THE EVOLUTION OF WORLD WELFARE INEQUALITY

Davide Fiaschi[*] *and Marzia Romanelli*[**]

*The paper proposes a measure of countries' well-being based on individuals' lifetime utility and applies it to a large sample of countries in the period 1960-2011. Together with a decreasing trend in welfare inequality across world populations, we find clear evidence of polarization with the formation of three groups: those with high welfare levels, those in transition towards the upper part of the distributions and those "trapped" at medium-low levels. Such tendencies to polarization shall strengthen in the future, jointly with an increase in the world welfare inequality. We also suggest a method to take into account within country-inequality along the two relevant dimensions of welfare we are considering, namely income and health (*i.e., *life expectancy). The analysis not only confirms the evidence in favour of polarization but also points to a level of inequality remarkably higher.*

1 Introduction

Despite the wide consensus on the multidimensionality of human well-being, most of the studies that analyse the dynamics of world inequality mainly focus on the distribution of income or consumption alone. Bourguignon and Morrisson (2002) and Becker *et al.* (2005), however, have argued how a more meaningful analysis of the evolution of welfare inequality across countries/among world citizens should jointly consider at least the dynamics of income and life expectancy, even by simply looking at some composite indicator of welfare such as lifetime income or utility.[1] In particular, Bourguignon and Morrisson (2002) observe that inequality in the per capita GDP across the world population increased from the beginning of the 19th century to World War II, and then stabilized (or slightly increased). On the contrary inequality in life expectancy decreased markedly after 1920-1930. Taking lifetime income as a proxy of welfare, they conclude that the decreasing trend observed in welfare inequality since 1950 has stopped since the main determinant of such dynamics, *i.e.*, the pronounced drop in life expectancy disparities, has lost its momentum or even reversed its path. Becker *et al.* (2005) propose a more sophisticated approach to the measurement of welfare based on the concept of lifetime utility as previously discussed in Rosen (1988), computing the countries' "full income" growth rates, *i.e.*, growth rates which include the monetary value of the gains in longevity experienced by countries' populations. They conclude in favour of an even stronger convergence in the world welfare distribution over the period 1960-2000 – with the partial exception of the populations from Sub-Saharan countries – than the one that would emerge looking at income alone.

In this paper we make both a theoretical and an empirical contribution to the current literature: we propose a methodology to measure welfare based on the lifetime utility of individuals; we then apply it to a large sample of countries to assess the evolution of world inequality of well-being using non-parametric techniques to identify the possible emergence of polarisation.

---

[*] Department of Economics, University of Pisa.

[**] Bank of Italy, Economic Research and Statistics Directorate. The views expressed in this paper are those of the authors and do not necessarily reflect those of the institutions to which they are affiliated. The usual disclaimer applies.

[1] Many different approaches to the measurement of multidimensional well-being have been proposed in the literature so far, based on very diverse definitions of well-being itself. Those measures range from the identification of a single (usually utility-based) "sufficient statistics" for welfare to a dashboard of non-comparable dimensions. For a critical review of the literature, see Aaberge and Brandolini (2015).

Starting from the concept of lifetime utility of Becker *et al.* (2005), we directly consider the indirect utility function as a cardinal index of welfare. Our approach brings some advantages, among which the potential inclusion of the expected income growth rates on the determinants of welfare, and, mainly, the possibility to directly compare welfare across populations.[2]

Using such an index, we find evidence of a decreasing trend in welfare inequality (as Becker *et al.* (2005)), but also of a strong pattern of polarization. Polarization in welfare is more pronounced than the one characterising income distribution, and is expected to persist in the future. In particular, we first consider as a proxy of the world distribution of welfare, the population-weighted cross-country distribution (in the following, "cross-population distribution" or "PWCC") in the period 1960-2011. We find a clear pattern of polarization with the emergence of three clusters in 2011, together with of a fall in welfare inequality. The populations of Sub-Saharan countries represent the poorest part of the low-welfare cluster, but the most of the mass is constituted by the populations of South Asia. The second cluster is in a relatively higher position and, with the notable exception of China, is constituted by Latin American populations. The upper cluster is instead mainly formed by populations from Western Europe and Western Offshoots and some Asian Tigers (Hong Kong, Korea and Singapore).

The clusters differ by some typical features discussed in literature as determinant of poverty traps, as the level of life expectancy, the degree of social conflict, the quality of institutions and the quality of labour force (human capital).[3]

The estimate of the long-run tendencies suggests that polarization should be a persistent phenomenon, while welfare inequality is expected to increase in the future. This expected pattern results from a stop in income and life expectancy convergence across the medium and high-income populations, and from a divergent dynamics of the lower income group.

We show how the polarization in welfare appears the result of the polarization in the cross-population distribution of income,[4] jointly with the positive relationship between income and life expectancy also at high levels of per capita GDP.[5] The complementarity between life expectancy and income in our welfare index implies that under a general upward trend of per capita income, a constant absolute difference in life expectancy, as the one observed between medium and high-income countries, leads to an increasing gap between welfare levels. In our sample this divergent dynamics is indeed only partially counterbalanced by the (recent) higher income growth rates of medium-income countries.

The estimate of the cross-population distribution disregards within-country disparities. However, there is increasing evidence that such inequalities both in income and life expectancy can be sizeable and changing over time. A more comprehensive approach should aim at directly considering the entire world population, ranking the individuals from the poorest to the richest irrespective of their nationality.[6] We then make a further step ahead with respect to Becker *et al.*

---

[2] The methodology proposed by Becker *et al.* (2005) allows to compute only variations in "full income"; absolute welfare levels can be computed only if income and "full income" are assumed to coincide in a chosen base year (Becker *et al.* (2005) consider 1960 as base year, p. 283). Fleurbaey (2005) discusses how the income equivalent variations may depend on the choice of the base year, and, in turn, how it may lead to intransitive comparisons. An alternative method directly compute money-metric indices, which however need arbitrary references for both the income and non-income dimensions to be fixed. Fleurbaey and Gaulier (2009) applied such a method to 24 OECD countries in 2004, obtaining a ranking which strongly differs from the one based only on per capita GDP. Similarly the work by Jones and Klenow (2010) considers 134 countries and takes into account consumption, life expectancy as well as leisure and inequality.

[3] See Durlauf *et al.* (2005).

[4] See Quah (1997) and, for a more recent evidence, Vollmer *et al.* (2010)).

[5] This evidence partially contrasts with the so called "Preston curve" (see Preston (2007)), which points to convergence in life expectancy for medium/high-income countries.

[6] In terms of Milanovic's taxonomy the cross-population distribution corresponds to *Concept 2 inequality*, while this second approach is labelled as *Concept 3 inequality*.

(2005), estimating the world population distribution (in the following, “WP”) of welfare by taking into account also within-country inequalities in the period 1993-2005.[7] With respect to the estimates based on the cross-population distribution, the world welfare inequality as measured by the Gini index appears to be remarkably higher (by 8 percentage points on average), but the qualitative pattern of its dynamics is confirmed: decreasing inequality over time and evidence of polarization.

Bourguignon and Morrisson (2002) and Becker *et al.* (2005) are the main sources of inspiration of the paper. Our theoretical model follows the approach in Rosen (1988), while the empirical analysis is inspired by the work of Danny Quah on income distribution and club-convergence dynamics (see, for example, Quah (1997)).

Our methodology is strongly related to the recent literature which proposes a more theoretically grounded approach towards the analysis of non-market dimensions of inequality and the evaluation of gains in quality and quantity of life.[8]

In the estimate of individual welfare by lifetime utility we are close to Murphy and Topel (2006); their goal, however, is different, since they aim at valuing improvements in overall longevity and health care. From a theoretical point of view Anderson (2005) presents a similar framework: however, no randomness in the length of life is considered; moreover, the empirical analysis is limited to African countries. Finally, Nordhaus (2003) and Hall and Jones (2007) provide stimulating discussions on the evaluation of welfare associated to extensions in life expectancy.

The non-parametric methodology used in the empirical analysis is based on Fiaschi and Lavezzi (2003). The estimate of the long-run distribution follows Johnson (2005), thus avoiding the discretization of the state space. In addition, we propose a novel bootstrap procedure to identify confidence intervals for the estimation of the long-run distributions.

The paper is organized as follows: Section 2 presents the theoretical measure; Section 3 reports and discusses the empirical results; Section 4 concludes. The appendices contain proofs, some extensions of the analysis, and other technicalities.

2 A measure of individual welfare

The measure of individual welfare we propose is based on the model in Rosen (1988) with state dependent utility. In particular, we apply it in a framework with long-run growth and CIES instantaneous utility function, in order to calculate an explicit formulation of the lifetime utility of agents. Consider an agent born at time 0 with a maximum length of life equal to T and a positive probability of dying before $T > 0$. Given her initial wealth, $\bar{p}_0$, and a flow of potential labour incomes ($yl_0, yl_1, ..., yl_T$), the intertemporal budget constraint of the agent is:

$$\int_0^T c_t \exp(-rt) S_t dt \leq w, \tag{1}$$

where r is the interest rate, S_t the probability to survive at age t, and w is the lifetime wealth of the agent, given by:

$$w = \bar{p}_0 + \int_0^T yl_t \exp(-rt) S_t dt. \tag{2}$$

7 Limitation in data availability constraints the time span we can consider.

8 For a review on these issues, cf. Decancq *et al.* (2015) and Weil (2014).

We assume that r is constant over time and non-negative.

Budget constraint (1) assumes full annuity insurance, or the existence of a complete contingent claims market (see Becker *et al.* (2005)): the agent can borrow in perfect capital markets all her potential future labour incomes at the current interest rate r, and the survival function S is common knowledge across all the agents in the economy.

When the agent is alive, her preferences are described by the following *CIES* instantaneous utility function:[9]

$$u(c) = \begin{cases} \frac{c^{1-\sigma}}{1-\sigma} - M & \text{for} \quad \sigma > 0 \text{ and } \sigma \neq 1; \\ \log(c) - M & \text{for} \quad \sigma = 1, \end{cases} \tag{3}$$

Preferences (3) depend on two additive components: a constant term, M, which represents the utility of the state "dead",[10] and the term $c1-\sigma$ / $(1 - \sigma)$ describing the utility of the state "alive".[11] Subtracting M from utility in each state (both "dead" and "alive") normalizes the utility of non-survival to zero.

If $\sigma \in (0, 1)$ and $M < 0$ being alive has a positive utility *per se*; the agent would prefer a longer life independently of her consumption level. On the contrary, if $\sigma > 1$, then M should be negative, otherwise $u(c) < 0$ for all c and therefore "dead" would be always the preferred state of the agent. We therefore assume that:[12]

$$\begin{aligned} &1. \quad \text{if } \sigma \in (0,1) \text{ then } M > 0; \\ &2. \quad \text{if } \sigma = 1 \text{ then } M \in (-\infty, +\infty); \text{ and} \\ &3. \quad \text{if } \sigma > 1 \text{ then } M < 0. \end{aligned} \tag{4}$$

Under Assumption (4) there exists a *zero utility consumption*, c^{ZUC}, such that $u(c^{ZUC}) = 0$, *i.e.*:

$$c^{ZUC} = [(1-\sigma) M]^{\frac{1}{1-\sigma}}. \tag{5}$$

The expected utility of the agent is given by:[13]

$$E[U] = \int_0^T \left(\frac{c^{1-\sigma}}{1-\sigma} - M \right) \exp(-\rho t) S dt, \tag{6}$$

where ρ is the discount rate.

Assume that:[14]

$$\dot{S}/S = -\pi^D, \tag{7}$$

---

9 The form of the utility function for $\sigma \to 1$ in Eq. (3) is obtained by adding the constant term $-1/(1-\sigma)$ to the term $c^{1-\sigma}/(1-\sigma)$.

10 The presence of the constant term M allows the utility elasticity to decline with consumption. Under reasonable assumptions on the parameters' values, this implies that an agent would eventually prefer to substitute consumption with additional years of life.

11 The latter term is commonly used in the literature on economic growth, because it ensures constant growth rates in steady state.

12 Rosen (1988), p. 287, argues that the economically interesting cases are those for which the elasticity of the instantaneous utility function $\varepsilon \in (0, 1]$. This corresponds to the cases: i) if $\sigma \in (0, 1)$ then $M > 0$ or ii) if $\sigma > 1$ then $M < 0$.

13 In the following, we omit the time index whenever it does not cause confusion.

14 See Nordhaus (2003) for a similar framework

where $\pi^D > 0$ is the mortality rate. Under Assumption 7 life expectancy at birth (*i.e.*, at time $t = 0$) is given by:

$$LE = \frac{1 - \exp\left(-\pi^D T\right)}{\pi^D}. \tag{8}$$

If $T \to \infty$ then $LE = 1/\pi^D$, while if $\pi^D = 0$ then $LE = T$.

We also assume that the agent's expected labour income grows at a rate equal to the steady-state growth rate g, *i.e.*:[15]

$$yl_t = yl_0 \exp\left(gt\right) \text{ for } t \in [0, T]. \tag{9}$$

When the agent has no initial wealth, *i.e.*, $\bar{p}_0 = 0$, her indirect lifetime utility is given by:[16]

$$V\left(T, yl_0, g\right) = \left(\frac{1}{1-\sigma}\right)\left\{yl_0^{1-\sigma}\left[\frac{\exp\left(\left(g-\hat{r}\right)T\right)-1}{g-\hat{r}}\right] + \frac{\left(1-\sigma\right)M\left[\exp\left(-\hat{\rho}T\right)-1\right]}{\hat{\rho}}\right\}, \tag{10}$$

where $\hat{r} = r + \pi^D$ and $\hat{\rho} = \rho + \pi^D$ are respectively the interest rate and the discount rate adjusted for the instantaneous probability of dying before T.[17]

In our analysis, V is considered as a direct index of human well-being. We depart from Becker *et al.* (2005), whose index of well-being is the sum of per capita GDP in 1960 plus the gains in both material income and longevity expressed in "full income" variations, assuming that in 1960 "full income" and income coincide. In the empirical analysis, under the hypothesis of equal preferences across world population, the two approaches lead to the same results.

