

The Global Joint Distribution of Income and Health

Ximing Wu* Andreas Savvides[†] Thanasis Stengos[‡]

March 7, 2008

Abstract

We investigate the evolution of global welfare in two dimensions: income per capita and life expectancy. First, we estimate the marginal distributions of income and life expectancy separately. More importantly, in contrast to previous univariate approaches, we consider income and life expectancy jointly and estimate their bivariate global distribution for 137 countries during 1970 - 2000. We reach several conclusions: the global joint distribution has evolved from a bimodal into a unimodal one, the evolution of the health distribution has preceded that of income, global inequality and poverty has decreased over time and the evolution of the global distribution has been welfare improving. Our decomposition of overall welfare indicates that global inequality would be underestimated if within-country inequality is not taken into account. Moreover, global inequality and poverty would be substantially underestimated if the dependence between the income and health distributions is ignored.

JEL Code: 01; 015; 057

Keywords: Income; Health; Global Distribution; Inequality; Poverty

*Department of Agricultural Economics, Texas A&M University, College Station, TX 77843; xwu@ag.tamu.edu

[†]Department of Commerce, Finance and Shipping, Cyprus University of Technology and Department of Economics, Oklahoma State University; andreas.savvides@cut.ac.cy

[‡]Department of Economics, University of Guelph; tstengos@uoguelph.ca

1 Introduction

The extent of global inequality has long been a controversial issue. It has been the subject of numerous academic studies, dwarfed, however, by discussions in the financial and popular press. Debate on this issue has renewed with increased vigor during the current wave of globalization. The academic debate has been concerned with a number of different issues that, at the risk of oversimplification, can be condensed into two fundamental questions: what should global inequality measure and how should global inequality be measured?

Concerning the first question, it is natural to argue that global inequality should measure an indicator of global welfare. In practice, the concept of welfare has most frequently been identified with a single attribute, income per capita, and its evolution over time. It has long been recognized that welfare encompasses not only income but also other attributes such as access to adequate health facilities or a basic level of education. In this respect, recent research has focused attention on other attributes of welfare and the realization that a global distribution of welfare must, in addition to income, incorporate a measure of health. Becker *et al.* (2005) estimate a monetary equivalent to improvements in life expectancy and calculate the growth rate of “full income” (that includes improvements in life expectancy). They conclude (p. 284) that changes in life expectancy that took place “...in the period between 1960 and 2000 worked towards reducing the disparity in welfare across countries.”

As for the second question, how to measure world or global inequality, the concept has been interpreted variously. At one level, the question concerns whether poorer countries tend to grow faster than richer countries and the investigation of this question has generated a voluminous literature on convergence.¹ One important drawback of interpreting the literature on convergence as an indication of the evolution of world inequality is that there is no account of differences in population and countries of differing population sizes are treated as individual observations. Following up on this criticism, researchers have employed a second measure, weighting each country by its population size. This measure of inequality is also subject to criticism in that each individual in a specific country is assumed to possess the same attribute (e.g. per capita income or life expectancy) as every other individual. In other words, this measure looks at *between-country* inequality but ignores *within-country*

¹The two most popular concepts are β -convergence according to which countries that begin at a lower level of income per capita grow at faster rates and σ -convergence according to which the dispersion of country growth rates decreases over time. For a discussion see, for example, Maasoumi *et al.* (2007) and references therein.

inequality. Therefore, a third measure of inequality, termed global or world inequality, has been introduced that encompasses both between- and within-country differences.²

This paper addresses both questions. We recognize that global inequality refers not only to income, but an estimate of the distribution of global welfare must include a consideration of differences in health levels. Increases in welfare come about not only from increases in income but also from increases in longevity that allow individuals to enjoy income over longer periods. We also recognize that a measure of global inequality must adopt the third interpretation of the concept, accounting for both between- and within-country differences. In what follows, we use the term global or world inequality in the sense of the third definition.

The focus of this study is to estimate the joint global distribution of income and life expectancy. Estimating the joint distribution is especially important for studies that involve multiple attributes because welfare indices, such as inequality or poverty, are generally not uniquely defined for a multivariate distribution. Arbitrarily chosen welfare indices capture only certain aspects of the joint distribution of the attributes in question. To the best of our knowledge, this is the first study to estimate the joint distribution of global income and life expectancy. Existing studies that examine multiple dimensions of global welfare look at each attribute separately and therefore do not take into account the interdependence among the attributes in question.

We propose a new methodology for estimating the joint global distribution from a limited amount of information. The marginal and joint distributions of income and health are estimated using an information-theoretic approach that is based on summary statistics by intervals for each country. Our data cover the years 1970, 1980, 1990 and 2000 for 137 countries (accounting for 95 percent of global population). For each country, interval data on income and health provide information on their respective distributions. However, because income and health information is reported separately within each country, we rely on aggregate country level data to infer the dependence structure between the global distributions of income and health. Due to the specific structure of our data, we propose a novel three-step estimation strategy that uses the device of copula to link estimates based on information at different levels of aggregation.

Estimation of the joint distribution allows a researcher to compute any feature of the distribution and thus facilitates various welfare inferences. Moreover, visual examination of

²Milanovic (2005) discusses the issues and controversies surrounding the definition of the term global inequality.

the joint distribution provides important insight into the joint evolution of the distribution of global income and health that is otherwise difficult to detect from statistical inferences. Our estimates demonstrate a remarkable evolution of the global joint distribution of income and health from a bimodal to a unimodal distribution. Furthermore, they suggest that the distribution of global income and health during the past four decades has been welfare improving. This finding is confirmed by rigorous inequality and poverty inferences. A noteworthy feature of the joint distribution is that the (welfare-improving) evolution of income has lagged behind that of life expectancy.

Finally, we conduct two decompositions of the joint distribution of global welfare in terms of income and health. We decompose overall inequality into within- and between-country components and note that global inequality will be substantially underestimated if one does not take into account the within-country component. The second decomposition computes various welfare indices under the (false) assumption of independence between the income and health distributions. Our results suggest that welfare indices will be generally underestimated if one does not consider the dependence between income and health distributions. In particular, the lower bound of the poverty index of the joint distribution (an indicator of extreme poverty) would be underestimated by an average of over 60%.

The following section provides a short discussion of the relevant literature and relates it to our work. Section 3 discusses data issues. Section 4 introduces the methodology for estimating the joint bivariate distribution of income and life expectancy. Section 5 discusses our estimates of the world distribution of welfare. The final section summarizes and concludes.

2 Literature

This section reviews briefly the literature on the global distribution of welfare and some related empirical issues. As far as income per capita is concerned, there has been a lot of recent interest in estimating its global distribution. The more recent contributions, including Bhalla (2002), Bourguignon and Morrisson (2002), Chen and Ravallion (2001, 2004), Milanovic (2005), and Sala-i-Martin (2006), have computed various measures of global income inequality that take into account both within- and between-country inequality. The indicators of inequality most frequently encountered in this literature are the Gini coefficient, the Theil index and the mean logarithmic deviation. These studies find that the global Gini

coefficient, the most popular index of income inequality, shows either no appreciable trend or a small reduction during recent years. Its estimated value is in the 0.64-0.68 range; it is noteworthy that this range is higher than income inequality of all (except the most unequal) individual societies reflecting substantial inter-country income differences. Some of these studies have also estimated the extent of global poverty, defined as the proportion of global population that falls below a specific level of income or consumption per capita.

There has been less attention devoted to other aspects of welfare and very few studies on measuring their global distribution. Some recent studies have taken into consideration the multidimensional nature of global welfare and have looked into individual attributes other than income. Gakidou *et al.* (2000) discuss some of the conceptual problems associated with measuring health inequality but provide no estimates either for individual economies or globally. Pradhan *et al.* (2003) look at the distribution of height of pre-school age children across countries as an indicator of health inequality. Goesling and Firebaugh (2004) compute various measures of inequality in average life expectancy between countries and discuss how these have evolved during 1980-2000. They, however, ignore differences in life expectancy within countries. Another welfare attribute is education and Thomas *et al.* (2000) provide estimates of the degree of inequality (Gini coefficients) in educational attainment (mean years of schooling) for individual countries during 1960-1990. Morrisson and Murtin (2006) estimate and compare the global marginal distributions of income and education.

