes?

GIUSEPPE GABRIELLI
ANNA PATERNO
SILVANA SALVINI
ISABELLA CORAZZIARI

Abstract:

Recent years are characterized by both a rise in life expectancy and a further fall in fertility in the developing countries (DCs). These processes coexist with large heterogeneity according to the specific living conditions of countries. The aim of our research is to analyze the trends of specific demographic parameters regarding mortality and fertility, jointly with some socio-economic characteristics of more than 100 DCs, to assess if convergence patterns in demographic behaviors 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. Thus, we apply Dynamic Factor Analysis 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.

Giuseppe Gabrielli, corresponding author
Assistant Professor of Demography, Università degli Studi di Napoli Federico II, e-mail: giuseppe.gabrielli@unina.it

Anna Paterno,
Professor of Demography, Università degli Studi di Bari, e-mail: anna.paterno@uniba.it

Silvana Salvini,
Full Professor of Demography, Università di Firenze, email: salvini@disia.unifi.it

Isabella Corazziari,
Researcher of Statistic, Istituto Nazionale di Statistica (ISTAT), email: corazzia@istat.it

Demographic Trends in Developing Countries: Convergence or Divergence Processes?

In recent years, many developing countries (DCs) have witnessed both a rise in life expectancy and a further decline in fertility. These demographic trends have changed together with many other social aspects. Based on these empirical findings, various theories of social change shared the assumption that demographic patterns in various societies are converging (Coughlin, 2000; La Croix, Mason and Shigeyuki, 2002).

According to convergence theory, as countries achieve similar levels of economic development, they become more similar in terms of their demographic characteristics and other aspects of social life. This theory hypothesizes a link between economic development and concomitant changes in social organization. Nevertheless, these processes may coexist with considerable heterogeneity in living and economic conditions in various countries.

This convergence hypothesis has been joined by a theory of modernization. In most population studies, the concept of convergence is linked to the demographic transition theory (Chesnais, 1986). Based on observations of the experiences of developed countries, this theory anticipates that fertility and mortality rates will vary over time in a predictable and uniform manner. In this approach, the transition from a high fertility and mortality situation to a situation characterized by low vital rates may be seen as an example of convergence; thus, the modernization theory describes a world that is moving toward a new "demographic equilibrium" (Wilson, 2001).

The aim of our research is to analyze the trends of specific demographic parameters regarding mortality and fertility, together with some of the socioeconomic characteristics (e.g., living conditions, socio-sanitary conditions) of more than 100 DCs, in order to assess whether demographic behaviors are indeed converging, or whether marked differences persist. As the paths of mortality and fertility differ significantly over space and time, we need a specific statistical multi-way analysis technique that considers the time dimension.

Since the 1970s, researchers have become increasingly interested in multi-way data, or data which can be classified according to more than two dimensions (the classic units x variables). Thus, many methods for using such data have been developed (Coppi, Bolasco 1989, Coppi 1994). However, few of these methods allow for a specific statistical treatment of the third dimension (be it time or space, or another criterion of classification), which is usually considered symmetrically in relationship to the other two dimensions. Moreover, very few of the methods that treat the third dimension asymmetrically consider the ordering feature of time (Corazziari 1999). The Dynamic Factor Analysis (DFA), which was proposed and developed in 1970s (Coppi, Zannella 1979) for use with multi-way data of the type unit x variables x time, explicitly considers 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; and thus allows for a descriptive and explorative analysis of data.

We apply Dynamic Factor Analysis and Cluster Analysis of trajectories in order to evaluate at the macro level the main demographic trends in DCs in the 1995-2010 period. Our proposed

contribution to the research on these issues is based on the inclusion in the models of the temporal trajectories of demographic parameters. Our results allow us to reconsider the processes of convergence, and to shed light on the heterogeneity among clusters. First, we provide an overview of the theoretical and empirical literature on convergence and demographic transition. Second, we describe the data and the methods used to analyze the convergence/divergence trends in DCs. The results of our research are displayed separately for mortality and fertility. We conclude with a discussion of our results which takes into account our initial considerations regarding convergence.

Convergence and demographic transition: an overview of the theoretical and empirical literature

Theoretical background

In the past two hundred years, many social science scholars have hypothesized that the differences between societies would decrease over time, and that demographic behaviors would converge in the future. In the 1950s and 1960s, the hypothesis of convergence was associated primarily with the theory of "demographic transition", which was driven by modernization. This theory generally assumed that developing countries would follow a path of economic and social progress similar to the one already observed in developed countries. The term "modernization" refers to the set of large-scale processes of change that take place in a particular society, profoundly transforming its structures and patterns of social organization. This concept refers more specifically to the tendency of a society affected by these processes to acquire the economic, political, social, and cultural characteristics typical of modernity, which reflects aspects such as individualism and rationalism. 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 in which these levels are low. This "demographic transition" is related to the sociological concept of convergence, and also implies the transformation of the status of women, or "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). Chesnais (1986; 1997) and Oeppen (1999) were among the first scholars of population dynamics to deal with convergence explicitly, while Heuveline (1999) considered the consequences of convergence on a regional and on a global scale. Similarly, the United Nations (http://esa.un.org/wpp/index.htm) base their projections on the assumption of convergence, forecasting a homogeneous world in which almost all demographic variability has disappeared (Wilson 2001).

From a methodological point of view, demographers draw upon theoretical, methodological, and empirical literature—including studies by economists—based on modern growth theories that emphasize convergence (Barro and Sala-i-Martin, 1992; 1995). Thus, demographers have applied a variety of statistical methods to test for convergence within and between countries.