A key feature of lifetime utility in Eq. (10) is that income and life expectancy are complements, which means that the same gain life expectancy is valued more by rich individuals than by poor ones (both in absolute and relative terms). This element has been partially embodied also in the new formulation of the *Human Development Index* (HDI), which before the revision showed the opposite (and mostly criticized) feature, *i.e.*, income and life expectancy were pure substitutes. The HDI retains however the drawback of the lack of a clear microfoundation (cf. Weil (2014), p. 668). The same objection applies to the more recent *OECD Better Life Index*. In this regard our index based on lifetime utility overcomes this limit.[18] As we will discuss below, this has relevant implication for the analysis: under a general upward trend of per capita income, a constant absolute difference in life expectancy, as the one we will observe between medium and high-income countries, leads to an increasing gap between welfare levels.

3 Empirical evidence

This section studies the evolution of world inequality in welfare, per capita GDP and life expectancy and their distribution dynamics. Ideally, in order to derive the proper distribution we

[15] For the sake of simplicity, in Eq. (9) we are considering that the agent works over her whole life; however, the analysis could be easily extended to the case in which the agent retires at age T^R, with $T^R \in (0, T]$.

[16] See Appendix A for the details.

[17] Lifetime utility V can be a non-monotonic function of life expectancy. The parameters' setting adopted in the paper excludes such possibility. We refer to Fiaschi and Romanelli (2010) for a more detailed analysis of this issue.

[18] On the other side, we do not incorporate dimensions other than health and income (consumption) in measuring welfare while both of those indices include other aspects, more or less correlated with income, such as education, environmental quality, civic engagement, etc.

should estimate the welfare of each individuals in the world. This would require a tremendous amount of microdata which is so far not available. Our *first* approximation is then deriving an estimate of the population-weighted welfare distribution among countries.

3.1 Methodology of the empirical investigation

As in Becker *et al.* (2005), the welfare of a population in a given country is assumed to be equal to the (indirect) lifetime utility of a representative agent with no initial wealth, *i.e.*, $\bar{p}_0 = 0$, whose first yearly income, yl_0, is proxied by the per capita GDP of that country and whose life expectancy, *LE*, is equal to the average life expectancy at birth of its citizens; it is therefore equal to the utility of a representative newborn.

Given our welfare indicator for each country in each year of the considered time-span (1960-2011), we can estimate the population-weighted distributions over time. Such analysis provides a picture of the dynamics of inequality across individuals and possibly allows to identify the emergence of clusters of populations. Such estimates contain a bias since they neglect the *within-country* distribution of welfare.[19] However, in Section 3.4 we will show for the period 1993-2005 how the inclusion in the analysis of the *within-country* distribution of welfare substantially confirm our findings.

Concerning the empirical analysis, we depart from Becker *et al.* (2005) in a key methodological aspect: the use of non-parametric techniques, which crucially affects the results because of the presence of non-linearities in the distribution dynamics. As discussed by Durlauf and Johnson (1995) the presence of σ (or, in our case, Gini) and β (absolute) convergence does not exclude the existence of multiple equilibria, *i.e.*, polarization.

Finally, Eq. (10) shows that a proper estimate of the welfare distribution should take into account all the non-linearities between growth rates, income and life expectancy, especially in presence of high cross-country heterogeneity in income growth rates. However, estimating g for a given country in a given year is not a simple task, because it should represent the *expected* income growth rate for a newborn in that country in that year. This suggests to limit the analysis to the baseline case of $g = 0$.[20] We checked the sensitivity of our results to the assumption $g = 0$ by considering non-null country-specific growth rates. The picture is qualitatively confirmed, that is the presence of polarization, even tough with higher welfare inequality. Therefore, cross-country heterogeneity in income growth rates does seem only to exert a second-order effect on the dynamics of welfare inequality.[21]

3.2 Calibration of individual welfare

The sample in the empirical analysis includes 103 countries, for which we have complete information on per capita GDP, life expectancy and population size for the period 1960-2011. Countries' GDP is measured by the expenditure-based real GDP at chained PPPs in 2005

---

19 Bourguignon and Morrisson (2002) show that in modern economic history the within-country component was the main source of inequality in per capita GDP until World War II, accounting for almost 3/4 of total inequality on average. However, since the 1950s, its contribution to world inequality has been halved, given that the dynamics of between-country inequality is the leading factor in determining inequality across world citizens.

20 In fact, the decomposition of changes in welfare into additive separable components, namely changes in income and changes in life expectancy or in other non-income dimensions, as for example in Becker *et al.* (2005) or in Jones and Klenow (2010), relies on such assumption.

21 For the sake of brevity, we omit to report such robustness check here. For more details, we refer the interested reader to Section 4 in Fiaschi and Romanelli (2009).

international prices (I$) drawn from Penn World Table 8.1 (PWT 8.1); the population is taken from the same dataset, while life expectancy at birth comes from the 31st January 2015 release of the World Development Indicators (WDI 2014).[22]

For the model parameters, we use almost the same set as in Becker *et al.* (2005); in particular $\rho = 0.005$, $\pi^D = 0$, so that $LE = T$,[23] and $\sigma = 1/1.250$. For what concerns the estimation of M, we derive it from Eq. (5), setting c^{ZUC} equal to the minimum level of per capita GPD observed in our sample (*i.e.*, I$225.2 for Nigeria in 1995; which implies that $M = 14.8$). This setting represents a lower bound: indeed, no country (not even Nigeria) displays a per capita GDP permanently lower than that (remind that no agent in any case would be willing to consume permanently less than c^{ZUC} and still survive). An alternative specification is proposed by Becker *et al.* (2005), who calibrate M using parameters values estimated from the U.S. economy: specifically, $\varepsilon = u'(c)\, c/u(c) = 0.346$ and $c = 31,439$I$ in 1990, from which $M = 16.7$.[24] The implied zero utility consumption, c^{ZUC}, would be equal to I$419 (see Eq. (5)): an individual whose consumption in every period is equal to I$419 would be indifferent between living or dying independently of her life expectancy. In our sample there are 3 countries for which per capita GDP would be lower than I$421 for at least 20 per cent of the time span (Mozambique, Democratic Republic of the Congo, Nigeria). This leads us to focus on the first and more conservative calibration for *M*. However, the findings discussed below appear robust to alternative specification of the model's parameters.[25] What could make a difference is a (implausible) value of c^{ZUC} close to zero, which would determine a collapse in the Gini index.

As discussed above, a country's welfare is computed by Eq. (10) assuming $g = 0$.[26]

In order to gain an intuition of the relationships between per capita GDP, life expectancy and welfare, Figure 1 displays a series of level curves for welfare in the space *(per capita GDP*, *life expectancy)*. It also reports the positions of some representative countries in 1980 (diamond) and in 2011 (grey circle).

Between 1980 and 2011, for example, Cote d'Ivoire and Democratic Republic of the Congo show a decrease in their welfare, while China and India a large increase. Some developed countries present a relatively high increase in their life expectancy (Italy and Japan), while others a relatively marked increase in their per capita GDP (*i.e.*, the United States). The numbers reported in the three triangles along the dashed line (which represents an estimation of the "Preston Curve" in 2011) are the marginal rates of substitution (MRS) between life expectancy and per capita GDP (expressed in ten 2005 international dollars). As expected, at very low levels of life expectancy and per capita GDP (respectively around 35 years and I$440), individuals value income relatively more than life expectancy (*i.e.*, individuals value I$10 more in each year of their life equal to 1.4 years of life expectancy at birth). Instead, at very high level of life expectancy and per capita GDP (respectively 82 years and I$39100), the opposite occurs (*i.e.*, individuals value I$10 per year equal

---

22 Appendix B reports the country list; expenditure-side real GDP at chained PPPs in 2005 international prices: variable *rgdpe* in Penn World Table 8.1, see http://www.ggdc.net/pwt; population: variable *pop* in Penn World Table 8.1; life expectancy at birth: variable *SP.DYN.LE00.IN* in World Development Indicators, see http://databank.worldbank.org/data/home.aspx/

23 An alternative specification could consider $T \to \infty$, from which $LE = 1/\pi^D$, thus setting π^D equal to the inverse of the observed life expectancy, in the estimates of the agent's utility. All the empirical results reported below are robust to this alternative specification.

24 Indeed, from Eq. (3) $M = c^{(1-\sigma)}\,[1/(1-\sigma) - 1/\varepsilon]$.

25 See Fiaschi and Romanelli (2009) for a broader discussion.

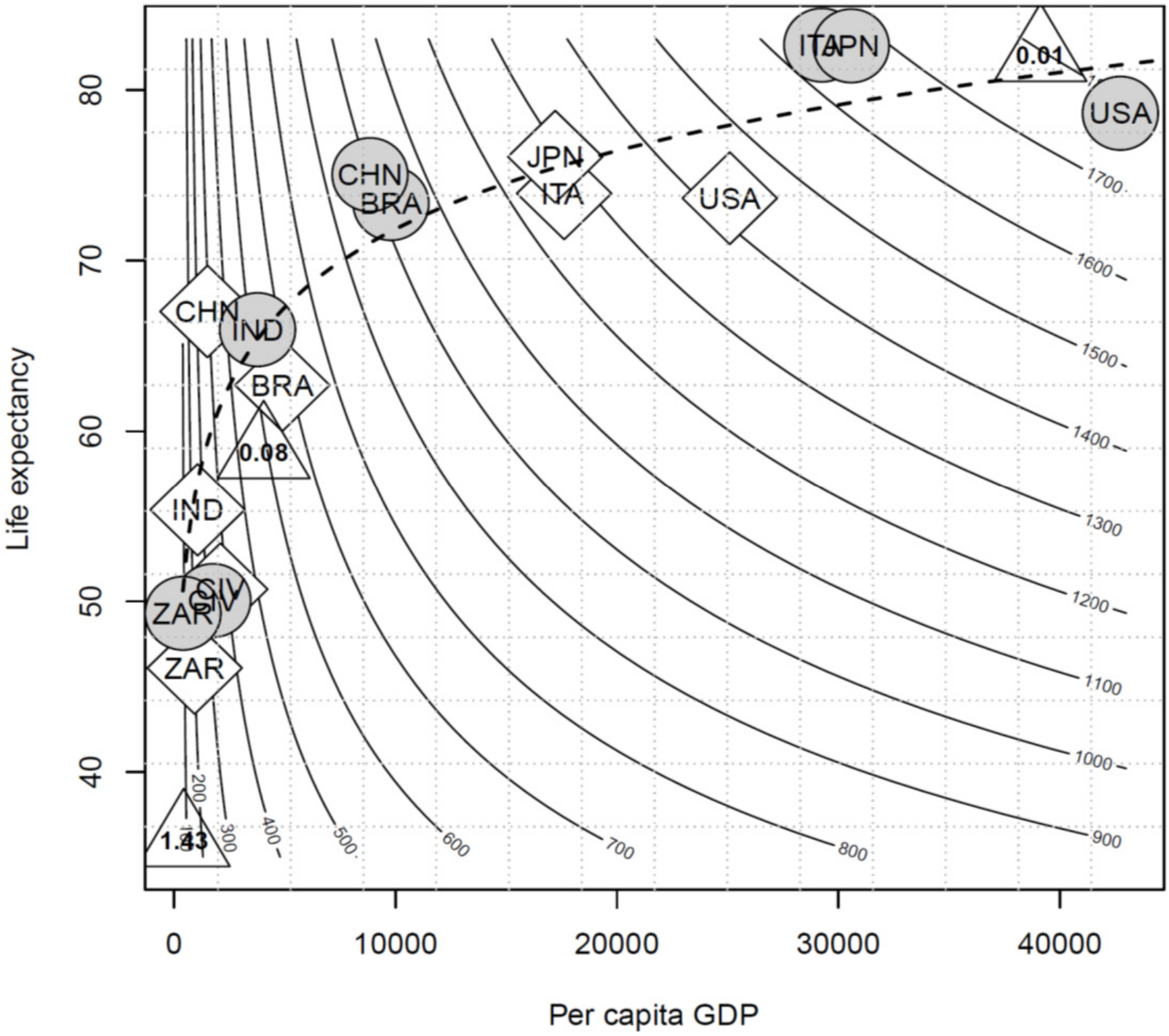
26 For example, the expected welfare of a representative American newborn in 2011 is:

$$V_{US} = \left(\frac{1}{1-\sigma}\right)\left\{\frac{\exp\left(-\rho LE_{US}\right)-1}{\rho}\left[(1-\sigma)\,M - yl_{US}^{1-\sigma}\right]\right\} = 1782.1,$$

where yl_{US} = I$42734 and $LE_{US} = 78.64$.

Figure 1

Welfare Calculated with *g* = 0 for a Sample of Countries in 1980 (Diamond) and in 2011 (Grey Circle)
(ten international dollars)



Country codes: Brazil (BRA), China (CHN), Cote d'Ivoire (CIV), India (IND), Italy (ITA), Japan (JPN), Democratic Republic of the Congo (ZAR), United States (USA).
Numbers in Triangles are the Marginal Rate of Substitution Between Life Expectancy and per capita GDP.

to 0.01 years of life expectancy at birth).[27] The marginal rate of substitution in the bottom part of the distribution clearly depends on c^{ZUC}: for example, if c^{ZUC} is around I$100 at the same low level of life expectancy and per capita GDP (35 years and I$440)individuals value I$10 per year equal to 0.7 years of life expectancy at birth, while at high level of per capita GDP and life expectancy (82 years and I$39100) the MRS remains unchanged (at 0.01). The latter finding is not surprising, given that for rich people the level of c^{ZUC} is almost irrelevant.

[27] This feature stems from the fact that while marginal utility of consumption decreases, that of life expectancy does not. Hall and Jones (2007) discuss such element as an explanation for the increasing size in health care expenditure the richer the country.