Becker *et al.* (2005) look at two measures of inequality between countries: income per capita and life expectancy. To begin with, they treat each attribute separately and show that the two have evolved differently over the 1960-2000 period: there has been a tendency for increased income inequality between countries and a tendency for reduced life expectancy inequality. They caution against drawing conclusions on the world distribution of welfare based on income alone. To address this issue, they calculate a monetary equivalent to improvements in life expectancy, defined as the infra-marginal income (per person and per year) that would provide the same level of utility in the first period (e.g. 1960) but with the mortality rate experienced in the second period (e.g. 2000). Thus, they are able to calculate a growth rate of 'full' income (including the valuation of improvements in longevity) which they compare to 'conventional' income (in PPP dollars) growth rates of 96 countries. They find that, for the poorer half of their sample, the growth rate of full income was much faster than for the richer half. Moreover, they find significant evidence of reversion to the mean for growth rates of full income but not for conventional growth rates. They compute various

measures of the inequality of income (both conventional and full) weighted by each country's population. They, however, abstract from within-country differences.

Considering the joint distribution of individual attributes of welfare (income and life expectancy) is very important because, as Becker *et al.* (2005) point out, income and life expectancy may evolve differently over time and so conclusions about the distribution of global welfare must allow for this. Along the same lines, Bourguignon and Morrisson (2002) trace the global distribution of income (both within and between countries) over the last two centuries. In the final section of their paper, they acknowledge the importance of differences in life expectancy in arriving at a measure of world inequality and compute a measure of between-country inequality in life expectancy. Moreover, they combine income and life expectancy data to derive a measure of between-country inequality in lifetime income.³ Bourguignon and Morrisson (2002, p.741) point out an important caveat to their measure of inequality: "... all individuals within a country or group of countries are assumed to have the same life expectancy. Thus, unlike for income, within-country inequality is assumed to be zero."

In this paper we trace the evolution of the global distribution of per capita income and life expectancy over the period 1970-2000 and draw several conclusions about global welfare over the last four decades. Unlike previous studies, we take into account both between- and within-country inequality in income and life expectancy. Our study is in line with recent developments in welfare economics that place increasing emphasis on multidimensional indicators of welfare (e.g. Maasoumi, 1999). This literature recognizes that analysis based on a single attribute does not capture adequately the different dimensions of welfare. A basic needs approach contends that individual wellbeing and social welfare depend on the joint distribution of various attributes (such as income and health). Traditionally, welfare analysis of multiple attributes is undertaken by examining each individual attribute separately. Evidently, this approach fails to account for the dependence among various attributes. An alternative is to construct a single welfare index as an aggregation of multiple indices of individual attributes. This has led to the calculation of multidimensional indices of welfare such as the Human Development Index (HDI) or the Human Poverty Index (HPI) of the United Nations Development Programme (UNDP, 2006).⁴ Although carefully constructed,

³Lifetime income is defined as $income \times (1 - \exp^{-\delta \times life}) / \delta$ where *income* is a country's per capita income, *life* is life expectancy at birth and δ is the discount rate (set equal to 2 percent).

⁴The HDI is an equally weighted average measure of three indicators of achievement: per capita income, life expectancy, and educational attainment (itself a weighted average of the rates of literacy and school

this kind of ‘hybrid’ index shares a common shortcoming with separate analysis of individual attributes: it does not consider the dependence among various attributes.⁵

In a multivariate framework, one needs to account for the dependence among various attributes to arrive at reliable welfare assessments. Regardless of the choice of welfare index, general inferences should be based on the *joint* distribution of the multiple attributes in question. In this paper we undertake a formal bivariate analysis of the global distribution of income and life expectancy that incorporates explicitly the interdependence between these two attributes.

The estimation of the joint global distribution requires data on the distribution of income and life expectancy for each country. As far as the world distribution of income is concerned, there are two main approaches in the literature. One approach uses data on the distribution of income (or expenditure) within each country from surveys of income (or expenditure) and combines that with estimates of the mean level of income per person from National Account Statistics to arrive at an estimate of income per person for different subgroups of a country’s population (e.g. Bourguignon and Morrisson, 2002; Sala-i-Martin, 2006). This approach has become known as anchoring the income distribution to national account statistics. Some researchers (e.g. Chen and Ravallion, 2004; Milanovic, 2005) have opted for an alternative approach whereby data for both mean income (or consumption) and its distribution within each country are drawn from surveys (see Deaton, 2005, for a discussion and comparison of the two approaches). In this study, we adopt the first method for practical reasons. Income (or expenditure) surveys are not available prior to the late 1980s for a sufficiently representative global sample of population. As the World Bank (2006) argues, if the objective is in tracing the world inequality of income over a longer period, as in our study, researchers must resort to anchoring to national account statistics. Moreover, as Sala-i-Martin (2006) acknowledges, anchoring estimates of income shares to national account statistics enables comparisons with extant estimates of the world distribution of income.

Characterizing the world distribution of health presents a challenge because, as Deaton

enrollment). The HPI is also an equally weighted average measure of three indicators of deprivation: probability of survival to age 40, adult illiteracy, and lack of a decent standard of living (itself an equally weighted average of the proportion of the population without access to improved water source and the proportion of under weight children).

⁵For example, suppose there exists a simple economy with two individuals A and B , each endowed with two attributes X and Y . Consider the following two scenarios: (i) $(X_A = 2, Y_A = 2)$ and $(X_B = 0, Y_B = 0)$ and (ii) $(X_A = 2, Y_A = 0)$ and $(X_B = 0, Y_B = 2)$. Since the marginal distributions of X and Y are identical for these two scenarios, any ‘hybrid’ index that fails to account for the dependence between marginal distributions will conclude that the two share the same level of welfare.

(2006) points out, unlike income there is no natural metric for health. The previous literature offers very little guidance. For instance, there is a lengthy literature on measuring health inequality by ‘gradients’, according to which mortality or life expectancy are related to measures of economic or social status. For example, Cutler *et al.* (2006) report that, in 1980, persons at the bottom 5 percent of the U.S. income distribution had a life expectancy that was about 25 percent lower than those at the top 5 percent of the income ladder, irrespective of age. Such data, however, are only available for selected countries during specific years, a fact that precludes their use in this study. One study that aims to provide a measure of world health inequality is Pradhan *et al.* (2003), who use the standardized height of children aged up to 36 months to measure the within-country component of health inequality. Their data cover 72% of the non-OECD population (for the OECD they have no data and assume there is no stunting of children and height is normally distributed). Moreover, their data are available for a single year mostly during the 1990s. Bourguignon and Morrisson (2002) and Goesling and Firebaugh (2004) use life expectancy as the indicator of health and calculate an index of between-country inequality, assuming zero within-country inequality.

In this study we also use life expectancy as our indicator of health but allow for within-country variation. To accomplish this, we divide each country’s population into age groups and assign to each group the life expectancy corresponding most closely to their year of birth. We treat life expectancy as a summary measure of the various forces (economic or otherwise) that lead to an improvement (or deterioration) of health conditions, as they are reflected in survival rates. It is important to stress that we treat the estimate of life expectancy at birth as a general indicator of the impact of various forces, such as the dissemination of medical knowledge, public health programs to combat the spread of infectious diseases etc., have had on improving health over time or the devastating effect of the HIV/Aids pandemic on the deterioration of health in more recent years.⁶

3 Data

To estimate the global joint distribution of income and life expectancy we collected data on 137 countries for four years: 1970, 1980, 1990 and 2000. These countries account for approximately 95 percent of global population. In this section, we describe the construction

⁶For a comprehensive discussion of the forces that have shaped changes in mortality and life expectancy during the post-WWII period see Cutler *et al.* (2006).

of data used in our estimation.