Among the methods most commonly used in this context are beta-convergence and sigma-convergence. Applied to demographic behaviors, beta-convergence occurs when countries that are laggards in the demographic transition show more movement toward convergence than the

countries that are farther along in the process; while sigma-convergence describes the overall spread of the observed distribution, and refers to a reduction in the disparities between countries in time (Sala-i-Martin 1996; Neumayer 2004). Referring to fertility and/or mortality, sigma-convergence suggests that if the repeated cross-sectional standard deviation increases faster relative to the mean of the observed phenomenon, the countries are diverging; whereas if the standard deviation declines faster relative to the mean, the countries are converging.

A third method, which uses measures of inequality, estimates the spread of a distribution (Dorius, 2008; Firebaugh, 2003). Inequality in fertility levels is measured by examining the relative differences in the national Total Fertility Rate (TFR) estimates across countries, based on the idea that an absolute decline in the fertility gap is neither a necessary nor a sufficient condition for a decline in fertility inequality. Dorius (2008) estimated the 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.

Finally, the so-called "convergence clubs" approach, derived from the Solow model in economics, refers to groups of countries in which the trends are similar, even if they differ from the more general patterns of convergence (Sala-i-Martin, 1996; Solow, 1956). This approach relies on the idea of conditional convergence. The equilibrium that each "club" eventually reaches 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 homogenous demographic samples of countries, and that economic growth should be sensitive to demographic growth. To the best of our knowledge, none of these approaches explicitly includes time in the models, and the analyses focus exclusively on differences in temporal series.

Empirical findings

While the convergence hypothesis has increasingly attracted the interest of researchers, studies based on the concept have so far produced intriguing, but sometimes conflicting results. Here we discuss the main empirical results related to our topic which have been produced in recent years

At the start of the new millennium, Wilson (2001) examined the TFRs and the trends in life expectancy at birth in DCs. He noted that in the second half of the 20th century, the share of the world's population who were living under conditions of declining fertility and rising life expectancy increased steadily; a process he described as "global demographic convergence." Specifically, Wilson observed that social and demographic change had progressed more rapidly than economic development, and he saw demographic convergence as being one aspect of the increase in social similarity.

In terms of mortality, empirical researchers have generally confirmed the assumption that life expectancy levels are rising and converging. In fact, a number of studies have demonstrated that between 1920 and 2000 a convergence in life expectancy levels was occurring in many of these 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). McMichael et al. (2004) suggested suggested a recurring transition process of health. Specifically, they observed that setbacks 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) found that despite the overall improvements in global life expectancy at birth during the second half of the 20th century, this long period of global convergence in life expectancy at birth has been replaced

since the late 1980s by a divergence of the abovementioned regions of the world (i.e., many Sub-Saharan countries and many Eastern European countries, such as Russia). However, they showed that the shift away from global convergence toward divergence has been driven by reversals in adult mortality rates, and that the divergence trend appears to be relatively small and of limited duration compared with the earlier convergence trend (Grigoriev et al., 2014).

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 inter-country and intra-regional variation in this process. Wilson and Pison (2004), based on their investigation of the cumulative distribution of the world's population by fertility level in 1950-2003, suggested that despite significant changes in the middle of the distribution, the overall range did not decrease.

Dorius (2008) argued that the observed variation in inter-country fertility decline for much of the last half-century pointed to divergence, rather than convergence; and that the fertility rates of countries did not begin to converge until around 1995. He showed that the decrease in fertility rates around the world did not necessarily mean that fertility rates were converging, and argued that the delayed onset of the fertility transition in many DCs is the single biggest source of divergence in the TFR. He further speculated that the convergence in health, wealth, and life expectancy may be explained by the consistent link between economic and social development, while noting that fertility is less consistently linked to development. We will pay particular attention to this issue in the following sections.

Other researchers have shown, in a cross-country perspective, that convergence in demography might occur mainly among countries with comparable socioeconomic and environmental characteristics (Mishra, Ouattara and Parhi, 2011). Using different variables, Angeli and Salvini (2009) carried out a descriptive analysis of the population characteristics of countries with low and medium Human Development Index (HDI) levels. While acknowledging that some exceptions in the convergence process emerged in terms of the mean values of the parameters, the authors emphasized the strong link between social, economic, and demographic development; and the need to examine the paths taken by different countries toward the globalization of behaviors.

Recently, Wilson (2011) argued that most demographic change over the past half-century has occurred along a "main sequence" of demographic transition, and that the great majority of the world's population are engaged in a process of convergence. According to Wilson, the principal differences between the regions of the developing world are identified when they enter this main sequence, and are based on how rapidly they move along it. Noting the similarities in the regional trends along the main sequence, he suggested that health and fertility transitions are tightly connected. He further observed that the fertility transition is a truly global process, pointing out that there is no evidence of significant reversals, and that only a few countries have yet to embark upon the transition. At the same time, however, he argued that the health transition is progressing more slowly, and cited "disturbing" evidence of its fragility, as stagnation and reversals in health affect hundreds of millions of people. He also pointed out that when it comes to the health transition, the world is not a single demographic system, but is instead divided by deep faults into a number of blocks, each of which has its own distinctive life expectancy trajectory.