Table 1

Descriptive Statistics for the Sample's Variables

Year		**1960**	**1980**	**2000**	**2007**	**2011**
Per capita GDP						
	Mean	3536	5779	8612	10493	10962
	Gini	0.56	0.60	0.60	0.54	0.51
	Top 5%	0.323	0.272	0.254	0.232	0.222
	Bottom 5%	0.005	0.005	0.003	0.004	0.004
Life expectancy						
	Mean	51	63	68	70	71
	Gin	0.12	0.07	0.07	0.06	0.06
Welfare (g = 0)						
	Mean	425	588	753	872	917
	Gini	0.40	0.35	0.32	0.28	0.26
	Top 5%	0.203	0.152	0.128	0.119	0.117
	Bottom 5%	0.009	0.012	0.006	0.01	0.011
Pop						
	Total (*millions*)	2548	3789	5272	5752	6024

Table 1 reports some descriptive statistics of the sample, including a set of inequality indices for selected years (1960, 1980, 2000, 2007 and 2011). Following the standard in the literature on income distribution, inequality is measured in relative terms, even though we are aware of the possible important consequences of such choice in our analysis with variables generally growing over time. For example, if the average welfare is increasing over time a constant relative inequality would mean an increasing absolute inequality.[28]

Inequality in both per capita GDP and life expectancy across populations decreased markedly from 1960 to 2011, with the inequality of per capita GDP always higher than the one of life expectancy. Accordingly, we can also observe a strong reduction in the inequality of welfare and a level that is systematically lower than that of income inequality. However, looking at two sub-periods, namely 1960-1980 and 1980-2011, per capita GDP and life expectancy seem to follow two different patterns: inequality in income first rose and then started declining, while disparities in life expectancy shrank dramatically in the first sub-period and then remained substantially constant. This is consistent with Ram (2006) who finds in fact a reversal in the convergence dynamics of life expectancy at the country level after 1980 (see also Bloom and Canning (2007) and Becker *et al.* (2005) for similar findings). This is also the reason why we will focus on such two sub-periods to elicit long-run tendencies.

---

28 See Anand and Segal (2008) for a discussion on this issue.

3.3 Distribution dynamics of welfare

To further investigate the evolution of welfare inequality over time, we use the non-parametric methodology proposed in Fiaschi and Lavezzi (2003). In particular, Section 3.3.1 reports the estimated growth path of welfare so to detect possible non-linearities, a necessary condition for the presence of polarization; Section 3.3.2 then analyses how the distribution of welfare has changed, estimating also the evolution of the joint dynamics of per capita GDP and life expectancy over time and the related stochastic kernels; and, finally, Section 3.3.4 discusses the long-run tendencies by comparing the actual distributions and the estimated ergodic distributions.

3.1.1 Con(Di)vergence in welfare

Figures 2-3 report the population-weighted estimate of the growth paths of welfare. In particular, they show the estimate of Model (11) over different time-spans, where *x* is the log of welfare level.

$$\overline{GR}_i^x = m\left(x_i^{INI}\right) + \epsilon_i; \tag{11}$$

$\overline{GR}_i^x$ is the average growth rate of *x* in country *i* in a given period, x_i^{INI} is the initial value of *x* and ε_i is an independently distributed random variable with zero mean.[29]

The estimate of *m* (.) is made using the Nadaraya-Watson estimator with the optimal normal bandwidth.[30]

A note of caution is needed. It is well-known that in presence of measurement errors related to the initial value of *x*, the linear estimate of Model (11) can be biased in favour of convergence (*i.e.*, at low level of x is associate a higher growth rate). Heuristically, non-parametric regressions, given their nature of "local" regressions, should be more robust to the presence of non-classical measurement errors, in particular larger errors in the lower tail of the distribution, because they would not affect the whole range of the variable; however, the problem still remains.[31]

The growth path welfare is estimated for the whole period 1960-2011 and for the subperiod 1980-2011.[32] The figures report the cross-population estimates, where the weights used are the population sizes at the initial year. Dotted lines represent the pointwise confidence intervals at 95 per cent (see Ha¨rdle *et al.* (2004)) and the red line signals the overall annual average growth rate. We also report countries' observations by circles, whose area is proportional to their population at the initial year (the country-codes reported in the figures refer to the top ten countries by population size). Finally, Sub-Saharan countries are represented by grey circles.

---

29 Usually, the relationship between the income growth rates of a cross-section of countries and their levels of income is called "growth path" because, under the assumption of an equal stochastic process governing income growth in all countries, this relationship should represent the path followed by each country in its development. With a slight abuse we use the same denomination for the case of our welfare measure.

30 All the calculations and estimates in the paper are made using R. The estimates of nonparametric regressions are made using the package *sm* (see Bowman and Azzalini (2005)).

31 For example, one of the two main components of the welfare measure, that is life expectancy, can suffer from an upward bias particularly relevant at lower levels of the variable and which could decline over time, affecting the estimates for poor countries, see Becker *et al.* (2005), p. 278.

32 We also performed the same estimation for per capita GDP and life expectancy separately. In the case of life expectancy, the growth rate is replaced by the average difference. The estimates for the subperiod 1960-1980 and all the estimates for per capita GDP and life expectancy are available upon request.

Figure 2

Growth Path for Welfare in 1960-2011

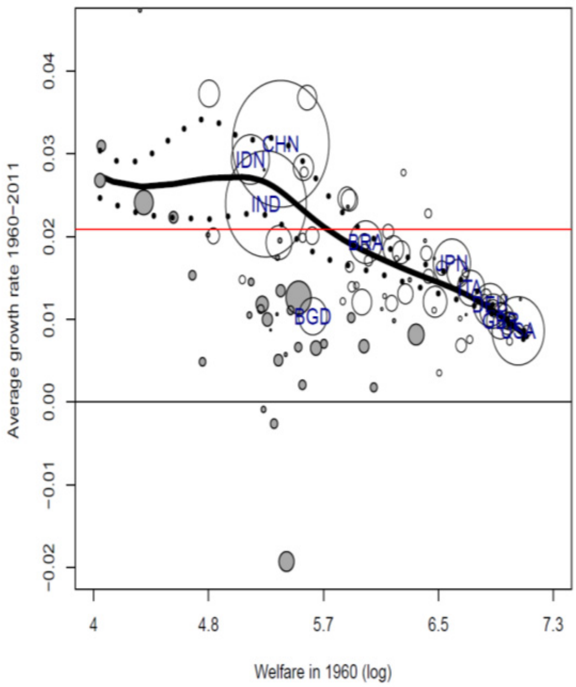
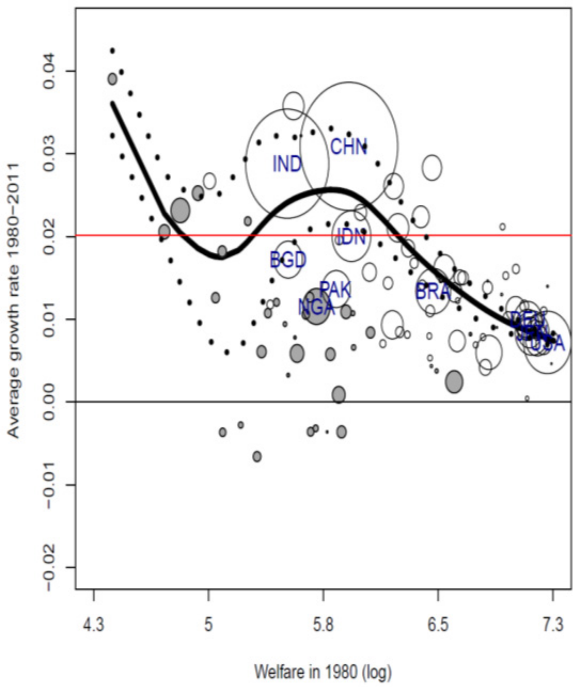


Figure 3

Growth Path for Welfare in 1980-2011



For the entire timespan 1960-2011, the growth path of welfare points to convergence, at least at medium-low levels of welfare (see Figure 2). The main driver of such dynamics is the evolution characterising the populations of some of the largest (and still poor in 1960) countries in the world, such as China, India and Indonesia,[33] and in particular their spectacular performance both in terms of income growth rates and life expectancy gains.

However, at lower levels of welfare the caching up process seems less robust and some populations appear instead to be getting trapped into middle-welfare levels. Focusing on the period 1980-2011 (Figure 3), such club-convergence dynamics appears clearer. As some of the Asian largest countries continue along their convergence path, other large populations with similar welfare levels (for example, those from Bangladesh or Pakistan) get relatively stuck. Indeed they have not over-performed compared to the people from high-income countries, so that the gap between those populations and the rich people has been growing in absolute terms.

A specific case can then be made for the populations of Sub-Saharan countries, whose wellbeing is rather diverging, with general stagnant or even negative growth rates. This is owed both to their gloomy performance in terms of GDP growth rates compared with that of China and India and to their very small increases in life expectancy mainly due to AIDS epidemics which had, and, unfortunately, continue to have, a devastating impact on mortality rates in the area (see, e.g., Bloom and Canning (2007)). Such evidence is not substantially reverted even when we take a look at the years of the "African growth miracle", which has characterised African countries' income growth rates in the first decade of the XXI century (Rodrik (2014)): even though a light convergence toward the medium-welfare club could be detected (Figure 4), it seems to have lost its momentum after the beginning of the Great Recession (Figure 5).

---

[33] They represent almost 50 per cent of the total population in the sample in 2011.

Figure 4

Growth Path for Welfare in 2000-2011

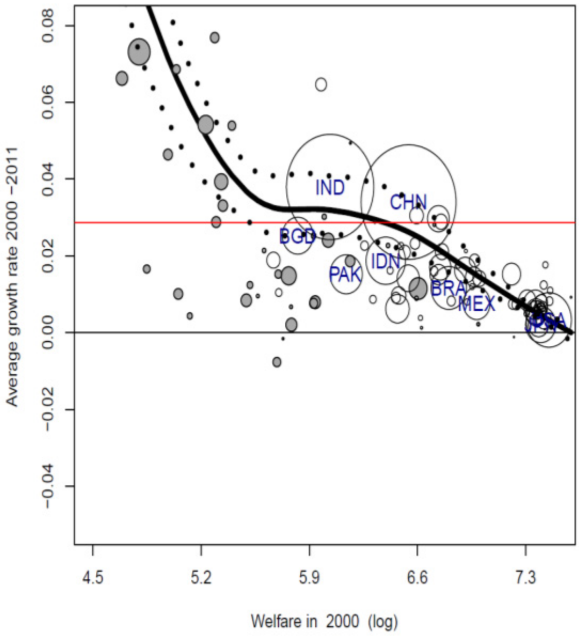
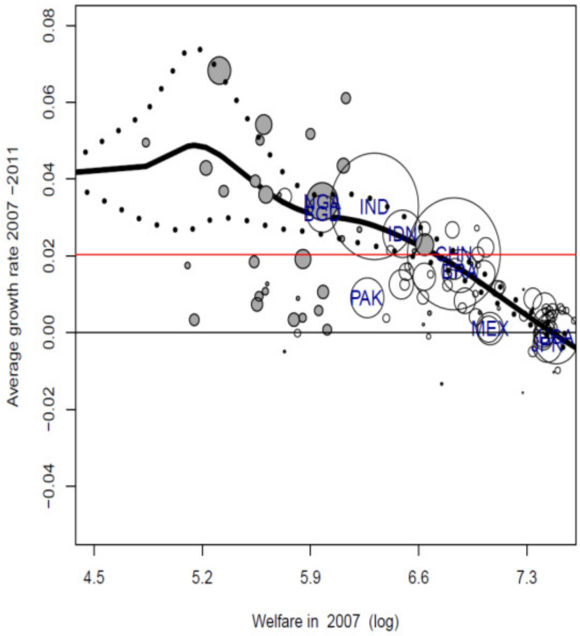


Figure 5

Growth Path for Welfare in 2007-2011



Looking separately at the two welfare components, namely income and life expectancy, a clear convergence path across the world population over the entire period considered (1960-2011) is observable only at very high levels of per capita GDP. However, focusing on the period 1980 onwards, the picture is rather different, with convergence regarding populations only at low-medium level of income. This pattern is mainly due to the high growth rates of four big Asian countries, Bangladesh, China, India and Indonesia, as already mentioned. Instead, very poor people, *i.e.*, people from Sub-Saharan countries, tends to diverge.

On the contrary, life expectancy across population shows a clear path of convergence between 1960 and 2011, driven by large gains in life expectancy of highly populated Asian countries. Again, since 1980 things seem to change and convergence stops. The population of the Sub-Saharan countries are left behind, and no convergence of the people with medium life expectancy to those with high expectations occur. Also high life expectancy countries stop converging. Various explanations have been proposed, among which the increasing difficulties in transferring medical technology among countries with respect to the past (e.g., immunization and antibiotics), and the different role of the governments in the health system.[34]

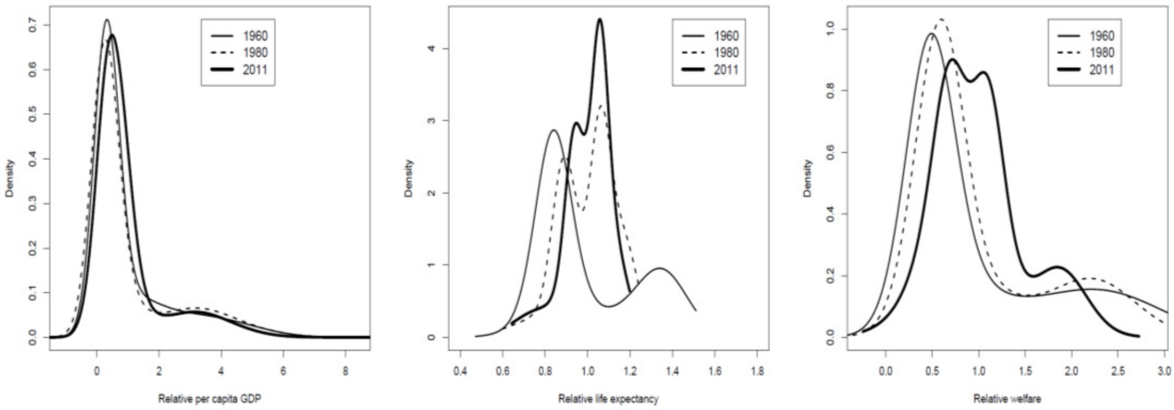
Overall the evidence suggests the presence of polarization across world population. In particular, the ’80ies seem to mark a change in the dynamics of convergence. The evolution of welfare appear highly non-linear and affected by a strong cross-country heterogeneity. The next section discusses the implications for the distribution dynamics.

[34] We refer to Easterlin (2004), Cap.7, and Becker *et al.* (2005) for a more detailed discussion of the possible causes.