Data on income per capita for 1970, 1980, 1990 and 2000 are in PPP dollars from the *Penn World Tables 6.1* (PWT 6.1). This database provides estimates in 1996 international prices for most countries beginning in 1950 until 2000. For a few countries, PWT 6.1 does not provide estimates for 2000. For those countries, we used the estimate for the year closest to 2000 (usually late 1990s) and updated it to 2000 using growth rates of real per capita GDP from the *World Development Indicators* (WDI) of the World Bank. For 14 countries in our data set PWT 6.1 does not provide estimates of income per capita and for another 4 (members of the former Soviet bloc) it provides estimates only after 1990. Four of these 18 countries (with a combined population of 291 million in 2000) have a sizeable (more than 10 million) population: Myanmar, the Russian Federation, Sudan and Syria. Rather than exclude these 18 countries from consideration, we resorted to estimates of income per capita from an earlier version of Penn World Tables (PWT 5.6) and adjusted the estimates of per capital income from base year 1985 (the base year of PWT 5.6) to base year 1996. For the Russian Federation, we used the income estimate for the U.S.S.R. from PWT 5.6.⁷

For each country in our sample, income information is reported in the form of interval summary statistics obtained by combining estimates of income per capita in PPP dollars with estimates of its distribution. In particular, the frequency and average income of each interval are reported. The number of income intervals differs between the first three years (1970, 1980, and 1990) and the final year (2000). Our estimation method does not require that the number of intervals be the same in each year. For 1970, 1980, and 1990, we used income interval data from Bourguignon and Morrisson (2002) who provide interval data by decile, except for the top decile where they provide additional data for the top two vintiles. Thus for each country, we were able to partition income data into 11 intervals. An alternative source of data for the distribution of income for these years would have been the WDI. There are two reasons for using the Bourguignon/Morrisson data set: first, it provides a greater number of intervals and thus more detailed information on income distribution, and, second, our results on income distribution can be compared to earlier studies.⁸ For 2000, Bourguignon/Morrisson

⁷The conversion is discussed below in connection with the definition of the poverty line. PWT 5.6 provides estimates of per capita income up to and including 1992. In order to generate an estimate for 2000, we used data on the growth of real per capita GDP from the WDI.

⁸Bourguignon and Morrisson (2002) provide data on income distribution for almost two centuries, the last three years being 1970, 1980 and 1992. We used their 1992 income data to represent 1990 in our data set (see also the next footnote). They provided data for few individual countries but in most cases for geographic groups of countries (see their study for group definitions). Our study is based on country-

do not provide data and we used income interval data from the WDI.⁹ These data are based on household surveys of income (in some cases consumption) from government statistical agencies and World Bank country departments. The WDI provide data for the share of income of the middle three quintiles and for the top two and bottom two deciles. Thus, we were able to partition income into seven intervals.¹⁰

Data on life expectancy *at birth* are also in the form of interval statistics. The most detailed division of each country's population by age is in 5-year intervals from the *World Population Prospects* compiled by the Population Division of the United Nations Department of Economic and Social Affairs (2005). Therefore, we divided each country's population into seventeen age groups as follows: ages 0-4, 5-9, ... 75-79, and 80+. We collected data on the number of people in each group and, consequently, the weight of each group is each group's share in the country's population. The U.N. Population Division defines life expectancy as "...the number of years a newborn infant would live if prevailing age patterns of mortality at the time of its birth were to stay the same throughout its life." For each country, we assume that, for a specific year, all persons in the 0-4 age bracket have the same life expectancy, that of a newborn; they have benefitted (or been affected adversely) by conditions prevailing to that date. For the same year, the life expectancy of all persons in the 5-9 age group is that prevailing in the country five years previously. In the same way, the general forces that shaped the health conditions of persons ages 5-9 were those prevailing in the country five years prior to the specific date. As an example, for a specific country in 2000, all persons in age bracket 0-4 are associated with life expectancy at birth of the country in 2000, persons in the 5-9 age bracket are associated with life expectancy of the country in 1995, persons in the 10-14 age bracket with life expectancy in 1990, and so on. This way we can trace life expectancy of each age group to 80+. In general, in our data younger groups have higher life expectancy due to the overall improvement in health conditions over time.

level data. Therefore, where individual-country interval data were unavailable we used the corresponding geographic-group data.

⁹Income interval data from the WDI are available only for selected years. When referring to data for 2000, we chose the year closest to 2000 with available data (in most cases the late 1990s). This practice is widely adopted in the literature as a practical matter because distribution data are sparse. Many researchers acknowledge that it would not affect results much because income share data do not show wide fluctuations from year to year.

¹⁰For some mainly small countries in our sample (35 of the 137), the WDI does not provide share data; we assumed that the distribution of income did not change between 1990 and 2000. These countries accounted for only 5 percent of global population in 2000.

4 Methodology

The main obstacle to estimating the joint global distribution of income and health is the scarcity of data. If a reliable, representative and sufficiently large sample of the world population with income and health information were available, one can use a nonparametric bivariate density estimator, such as a kernel estimator, to obtain a consistent estimate of the joint distribution. Lacking such data, existing studies estimating the global distribution of income or health use interval data. Moreover, most of these studies use univariate kernel estimation methods. It is well known that kernel estimators require a relatively large sample size and the data requirements for multivariate distributions are increasingly demanding due to the “curse of dimensionality”.

In this section, we propose a novel three-step procedure to estimate the joint distribution of income and health based on available interval data. We use an information-theoretic approach that constructs density estimates based on a set of moment conditions. To estimate a bivariate distribution, we require moments for both marginal distributions and their cross moments. The moments of the marginal distribution of income and health are calculated using country-specific interval data. Because the interval data of each country are reported separately for income and health, they do not provide information on their joint distribution within each country.¹¹ Therefore, we rely on cross-country variation to infer their dependence structure and construct the cross moments using country averages.

In the first step of the procedure, we estimate the two marginal distributions, and their joint distribution, using moments calculated from country average data. Second, we re-estimate the two marginal distributions using moments calculated from interval data for each country. In the third step, we take advantage of the fact that a joint distribution can be expressed as the product of two marginal distributions and a copula function. Consequently, we use the device of copula to combine estimates from the first two stages. Because they use more detailed information, the marginal distribution estimates of the second stage are preferred to those obtained in the first stage. We extract the copula function from the first stage joint distribution and combine it with the two marginal distributions from the second

¹¹To see this, consider again the simple example given in footnote 5. If the data are in the format of $(X_A = 2, Y_A = 2)$ and $(X_B = 0, Y_B = 0)$, then both marginal distributions (of X and Y) and their joint distribution are identified. In contrast, if the data are reported in the format $X = (2, 0)$ and $Y = (2, 0)$, then only the marginal distributions of X and Y are identified, but not their joint distribution. Our data on income and health for each country are of the second format.

stage to form our final estimate of the distribution. The rest of this section describes the proposed estimation procedure.

4.1 Construction of moments

Because our density estimation is based on moments, first we compute the moments of the marginal and joint distributions. The calculation of sample moments is straightforward given a sample of individual data. A complication in our case is that we only have summary statistics on the distribution of income and life expectancy by intervals.¹² Suppose in our data set we have J countries. For the j^{th} country, $j = 1, \dots, J$, denote the density function of life expectancy and income by $f_{X,j}$ and $f_{Y,j}$ respectively. The support of $f_{X,j}$ is partitioned into K_x intervals $(\mathcal{X}_1, \dots, \mathcal{X}_{K_x})$. The population share and average level of life expectancy of the k^{th} interval are, respectively,

$$w_{x,j,k} = \int_{x \in \mathcal{X}_k} f_{X,j}(x) dx,$$

$$\nu_{x,j,k} = \int_{x \in \mathcal{X}_k} x f_{X,j}(x) dx,$$

where $1 \leq k \leq K_x$ and $\sum_{k=1}^{K_x} w_{x,j,k} = 1$. Similarly, the support for the income distribution $f_{Y,j}$ is partitioned into K_y intervals $(\mathcal{Y}_1, \dots, \mathcal{Y}_{K_y})$. The population share and average income of the k^{th} interval are respectively

$$w_{y,j,k} = \int_{y \in \mathcal{Y}_k} f_{Y,j}(y) dy,$$

$$\nu_{y,j,k} = \int_{y \in \mathcal{Y}_k} y f_{Y,j}(y) dy,$$

where $1 \leq k \leq K_y$ and $\sum_{k=1}^{K_y} w_{y,j,k} = 1$. Assuming that each person within a given interval has the interval average income or life expectancy, we compute the approximate moments of

¹²Wu and Perloff (2005) use the maximum entropy density approach to estimate China's income distributions from interval summary statistics.

the life expectancy and income distributions of the global population, respectively, by

$$\begin{aligned}\tilde{\mu}_{x,r} &= \sum_{j=1}^J p_j \sum_{k=1}^{K_x} w_{x,j,k} \mathcal{V}_{x,j,k}^r, \\ \tilde{\mu}_{y,s} &= \sum_{j=1}^J p_j \sum_{k=1}^{K_y} w_{y,j,k} \mathcal{V}_{y,j,k}^s,\end{aligned}$$

where r, s are positive integers, p_j is the population share of the j^{th} country in the world and $\sum_{j=1}^J p_j = 1$.