In these studies, time remains an "exogenous variable," and the results are generally based on a comparison of temporal differences in parameter series. Our research fits into this theoretical and empirical framework, but our original contribution lies in the fact that time is included in the models of analysis. We will examine recent trends in the socio-demographic parameters of DCs in order to determine whether a process of convergence is occurring, and whether clusters of countries based on time trajectories of demographic behaviors can be identified.

Methods and data

The methodological approach

To study the process of convergence/divergence of the temporal dynamics of fertility and mortality, we use a method called Dynamic Factor Analysis (DFA). This method, which allows us to handle multi-way data through the joint application of a factorial analysis and regression over time, was developed in the 1970s by Coppi and Zannella (1979), and improved in the 1990s by Corazziari (1999).

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

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

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 $X(I, J, T)$, three sources of variation can be modeled, each of which depends 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, which generates a sort of structural variability, or *static*. This is the undertone of the overall variability subject to time changes due to the interaction of time with variables and with units. The second and the third sources of variation refer to time and its relationship 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 this work is on fertility and mortality indicators, more relevance is given to variables and their dynamics[1]. Thus, the dynamic of each unit over time is considered differentially: given the mean time changes of each variable, each unit will be observed in its net time variation; that is, we will observe whether it strengthens the changes in the variables; or whether it moves in other directions over time, thereby weakening or even reversing the overall dynamics.

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

$$S = {}^{*}S_i + {}^{*}S_t + S_{it}$$

where S is the overall covariance matrix of the array $X(I, J, T)$, ${}^{*}S_i$ is the covariance matrix of the centers $x_{ij.}$ representing the mean structure over time, ${}^{*}S_t$ is the covariance matrix of $x_{.jt}$, and 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 of four models, each of which employs a specific strategy in approaching the three sources of variation. Models 1, 2, and 3 are more similar with regard to the type of analysis undertaken: they all combine factorial and regression analysis to describe the three sources of variation. Model 4 analyzes the three sources of variation considering a strategy similar to that of index numbers, considering one time matrix of data as a reference to be analyzed by factorial analysis, and the obtained factorial space as the reference space on which the other occasions can be plotted.

With regard to the time evolution of the centers $x_{.jt}$, all of the first three DFA 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 zero otherwise.

The variability of the centers x_{ij} is analyzed using a factorial analysis of the specific covariance matrices in each of the three DFA models. In Model 1, a factorial analysis is applied to the covariance matrix $S_t = *S_i + S_{it}$, and it can be easily shown that S_t is also obtained as the sum of the covariance matrix of variables in each occasion, divided by the total number of times. Thus, by projecting the matrices X_t centered in each time, we obtain the factorial representation of each unit in each time—i.e., their dynamic—net of the overall mean time dynamic. The representation of the centers $x_{ij.}$ is obtained by projecting their matrix $*X_i$ centered on the factorial plane due to the decomposition of $S_t = *S_i + S_{it}$, as $*S_i$ is the covariance matrix of $*X_i$. Models 2 and 3 consider different matrixes for the analysis of the static source of variation, and for the analysis of the differential dynamic of units (Corazziari 1997;1999).

With regard to differential time evolution of the units, in Model 1 it is described by comparing the projection of each unit in each time, with the projection of the corresponding center $x_{ij.}$, according to the decomposition $S_t = *S_i + S_{it}$. In Models 2 and 3, regression techniques are used to analyze the differential time evolution of units.

In our research, we apply Model 1 of the DFA, as previously described. The projected trajectories of the units on the factorial plane can be better described by applying a cluster analysis of trajectories (Carlier 1986).

The main feature of the cluster analysis of trajectories is the type of distance between units. When studying trajectories, two types of distance between couple of units can be considered: i.e., 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 distances considered 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 are applied based on the defined distance between clusters and between units and clusters. In the present work, for both the mortality and the fertility analysis, we use the mean of the two distances, as proposed by Carlier (1986). Of the hierarchical cluster analysis methods we chose the Ward method. This choice has been confirmed by a final K-means cluster analysis based on the barycenter of the clusters of the better Ward partition.

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 modeled covariance matrix of the specific source of variation, and the corresponding observed trace, for each of the covariance matrices described above.

Data

The analyses are carried out on 103 countries with a population of at least one million in 2010, and which have been identified by the UN as being among the "less" or "least" developed countries. The data used in the analyses come from major international sources (World Health Organization, World Bank, United Nations). They, include information on the main vital events and socioeconomic indicators of the DCs. Unfortunately, it was not possible to consider the years prior to 1995 because there was too much missing data in the considered variables.

We analyze the fertility and the mortality processes separately through their main indicators. The analyses are conducted on the same set of countries for which complete data are available. As both fertility and mortality are associated in the analysis with the most commonly correlated socioeconomic variables, we can discern different "patterns" in the DCs studied. For the analysis of mortality, the variables consist of eight yearly indicators: probability of death at ages 15-60, separately for males and females; mortality of children under age five; prevalence of HIV-AIDS; Human Development Index (HDI); immunization for DPT3 (diphtheria, pertussis, and tetanus); access to improved sanitation; and a macro area variable indicating the country's geographical region[3]. For the analysis of fertility, six yearly indicators have been selected: adolescent fertility rate, TFR, contraceptive prevalence, HDI, GPI, and the macro area.

The units are the countries, and the times are the four years 1995, 2000, 2005, and 2010. Each application identifies a factor plane, the interpretation of which 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 analyzed by a cluster analysis of trajectories (Carlier 1986). As this analysis generates clusters of countries which are homogeneous in terms of the levels and the dynamics of the considered variables, the positions of these countries on the plane are easy to be interpreted. If the clusters move toward the center of the plane - which characterizes the overall dynamic of the system of data - homogeneity is increasing; i.e., a convergence process is underway. By contrast, if the clusters move away from the center, heterogeneity is increasing.