Figure 6

Estimated Distributions

(a) Relative GDP **(b) Relative Life Expectancy** **(c) Relative Welfare**



3.3.2 The evolution of the distribution of per capita GDP, life expectancy and welfare from 1960 to 2011

In the following we first report estimates of the distribution of welfare in three significant moments – at the beginning, in the middle and at the end of the period considered (Figure 6) – and then we analyse the dynamics of such distribution focusing on the period 1980 onwards. In particular, for this second step we estimate the evolution of the joint distribution of per capita GDP and life expectancy and then the stochastic kernel for welfare, so to take into account non-linearities.[35]

In estimating densities, we use the *adaptive kernel estimation* with the Gaussian kernel as suggested by Silverman (1986).[36]

Turning to the results, we already noted that inequality of per capita GDP among the world population decreased between 1960 and 2011 (actually the declining trend started in 1980). The Gini index indeed falls slightly but significantly from 0.56 in 1960 to 0.51 in 2011 (see Table 2). Looking at the distributions of relative GDP (Figure 6a), apparently they seem to be always single-peaked (around 0.5) with a thick right tail in all three years, even though as time goes by a second peak around 3.5 becomes more and more evident: indeed tests for the presence of multimodality in the per capita GDP distributions suggest that while unimodality cannot be rejected for the distribution in 1960, bimodality is instead a likely feature already in 1980 (see Table 3).[37] This in turn points to a stronger identification of at least two clusters of populations.

The picture for life expectancy is slightly different (Figure 6b). Inequality decreases from 1960 to 1980, and then remains steady. The Gini index almost halves in the first twenty years considered (from 0.12 in 1960 to 0.07 in 1980; see Table 2) and then stops. Polarization is clearly present since 1960, as suggested by the multimodality tests which support the presence of multiple modes in the distribution from the very beginning (see Table 3). However, the two groups (*i.e.*, the two modes), although neatly separated, tend to be closer over time.

---

[35] The stochastic kernels of per capita GDP and life expectancy are not reported for the sake of brevity. They are all available upon request.

[36] See Appendix C.

[37] For the 1980 distribution, the null hypothesis of unimodality is rejected with a p-value of 0.024, while the null hypothesis of bimodality would be rejected only with a p-value equal to 0.346. Details on the tests of multimodality are presented in Appendix D.

Table 2

The Gini Index of the Distributions of Per Capita GDP, Life Expectancy and Welfare (g = 0)

Year	GDP	Life exp.	Welfare ($g = 0$)
1960	0.56 ** (0.015)	0.12 *** (0.006)	0.40 *** (0.011)
1980	0.60 *** (0.013)	0.07 ** (0.005)	0.35 *** (0.012)
2011	0.51 (0.022)	0.06 (0.005)	0.26 (0.017)

Note: Standard errors are in parentheses. The results of the test on the equality between Gini indices (base-year 2011) are reported as follows: ”#”15 per cent significance level, ”*” 10 per cent significance level, ”**” 5 per cent and ”***” 1 per cent.

Table 3

P-value of the Null-hypothesis of Unimodality and Bimodality of the Cross-population Distribution of Per Capita GDP, Life Expectancy and Welfare

	Unimodality Test			Bimodality Test		
Year	**GDP**	**Life Expectancy**	**Welfare**	**GDP**	**Life Expectancy**	**Welfare**
1960	0.722	0.000	0.022	0.374	0.194	0.528
1980	0.024	0.026	0.016	0.346	0.036	0.272
2011	0.020	0.028	0.016	0.342	0.130	0.000

As a result of the dynamics of per capita GDP and life expectancy, the inequality of the cross-population distribution of welfare decreases remarkably, while clusterization strengthens over time. Not only all the distributions are two-peaked, but the 2011 distribution seems to be characterised by the emergence of a third peak (supported also by the tests for multimodality), made of some of the populations in the lower welfare group who turn out to be less able to catch up (Figure 6c).

Both the growth paths and the distribution estimates support the idea of a polarization or club convergence dynamics of welfare across populations, besides an overall reduction in inequality. In particular, such dynamics starts realizing in the ’80ies. Per capita GDP tends to polarize and life expectancy stops converging. The evidence is towards the formation of at least two clusters over the last 3 decades, with the possibility of a third cluster in the lower part of the distribution of welfare. Analysing the evolution of the joint dynamics of (relative) per capita GDP and life expectancy between 1980 and 2011 across populations can shed some lights on what forces are at play.

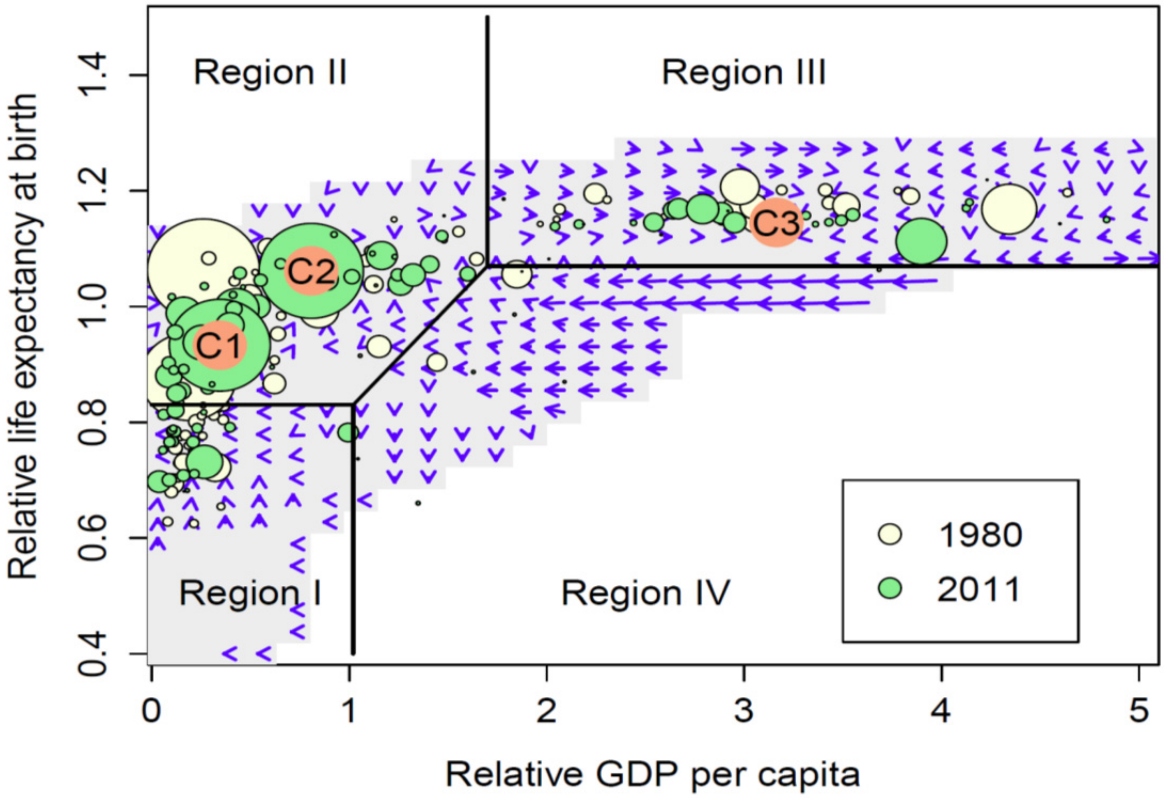
3.3.3 *Club-convergence in welfare*

Figure 7 depicts a vector field, where the arrows indicate the direction and magnitude of

Figure 7

The Joint Dynamics of Relative Per Capita GDP and Relative Life Expectancy, 1980-2011

(circles represent countries in 1980 (light yellow) and in 2011 (light green) and their size is proportional to countries' populations)



the joint dynamics of per capita GDP and life expectancy at different points in the space *(per capita GDP*, *life expectancy)*.[38]

Circles, representing countries observations in 1980 (light yellow) and in 2011 (light green), are proportional to the size of the countries population.

Four regions are defined on the basis of the pattern of the arrows: in particular, the frontiers of the regions are drawn where the vector field displays divergent dynamics. The Sub-Saharan countries lay in Region I, the highly populated countries (*i.e.*, China and India) are located in Region II and the OECD countries in Region III. Basically, no country is located in Region IV, suggesting that a high per capita GDP is always associated with a long life expectancy.

From 1980 to 2011 the distribution of populations across the four regions changes in favour of Region II: the probability mass varies from (0.1, 0.72, 0.16, 0.02), respectively, in Region I, II, III and IV in 1980 to (0.09, 0.75, 0.16, 0.0) in 2011. The change mainly reflects the transition into Region II of some large Sub-Saharan populations, such as Ethiopian and Tanzanian. Mobility across regions however is very low (with the obvious exception of Region IV, which is basically empty): the probabilities that an individual in Region I, II, III and IV were in the same region in 1980 and in 2011 are respectively equal to (0.64, 0.97, 1, 0).

[38] For the methodology used, refer to Appendix E.

In terms of per capita GDP at least two clusters of populations seem to exist in 2011, one in Region II (*i.e.*, populations with relative per capita GDP of around 0.5) and the other one in Region III (*i.e.*, populations with relative per capita GDP of around 3).[39] Similarly, the distribution of life expectancy shows at least two clusters in 2011, one in Region II (around 1.0) and one in Region III (around 1.15). The joint distribution of life expectancy and per capita GDP, therefore, suggests the existence of (at least) two clusters of populations also in terms of welfare. However, looking at both lower levels of per capita GDP (around 0.10.2) and life expectancy (around 0.8) a non-negligible mass of countries can be detected, pointing to the possible presence of a third cluster (in line with the observation drawn from the analysis of the welfare distribution in 2011, cf. Figure 6c).

For descriptive purposes only, we applied the *k-medians algorithm* to the observations in 2011 assuming the existence of such three clusters;[40] the centroids of these three possible clusters are located in C1 = (0.34, 0.93), C2 = (0.81, 1.06) and C3 = (3.16, 1.15): we refer to Appendix B for the list of countries (and their share of the world population) in the different clusters.

Cluster 1 is centred at low levels of per capita GDP (about 34 per cent of the average) and life expectancy (about 93 per cent of the average); it is mainly composed by populations from Sub-Saharan countries, some very large countries in South Asia, like India, Indonesia and Bangladesh, and few North African countries, like Egypt and Morocco. All low-income populations present on average also a low life expectancy, as suggested by the Preston curve. Cluster 2 is centred at relatively low levels of per capita GDP (about 80 per cent of the average) and medium-high levels of life expectancy (around 106 per cent of the average); apart from China, the cluster is mainly composed of Latin American populations and people from Central Asia. Finally, Cluster 3 is centred at high levels of per capita GDP and life expectancy (both variables are well above the average, *i.e.*, 316 and 115 per cent); the cluster is formed by OECD countries located in Western Europe, Western Offshoots and by some Asian Tigers, like Hong Kong, Korea and Singapore. The three clusters therefore appear to have a strong regional characterization.

Table 4 reports some descriptive statistics of the three clusters.[41] Cluster 1 only partially fits the description made by Collier (2007) of the poverty trap: even tough it is characterised by very low income and life expectancy levels, a relatively high level of social conflict, lowquality institutions and governance and the lowest level of human capital, income growth rates are not dissimilar to that of Cluster 3 and output does not seem to rely only on natural resources (as suggested by the even share of manufactures exports of the total of merchandise exports). Moreover, saving rates are substantial.

Also Cluster 2 seems to be partially plagued by high political instability and social conflict, as well as by low-quality institutions. However, it presents a very high level of savings, a higher stock of human capital, a higher share of output deriving from manufactures and a lower population growth rate. Moreover, the growth rate on average is by far the largest. Overall this results in substantially higher levels of both per capita income and life expectancy with respect to Cluster 1. Finally, Cluster 3 is, by far, the cluster with the highest living standards under several points of view (e.g., not only for the high level of per capita income and life expectancy, but also for less growth volatility, low intensity of social conflicts, etc). Moreover, remarkably larger than

---

39 Quah (1997) finds a similar feature.

40 See Leisch (2006) for details. We choose the *k-medians algorithm* since its objective is to minimize the total intra-cluster absolute distance; it thus appears more robust to outliers than *k-means algorithm*.

41 In particular, we report some average characteristics of the countries belonging to the three different clusters in 2011, weighted by populations' size. Apart from average income, average life expectancy and population growth, we consider the average volatility of the income growth within the clusters, indicators of capital accumulation (gross fixed capital formation) and human capital accumulation (the share of the labour force with at least a secondary or a tertiary degree) and measures of the quality of governance and political instability.

Table 4

Descriptive Statistics for the Three Clusters of Countries in 2011 Selected Variables

	Cluster 1	**Cluster 2**	**Cluster 3**
Average per Capita GDP (PPP $2005)	3,240	9,944	35,792
Average Growth Rate of per Capita GDP 1980-2011 *(annual percent)*	2.73	4.5	2.15
Stand. Dev. of the Growth Rate of per Capita GDP 1980-2011	1.76	1.92	1.04
Average Life Expectancy	64	74	81
Gross capital formation *(percent of GDP)*	27.60	36.61	20.43
GG public expenditure *(percent of GDP)*	23.32	28.26	41.78
Total health expenditure *(percent of GDP)*	4.22	5.71	12.28
Labor force with secondary education *(percent of total)*	28.68	36.79	44.55
Labor force with tertiary education *(percent of total)*	14.50	20.27	29.55
Manufactures exports *(percent of merchandise exports)*	50.41	74.36	70.29
Political Stability and Absence of Violence *(percentile rank)*	13.78	29.04	69.03
Regulatory Quality *(percentile rank)*	35.62	47.04	88.24
Rule of Law *(percentile rank)*	37.88	44.42	88.41
Population growth *(annual percent)*	1.71	0.69	0.56
Population *(percent of total)*	47.28	36.52	16.20