Since the interval data for income and health are reported separately, they do not provide information on the joint distribution of income and health within each country. Hence, we rely on country averages to compute the approximate cross moments of the joint distribution, which are given by

$$\hat{\mu}_{r,s} = \sum_{j=1}^J p_j \left(\sum_{k=1}^{K_x} w_{x,j,k} \mathcal{V}_{x,j,k} \right)^r \left(\sum_{k=1}^{K_y} w_{y,j,k} \mathcal{V}_{y,j,k} \right)^s \quad (1)$$

where r, s are positive integers. When $r = 0$ or $s = 0$, we obtain the marginal moments of life expectancy or income distribution using country averages. In this study, we compute moments up to order $r + s \leq 4$.

4.2 Information-theoretic density estimation

Given the estimated moments, we use an information-theoretic approach to estimate the underlying densities. Shannon's information entropy is the central concept of information theory. The differential entropy of a continuous univariate random variable x with density function f is defined as

$$W = - \int f(x) \log f(x) dx.$$

Suppose for an unknown univariate density function $f_0(x)$, the only information available is a set of moments $\mu = (\mu_1, \dots, \mu_M)$, where $\mu_m = \int x^m f_0(x) dx, 1 \leq m \leq M$. There may exist an infinite number of distributions satisfying these moment conditions. The principle of maximum entropy (Jaynes, 1957) prescribes a method for constructing a density from given moment information: one chooses the distribution $f(x)$ that maximizes the entropy among all distributions consistent with the given moment information. The resulting maximum

entropy density “... is uniquely determined as the one which is maximally noncommittal with regard to missing information, and that it agrees with what is known, but expresses maximum uncertainty with respect to all other matters.” (Jaynes, 1957).

Formally, the maximum entropy (maxent) density is defined as

$$f = \arg \max_f \left\{ - \int f(x) \log f(x) dx \right\}$$

such that

$$\begin{aligned} \int f(x) dx &= 1, \\ \int x^m f(x) dx &= \mu_m, m = 1, \dots, M. \end{aligned}$$

The resulting maxent density takes the form of a natural exponential family

$$f(x) = \exp \left(\sum_{m=1}^M \theta_m x^m - \psi(\theta) \right) \quad (2)$$

where $\theta = (\theta_1, \dots, \theta_M)$ and $\psi(\theta) = \log \int \exp \left(\sum_{m=1}^M \theta_m x^m \right) dx < \infty$ ensures that $f(x)$ integrates to unity. The θ are the Lagrangian multipliers of the optimization procedure and can be interpreted as the “shadow value” of the constraints. Generally, the maxent density has no analytical solution, and a non-linear optimization method is used to solve for θ numerically.¹³

The principle of maximum entropy generalizes naturally to multivariate distributions. The maxent density for a d -dimensional random vector (X_1, \dots, X_d) from a distribution $f_0(x_1, \dots, x_d)$ is obtained by maximizing

$$W = - \int f(x_1, \dots, x_d) \log f(x_1, \dots, x_d) dx_1 \cdots dx_d$$

¹³Wu (2003) provides details and additional references. Most of the commonly used mathematical distributions can be characterized or closely approximated by a maxent density. For example, if the first two moments of a distribution are known, maximizing the entropy subject to these two moments yields a normal probability density function.

subject to

$$\int f(x_1, \dots, x_d) dx_1 \dots dx_d = 1$$

$$\int x_1^{m_1} \dots x_d^{m_d} f(x_1, \dots, x_d) dx_1 \dots dx_d = \mu_{m_1, \dots, m_d},$$

where $\mu_{m_1, \dots, m_d} = \int x_1^{m_1} \dots x_d^{m_d} f_0(x_1, \dots, x_d) dx_1 \dots dx_d$.¹⁴ See Wu (2007) for the general d -dimensional case and its asymptotic properties.

In this study, the maxent density for the bivariate vector of health and income (X, Y) is specified as

$$f(x, y) = \exp\left(\sum_{r+s=1}^M \theta_{r,s} x^r y^s - \psi(\theta)\right) \quad (3)$$

where r and s are non-negative integers, and $\psi(\theta) = \log \int \exp\left(\sum_{r+s=1}^M \theta_{r,s} x^r y^s\right) dx dy$ is the normalization factor. This density is characterized by the set of moments $\mu = \{E[X^r Y^s] : 1 \leq r + s \leq M\}$. In other words, knowledge of μ suffices to identify $f(x, y)$ as defined in (3). The bivariate normal distribution is a special case of (3) when $M = 2$.

4.3 Construction of joint densities using a copula

The first stage of our estimation consists of estimating densities using country level data. We use (3) to estimate the joint distribution of income and health. We also obtain the marginal distributions, \hat{f}_X and \hat{f}_Y , for health and income, respectively using (2). In the second stage, we estimate the same marginal distributions, denoted by \tilde{f}_X and \tilde{f}_Y , using country specific interval data. In the third stage, we propose a method to combine density estimates from the first and second stage to form an improved estimate of the joint distribution.

The device that facilitates such a combination is called a *copula*. According to Sklar's Theorem, any bivariate density $f(x, y)$ can be expressed as

$$f(x, y) = f_X(x) f_Y(y) c(F_X(x), F_Y(y)),$$

where f_X and F_X are the marginal density and distribution function of X , and f_Y and F_Y are defined similarly. The third factor $c(F_X(x), F_Y(y))$ is called the copula density function.¹⁵

¹⁴To ease notation, we use $\int g(x_1, \dots, x_d) dx_1 \dots dx_d$ in place of $\int \dots \int g(x_1, \dots, x_d) dx_1 \dots dx_d$.

¹⁵See the monograph by Nelsen (1998) for a general treatment of copulas.

Sklar's Theorem states that a bivariate density can be decomposed into the product of two marginal densities and the copula density. The dependence structure between X and Y is completely summarized by the copula density. In the first stage of our estimation, we obtain $\hat{f}_X(x)$, $\hat{f}_Y(y)$ and $\hat{f}(x, y)$, based on which we can infer the copula function between X and Y using

$$\hat{f}(x, y) = \hat{f}_X(x) \hat{f}_Y(y) \hat{c}\left(\hat{F}_X(x), \hat{F}_Y(y)\right). \quad (4)$$

Because the second-stage estimates \tilde{f}_X and \tilde{f}_Y are constructed from country-specific interval data, they are subject to smaller aggregation biases compared to \hat{f}_X and \hat{f}_Y based on country averages. Hence, in the third stage, we combine the copula \hat{c} of the first stage with \tilde{f}_X and \tilde{f}_Y of the second stage to construct a *hybrid* joint density \tilde{f} .

Given the marginal distributions and their joint distribution, the copula density is uniquely defined for continuous variables. The construction of a copula function from the marginal and joint distributions, however, is not an easy task because an analytical solution is generally not available. Two methods are commonly used: fitting a parametric copula or a nonparametric empirical copula. The former involves strong assumptions about functional forms and often is insufficient to capture the dependence structure between variables because most parametric copulas used in practice allow only one or two parameters. The nonparametric empirical copula, although flexible, faces another difficulty: bandwidth selection in a multivariate set up.

To avoid these difficulties we propose a method that does not require an explicit construction of the copula function. Using $F^{-1}(F(x)) = x$, we have from (4)

$$\begin{aligned} \hat{c}\left(\hat{F}_X(x), \hat{F}_Y(y)\right) &= \frac{\hat{f}(x, y)}{\hat{f}_X(x) \hat{f}_Y(y)} \\ &= \frac{\hat{f}\left(\hat{F}_X^{-1}\left(\hat{F}_X(x)\right), \hat{F}_Y^{-1}\left(\hat{F}_Y(y)\right)\right)}{\hat{f}_X\left(\hat{F}_X^{-1}\left(\hat{F}_X(x)\right)\right) \hat{f}_Y\left(\hat{F}_Y^{-1}\left(\hat{F}_Y(y)\right)\right)}. \end{aligned}$$

A *hybrid* joint density can now be constructed by combining estimates from the first two

stages as follows:

$$\begin{aligned}\tilde{f}(x, y) &= \tilde{f}_X(x) \tilde{f}_Y(y) \hat{c}\left(\tilde{F}_X(x), \tilde{F}_Y(y)\right) \\ &= \tilde{f}_X(x) \tilde{f}_Y(y) \frac{\hat{f}\left(\hat{F}_X^{-1}\left(\tilde{F}_X(x)\right), \hat{F}_Y^{-1}\left(\tilde{F}_Y(y)\right)\right)}{\hat{f}_X\left(\hat{F}_X^{-1}\left(\tilde{F}_X(x)\right)\right) \hat{f}_Y\left(\hat{F}_Y^{-1}\left(\tilde{F}_Y(y)\right)\right)}.\end{aligned}\quad (5)$$

In summary, our method combines effectively information from two different levels of aggregation to obtain a final estimate of the joint distribution. The hybrid joint density estimate is a nonlinear combination of five densities linked via an implicit copula function. This final estimate is expected to have smaller aggregation bias than the joint distribution estimated from country average data alone.