Results

Mortality and health

The results of the DFA indicate that the first two components of the factor analysis explain a large portion of the variability (I_t=94.8%), and that the best-represented times are the second and the third ones (respectively 2000 and 2005, with percentages of 95.5 and 95.1, respectively.

With regard to the correlation between variables and factors (Table 1) we find a strong negative correlation of the first component with both the probabilities of death at adult ages (separately for men and women), and the prevalence of HIV-AIDS. Meanwhile, 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 with adult mortality (separately for men and women) is also positive, but the values are low. A strong negative correlation of the second component is shown with HDI, access to improved sanitation, and immunization for DPT3. In sum, the first component assumes the meaning of mortality and morbidity (increasing values mean decreasing mortality and morbidity rates), while the second component represents the sanitary conditions and the overall health status of observed countries (increasing values of the components mean a worsening of such conditions). 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 being 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
Human Development Index (HDI)	0.40	-0.83

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

Before commenting on the factorial results, we will discuss the overall dynamics of centers, in order to better explain 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

A polynomial in t (time) of the second order is required only for the variable for HIV-AIDS prevalence and for the two indicators of adult mortality; for all of the other variables (as well as for the variables in the analysis of fertility), a simple linear regression model in t is fitted. The average (over units) indicators of mortality decrease over the considered period, especially child mortality; while the indicators of HDI, DPT3 immunization, and access to improved sanitation increase. The mean prevalence of HIV-AIDS also 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 leads us to choose a partition formed by seven clusters (both when using the hierarchical Ward method, and then the K-means method for confirmation, very similar aggregations of countries and very few exceptions are found). 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

9

shown here (Table 3). A geographic concentration is quite evident for five of these clusters, while two (the first and the second cluster) include countries located in different macro areas.

Made up of 52 countries in various macro areas, Cluster 7 is the largest of the clusters. The main features of this cluster are that it has low levels of adult mortality and under-five mortality, and low HIV-AIDS prevalence levels. Thus, it appears to be the leading cluster in terms of immunization coverage, access to improved sanitation, and HDI values; and to include the countries with the best survival, health, and socioeconomic conditions. Cluster 7 may also be used as a reference group if the convergence dynamic is questioned. Cluster 6 includes 19 countries, which again belong to various macro areas. The main feature of this cluster is that its levels of adult and under-five mortality are higher than in Cluster 7, but lower than in the other groups of countries. Cluster 5, which contains 18 Sub-Saharan countries plus Afghanistan, displays significant delays in the health transition. This cluster has the lowest levels of immunization and access to sanitation. It is above all characterized by high levels of under-five mortality and some of the lowest HDI values. Cluster 4, which is made up of only five Sub-Saharan countries, has medium-high mortality levels, and intermediate positions for the other variables. Cluster 3, which includes five Sub-Saharan countries, has the second-highest under-five mortality rates, but medium levels of immunization and access to sanitation. The only country in Cluster 2 is Zimbabwe, which is characterized by very high adult mortality and a high prevalence of HIV-AIDS. This disease is also widespread in Cluster 1, which consists of only two countries (Botswana and Swaziland). Although the last three countries have roughly similar mortality and morbidity conditions, they are divided into different clusters due to the different dynamics these variables display over the time, as can be seen in Figure 1.

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, 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 P.D.R., Madagascar, Mauritania, Mongolia, Myanmar, Nepal, Pakistan, Papua New Guinea, 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

Based on the DFA method, the differential dynamic of median centers of clusters is represented in Figure 1. Generally speaking, the trajectories of the clusters on the factor plane show a trend toward the average situation; indeed, they show weak differential dynamics toward the center of the axes. Cluster n. 7, 6 and 5 are located in different quarters of the factorial plan, based on the

different values of the active variables. However, when we look at the temporal trends of each cluster, we can see that they have a slight tendency to converge toward the center of the axes of the factorial plan. In other words, the dynamics of the countries included in such clusters (which represent almost 90% of the countries included in the analysis) show a slow reduction in their differentials as they move toward the centers of the variables under investigation.

The pattern of convergence toward the center of the axes is much more evident in Clusters 2 and 4. HIV-AIDS prevalence decreased significantly in these clusters in the period studied: i.e., from 25% to 15.2% in Cluster 2, and from 9.3% to 5.8% in Cluster 4.