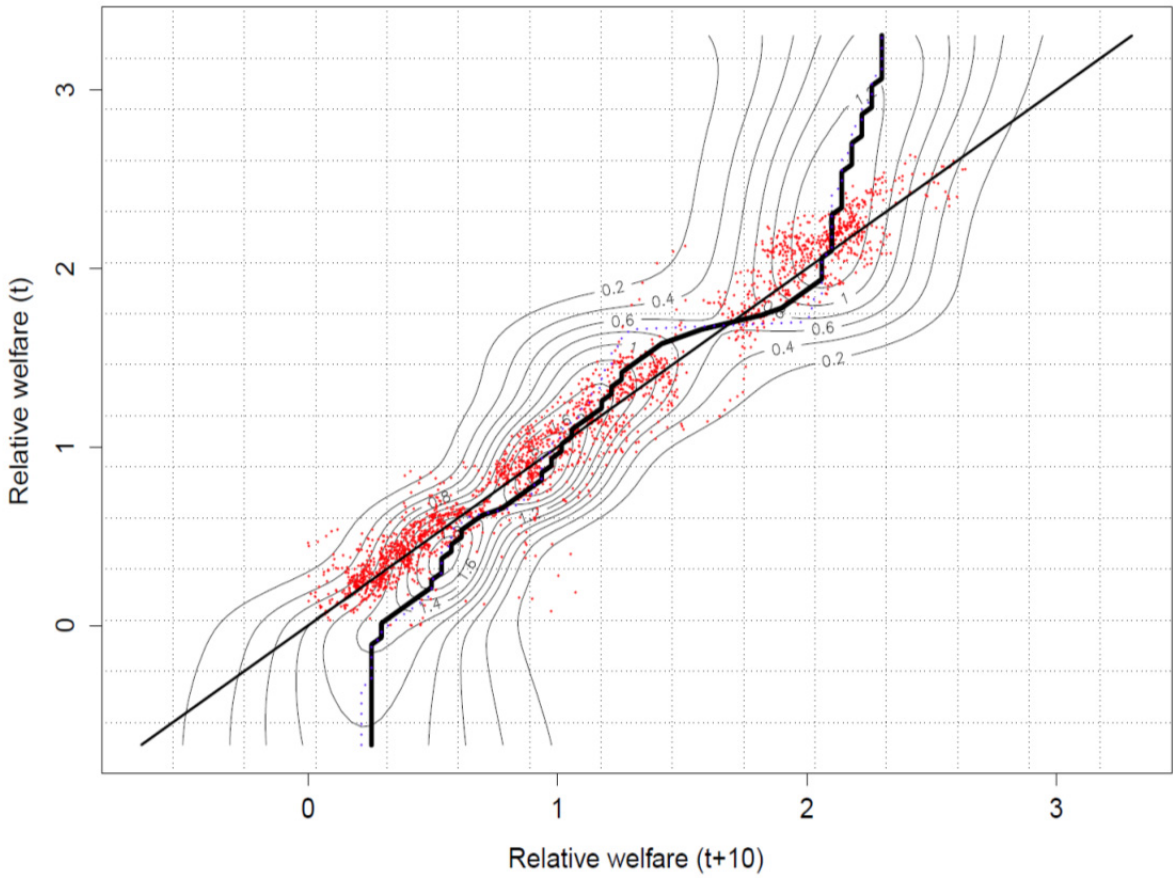
Source: PWT 8.1, World Development Indicators 2014 (January 2015 release), World Economic Outlook (April 2015) and Worldwide Governance Indicators (www.govindicators.org).

in the other two clusters are also the size of the public sector and the resources (both public and private) devoted to health care.

Given the evolution of the joint distribution of income and life expectancy, a clearer picture on how welfare (that is on how the non-linear combination of per capita GDP and life expectancy) evolves can be given by the estimation of its stochastic kernel over the period 1980-2011, which overcomes the bias in the estimates of the growth paths caused by the presence of cross-country heterogeneity.

Figure 8

Stochastic Kernel Estimation of the Relative Welfare (*g* = 0)



The stochastic kernel indicates for each level of *x* at time t the probability distribution of *x* at time $t + \tau$.[42] In the estimate, τ is set at ten years to reduce the influence of short-run fluctuations. The total number of observations is 2163).

Figure 8 reports also a solid line representing the estimated median value at $t + \tau$ conditional on the value at time *t*, a dotted light-blue line indicating the “ridge” of the stochastic

kernel (which is the mode at t + τ conditional on the value at time t), and the 45° line. The red dots represents observations.

Two clusters of populations are located around 1 and slightly above 2 can be clearly detected (see Figure 8) even though a third substantial mass can be noticed at lower level of welfare. Accordingly, in terms of relative welfare in 2011, Centroid C1, C2 and C3 of Figure 7 correspond to around 0.5, 1 and 2.1 respectively.

[42] More formally, let $q(x_t, x_{t-\tau})$ be the joint distribution of $(x_t, x_{t-\tau})$ and $f(x_{t-\tau})$ be the marginal distribution of $x_{t-\tau}$, then the stochastic kernel is defined as $g_\tau(x_t|x_{t-\tau}) = q(x_t, x_{t-\tau})/f(x_{t-\tau})$. The ergodic distribution $f_\infty(x)$ is implicitly defined as $f_\infty(x) = \int_0 g_\tau(x|z) f_\infty(z)\, dz$

Table 5

The Gini Index of the Estimated Ergodic Distributions of Per Capita GDP, Life Expectancy and Welfare (g = 0)

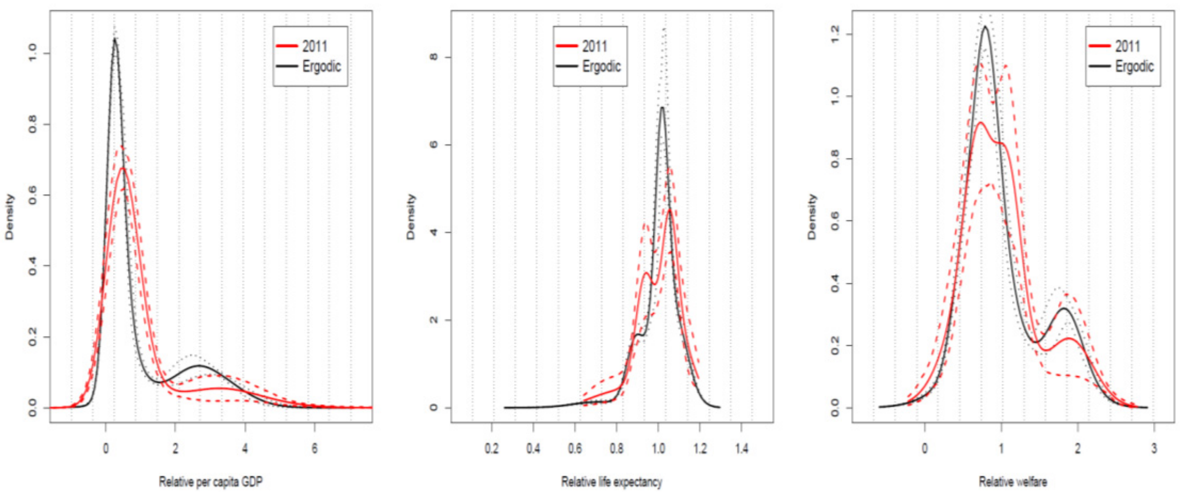
Year	GDP	Life Expectancy	Welfare (g = 0)
2011	0.51	0.06	0.26
	(0.022)	(0.005)	(0.017)
Ergodic	0.61	0.05	0.27
	(0.006)	(0.001)	(0.005)

Note: Standard errors in parentheses; those relative to the ergodic distribution are calculated by the bootstrap procedure described in Appendix G.

Figure 9

Estimated Ergodic Distributions

(a) Relative GDP **(b) Relative Life Expectancy** **(c) Relative Welfare (g = 0)**



3.3.4 The ergodic distribution of per capita GDP, life expectancy and welfare

The estimate of the ergodic distribution of per capita GDP, life expectancy and welfare by stochastic kernel aims at assessing the long-run tendencies resulting from the distribution dynamics just discussed. In other words, the ergodic distribution shows if the estimated distribution dynamics in the period 1980-2011 have completely exhausted their effects on the distribution in 2011 or, instead, whether significant distributional changes are embedded in the ongoing process. Clearly, such estimate of the long-run tendencies does not take into account structural shocks which could lead to non-stationarity.

The ergodic distributions are estimated following the procedure in Johnson (2005), adjusted for the use of normalised variables (with respect to the average) in the estimate.[43] Both the ergodic distribution and the distribution in 2011 are depicted with their confidence intervals at 95 per cent significance level, computed via a bootstrap procedure suggested in Bowman and Azzalini (1997) (Figure 9).[44]

Both inequality and polarization of the cross-population distribution of per capita GDP would increase. The Gini index of the ergodic distribution is indeed equal to 0.61 versus 0.51 in 2011 (see Table 5). The presence of two clearly identified group of populations becomes neater and neater.

By contrast, inequality in life expectancy would continue to stay stable (the Gini index of the ergodic distribution is substantially unchanged with respect to 2011: 0.05 vs. 0.06; see Table 5), while polarization would probably slightly decrease.

As a result, inequality of the cross-population distribution of welfare will stop decreasing (or even increase; see Table 5). The high-welfare peak already present in 2011 is more and more evident and identified, while the two lower peaks tend to merge and locate at a relative welfare level lower than 1 (see Figure 9c).

3.4 *The world distribution of welfare*

So far we have neglected within-country inequality in welfare; however, several contributions related to the world distributions of life expectancy and, mainly, income suggest that such source of inequality can be sizeable and changing over time (see Anand and Segal (2008) for a survey of the literature on the world distribution of income and Pradhan *et al.* (2003) and Ryan (2010) for the world distribution of life expectancy).

In order to have a proper estimate of within-country welfare inequality we need information on the joint distribution of income *and* life expectancy, which could be calculated starting from the two single distributions by a random-matching procedure if the variables were independently distributed. Unfortunately, there is strong evidence which points to the existence of a within-country negative correlation between mortality and socio-economic conditions (see, e.g., Cutler *et al.* (2006) for developed countries, and Grimm *et al.* (2010) for the poor ones). The variability of life expectancy among different income groups can therefore be quite large.[45]

Several works estimate the joint distribution of life expectancy and socio-economic indicators,[46] but very few (three to our knowledge) directly put into relation income and life expectancy. In particular, Gerdtham and Johannesson (2000) estimate life expectancy by income deciles in Sweden, McIntosh *et al.* (2009) make the same for a sample of Canadian population, and, finally, Khang *et al.* (2010) quantify the differences in life expectancy by income quartiles for 4 million public servants in South Korea. Visual inspections of the data supplied by these three

43 See Appendix F for more details.

44 See Appendix G for more details.

45 For example, Marmot (2004) calculates a difference of almost 15 per cent in the life expectancy at 45 years of age between the lowest and the highest employment grades among the British civil servants.

46 For example Grimm *et al.* (2010) apply a principal component analysis on data collected in the Demographic and Health Surveys to proxy income at household level for 32 countries and use life tables and the survival status information on all children born in the 5 years preceding the surveys to estimate life expectancy; a very similar analysis is made by Harttgen and Klasen (2010) on a smaller sample of developing countries; Singh and Siahpush (2006) study changes in the extent of inequalities in life expectancy at birth in US between 1980-2000 by socio-economic deprivation status computed at counties' level (it is worth to notice that their deprivation index relies, among other things, on the *m*edian incomes of the counties); other studies proxy socio-economic status by education attainment (see, among others, Brønnum *et al.* (2008) for Denmark, Leinsalu *et al.* (2003) for Estonia and Hoi *et al.* (2009) for Vietnam).

studies suggests to model the within-country relationship between relative life expectancy and income in the following way:

$$\frac{LE_i}{\overline{LE}} = \beta_0 + \beta_1 \log\left(\frac{y_i}{\bar{y}}\right), \tag{12}$$

where LE_i and y_i are respectively the life expectancy and the average income of the *i*-th income quantile, and $\overline{LE}$ and $\bar{y}$ the sample averages of life expectancy and income respectively.[47] Indeed, the estimation of Model (12) on the data of Canada, Sweden and South Korea results in an adjusted R-squared which ranges from 0.95 up to 0.98, which provides a strong support in favour of the proposed specification (see Appendix H).

In light of the parameters' estimates reported in Appendix H, we set β_0 = 1.009 and β_1 = 0.054 in building the joint distribution of life expectancy and income for all the countries in the sample.[48] In general, such assumption could appear to be very strong, since it implies that the relationship between relative life expectancy and relative income is invariant across countries (notwithstanding, e.g., possible heterogeneity in their health systems) and over time. However, and surprisingly, the differences in the parameters' estimates across countries, also with very different levels of per capita income and life expectancy, are quite modest (see Appendix H), and the analysis will concern a quite short time-span for data unavailability.

The second piece of information we need to estimate the joint distribution of life expectancy and income is the world distribution of income. In this respect we exploit the WYD (World Income Distribution) dataset built and used by Milanovic (2012), which contains income distribution by quantiles drawn from nationally representative household surveys for a large set of countries (covering up to around 95 per cent of the world population).[49] Unfortunately, so far data are available only for a relatively small time period. In particular, we will use the data labelled "1993" for the estimate of the world income distribution in 1993 and those labelled "2005" for the 2005 estimate.[50] The data on within-country inequality are then combined (scaled) with countries' per capita GDP (for consistency with respect to the previous analysis),[51] for deriving the world income distribution; finally, by Model (12) we calculate the joint world distribution of income and life expectancy.

As expected the estimate of the world inequality by the population-weighted crosscountry distribution (Milanovic's Concept 2 inequality) that we have previously discussed,

47 In the estimate of Model 12 we expect β_0 = 1 by definition, and β_1 positive, but strongly less than 1. It is indeed reasonable to expect that an increase in the variance of income positively affects the variance of life expectancy but to some limited extent.

48 Such values corresponds to those for Canada, but the use of alternative parameters do not appear to alter the results. In Appendix H we also report the estimates of Model 12 using data for the US, taken from Singh and Siahpush (2006), and for other 15 developing countries, taken from Harttgen and Klasen (2010). Model (12) seems to fit very well all the datasets (expect for Armenia's), according to the adjusted R-squared (the lowest value being equal to 0.74).

49 Other approaches have been followed in the literature. For example, Bourguignon and Morrisson (2002) and Sala-i-Martín (2006) overcome the lack of data on the within-country distribution of income by assuming that similar countries have similar income distributions. Other scholars (e.g. Chotikapanich *et al.* (1997), Schultz (1998) and, for recent estimations, Holzmann *et al.* (2007) and Vollmer *et al.* (2010)) estimate the countries' income distributions assuming a lognormal density function whose first two moments are inferred by the countries' mean income (or per capita GDP) and by a summary of inequality statistics as Gini index. Milanovic (2012) dataset is available at http://go.worldbank.org/IVEJIU0FJ0 to which we refer for more details.

50 In fact, the surveys composing the dataset are not available at annual intervals for most countries. Milanovic (2012) aggregates them around benchmark years, spaced approximately at 5-year intervals so that all countries that have had surveys within that interval are included.

51 This implies that the discrepancy between national accounts GDP and household surveys aggregate income is evenly spread across the distribution. Other approach could be followed: for example allocating the entire gap to the top tail of the distribution or making correction according to a Pareto tail estimate of right tail. Our assumption however is more conservative and reduces the risk of artificially inflating inequality. See Anand and Segal (2008) for a discussion of the issues related to this choice of rescaling.