4.4 Estimation

For each year in our sample, first we compute the moments of the income and health distribution and their cross moments, up to degree 4. We estimate the first-stage marginal distributions of income and health with country-level aggregate data by

$$\hat{f}_X(x) = \exp\left(\sum_{r=1}^4 \hat{\theta}_{x,r} x^r - \psi\left(\hat{\theta}_x\right)\right) \text{ and } \hat{f}_Y(y) = \exp\left(\sum_{r=1}^4 \hat{\theta}_{y,r} y^r - \psi\left(\hat{\theta}_y\right)\right) \quad (6)$$

The joint densities are estimated in the first stage by

$$\hat{f}(x, y) = \exp\left(\sum_{r+s=1}^4 \hat{\theta}_{r,s} x^r y^s - \psi\left(\hat{\theta}\right)\right) \quad (7)$$

Although relatively simple in functional form, these maxent densities are remarkably flexible. For both the marginal and joint distributions, the density is allowed to be fat-tailed, skewed or multimodal. In fact, we show below that the estimated income distributions and joint distributions are bimodal in the earlier years of our sample.

The second step uses the same specification, i.e. (6), to estimate the marginal densities \tilde{f}_X and \tilde{f}_Y using marginal moments based on country specific interval data. The final estimates of the joint density \tilde{f} are then constructed using (5).

5 Results and Discussion

In this section, first we examine the estimated marginal distributions of income and health and then look at their joint distributions. Estimating the density function rather than some inequality or poverty index allows us to compute any welfare characteristic. Moreover, we can examine the density visually to uncover important features that are otherwise difficult to detect by statistical inference alone. For each marginal and joint distribution, we offer a visual presentation of the estimated densities followed by rigorous tests of stochastic and Lorenz dominance as well as inferences concerning inequality and poverty. It is important to note that upon estimating the joint distribution, we are able to compute any of its features. To save space, we only report and discuss the main results below. Detailed results are available from the authors upon request. The Appendix provides a brief description of the univariate and bivariate welfare measures and the associated stochastic dominance and Lorenz dominance tests that are used in this study.

5.1 The evolution of the global income distribution

The upper plot in Figure 1 shows the estimated (univariate) marginal densities that describe the evolution of the world distribution of income per capita over 1970 – 2000. The horizontal axis plots the logarithm of income per capita. Between 1970 and 2000, a distinct rightward movement of the marginal distribution of income is evident. Furthermore, the evolution of the distribution of global income reveals a bimodality in earlier years with a prominent mode at low levels of income and a less prominent mode at higher levels of income representing those of North America and Europe. Over time, however, the distribution becomes unimodal, especially that of 2000. A similar evolution of the world income distribution, especially the disappearing bimodality over time, is reported in Bourguignon and Morrisson (2002) and Sala-i-Martin (2006).¹⁶

The estimated median income for 1970 was \$639 and the mean \$4,191. The distribution did not experience any noticeable change in shape between 1970 and 1980, except for a slight movement to the right. The median and mean increased to \$721 and \$5,119 respectively. In 1990, the median increased to \$1,044 and the mean to \$6,038. There was a further movement

¹⁶Bourguignon and Morrisson (2002) plot the estimated density only for 1950 and 1992 and work with relative income (per capita income relative to that of the richest country for each year). Nonetheless, a bimodality is present for 1950 in their graphical analysis as well as a rightward shift of the distribution between 1950 and 1992. A similar ‘Twin Peaks’ phenomenon is reported in Quah (1996).

of the distribution to the right and no indication of a secondary mode. In 2000, however, a significant shift of the entire distribution to the right is evident. The median was \$2,780 and moreover, the distribution appears to be considerably more symmetric and the mean level of income was \$7,282. Overall, the income distribution for 2000 represents a departure from earlier years and suggests a movement towards a more equitable distribution of world income.

In order to provide a more rigorous welfare comparison, we use the stochastic dominance test. Stochastic dominance provides unambiguous welfare inferences without restrictive assumptions on the utility function of income. For instance, first order stochastic dominance only requires that marginal utility from income is positive, while second order stochastic dominance requires additionally that the utility function be concave. Our results suggest that the 1980 distribution first-order stochastically dominates 1970 and similarly the 1990 distribution first-order stochastically dominates that of 1980. The 2000 distribution does not first-order stochastically dominate that of 1990. It does, however, second-order stochastically dominate the 1990 distribution.¹⁷ Thus stochastic dominance analysis indicates that there is an unambiguous welfare improvement in the distribution of income from 1970 to 2000.

Estimation of the world distribution of income allows direct measurement of global income inequality and poverty, regardless of the choice of welfare measures. To begin with, we compute several popular measures of income inequality including the Gini coefficient, the Mean Logarithm Deviation (MLD) and Theil's Entropy index. The results are reported in Table 1. For all three indices, the results suggest a relatively stable level of income inequality for the three decades up to 1990, but a reduction in inequality between 1990 and 2000. Our estimate of the Gini is consistent with previous estimates of global inequality that generally fall in the 0.64-0.68 range. For example, Bourguignon and Morrisson (2002) report a value of 0.65 for 1970 and 0.66 for 1992 while Sala-i-Martin's (2006) estimate is 0.65 in 1970 and 1990 but declines to 0.64 in 2000. Bhalla (2002) and Berry and Serieux (2006) also report a reduction in inequality from 1980 to 2000 (from 0.68 to 0.65 and from 0.65 to 0.64, respectively). Milanovic (2005) reports an increase in inequality from 1988 to 1993 but a reduction from 1993 to 1998 when it is equal to 0.64.

For each pair of adjacent years we find that standard Lorenz dominance (a concept closely related to the Gini index) of a more recent year is rejected against an earlier one, but gen-

¹⁷First order stochastic dominance implies second order stochastic dominance, but not vice versa.

eralized Lorenz dominance is established for each pair. Since generalized Lorenz dominance accounts for changes in both average income and the distribution of income, this result indicates that the overall evolution of income distribution has been welfare improving. It should be noted that generalized Lorenz dominance implies second order stochastic dominance and, therefore, these results confirm earlier findings from stochastic dominance. Our general conclusion from welfare analysis is a reduction in income inequality in 2000 compared with earlier decades, a result that supports the graphical analysis.

In addition to measuring the degree of income inequality, estimation of the world distribution of income allows measurement of the extent of global poverty. Defining a criterion that must be met for a person to be classified as poor has been widely debated. One of the earliest attempts to measure global poverty is the 1990 *World Development Report* (World Bank, 1990, p. 26) that defines poverty as "... the inability to attain a minimal standard of living." While acknowledging the multidimensionality of the concept of poverty, in order to arrive at a measure the WDR defines poverty in terms of a (consumption) threshold that an individual must meet. The threshold was expressed in dollars at 1985 international prices and was meant to be representative of the poverty line found in some of the poorest countries. Two thresholds were defined: \$370 per year to represent poverty and \$275 per year to represent extreme poverty. The first (corresponding to \$1.02 per day) has frequently been referred to as the one-dollar-per-day (\$1/day) line.