Meanwhile, conditions are worsening in Clusters 3 and 1 that moves leftward in the plane, and thus toward the highest levels of the aforementioned variables. This trend appears to be moving in the opposite direction with respect to the decreasing mean of the general mortality and 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
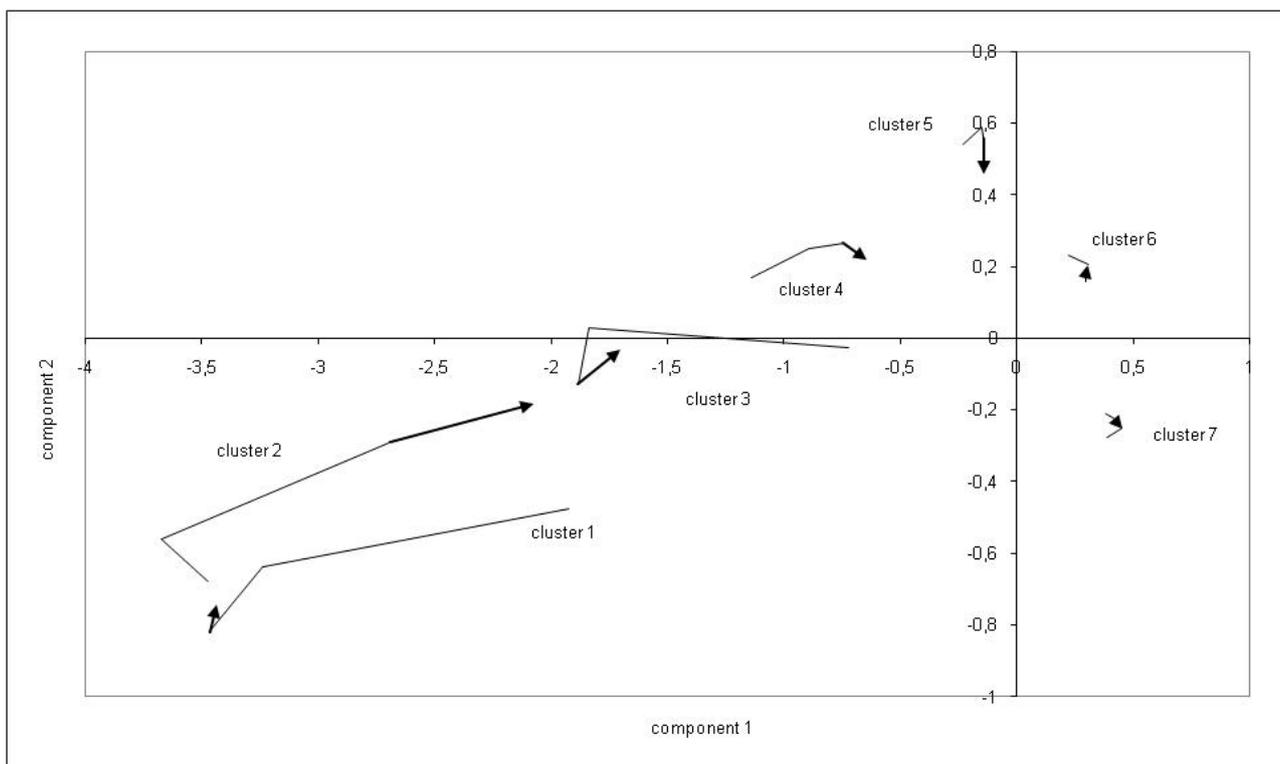
Note: cl_i means cluster i, t_i means time i.
SOURCE: Our elaboration on data of World Health Organization, World Bank, United Nations

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, or a little less than in the mortality analysis, in which the variables are slightly more numerous. For fertility, the third and the fourth times are better represented (2005 and 2010, respectively, with percentages of explained variability of 86.5 and 85.6).

In terms of the correlation of the two components with the active variables (Table 4), positive values for the first component indicate better situations in terms of human development, contraceptive prevalence, and GPI; while negative values indicate higher total and adolescent fertility rates.

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

Looking at the correlations between factor components and variables shown in Table 5, we can see the division outlined above between adolescent fertility and TFR on the one side; and GPI, HDI, and contraceptive prevalence on the other. In sum, the values for variables measuring development are the opposite of those for variables measuring fertility behavior, as at increasing levels of development, fertility generally declines.

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 explain the trajectories of countries over the factorial plane, a look at the overall dynamic of the data is necessary. Table 5 shows the time regression parameters of the indicator means. The two indicators of fertility are decreasing over time, while the indicators that describe development are increasing on average in the considered period. Based on these dynamics, the positions and trajectories of the aggregated clusters are to be described (Table 6).

For fertility, the cluster analysis of the countries trajectories projected on the plane have again led us to choose a partition formed by seven clusters (again, both the hierarchical Ward method and the K-mean method for confirmation produced very similar aggregations of countries, and with very few exceptions). The number of the clusters expresses the ranking based on the mean value of TFR (from the highest to the lowest), and thus generally describes 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, Congo, Côte d'Ivoire, Ethiopia, Gambia, Guinea-Bissau, Madagascar, Malawi, Mozambique, Nigeria, Senegal, Tanzania, Yemen, Zambia..
3	Burundi, Cambodia, Ghana, Lao P.D.R., Mauritania, Pakistan, Rwanda, Sudan, Togo.
4	Bangladesh, Gabon, Guatemala, Haiti, India, Iraq, Kenya, Namibia, Nepal, Papua New Guinea, 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, Saudi Arabia, Tajikistan, Tunisia, Ukraine, Uzbekistan.
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 clusters chosen for the partition are well separated, and are of comparable sizes (e.g., clusters 2, 5, 6, and 7 are, respectively, made up of 17, 18, 19, and 20 countries). Clusters 1 and 2 are the most distinct geographically, as they include only Sub-Saharan countries, with the exception of Afghanistan and Yemen. These countries are characterized by a strong delay in the demographic transition: i.e., they have high levels of total fertility and adolescent fertility, and low HDI values. Adolescent fertility, which is strongly associated with women empowerment and female human capital, may be considered a proxy for women's status, and thus for the degree of growth of the country. Cluster 3, which includes countries in southern Asia and Sub-Saharan Africa, represents an intermediate stage on the path toward modernization. Cluster 4 is geographically more heterogeneous: it includes three Asian countries (India, Bangladesh, and Nepal), some African countries, a Central American country (Guatemala) and a Caribbean country (Haiti). This cluster has a high level of adolescent fertility. Cluster 5, which is made up of Latin American countries, has lower fertility, higher contraception use, and relatively high HDI and GPI values. Clusters 6 and 7 include countries located on different continents, but which nevertheless have similar demographic characteristics. Cluster 6 includes several Asian countries that were previously part of the Soviet Union, and which are therefore currently undergoing processes of economic and demographic transition. This cluster also includes countries located along the South-Eastern shore of the Mediterranean with demographic characteristics that approach those of European countries, but which are less advanced in terms of women's status and stage of development, as measured by their HDI and GPI scores. In Cluster 7, China and Iran are examples of countries that have completed the fertility transition, with TFR values that are currently below replacement level. This cluster, which also includes some countries located on the South-Eastern side of the Mediterranean, has low fertility, high levels of contraception use, and high HDI and GPI values. In sum, Clusters 5, 6, and 7