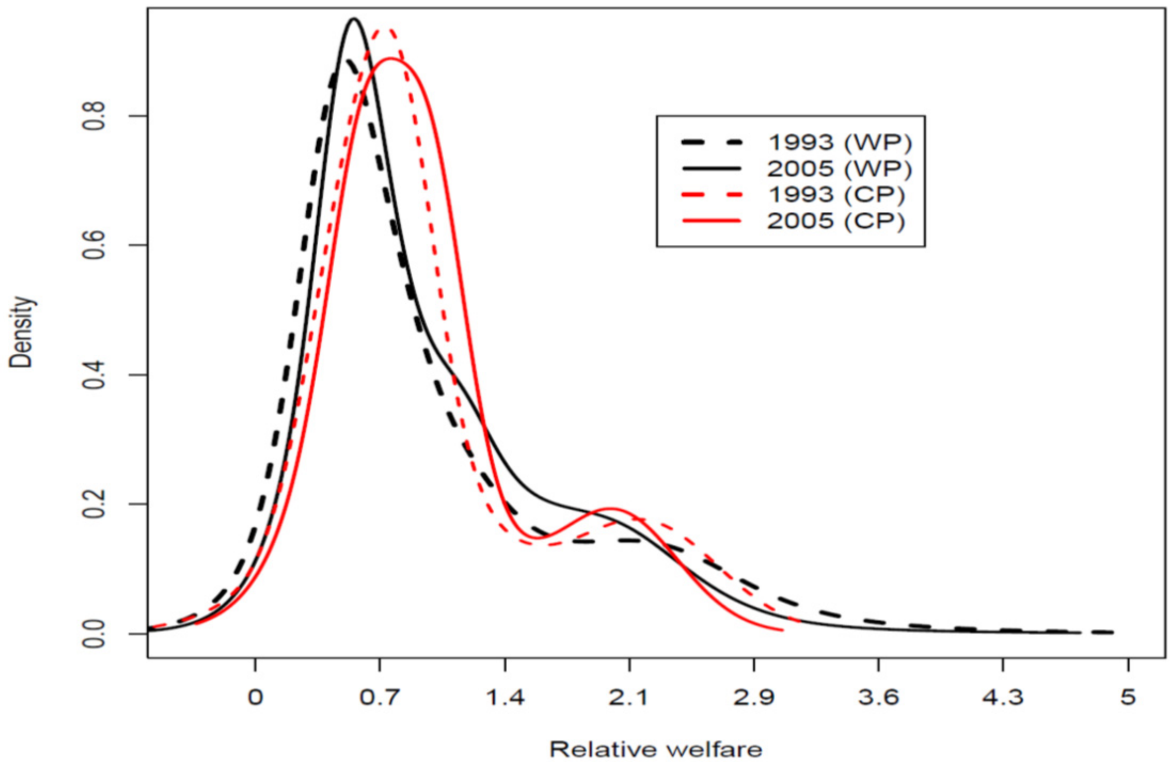
Table 6

A Comparison of Different Levels of Inequality for the World Distribution of Welfare (*g* = 0)

Year	**1993**	**2005**
Cross Population	0.34	0.30
World Population	0.42	0.38

Figure 10

The World Distributions and the Cross-population Distribution of Relative Welfare (*g* = 0; with Respect to the Average of the Period) in 1993 and 2005



leads to a substantial undervaluation (by about 8 percentage points in terms of Gini index, see Table 6). However, the downwards trend in welfare inequality is confirmed also for the world distribution of welfare, with a fall of around 4 percentage points from 1993 to 2005, which contrasts with the substantial stability of the Gini index of the world income distribution that several studies observed for the same period (see Anand and Segal (2008) and Milanovic (2012)).

Figure 10 displays the estimates of the world distribution of welfare in 1993 and 2005 and the analogue cross-population distributions.

In 2005 the world distribution presents a two-peaked distributions, with the peaks approximately around the same position (around 0.7 and 2) as in the cross-population distribution, but with a larger mass at the center of distribution. The within-country inequality, therefore, seems to mainly affect the middle-welfare individual.

The comparison of the estimates in 1993 and 2005 gives less clear-cut conclusions for the evolution of world distribution. The two modes are better identified in 2005, but also closer to each other; moreover a non-negligible probability mass is shifting away from the upper tail toward the center of distribution.

4 Concluding Remarks

This paper presents two main contributions to the existing literature on growth empirics: i) it provides a methodology to measure the welfare of a country/individual and ii) it finds evidence of polarization across world population; moreover, such pattern is expected to be persistent in the future.

This evidence is not in contrast with the recent observed impetuous income growth of some large Asian countries nor with the so called ”African growth miracle”; indeed, those countries appear to be converging in terms of populations to a cluster of medium-welfare countries; but these, in turn, are not converging to the high-welfare cluster, because their higher growth of per capita GDP is counterbalanced in terms of welfare growth by the relative large increase in life expectancy experienced by the countries in the high-welfare cluster (equal *absolute variations* in life expectancy have a higher impact in terms of welfare at higher levels of income).

This suggests the existence of middle-welfare traps with relevant policy implications. For countries in the medium-welfare cluster welfare-enhancing policies should complement income growth with health-improving measures. Moreover, assuming the “Preston curve” as a causal relationship at low levels of income (higher levels of income lead to higher levels of life expectancy) implies that the best welfare-improving policy for the very poor countries should be mainly income-growth oriented (as suggested by Sachs (2005) and Collier (2007)). On the contrary, if the causality run from life expectancy to income, as suggested by Lorentzen *et al.* (2008), the limited diffusion of recent medical technology (see Easterlin (2004) and Becker *et al.* (2005)) should increase the concern to provide an appropriate support to health-enhancing policies in poor countries.

Finally, our findings on the distribution dynamics of welfare can integrate the analysis of the effect of globalization on income distribution provided by Milanovic (2012). For example, we can account for many phenomena, such as the migration of the relatively poor people, where the difference in living standards is one of the crucial explanatory factors (see Anand and Segal (2008)).

Four aspects need to be further investigated. First, the methodology used to measure welfare might be extended to account for factors which appear very different across countries, such as the labour market structure, the provision of public goods and the level of taxation, and the market incompleteness. Second, in our approach we do not distinguish between changes in life expectancy at birth due to changes in infant mortality or changes in adult mortality. Given an increase in life expectancy at birth, welfare could differently change if such increase is the result of a fall in infant or in adult mortality. So far the lack of data for a sufficiently large set of countries and years makes this extension problematic; however the observation of the dynamics of infant mortality and adult mortality, that have recently shown opposite trends (being the first one characterized by a strong convergence pattern across countries, while for the second one divergence prevails, see Edwards (2010)), suggests that such extension could provide further support to the conclusion of polarization in welfare. Third, the methodology related to the inclusion of the within-country distribution should be refined. Indeed, taking into account the within-country inequality seems to have a non-negligible impact on the magnitude and the dynamics of the world welfare distribution; however, the non-availability of comparable microdata on the relationship between income and life expectancy for a large sample of countries represents a serious obstacle.

Finally, for a more thorough picture of the world welfare inequality, it could be interesting to consider cases where inequality is measured in absolute terms, and/or where inequality by itself produces a welfare loss (see Gruen and Klasen (2008) and Atkinson and Brandolini (2010), who respectively discuss the same issues but related to the world income inequality).

APPENDIX A
SOLUTION OF THE AGENT'S PROBLEM

The agent solves the following problem:

$$V = \max_{\{c_t\}_{t=0}^{T}} \int_0^T \left(\frac{c^{1-\sigma}}{1-\sigma} - M \right) \exp(-\rho t)\, S dt \tag{13}$$

$$s.t. \begin{cases} \dot{p} = p\hat{r} + yl - c; \\ p_0 = \bar{p}_0; \\ \lim_{t \to T} p \exp(-\hat{r} t) \geq 0; \end{cases}$$

where $\hat{r} = r + \pi^D$ is the interest rate adjusted for the instantaneous probability of dying before T. Dynamic constraint $\dot{p} = p\hat{r} + yl - c$ in Problem 13 is derived directly from the intertemporal budget constraint given in Eq. (1).

The Hamiltonian of Problem (13) is given by:

$$\mathcal{H} = \left(\frac{c^{1-\sigma}}{1-\sigma} - M \right) \exp(-\rho t)\, S + \lambda (p\hat{r} + yl - c) \tag{14}$$

and the necessary and sufficient conditions of Problem (13) are the following:

$$\lambda = c^{-\sigma} \exp(-\rho t)\, S; \tag{15}$$

$$\dot{\lambda} = -\lambda \hat{r}; \tag{16}$$

$$\lim_{t \to T} \lambda p = 0, \tag{17}$$

from which:

$$\frac{\dot{c}}{c} = \frac{r - \rho}{\sigma} = g. \tag{18}$$

Given $\lambda(0) > 0$ and the constraints in Problem 13, Eq. (17) is always satisfied. Since r is assumed constant over time, we have:

$$c_t = c_0 \exp(gt). \tag{19}$$

The growth rate of consumption g is independent of T and S and it represents the steadystate growth rate.

Because of the strict monotonicity of $u(c)$, budget constraint (1) holds with strict equality. Hence, the initial consumption level c_0 is given by:

$$c_0(T, w) = w \left[\frac{g - \hat{r}}{\exp((g - \hat{r})T) - 1} \right]. \tag{20}$$

Substituting Eq. (19) into Eq. (13) yields the agent's (indirect) utility:

$$V(T, w) = \frac{1}{(1-\sigma)} \left\{ c_0(T, w)^{1-\sigma} \left[\frac{\exp[((1-\sigma)g - \hat{\rho})T] - 1}{(1-\sigma)g - \hat{\rho}} \right] + \frac{(1-\sigma) M [\exp(-\hat{\rho}T) - 1]}{\hat{\rho}} \right\} \tag{21}$$

where $\hat{\rho} = \rho + \pi^D$. V in Problem (13) is an improper integral for $T \to +\infty$ if $(g - \hat{r}) \geq 0$ Therefore if $T \to +\infty$ we must assume that $(g - \hat{r}) < 0$ in order to have a well-defined maximisation problem.

The agent's lifetime wealth w is therefore given by:

$$w = \frac{yl_0 \left[\exp\left(\left(g - \hat{r}\right)T\right) - 1\right]}{g - \hat{r}} + \bar{p}_0, \tag{22}$$

which substituted in Eq. (21) yields:

$$\begin{aligned} V\left(T, yl_0, g\right) = \frac{1}{1-\sigma} \Bigg\{ & \left(\frac{yl_0 \left[\exp\left(\left(g - \hat{r}\right)T\right) - 1\right]}{g - \hat{r}} + \bar{p}_0 \right)^{1-\sigma} \left(\frac{\exp\left(\left(g - \hat{r}\right)T\right) - 1}{g - \hat{r}} \right)^{\sigma} \\ & + \frac{(1-\sigma) M \left[\exp\left(-\hat{\rho}T\right) - 1\right]}{\hat{\rho}} \Bigg\}. \end{aligned} \tag{23}$$

APPENDIX B
COUNTRY LIST WITH THE INDICATION OF CLUSTERS

Table 7

Country List with the Indication of Clusters

Country Name	Population 2011 *(million)*	Cluster in 2011	Country Name	Population 2011 *(million)*	Cluster in 2011
Bangladesh	150.49	1	Chile	17.27	2
Benin	9.10	1	China	1324.35	2
Bolivia	10.09	1	Colombia	46.93	2
Burkina Faso	16.97	1	Costa Rica	4.73	2
Burundi	8.58	1	Dominican Republic	10.06	2
Cabo Verde	0.5	1	Ecuador	14.67	2
Cameroon	20.03	1	Equatorial Guinea	0.72	2
Central African Republic	4.49	1	Gabon	1.53	2
Chad	11.53	1	Iran, Islamic Rep.	74.8	2
Comoros	0.75	1	Jordan	6.33	2
Congo, Dem. Rep.	67.76	1	Malaysia	28.86	2
Congo, Rep.	4.14	1	Mauritius	1.31	2
Cote d'Ivoire	20.15	1	Mexico	114.79	2
Egypt, Arab Rep.	82.54	1	Panama	3.57	2
El Salvador	6.23	1	Peru	29.4	2
Ethiopia	84.73	1	Romania	21.44	2
Fiji	0.87	1	South Africa	50.46	2
Gambia, The	1.78	1	Sri Lanka	21.05	2
Ghana	24.97	1	Thailand	69.52	2
Guatemala	14.76	1	Trinidad and Tobago	1.35	2
Guinea	10.22	1	Tunisia	10.59	2
Guinea-Bissau	1.55	1	Turkey	73.64	2
Honduras	7.75	1	Uruguay	3.38	2
India	1241.49	1	Venezuela, RB	29.44	2
Indonesia	242.33	1	Australia	22.61	3
Jamaica	2.75	1	Austria	8.41	3
Kenya	41.61	1	Belgium	10.75	3
Lesotho	2.19	1	Canada	34.35	3
Madagascar	21.32	1	Cyprus	0.82	3
Malawi	15.38	1	Denmark	5.57	3
Mali	15.84	1	Finland	5.38	3
Mauritania	3.54	1	France	65.09	3
Morocco	32.27	1	Germany	82.16	3
Mozambique	23.93	1	Greece	11.39	3
Namibia	2.32	1	Hong Kong SAR, China	7.12	3
Nepal	30.49	1	Iceland	0.32	3
Niger	16.07	1	Ireland	4.53	3
Nigeria	162.47	1	Italy	60.79	3
Pakistan	176.75	1	Japan	126.5	3
Paraguay	6.57	1	Korea, Rep.	48.39	3
Philippines	94.85	1	Luxembourg	0.52	3
Rwanda	10.94	1	Malta	0.42	3
Senegal	12.77	1	Netherlands	16.66	3
Syrian Arab Republic	20.77	1	New Zealand	4.41	3
Tanzania	44.92	1	Norway	4.92	3
Togo	6.15	1	Portugal	10.69	3
Uganda	34.51	1	Singapore	5.19	3
Zambia	13.47	1	Spain	46.45	3
Zimbabwe	12.75	1	Sweden	9.44	3
Argentina	40.76	2	Switzerland	7.7	3
Barbados	0.27	2	United Kingdom	62.42	3
Botswana	2.03	2	United States	313.09	3
Brazil	196.66	2			