More recent estimates of poverty for the World Bank by Chen and Ravallion (2001, 2004) have used 1993 prices. Chen and Ravallion define the poverty line in a similar way, as that representing the median of the lowest ten poverty lines (in 1993 prices) amongst developing countries. Their poverty threshold (in 1993 prices) was \$1.08/day (or \$32.74/month), referred to also as the \$1/day poverty line. In addition to the one-dollar-per-day line, Chen/Ravallion also present estimates of poverty based on \$2.15/day in 1993 prices (referred to as the two-dollars-per-day poverty line). These two poverty lines are regularly presented in discussions of the extent of global poverty. The Chen/Ravallion lines are intended to represent a minimum level of consumption. As indicated previously, several researchers have anchored distribution data to estimates of income per person. Sala-i-Martin's (2006) estimates of income are from PWT 6.1 and are expressed in 1996 prices (as are our estimates). He defines several poverty lines, the first of which is \$495 per year. This is obtained by converting the original World Bank poverty threshold of \$1.02/day (or \$370 per year in 1985 prices) to 1996 prices. He defines another poverty line at \$730 per year or \$2/day in 1996

prices and refers to it as the two-dollars-per-day line. Sala-i-Martin argues that because his threshold is based on income while the Chen/Ravallion (World Bank) is based on consumption, his two-dollar-per day estimates of poverty are more akin to the one-dollar-per day estimates of Ravallion/Chen.¹⁸ In our study we report poverty rates based on the two thresholds of Sala-i-Martin (2006), noting the caveat about thresholds expressed in terms of income and consumption.¹⁹

Estimates of the extent of global poverty are reported in Table 1. We compute three popular measures of global poverty: (i) the poverty count (or head count) is the number of persons below the poverty line; (ii) the poverty rate (or head count ratio) is the percentage of the global population that falls below the poverty line; and (iii) the poverty gap is the percentage of global income that is needed to bring the income of persons below the poverty line up to the line. For the \$1/day poverty line, the poverty rate decreased from 19.20% in 1970 to 5.69% in 2000 (or a drop of some 70%) and the poverty gap decreased by 68%. Our estimates of the global poverty rate are close to those of Sala-i-Martin's (2006) who finds the poverty rate to be 20.2%, 15.9%, 7.3% and 5.7% (for the four years of our sample, respectively). At the same time, the poverty head count, even without adjusting for the sizable increase in population, decreases by 51%. Similar patterns are observed for the \$2/day poverty line. We note, however, that if we adopt the definition of extreme poverty (\$275 per year), the poverty rates (not reported in the table) are 3.7%, 2.2%, 0.9% and 1.5%, indicating that the initial improvement in extreme poverty of earlier decades did not persist, and was indeed reversed between 1990 and 2000.

5.2 The evolution of the global health distribution

The lower panel of Figure 1 plots the univariate world distribution of life expectancy estimated for the same four years as income. The evolution of the health distribution over time appears welfare improving. There is a distinct shift of the distribution from 1970 to 1980 as the main and secondary modes switch positions with the main mode moving from lower towards higher life expectancy. The improvement may be indicative of the success of public health programs in improving life expectancy in the developing world. The bimodality

¹⁸On this issue, see the debate through the pages of *The Economist* of July 18 2002, March 11 2004, and April 7 2004.

¹⁹Results using the Chen/Ravallion poverty line and other reasonable poverty lines are qualitatively similar in terms of their evolution over time. These additional results are available from the authors upon request.

remains in the 1980 distribution with the secondary mode around 36 years, the same life expectancy level as the main mode in 1970. In just one decade, the mean life expectancy increases by four years and the median by six. Increases in mean and median life expectancy continue in later years (1990 and 2000) but at a slower pace.

Stochastic dominance analysis suggests that each year first-order stochastically dominates the previous one. Lorenz dominance is also established for each pair of adjacent years. Correspondingly, inequality in global health decreases substantially during the sample period as shown by the inequality indices in Table 2. For example, the Gini index decreases from 0.167 to 0.122 or by 27 percent between 1970 and 2000.

When it comes to defining a measure of ‘poverty’ for the life expectancy distribution, there is no definition equivalent to income poverty. In order to derive an ‘equivalent’ measure to income poverty, which can be termed health deprivation, we follow the World Bank methodology in defining an income poverty line. The poverty line chosen by the World Bank represents the median poverty line of the ten poorest countries (expressed in 1985 or 1993 PPP dollars - see Chen and Ravallion, 2001 and 2004). In a similar manner, we focus on life expectancy at birth of the ten countries with the lowest levels of life expectancy in 1970-2000. The median life expectancy of the ten most deprived countries was 38.4 in 1970, 42.7 in 1980, 42.4 in 1990 and 39.4 in 2000. Thus a life expectancy of 40 years may serve as a useful threshold measure of health deprivation. The three indicators of health deprivation (defined in a manner similar to income poverty) based on this threshold are shown in Table 2. All three show a steady improvement in health deprivation. In particular, the percentage of population with life expectancy below 40 decreases from 37 percent in 1970 to 12 percent in 2000 or by 68%. Similarly, the ‘poverty gap’ in health decreases by 77%.

In sum, the univariate (marginal) income and health distributions display a similar pattern of lower inequality and welfare improvement between 1970 and 2000. The health distribution shows somewhat more marked changes than those exhibited by the income distribution.

5.3 The joint global distribution of income and life expectancy

The estimated joint distributions of income and life expectancy are plotted in Figures 2 (for 1970 and 1980) and 3 (for 1990 and 2000). The contours of the joint distributions are also plotted to help visualize the overall pattern. There are a number of interesting

features revealed by the bivariate distribution graphs. The overall evolution of the joint distributions from a bimodal distribution to a unimodal one is consistent with convergence of income and health during the past four decades. The 1970 distribution has two significant modes, one at low income and low life expectancy levels and the other at high income and high life expectancy. In effect, the joint distribution consists of separate ‘higher’ and ‘lower’ distributions, reflecting the substantial gap in terms of income and health between the developed countries and the rest of world. The distribution at higher levels of income and life expectancy is more concentrated, indicating a relatively high degree of homogeneity within the developed countries. Its narrow profile along the diagonal of the income-health space suggests a strong positive correlation between these two attributes for these countries. The distribution at lower levels is clearly more dispersed, reflecting a substantial degree of heterogeneity among developing countries. The round profile of this distribution also suggests a weaker correlation between income and health among countries in this group.

The 1980 distribution exhibits a similar bimodal pattern. There are, however, a few noteworthy differences. The lower distribution moved ‘up’ in terms of both income and health. At the same time, the higher distribution also improved in both dimensions. By comparing the two contours, we note that the gap between the two distributions in terms of life expectancy was reduced, while the income gap changed little. In particular, the mode of the lower distribution improved considerably in terms of life expectancy.

The lack of clear dependence (as portrayed by the wider contours of the joint distribution) between income and health in the developing world for the earlier years can be explained, to some degree, by its composition of two different country groups, the first with a failing health system and the second with a relatively advanced health system. The former group includes countries in Sub-Saharan Africa and parts of Asia plagued by infectious diseases, while the latter group includes countries (at the time) in the socialist group with government-run universal health care systems and some countries in South and South East Asia (e.g. Bangladesh, Indonesia, Pakistan, Sri Lanka) with relatively well run systems despite a relatively low income level.

The most marked change is observed in 1990, when the previously separate higher and lower distributions merge into a nearly unimodal one. There is a dominant mode around \$1,900 and life expectancy of 59 years. The joint distribution also shows a minor mode at higher income and health level, but the two modes are not completely separated. The movement toward a unimodal distribution is further solidified during the 1990s as evidenced

by the 2000 distribution when the joint distribution is clearly unimodal, with a single significant mode at \$7,100 and life expectancy of 67 years. The ‘volume’ at the lower tail of the distribution is substantially reduced compared to earlier years. At the same time, the upper end of the distribution becomes more concentrated (as shown by the contours), suggesting an exceptionally high level of income and longevity enjoyed by a rather small proportion of the global population.

The observed evolution of income and health has certainly been shaped by a whole host of forces. One major contributing factor is the rapid growth of China and India during this period. China’s and India’s share of the global population is approximately 23% and 17% during the sample period. Both countries experienced rapid and substantial improvements in terms of income and health. The per capita income of China quadrupled during 1970-2000, from \$820 in 1970 to \$3747 in 2000; the corresponding average life expectancy was 44, 50, 54 to 58 for the four sample years, respectively. India experienced a similar, albeit, less remarkable improvement in living standards with the corresponding levels for income at \$1,077, \$1,162, \$1,675 and \$2,480, and for life expectancy at 39, 43, 48 and 52. Such substantial and steady improvements in terms of income and health in about two-fifths of the global population has been a major driving force behind the evolution of the global joint distribution observed in our findings.

The general conclusion that emerges from the estimated joint distribution is the gradual evolution from a bimodal to a unimodal one. At the same time, however, there are substantial observed differences in life expectancy between relatively similar (in terms of income) countries, especially seen in the dispersion of the contours of the distribution at the lower end during earlier years. The dispersion at the lower end narrows over time, but the contours remain more dispersed compared to those at the higher end. The observation that countries at roughly the same level of income can display substantial differences in life expectancy has been raised in the literature in connection with the “Preston curve”. The curve portrays the cross-sectional relationship between life expectancy and per capita income.²⁰ One of the important characteristics of this curve is that it shifts vertically upward (along its entire range) when plotted for successive years during the period 1960-2000. Our estimate of the joint distribution is consistent with the curve and especially for countries at lower levels of

²⁰Deaton (2006), Soares (2005), and World Bank (2006) discuss the shape of this curve. With per capita income plotted along the horizontal and life expectancy along the vertical axes, at low levels of income the curve is steeply upward sloping and tapers off at higher levels of income.

income the greater dispersion of the contours is consistent with the steep upward slope of the curve.