have higher values in terms of development and the use of contraception, while Clusters 1 and 2 have lower values in terms of development.

In Figure 2, we show 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. Clusters 4, 6, and 7 show a slight tendency to converge toward the average of all of the countries in the period we examine, represented by the barycenter of the axes. This means that, to some extent, the differential dynamic of countries included in these clusters tends to be similar to the dynamic of the centers of variables that determine the factorial plan. Cluster 1 starts from the worst position and goes toward the barycenter. Clusters 2, 3, and 5 do not show a high degree of variability over time.

FIGURE 2 Differential dynamics of median centers of clusters – Fertility
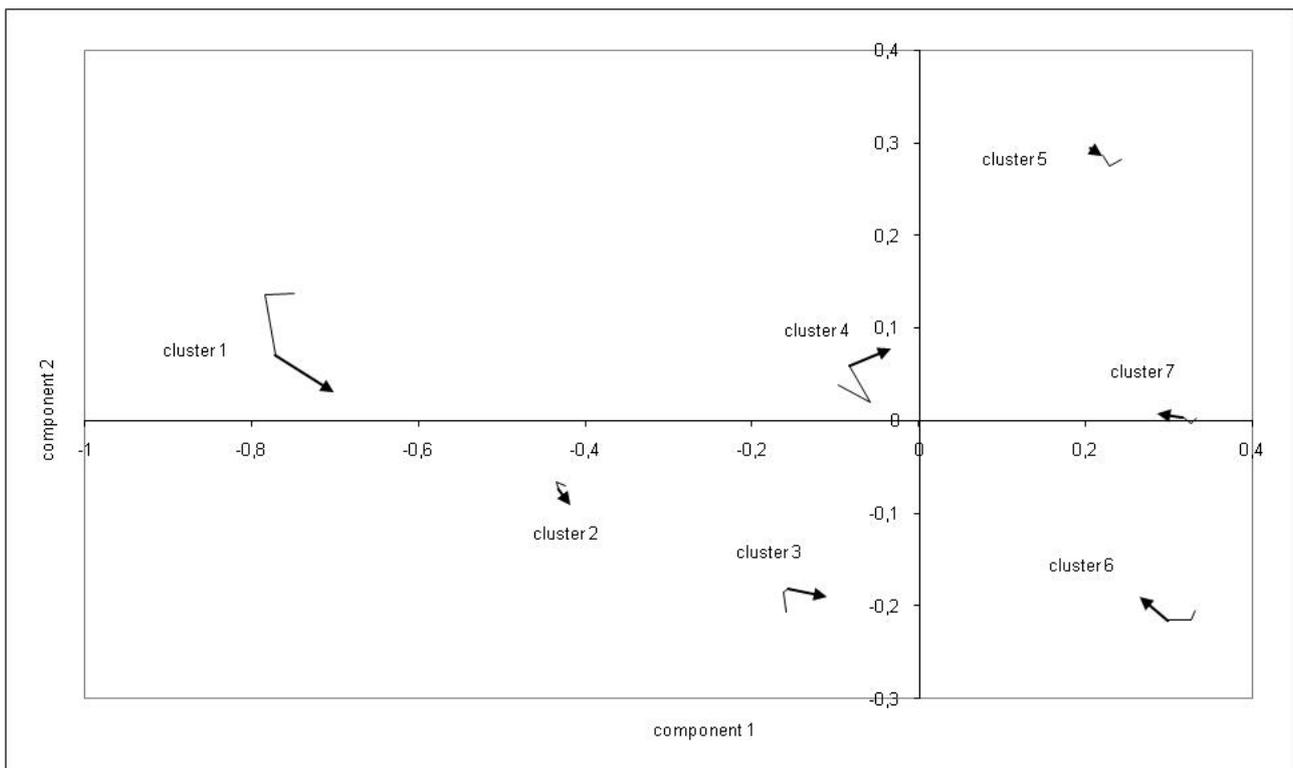
Note: cl$_i$ means cluster i, t$_i$ means time i.
SOURCE: Our elaboration on data of World Health Organization, World Bank, United Nations

Discussion

In this study, we have examined the trends of specific demographic parameters regarding fertility and mortality in DCs, together with some socioeconomic variables, in order to evaluate whether demographic behavior in these countries is converging, or whether 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 significantly over space and time, we applied DFA and Cluster Analysis of trajectories in order to evaluate at the macro level the main demographic trends of DCs between 1995 and 2010.