APPENDIX C
ADAPTIVE KERNEL ESTIMATION

When observations vary in sparseness over the support of the distribution, the adaptive kernel estimation is a two-stage procedure which mitigates the drawbacks of a fixed bandwidth in density estimation (see Silverman (1986), p. 101). In general, given a multivariate data set $X = \{X_1, ..., X_n\}$ and a vector of sample weights $W = \{\omega_1, ..., \omega_n\}$, where X_i is a vector of dimension d and $\sum_{i=1}^{n} \omega_i = 1$, we first run the pilot estimate:

$$\tilde{f}(\boldsymbol{x}) = \frac{1}{n \det(\boldsymbol{H})} \sum_{i=1}^{n} \omega_i k \left\{ \boldsymbol{H}^{-1} (\boldsymbol{x} - \boldsymbol{X}_i) \right\}, \tag{24}$$

where $k(u) = (2\pi)^{-1} \exp(-1/2u)$ is a Gaussian kernel and *bandwidth matrix H* is a diagonal matrix $(d \times d)$ with diagonal elements $(h_1, ..., h_d)$ given by the optimal normal bandwidths, *i.e.*,

$$h_i = [4/(d+2)]^{1/(d+4)} \hat{\sigma}_i n^{-1/(d+4)}$$

$\hat{\sigma}_i$ is the estimated standard error of the distribution of X_i. The use of a diagonal bandwidth matrix instead of a full covariance matrix follows the suggestions in Wand and Jones (1993). In the case of $d = 1$ we have:

$$H = \det(H) = (4/3)^{1/5} n^{-1/5} \hat{\sigma}$$

In the cross-population estimate we consider $W = \{p_i, ..., p_n\}$, where p_i is the population of country i. We then define local bandwidth factors λ_i by:

$$\lambda_i = \left[\tilde{f}(\boldsymbol{X}_i) / g \right]^{-\alpha}, \tag{25}$$

where $\log(g) = \sum_{i=1}^{n} \omega_i \log\left(\tilde{f}(\boldsymbol{X}_i)\right)$ and $\alpha \in [0, 1]$ is a sensitivity parameter. We set $\alpha = 1/2$ as suggested by Silverman (1986), p. 103. Finally the adaptive kernel estimate $\hat{f}(x)$ is defined as:

$$\hat{f}(\boldsymbol{x}) = \frac{1}{n \det(\boldsymbol{H})} \sum_{i=1}^{n} \lambda_i^{-d} \omega_i k \left\{ \lambda_i^{-1} \boldsymbol{H}^{-1} (\boldsymbol{x} - \boldsymbol{X}_i) \right\}. \tag{26}$$

The Gaussian kernel guarantees that the number of modes is a decreasing function of the bandwidth; this property is at the basis of the test for unimodality (see Silverman (1986), p. 139). In all the estimates we use package *sm* (see Bowman and Azzalini (2005)).

APPENDIX D
MULTIMODALITY TEST

The multimodality test follows the bootstrap procedure described in Silverman (1986), p. 146. Given a data set $X = \{x_1, \ldots, x_n\}$ and a vector of sample weights $W = \{\omega_1, \ldots, \omega_n\}$, we calculate the smallest value of bandwidth, $\hat{h}_0$, for which the estimated distribution is unimodal and the corresponding local bandwidth factors $\Lambda = \lambda_1, \ldots, \lambda_n$. We then perform a *smoothed bootstrap* from the estimated density of observed data set. Since we use the Gaussian kernel, it amounts to: i) draw (with replacement) a vector $I = \{i_1, \ldots, i_n\}$ of size n from $\{1, \ldots, n\}$, given the sample weights W; ii) define $Y = \{x_{i1}, \ldots, x_{in}\}$ and $W^* = \{\omega_{i1}, \ldots, \omega_i\}$, calculate:

$$x_j^* = \bar{\mathbf{Y}} + \left(1 + \left(\hat{h}_0\lambda_{i_j}\right)^2 / \hat{\sigma}^2_{\mathbf{Y}}\right)^{-\frac{1}{2}} \left(y_j - \bar{\mathbf{Y}} + \hat{h}_0\lambda_{i_j}\epsilon_j\right), \quad j = 1, \ldots, n; \tag{27}$$

where $\bar{Y}$ and $\hat{\sigma}^2_{\mathbf{Y}}$ are respectively the mean and the estimate variance of sample Y and ε_j are standard normal random variables; iii) find the minimum value of bandwidth, $\hat{h}^*$, for which the estimated density of X^* is unimodal; iv) repeat point i)-iii) B times in order to obtain a vector of critical

values of bandwidth $\left\{\hat{h}_1^*, \ldots, \hat{h}_B^*\right\}$. Finally, p-value of null-hypothesis of unimodality is given by:

$$\#\left\{\hat{h}_b^* \geq \hat{h}_0\right\} / B.$$

For testing the bimodality, point iii) has to be modified accordingly. We set $B = 1000$.

APPENDIX E
VECTOR FIELD ESTIMATION

Assume that the dynamics of economy j at period t only depends on (GDP_{jt} , LE_{jt}), *i.e.*, (GDP_{jt} , LE_{jt}) follows a *time invariant* and *Markovian* stochastic process.

The dynamics of the sample in the space (*GDP*, *LE*) can be therefore represented by a random vector field (RVF). In particular, given a subset L of the possible realization of (*GDP*, *LE*) (*i.e.*, a lattice, see small black points in Figure 7), a RVF is represented by a random variable $\Delta_\tau z_i$, where $\Delta_\tau z_i \equiv (\Delta_\tau GDP_i , \Delta_\tau LE_i) \equiv (GDP_{it+\tau} - GDP_{it}, LE_{it+\tau} - LE_{it})$, indicating the dynamics (*i.e.*, the dynamics from period t to period $t + \tau$ represented by a movement vector) at $z_i \equiv (GDP_i, LE_i) \in L$. For each point in the lattice z , with i = 1, ..., L, we therefore estimate the distribution of probability $Pr\ (\Delta_\tau z|z_i)$ on the $N\ (T - \tau\ )$ observed movement vectors $\Delta_\tau^{OBS} z$. In $\Pr\left(\Delta_\tau^{OBS} z_{jt}|z_i\right)$ measures the probability that the dynamics at z_i follow $\Delta_\tau^{OBS} z_{jt}$; this suggests $\Pr\left(\Delta_\tau^{OBS} z_{jt}|z_i\right)$ should decrease as function of the distance between z_{jt}^{OBS} d .

A convenient way to calculate these probabilities is to use a kernel function to measure the distance between z_i and z_{jt}^{OBS} . In particular:

$$\omega\left(z_i, z_{jt}^{OBS}\right) = \frac{K\left(\frac{(z_i - z_{jt}^{OBS})^T S^{-1} (z_i - z_{jt}^{OBS})}{h^2}\right) \frac{\det(S)^{-\frac{1}{2}}}{2h^2}}{\sum_{t=1}^{T-\tau} \sum_{j=1}^{N} K\left(\frac{(z_i - z_{jt}^{OBS})^T S^{-1} (z_i - z_{jt}^{OBS})}{h^2}\right) \frac{\det(S)^{-\frac{1}{2}}}{2h^2}} \quad (28)$$

is assumed to be an estimate of the probability that at *zi* dynamics follows observed movement vectors $\Delta_\tau^{OBS} z_{jt}$, where $K(\cdot)$ is the kernel function, h is the smoothing parameter and S is the sample covariance matrix of z^{OBS}. The kernel function $K(\cdot)$ is generally a smooth positive function which peaks at 0 and decreases monotonically as the distance between the observation z_{jt} and the point of interest z_i increases (see Silverman (1986) for technical details). The smoothing parameter h controls the width of the kernel function.[52] In the estimation we use a multivariate Epanechnikov kernel (see Silverman (1986) pp. 76-78), *i.e.*:

$$K(u^T S^{-1} u) = \begin{cases} \frac{2}{\pi}(1 - u^T S^{-1} u) & \text{if } u^T S^{-1} u < 1 \\ 0 & \text{if } u^T S^{-1} u \geq 1, \end{cases} \quad (29)$$

where $u \equiv \left(z_i - z_{jt}^{OBS}\right)/h$. Multivariate Epanechnikov kernel is particularly adapted to our scope because it assigns zero probability to observed movement vectors very far from z_i. Other possible kernels, as the Gaussian, does not allow such possibility. The exact quantification of "very far" is provided by bandwidth h, *i.e.*, higher bandwidth means higher number of observed movement vectors entering in the calculation of the movement at z_i.

Given Eq. (28) for each point in the lattice z_i we estimate the τ-period ahead *expected movement* $\mu_{\Delta_\tau} z_i \equiv E\ [\Delta_\tau z_i \mid z_i]$ using a *local mean estimator*, first proposed by Nadaraya (1964) and Watson (1964), where the observations are weighted by the probabilities derived from the kernel function, *i.e.*, (see Bowman and Azzalini (1997) for details):

$$\widehat{\mu}_{\Delta_\tau z_i} = \sum_{t=1}^{T-\tau} \sum_{j=1}^{N} \omega\left(z_i, z_{jt}^{OBS}\right) \Delta_\tau z_{jt}^{OBS} = \Pr\widehat{(\Delta_\tau z | z_i)} \Delta_\tau z^{OBS}. \quad (30)$$

---

[52] In all the estimation we use the optimal normal bandwidth; for a discuss on the choice of bandwidth see Silverman (1986).

The estimation of Eq. (30) strongly depends on the choice of τ . This choice is the result of a trade-off: from one hand, a too short τ can increase the noise in the estimation due to the possible presence of short-run fluctuations; on the other hand, a too long τ could contrast with the local characteristics of the estimate, increasing the probability that observed movement vectors very far from z_i affects the estimate of $\mu_{\Delta_\tau} z_i$. In the estimate we set $\tau = 10$.

Below we discuss in details how we have conducted the inference on the estimated expected movements by a bootstrap procedure, whose results is reported in Figure 7.

Given the observed sample of observations z_{jt}^{OBS} with $j = 1, ..., N$ and $t = 1, ..., T$, the bootstrap procedure consists of four steps.

1) Estimate the expected value of the τ-period ahead movement $\mu_{\Delta_\tau} z_i$ by Eq. (30) for each point of the lattice ($i = 1, ..., L$).
2) Draw B samples $z^b = \left(z_1^b, ..., z_{N(T-\tau)}^b\right)$ and the associated $\Delta_\tau^b z = \left(\Delta z_1^b, ..., \Delta z_{N(T-\tau)}^b\right)$, with $b = 1, ..., B$, by sampling with replacement from the observed z^{OBS} and the associated movement vectors $\Delta^{OBS} z$.
3) For every bootstrapped sample b and for each point of the lattice i estimate by Eq. (30) the expected value of the τ-period ahead movement $\mu_{\Delta_\tau z_i}^b$.
4) Calculate the two-side p-value of the estimated movement vector at point i in the lattice under the null hypothesis of no dynamics (note that null hypothesis of no dynamics is separately tested in the two directions y and Wy) as:

$$\widehat{ASL}_i = 2 \times \min\left(\sum_{b=1}^{B} \hat{\mu}_{\Delta_\tau z_i}^b \leq 0, \sum_{b=1}^{B} \hat{\mu}_{\Delta_\tau z_i}^b > 0\right) / B. \tag{31}$$

In the analysis we have set $B = 300$, and used the usual significance level of 5 per cent to decide which expected movements to report.

APPENDIX F
THE ESTIMATE OF ERGODIC DISTRIBUTION

The ergodic distribution solves:

$$f_\infty(x) = \int_0^\infty g_\tau(x|z)\, f_\infty(z)\, dz, \tag{32}$$

where x and z are two levels of the variable, $g_\tau(x|z)$ is the density of x, given z, τ periods ahead, under the constraint:

$$\int_0^\infty f_\infty(x)\, dx = 1. \tag{33}$$

Since in our estimates all variables are normalized with respect to their average, the ergodic distribution, moreover, must respect the additional constraint:

$$\int_0^\infty f_\infty(x)\, x dx = 1. \tag{34}$$

Following the methodology proposed by Johnson (2005) we first estimate the distribution $\tilde{f}_\infty(x)$, which satisfies Constraints 32 and 33, but not Constraint 34. We then calculate

$$f_\infty(x) = \tilde{\mu}_x \tilde{f}_\infty(x)$$

where:

$$\tilde{\mu}_x = \int_0^\infty \tilde{f}_\infty(x)\, x dx$$

which will satisfy all Constraints 32, 33 and 34. In particular, Theorems 11 and 13 in Mood *et al.* (1974), pp. 200 and 205 prove that if $\tilde{f}_\infty(x)$ satisfies Constraints 32 and 33 then $f_\infty(x)$ satisfies Constraints 32, 33 and 34. In fact, $g_\tau(z|x) = f_{z,x}(z, x) / f_x(x)$ and $f_{y,q}(y, q) = \mu_z \mu_x f_{z,x}(z, x)$, where $y = z/\mu_z$ and $q = x/\mu_x$. In all computations we set $\tau = 10$.

APPENDIX G
BOOTSTRAP PROCEDURE TO CALCULATE CONFIDENCE INTERVALS FOR DENSITY ESTIMATION

The following is a description of the bootstrap procedure used to calculate the confidence intervals for the estimates of densities and ergodic distributions; this is based on the procedure reported in Bowman and Azzalini (1997), p. 41. Given a sample $X = \{X_1, ..., X_n\}$ of observations and a vector of sample weights $W = \{\omega_1, ..., \omega_n\}$, where $\sum_{i=1}^{n} \omega_i = 1$ and X_i is a vector of d dimensions, the bootstrap procedure is as follows.

1) Construct a density estimate $\hat{\phi}$ from sample X, given the sample weights W.
2) Resample X with replacement, taking into account the sample weights W, to produce a bootstrap sample X^*.
3) Construct a density estimate $\hat{\phi}^*$ from X^*.
4) Repeat steps 2. and 3. B times in order to create a collection of bootstrap density estimates $\left\{\hat{\phi}_1^*, ..., \hat{\phi}_n^*\right\}$.

The distribution of $\hat{\phi}_i^*$ about $\hat{\phi}$ can therefore be used to mimic the distribution of $\hat{\phi}$ about ϕ, as discussed by Bowman and Azzalini (1997), p. 41, *i.e.*, to calculate confidence intervals for the estimates. In particular, the confidence interval for the distribution in 2000 corresponds to the case $\hat{\phi} = \hat{f}$, while for the ergodic distribution to the case $\hat{\phi} = \hat{f}_\infty$. In the bootstrap procedure $\hat{\phi}^*$ are calculated taking the bandwidth(s) equal to the bandwidth(s) calculated for the observed sample X, as suggested in Bowman and Azzalini (1997), p. 41. We set $B = 300$.