To facilitate formal comparisons of the joint distribution over time, we compute various inequality and poverty indices. The inequality results are presented in Table 3. They show a similar pattern to the univariate inequality results for income or life expectancy. The Gini coefficient for joint inequality remained roughly constant between 1970 and 1980 (at 0.45) but declined in 1990 with a further more marked decline to 0.395 in 2000. Joint world inequality is lower than income inequality alone but higher than health inequality. At 0.40-0.45, the Gini coefficient for joint world inequality is comparable to the Gini for income inequality of most middle-income countries.²¹

Finally, in Table 3 we also present the poverty rate results. There is no unique extension of the univariate poverty index to the higher dimensional case. What constitutes poverty in a multidimensional setting is a complex question and poverty is not uniquely defined when multiple attributes are considered. An individual with all her attributes below the poverty lines is certainly poor. It is not clear, however, if an individual with a strict subset of her attributes below the poverty lines should be counted as poor. To avoid an arbitrary choice of a single poverty definition, we chose to report the lower and upper bound of the joint poverty index. Given a poverty line for each attribute, the lower-bound poverty measure includes only those with all attributes below the poverty lines, while the upper bound measure includes those with at least one attribute below the poverty line.²² Both the lower-bound and upper-bound poverty rate decreased during the sample period. When \$1/day and 40 years are used as the poverty lines, the lower bound decreased from 11.94% to 4.68%, while the upper bound decreased from 40.37% to 13.04%. There is also a narrowing of differences between the upper and lower bound: in 1970 the gap was a factor of 3.5 and by 2000 it narrowed to a factor of 2.8. Moreover, despite the sizable increase in global population, the poverty head count (either lower or upper bound) declined steadily during the sample period. A similar pattern is observed when we use the \$2/day income poverty line.

One noteworthy result is that the correlation between income and health distribution

²¹There is no well-defined higher dimensional extension of Lorenz dominance. Stochastic dominance is still applicable. First and second order bivariate stochastic dominance is rejected for each pair of adjacent distributions. See Appendix for a brief discussion of relevant concepts.

²²Any reasonably defined poverty index resides within the range between the lower and upper bound. Atkinson (1987) discusses the relationship between the bounds of poverty measures and multidimensional stochastic dominance.

decreased during the sample period from 0.67 to 0.51, indicating lower importance of income in driving improvements in life expectancy (and vice versa). Our findings are consistent with existing arguments in the literature that suggest that major health improvements have been realized that are only weakly related to income improvements. For example, Soares (2005) argues that increases in longevity precede reductions in fertility as a country experiences demographic transition. Increases in longevity are largely unrelated to income improvements but come about mainly through advances in knowledge and technological breakthroughs in medical sciences. He demonstrates that the increase in longevity makes possible endogenous changes in fertility and investment in education that allows a country to experience sustained increases in per capita income.

5.4 Decomposing global welfare

In this section, we carry out two decompositions of the global joint distribution of income and health. The first looks at the relative contributions of between- and within-country to global inequality. The second looks at the impact of the two marginal distributions and their dependence structure on the distribution of global welfare. Evidently, the first is a decomposition by (sub)group while the second a decomposition by source.

First, we decompose overall inequality into between- and within-country inequality; this is done for the two marginal distributions and the joint distribution. The between- and within-country inequality decomposition identifies the structure of overall inequality by investigating the composition of some mutually exclusive subgroups (in our case, countries). The computational details, especially for the higher dimensional decomposition, are in the Appendix.

The decomposition results are in Table 4. Because the Gini index is not decomposable, we report results for the MLD and entropy index. The results from the two indices are rather similar and, thus, we restrict discussion to the MLD. For the marginal distribution of income, the contribution of between-country inequality is the dominant source of overall inequality but its relative importance decreased from 67% of the total in 1970 to 64% in 2000. This is consistent with previous evidence (e.g. Bourguignon and Morrisson, 2002). As far as the marginal distribution of life expectancy is concerned, the dominant source is within-country inequality and its contribution increased from 66% in 1970 to 73% in 1990 and 70% in 2000. The changes in the relative contribution of the two components are similar for both income

and life expectancy, displaying an increasing contribution of within-country inequality over time. The structure of overall inequality is different with around two thirds of overall income inequality due to differences in average income across countries and only one third of life expectancy inequality due to cross-country differences. Therefore, the computation of global inequality, especially that of life expectancy, will be substantially downward biased when within-country inequality is ignored.

As far as the decomposition of the joint distribution is concerned, the results are similar to those of income inequality: about two thirds of overall inequality can be attributed to between-country inequality. The relative contributions of between- and within-country inequality remain virtually constant. Hence, although the degree of overall inequality of global welfare (computed from the joint distribution of income and health) has changed considerably during the past two decades, the structure of inequality in terms of between- and within-country inequality has changed relatively little.

The second decomposition compares the inequality indices and poverty estimates computed from the estimated joint distribution to those indices computed under the (incorrect) assumption of independence between the two marginal distributions. In the latter case, the inequality indices are those computed using the product of the two marginal distributions. The results are in Table 5. We note that the Gini index of overall inequality will be underestimated by 10% and the other two indices by around 20% if we do not take into account the nature of the joint distribution. The bias is more severe for poverty estimates. For example, with \$1/day as the income poverty line, the lower bound is underestimated by 42% in 1960 and as high as 86% in 2000, while the upper bound is overestimated by 20%-30%. Thus without taking into account the dependence structure between income and health, the global level of extreme poverty (people with both income and life expectancy below the corresponding poverty lines) is understated severely. It is also noteworthy that the degree of underestimation rises over time and this suggests that extreme poverty is increasingly concentrated in a small group of the global population, although the ratio of people in this category is steadily declining. More precisely, although the lower tails of the two marginal distributions decreased over time, their co-movement (or lower tail dependence) actually increased during the sample period.

6 Conclusion

In this paper we have estimated the univariate distribution of world income per capita and computed various indicators of world income inequality and poverty over the 1970 – 2000 period. The results are in line with previous estimates of global income inequality. We also estimated the univariate distribution of life expectancy that encompasses both within- and between-country differences in life expectancy. More importantly, and the focus of our study is the bivariate distribution of these two attributes of welfare. The main obstacle to estimating the joint global distribution of income and health is the sparsity of data. We design and implement a novel three-step approach to estimating the joint distribution of income and health using an information-theoretic approach. In particular, estimates based on information at different levels of aggregation are combined through an implicit copula function.

Based on the estimated joint distributions, we computed various indices of global welfare inequality and also estimated the extent of global poverty in two dimensions: income and health. Several insights are obtained from the estimated joint distributions. We find a remarkable change in the global distribution of income and health. Whereas in 1970 the joint distribution had two modes, one at high income and life expectancy, and the other at low income and life expectancy, over time the joint distribution has evolved from a bimodal into a unimodal one. The evolution of the joint distribution indicates unequivocal improvement in the distribution of global welfare. The evolution of the income distribution lags behind that of life expectancy, a finding that is consistent with recent suggestions in the literature, such as Soares (2005).

Our decomposition of the overall joint inequality into within- and between-country components suggests that both have declined over time and their contribution to overall inequality has remained roughly constant. For the two marginal distributions, however, the relative contribution of within-country to overall inequality has increased over time. Finally, our results demonstrate the risk of undertaking welfare inferences on multiple welfare indicators separately. Ignoring the dependence between the income and health distributions underestimates substantially the degree of global inequality and poverty.