When we look at the results for mortality, we can see that the largest cluster includes 52 countries, and is characterized mainly by low levels of adult and under-five mortality, low HIV-

AIDS prevalence levels, high levels of immunization coverage and access to improved sanitation, and high HDI values. Moreover, the countries included in this cluster have the best survival, health, and socioeconomic conditions; and thus the highest development levels. In this cluster, a convergence in modernization characteristics is underway. Another cluster consisting of countries located in different areas also has low mortality levels, albeit higher than those of the previous cluster. All of the remaining clusters are more geographically homogeneous, and have higher levels of adult and under-five mortality. Three clusters in particular include Sub-Saharan countries and have high HIV-AIDS prevalence levels, and thus have the highest mortality levels.

The dynamic analysis of mortality therefore seems to confirm our assumption that a large number of DCs are converging toward a uniform model of health and mortality, and are leaving certain other countries behind. These countries, almost all of which are located in Sub-Saharan Africa, have high levels of HIV-AIDS prevalence, and thus have high levels of mortality. In fact, when we look at the trajectories of all the clusters, we observe that most of them have a slight tendency to converge toward the center of the axes of the factorial plan, and are thus moving toward uniformity.

The analysis of fertility indicators leads us to different considerations. Our findings generated well-separated clusters of comparable magnitude. The two clusters with a strong delay in the demographic transition were geographically distinct, consisting mostly of Sub-Saharan countries. In the relatively homogenous cluster that includes most of the Latin American countries, the demographic transition is well-established: i.e., fertility is low, contraception use as high, and the HDI and GPI development indicators are relatively high. We also identified a cluster with several Asian countries (previously part of the Soviet Union) in which the process of economic transition is currently underway. Some of the variable values for this cluster are similar to those of Latin America, but the mechanisms governing the path toward modernization are quite different. Finally, there is a more modernized, geographically heterogeneous cluster made up of countries like China and Iran, which have reached the last stage of the fertility transition, and which have TFRs 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 sum, although many countries show a process of convergence group by group, there is no common situation toward which they are converging. Moreover, a relatively large number of countries continue to follow models that are specific to certain territorial, cultural, and behavioral contexts.

If we compare the overall processes of convergence in mortality and in fertility, we can see that they differ markedly. Our mortality analysis identified a large cluster of countries that encompassed half of all of the DCs studied. This suggests that these countries are very similar in terms of their trajectories, and are thus undergoing a strong process of convergence. The other, less numerous clusters have a range of dynamic characteristics which diverge from those of the largest cluster. In terms of fertility, we did not find a large homogeneous group, but rather many clusters of similar sizes. Thus, fertility does not display a process of convergence similar to that of mortality. As mortality and fertility do not appear to have similar patterns, we can assume that health and mortality have reached a more advanced stage of convergence.

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

First, concerning mortality, our results are in line with those of researchers who noted the occurrence of rising and converging survival levels (Wilson, 2001). In line with Wilson's (2011) observation that the world is not a single demographic system, but is divided by deep faults into a number of blocks, each of which has its own distinctive trajectory of life expectancy, we identified heterogeneous clusters of countries. To this discussion, we add the insight that there is a large group

of countries which can be seen as a "point of arrival" when a convergence dynamic is questioned. Second, consistent with McMichael et al. (2004), we show some reversals of previous trends in some groups of countries, due in large part to the spread of HIV-AIDS, primarily in Sub-Saharan Africa. Nevertheless, in line with Moser, Shkolnikov, and Leon (2005), we found that divergence is far less common than convergence and that duration of periods of reversal mortality trend are shorter than those of the recovery of survival..

For fertility, our outcomes confirm Casterline's findings (2001) that there is inter-country and intra-regional variation in the pace of fertility decline. Furthermore, our outcomes are consistent with those of Dorius (2008) and of Wilson and Pison (2004), who found that the trends in 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, proven by several of the abovementioned studies (Mishra, Ouattara, and Parhi, 2011; Angeli and Salvini, 2009), between economic, social, cultural, and behavioral characteristics, as well as between levels of development and demographic convergence processes.

There is a further similarity between our results and the results of the analysis performed by Wilson (2001). For convergence in mortality, he found a consistent link between economic and social development; whereas for fertility he found that these variables are less consistently linked. In conclusion, our findings suggest that mortality depends above all on modernization, and that the convergence trends in mortality are heavily influenced by socioeconomic evolution. At the same time, and in line with Dorius (2008), our analysis shows that fertility convergence depends mostly on cultural variables, i.e., on attitudes and traditions, which are changing more slowly than economic development indicators.

Notes

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

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.

3 The macro area is conceived as a proxy variable summarizing other characteristics of each country not expressed by the chosen indicators, and which limits the clustering of countries that are geographically distant. It was conceived of as a variable that takes into account both the homogeneity of countries belonging to the same macro area, and, to a lesser extent, the proximities of different macro areas. Thus, for example, a macro area with a value of one is closer to a macro area with a value of than other macro areas. A disjunctive set of variables in which each refers to a single macro area was considered as an alternative. As it provided similar but more scattered results, the macro area unique variable has been used in the analysis.

References

Angeli, Aurora and Silvana Salvini. 2009. "Demographic Heterogeneity of Developing Countries between Reality and Convergence Myth," *Proceedings of XLII Scientific Meeting of Italian Association for Population Studies*, Milano, 2-4 february.