APPENDIX H
ESTIMATION RESULTS FOR THE RELATIONSHIP BETWEEN INCOME AND LIFE EXPECTANCY

Table 8

Estimation Results for Model (12) Various Countries

Country	β_0	β_1	Adjusted R^2	Source
Canada (2001)	1.009	0.054	0.95	McIntosh *et al.* (2009)
Sweden (1996)	1.004	0.039	0.98	Gerdtham and Johannesson (2000)
South Korea (2002)	1.006	0.033	0.97	Khang *et al.* (2010)
US (1981)	1.001	0.057	0.88	Singh and Siahpush (2006)
US (1990)	1.001	0.053	0.95	Singh and Siahpush (2006)
US (1999)	1.002	0.076	0.97	Singh and Siahpush (2006)
Armenia (2005)	1.000	0.003	–0.09	Harttgen and Klasen (2010)
Burkina Faso (2003)	1.001	0.086	0.86	Harttgen and Klasen (2010)
Bolivia (2003)	1.000	0.069	0.92	Harttgen and Klasen (2010)
Egypt (2007)	1.000	0.066	0.97	Harttgen and Klasen (2010)
Ethiopia (2005)	1.006	0.123	0.74	Harttgen and Klasen (2010)
India (2005)	1.000	0.024	0.81	Harttgen and Klasen (2010)
Indonesia (2003)	1.000	0.063	0.95	Harttgen and Klasen (2010)
Kyrgyz Republic (1997)	1.001	0.065	0.84	Harttgen and Klasen (2010)
Nicaragua (2000)	1.000	0.052	0.82	Harttgen and Klasen (2010)
Nigeria (2003)	1.000	0.107	0.97	Harttgen and Klasen (2010)
Pakistan (2007)	1.000	0.085	0.85	Harttgen and Klasen (2010)
Peru (2005)	1.000	0.059	0.89	Harttgen and Klasen (2010)
Senegal (2005)	1.000	0.116	0.91	Harttgen and Klasen (2010)
Vietnam (2002)	1.000	0.039	0.81	Harttgen and Klasen (2010)
Zambia (2002)	0.996	0.091	0.76	Harttgen and Klasen (2010)

Note: All the coefficients are significantly different from zero at 1 per cent level, with the exception of the β_1 for Armenia.

REFERENCES

Aaberge, R. and A. Brandolini (2015), "Multidimensional Poverty and Inequality", in A.B. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution*, Vol. 2, pp. 141-216, Elsevier.

Anand, S. and P. Segal (2008), "What Do We Know About Global Income Inequality?", *Journal of Economic Literature*, Vol. 46, No. 1, pp. 57-90.

Anderson, G. (2005), "Life Expectancy and Economic Welfare: The Example of Africa in the 1990s", *Review of Income and Wealth*, Vol. 51, No. 3, pp. 455-68.

Atkinson, T. and A. Brandolini (2010), "On Analyzing the World Distribution of Income", *World Bank Economic Review*, Vol. 24, No. 1, pp. 1-37.

Becker, G.S., T.J. Philipson and R.R. Soares (2005), "The Quantity and Quality of Life and the Evolution of World Inequality", *The American Economic Review*, Vol. 95, No. 1, pp. 277-91.

Bloom, D.E. and D. Canning (2007), "Mortality Traps and the Dynamics of Health Transition", proceedings of the National Academy of Sciences of the United States of America, Vol. 104, No. 41, pp. 16044-49.

Bourguignon, F. and C. Morrisson (2002), "Inequality Among World Citizens: 1820-1992", *The American Economic Review*, Vol. 92, No. 4, pp. 727-44.

Bowman, A.W. and A. Azzalini (1997), *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*, Oxford University Press, Oxford.

—————— (2005), Ported to R by B.D. Ripley up to version 2.0, later versions by A.W. Bowman and A. Azzalini. sm: Smoothing Methods for Non-parametric Regression and Density Estimation. R package version 2.1-0, available at: http://www.stats.gla.ac.uk/ adrian/sm

Brønnum-Hansen, H. and M. Baadsgaard (2008), Increase in Social Inequality in Health Expectancy in Denmark, *Scandinavian Journal of Public Health*, No. 36, pp. 44-51.

Chotikapanich, D., R. Valenzuela and D.S.P. Rao (1997), "Global and Regional Inequality in the Distribution of Income: Estimation with Limited and Incomplete Data", *Empirical Economics*, No. 22, pp. 533-46.

Collier, P. (2007), *The Bottom Billion*, Oxford University Press, Oxford.

Cutler, D., A. Deaton and A. Lleras-Muney (2006), "The Determinants of Mortality", *Journal of Economic Perspectives*, Vol. 20, No. 3, pp. 97-120.

Decancq, K., M. Fleurbaey and E. Schokkaert (2015), "Inequality, Income, and Well-being", in A.B. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution*, Vol. 2, pp. 67-140, Elsevier.

Durlauf, S.N. and P.A. Johnson (1995), "Multiple Regimes and Cross-country Growth Behaviour", *Journal of Applied Econometrics*, Vol. 10, No. 4, pp. 365-84.

Durlauf, S.N., P.A. Johnson and J.R.W. Temple (2005), "The Growth Econometrics", in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, Vol. 1, p. 555-677, Elsevier.

Easterlin, R.A. (2004), *The Reluctant Economist*, Cambridge University Press, Cambridge.

Edwards, R.D. (2010), "Trends in World Inequality in Life Span Since 1970", NBER, Working Paper, No. 16088.

Feenstra, R.C., R. Inklaar and M.P. Timmer, "The Next Generation of the Penn World Table", *American Economic Review*, available for download at www.ggdc.net/pwt (forthcoming).

Fiaschi, D. and A.M. Lavezzi (2003), "Distribution Dynamics and Nonlinear Growth", *Journal of Economic Growth*, Vol. 8, No. 4, pp. 379-401.

Fiaschi, D. and M. Romanelli (2009), "Nonlinear Dynamics in Welfare and the Evolution of World Inequality", Banca d'Italia, Temi di discussione, No. 724.

————— (2010), "Sheep for a Lifetime or Lion for a Day? Optimal Lifetime and the Non-monotonic Relation Between Life Expectancy and Individual Welfare", University of Pisa, Department of Economics, mimeo.

Fleurbaey, M. (2005), "Health, Wealth, and Fairness", *Journal of Public Economic Theory*, Vol. 7, No. 2, pp. 253-84.

Fleurbaey, M. and G. Gaulier (2009), "International Comparisons of Living Standards by Equivalent Incomes", *Scandinavian Journal of Economics*, Vol. 111, No. 3, pp. 597-624.

Gerdtham, U.G. and M. Johannesson (2000), "Income-related Inequality in Life-years and Quality-adjusted Life-years", *Journal of Health Economics*, Vol. 19, No. 6, pp. 1007-26.

Grimm, M., K. Harttgen, S. Klasen, M. Misselhorn, T. Munzi and T. Smeeding (2010), "Inequality in Human Development: An Empirical Assessment of 32 Countries", *Social Indicators Research*, No. 97, pp.191-211.

Grimm, M., K. Harttgen, S. Klasen and M. Misselhorn (2008), "A Human Development Index by Income Groups", *World Development*, Vol. 36, No. 12, pp. 2527-46.

Gruen, C. and S. Klasen (2008), "Growth, Inequality, and Welfare: Comparisons Across Space and Time", *Oxford Economic Papers*, Vol. 60, No. 2, pp. 212-36.

Hall, R.E. and C.I. Jones (2007), "The Value of Life and the Rise in Health Spending", *The Quarterly Journal of Economics*, Vol. 122, No. 1, pp. 39-72.

Härdle, W., M. Müller, S. Sperlich and A. Werwatz (2004), *Nonparametric and Semiparametric Models*, Springer, Berlin.

Harttgen, K. and S. Klasen (2010), "A Household-based Human Development Index", proceedings of the German Development Economics Conference, Hannover 2010 30, Verein fr Socialpolitik, Research Committee Development Economics.

Hoi, L.V., H.D. Phuc, T.V. Dung, N.T. Chuc and L. Lindholm (2009), "Remaining Life Expectancy Among Older People in a Rural Area of Vietnam: Trends and Socioeconomic Inequalities During a Period of Multiple Transitions", *BMC Public Health*, No. 9, p. 471.

Holzmann, H., S. Vollmer and J. Weisbrod (2007), "Income Distribution Dynamics and Pro-poor Growth in the World from 1970 to 2003", Ibero America Institute for Economic Research (IAI), Discussion Paper, No. 161, p. 471.

Johnson, P.A. (2005), "A Continuous State Space Approach to 'Convergence by Parts'", *Economic Letters*, No. 86, pp. 317-21.

Jones, C. and R. Hall (2007), "The Value of Life and the Rise in Health Spending", *Quarterly Journal of Economics*, Vol. 122, No. 1, pp. 39-72.

Jones, C. and P.J. Klenow (2010), "Beyond GDP? Welfare Across Countries and Time", NBER, Working Paper, No. 16352.

Khang, Y.H.K., S. Yong, H.J. Cho, K. Jung-Choi and S.C. Yun (2010), “Decomposition of Socio-economic Differences in Life Expectancy at Birth by Age and Cause of Death Among 4 Million South Korean Public Servants and Their Dependents”, *International Journal of Epidemiology*, No. 39, pp. 1656-66.

Leinsalu, M., D. Vågerö and A.E. Kunst (2003), “Estonia 1989-2000: Enormous Increase in Mortality Differences by Education”, *International Journal of Epidemiology*, Vol. 32, No. 6, pp. 1081-87.

Leisch, F. (2006), “Visualizing Cluster Analysis and Finite Mixture Models”, Humboldt Universität, Berlin, mimeo.

Luxembourg Income Study Database (LIS), www.lisdatacenter.org, Luxembourg: LIS (multiple countries, 1971-2013).

Lorentzen, P.J., J. McMillan and R. Wacziarg (2008), “Death and Development”, *Journal of Economic Growth*, No. 13, pp. 81-124.

Marmot, M. (2004), “Social Causes of Social Inequalities in Health”, in S. Anand, F. Peter and A. Sen (eds.), *Public Health*, *Ethics*, *and Equity*, Oxford University Press, Oxford.

McIntosh, C., P. Fins, R. Wilkins and M.C. Wolfson (2009), “Income Disparities in Health-adjusted Life Expectancy for Canadian Adults, 1991 to 2001”, *Statistics Canada: Health Reports*, Vol. 20, No. 4, pp. 55-64.

Milanovic, B. (2012), Global Inequality Recalculated and Updated: “The Effect of new PPP Estimates on Global Inequality and 2005 Estimates”, *The Journal of Economic Inequality*, No. 10, pp. 1-18.

Mood, A.M., F.A. Graybill and D.C. Boes (1974), Introduction to the Theory of Statistics, McGraw-Hill, New York (NY).

Murphy, K.M. and R.H. Topel (2006), “The Value of Health and Longevity”, *Journal of Political Economy*, No. 114, pp. 871-904.

Nadaraya, E.A. (1964), “On Estimating Regression”, *Theory of Probability and its Applications*, Vol. 9, No. 1, pp. 141-42.

Nordhaus, W.D. (2003), “The Health of Nations: The Contribution of Improved Health to Living Standards”, in K.M. Murphy and R.H. Topel (eds.), *Measuring the Gains from Medical Research: An Economic Approach*, University Chicago Press, Chicago (IL).

Pradhan, M., D.E. Sahn and S.D. Younger (2003), “Decomposing World Health Inequality”, *Journal of Health Economics*, Vol. 22, No. 2, pp. 271-93.

Preston, S.H. (2007), “The Changing Relation Between Mortality and Level of Economic Development”, *International Journal of Epidemiology*, No. 36, pp. 484-90.

Quah, D.T. (1997), “Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs”, *Journal of Economic Growth*, Vol. 2, No. 1, pp. 27-59.

R Development Core Team, R: (2006), A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Ram, R. (2006), “State of the ‘Life Span Revolution’ Between 1980 and 2000”, *Journal of Development Economics*, Vol. 80, No. 2, pp. 518-26.

Rodrik, D. (2014), “An African Growth Miracle?”, NBER, Working Paper, No. 20188.

Rosen, S. (1988), "The Value of Change in Life Expectancy", *Journal of Risk and Uncertainty*, Vol. 1, No. 3, pp. 285-304.

Rutstein, S.O. and G. Rojas (2006), "Guide to DHS Statistics", in *Demographic and Health Surveys Methodology*, No. 171, USAID.

Ryan, D.E. (2010), "Trends in World Inequality in Life Span Since 1970", NBER, Working Paper, No. 16088.

Sachs, J.D. (2005), *The End of Poverty: Economic Possibilities for Our Time*, Penguin Books.

Sala-i-Martín, X. (2006), "The World Distribution of Income: Falling Poverty and ... Convergence, Period", *The Quarterly Journal of Economics*, Vol. 121, No. 2, pp. 351-97.

Schultz, T.P. (1998), "Inequality in the Distribution of Personal Income in the World: How It Is Changing and Why", *Journal of Population Economics*, Vol. 11, No. 3, pp. 307-44.

Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Singh, G.K. and M. Siahpush (2006), "Widening Socioeconomic Inequalities in US Life Expectancy 1980-2000", *International Journal of Epidemiology*, Vol. 35, No. 4, pp. 969-79.

Vollmer, S., H. Holzmann and J. Weisbrod (2010), "Peaks vs. Components", Leibniz Universität Discussion Paper, No. 452, Hannover.

Wand, M.P. and M.C. Jones (1993), "Comparison of Smoothing Parametrizations in Bivariate Kernel Density Estimation", *Journal of the American Statistical Association*, No. 88, pp. 520-28.

Watson, G.S. (1964), "Smooth Regression Analysis", *Sankhya: The Indian Journal of Statistics*, A Series, Vol. 26, No. 4, pp. 359-72.

Weil, D.N. (2014), "Health and Economic Growth", in P. Aghion and S. Durlauf (eds.), *Handbook of Economic Growth*, Vol. 2, pp. 623-82.

World Bank (2014), *World Development Indicators*.