_____.1992. "Convergence". *Journal of Political Economy* 100, n. 2: 223-51.

Barro, Robert J., and Xavier Sala-i-Martin.1995. *Economic growth*, New York: McGraw-Hill.

Becker, Gary S., Philipson Tomas J. and Soares Rodrigo 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, François and Christian Morrisson. 2002. "Inequality among world citizens: 1890–1992," *American Economic Review* 92(4): 727–744.

Carlier, A. 1986. "Factor Analysis of Evolution and Cluster Methods on Trajectories," in Francesco de Antoni, Natale Lauro, and Alfredo Rizzi (Eds.), *Compstat*, Heidelberg: Physica Verlag: 140-145.

Casterline, John. 2001. "The pace of fertility transition: National patterns in the second half of the twentieth century," in Rodolfo Bulatao e John Casterline (eds.), *Global fertility transition, A Supplement to Population and Development Review* 27: 17-52.

Chesnais, Jean-Claude. 1986. *La transition démographique : étapes, formes, implications économiques*, Paris: Presses Universitaires de France - INED.

_____. 1997. "La transition démographique: trente ans de bouleversements," in Jean-Claude Chasteland and Jean-Claude Chesnais (eds.) *La Population du Monde: Enjeux et Problemes*. Paris: Presses Universitaires de France - INED.

Coppi, Renato. 1994. "An Introduction to Multiway Data and their Applications," *Computational Statistics and data analysis*, 18(4): 3-13.

Coppi, Renato and Franco Zannella. 1979. "L'analisi fattoriale di una serie temporale multipla relativa allo stesso insieme di unità statistiche," *Proceedings of the 29th Meeting of the Italian Statistical Society* 2, Bologna, 20-22 March 1978.

Coppi, Renato and Sergio Bolasco, 1989. *Multiway Data Analysis*, Amsterdam: North Holland.

Corazziari, Isabella. 1997. Dynamic Factor Analysis, Classification and data Analysis, *Proceedings of the Italian Statistical Society, Part III*, Pescara 3-4 luglio 1997: 171-178.

_____. 1999. *Analisi di Dati Longitudinali Multivariati*, Tesi di Dottorato XI ciclo, Napoli, aprile 1999.

Coughlin, Richard M. 2000. "Convergence theories," in Edgar F. Borgatta e Marie L. Borgatta (eds.), *Encyclopedia of sociology*, New York: Macmillan: 422-431.

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

Easterlin, Richard A. 2000. "The worldwide standard of living since 1800," *The Journal of Economic Perspectives* 14: 7-26.

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

Firebaugh, Glenn and Brian Goesling. 2004. "Accounting for the recent decline in global income inequality", *The American Journal of Sociology* 11: 283–312.

Grigoriev, Pavel, France Meslé, Vladimir M. Shkolnikov, Evgeny Andreev, Agnieszka Fihel, Marketa Pechholdova and Jacques Vallin. 2014. "The Recent Mortality Decline in Russia: Beginning of the Cardiovascular Revolution?", *Population and Development Review* 40(1): 107-129.

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

La Croix, Sumner J., Andrew Mason, and Abe Shigeyuki. 2002. "Population and Globalization," *Southeast Asian Studies* 40(3): 240-267.

Law Gail H., Conrad Wesley Snyder, Jr., Hattie, John A. and Roderick P. McDonald. (Eds.). 1984. *Research Method For Multimode Data Analysis*, Praeger, New York.

Lehmijoki, Ulla and Jenni Pääkkönen. 2006. "Demographic Clubs: Convergence within, divergence between," Helsinki Center of Economic Research Discussion Papers 136, in https://helda.helsinki.fi/bitstream/handle/10138/16714/demograp.pdf?sequence=1.

McMichael, Antony J., Martin McKee, Vladimir Shkolnikov, and Tapani Valkonen. 2004. "Mortality trends and setbacks: global convergence or divergence?" *The Lancet* 363(9415): 1155-1159.

Mishra, Tapas, Bazoumana Ouattara, and Mamata Parhi. 2011. "Cross-country convergence in demography under stochastic shocks," School of Business and Economics, Swansea University, Working Paper in Economics SBE-E/2011/01.

Moser, Kath, Vladimir Shkolnikov, and David A. Leon. 2005. "World mortality 1950–2000: divergence replaces convergence from the late 1980s," *Bulletin of the World Health Organization* 83(3): 202-209.

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

_____. 2004. "HIV/AIDS and cross-national convergence in life expectancy," *Population and Development Review*, 30: 727-742.

Oeppen, Jim. 1999. "The health and wealth of nations since 1820," *Paper presented at 1999 Meeting of the Social Science History Association,* Fort Worth, TX, November 13.

Pradhan, Memmo, David E. Sahn, and Stephen D. Younger. 2003. "Decomposing world health inequality," *Journal of Health Economics* 22 (2): 271-293.

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

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

Salvini, Silvana. 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, Italy, 9-11 June 2004, in http://old.sis-statistica.org/files/pdf/atti/RSBa2004p293-304.pdf.

Solow, Robert M. 1956. "A Contribution to the Theory of economic Growth," *Quarterly Journal of economics* 70: 65-94.

Tucker, Ledyard R. 1966. "Some Mathematical Notes On Three-Mode Factor Analysis," *Psychometrika*, 31-3: 279-311.

Wilson, Chris. 2001. "On the scale of global demographic convergence 1950-2000," Population and Development Review 27(1): 55-171.

_____. 2011. "Understanding global demographic convergence since 1950," *Population and Development Review* 37(2): 375-388.

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