References

- Atkinson, Anthony B.**, 1987, "On the Measurement of Poverty," *Econometrica*, 55, 749-764.
- Becker, Gary S., Tomas J. Philipson and Rodrigo R. Soares**, 2005, "The Quantity and Quality of Life and the Evolution of World Inequality," *American Economic Review*, 95, 277-291.
- Berry, Albert and John Serieux**, 2006, "Riding the Elephants: The Evolution of World Economic Growth and Income Distribution at the End of the Twentieth Century (1980-2000)," U.N. Department of Economic and Social Affairs, Working Paper No. 27.
- Bhalla, Surjit**, 2002, *Imagine There is No Country*, Washington DC: Institute for International Economics.
- Bourguignon, François and Christian Morrisson**, 2002, "Inequality Among World Citizens: 1820-1992," *American Economic Review*, 92, 727-744.
- Chen, Shaohua and Martin Ravallion**, 2001, "How Did the World's Poor Fare in the 1990s?" *Review of Income and Wealth*, 47, 283-300.
- Chen, Shaohua and Martin Ravallion**, 2004, "How Have the World's Poorest Fared since the Early 1980s?" *World Bank Research Observer*, 19, 141-169.
- Cutler, David, Angus Deaton and Adriana Lleras-Muney**, 2006, "The Determinants of Mortality," *Journal of Economic Perspectives*, 20, 97-120.
- Deaton, Angus**, 2005, "Measuring Poverty in a Growing World (or Measuring Growth in a Poor World)," *Review of Economics and Statistics*, 87, 1-19.
- Deaton, Angus**, 2006, "Global Patterns of Income and Health: Facts, Interpretations, and Policies," NBER Working Paper No. 12735.
- Foster, James, Joel Greer, and Erik Thorbecke**, 1984, "A Class of Decomposable Poverty Measures," *Econometrica*, 52, 761-766.
- Gakidou, E.E., C.J.L. Murray and J. Frenk**, 2000, "Defining and Measuring Health Inequality: An Approach Based on the Distribution of Health Inequality," *Bulletin of the World Health Organization*, 78, 42-54.
- Goesling, Brian and Glenn Firebaugh**, 2004, "The Trend in International Health Inequality," *Population and Development Review*, 30, 131-146.

- Jaynes, Edwin T.**, 1957 “Information Theory and Statistical Mechanics.” *Physics Review*, 106, 620-30.
- Maasoumi, Esfandiar**, 1986, “The Measurement and Decomposition of Multidimensional Inequality,” *Econometrica*, 54, 991-997.
- Maasoumi, Esfandiar**, 1999, “Multidimensional Approaches to Welfare Analysis,” ed. J. Silber. Dordrecht and Boston, Kluwer Academic Publishers.
- Maasoumi, Esfandiar**, Jeffrey Racine, and Thanasis Stengos, 2007, “Growth and Convergence: A Profile of Distribution Dynamics and Mobility,” *Journal of Econometrics*, 136, 483-508.
- Milanovic, Branko**, 2005, *Worlds Apart: Measuring International and Global Inequality*, Princeton NJ: Princeton University Press.
- Morrisson, Christian** and **F. Murtin**, 2006, “Education Inequalities Among World Citizens: 1870-2000,” Working Paper.
- Nelsen, R.B.**, 1998, *Introduction to Copulas*, Springer Verlag: Berlin.
- Pradhan, Menno, David E. Sahn** and **Stephen D. Younger**, 2003, “Decomposing World Health Inequality,” *Journal of Health Economics*, 22, 271-293.
- Quah, Danny T.**, 1996, “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics,” *Economic Journal*, 106, 1045-1055.
- Sala-i-Martin, Xavier**, 2006, “The World Distribution of Income: Falling Poverty and ... Convergence, Period,” *Quarterly Journal of Economics*, 121, 351-397.
- Soares, Rodrigo R.**, 2005, “Mortality Reductions, Educational Attainment, and Fertility Choice,” *American Economic Review*, 95, 580-601.
- Thomas, Vinod, Yan Wang** and **Xibo Fan**, 2000, “Measuring Education Inequality: Gini Coefficients of Education,” World Bank Working Paper, No. 2525
- United Nations Department of Economic and Social Affairs**, 2005, *World Population Prospects: The 2004 Revision*, available online at <http://esa.un.org/unpp/>
- United Nations Development Programme**, 2006, *Human Development Report 2006, Beyond Scarcity: Power, Poverty and the Global Water Crisis*, New York: Palgrave Macmillan.
- World Bank**, 1990, *World Development Report 1990: Poverty*, New York: Oxford University Press for the World Bank.

World Bank, 2006, *World Development Report 2006: Equity and Development*, New York: Oxford University Press for the World Bank.

Wu, Ximing, 2003, "Calculation of Maximum Entropy Densities with Application to Income Distribution," *Journal of Econometrics*, 115, 347-354.

Wu, Ximing, 2007, "Exponential Series Estimator of Multivariate Density," working paper, Texas A&M University.

Wu, Ximing and **Jeffrey, M. Perloff**, 2005, "China's Income Distributions, 1985-2001," *Review of Economics and Statistics*, 87, 763-775.

Appendix

In this appendix, we offer a brief description of various indices and tests for univariate and bivariate distributions used in arriving at welfare inferences.

- Gini index, Lorenz dominance, Generalized Lorenz dominance

Given a distribution function $F(x)$ defined on $(0, \infty)$, the Lorenz function is defined as $G(p) = \frac{1}{\mu} \int_0^p Q(t) dt$, where $0 < p < 1$, $Q(\cdot)$ is the quantile function of $F(x)$ and $\mu = \int_0^\infty x dF(x)$. The Gini index is then given as $g = 1 - 2 \int_0^1 G(p) dp$.

The distribution F_1 is said to Lorenz dominate F_2 if $G_1(p) > G_2(p)$ for all $0 < p < 1$. Denote μ_1 and μ_2 the mean of F_1 and F_2 respectively. Then F_1 generalized Lorenz dominates F_2 if $\mu_1 G_1(p) > \mu_2 G_2(p)$ for all $0 < p < 1$.

An alternative definition for the univariate Gini index is half of the relative mean difference or $g = \frac{1}{2} \int_0^\infty \int_0^\infty \frac{|x_1 - x_2|}{\mu} dF(x_1) dF(x_2)$. Given a bivariate distribution $F(x, y)$, the bivariate Gini index used in the text is defined as:

$$g = \frac{1}{4} \int_0^\infty \int_0^\infty \left(\left(\frac{x_1 - x_2}{\mu_x} \right)^2 + \left(\frac{y_1 - y_2}{\mu_y} \right)^2 \right)^{1/2} dF(x_1, y_1) dF(x_2, y_2),$$

where $\mu_x = \int_0^\infty \int_0^\infty x f(x, y) dx dy$ and $\mu_y = \int_0^\infty \int_0^\infty y f(x, y) dx dy$.

- Mean Log Deviation and Theil's Entropy Index

The mean log deviation and entropy index belong to the family of general entropy

(GE) index. Following Maasoumi (1986), the GE index is defined as

$$I_\gamma = \sum_{i=1}^N p_i [(s_i/p_i)^{1+\gamma} - 1] / [\gamma(1 + \gamma)],$$

where $s_i = X_i / (\sum_{i=1}^N X_i)$ and p_i is the population share of the i^{th} unit. I_0 and I_1 are then respectively the mean log deviation and entropy index with

$$I_0 = \sum_{i=1}^N s_i \log(s_i/p_i),$$

$$I_{-1} = \sum_{i=1}^N p_i \log(p_i/s_i).$$

The GE index can be decomposed into “between” and “within” group inequalities. Let there be G groups, $t_r, r = 1, \dots, G$, each containing N_r individuals, $\sum_{r=1}^G N_r = N$. Let $P_r = N_r/N$, $S_r = \sum_{i \in t_r} s_i$, $S = (S_1, \dots, S_G)$, and s_r^* be the N_r vector of relative share $s_{ir}^* = s_i/S_r$. Then

$$I_\gamma = I_\gamma(S) + \sum_{r=1}^G P_r^{-\gamma} S_r^{1+\gamma} I_\gamma(s_r^*).$$

For multivariate variable $X_i = (X_{i1}, \dots, X_{id})$, the GE family is defined similarly except that $s_i = h(X_i) / \sum_{r=1}^G h(X_i)$, where $h(X)$ is an “aggregate” attribute function. For the bivariate variable (X, Y) examined in this study, we define $h(X_i, Y_i) = (X_i / \sum_{i=1}^N X_i)^{1/2} (Y_i / \sum_{i=1}^N Y_i)^{1/2}$.

- Poverty rate, head count and poverty gap

The poverty indices examined in this study belong to the family of FGT index (Foster, *et al.*, 1984). For a univariate variable X and poverty line z , the FGT measure is defined as

$$P_\alpha = \frac{1}{N} \sum_{i=1}^N \left(\frac{z - X_i}{z} \right)^\alpha I(X_i < z).$$

P_0 and P_1 correspond to the poverty rate and poverty gap used in this study. The poverty count is defined as the total population multiplied by poverty rate.

The poverty index for a multivariate variable (X_1, \dots, X_d) is not uniquely defined as it

depends on the definition of the poor. Given a poverty line $Z = (z_1, \dots, z_d)$, one can generally define the poor as those with a subset of their attributes below the poverty line. We can, however, define the lower and upper bound of the poverty index. The former includes only those with *all* of their attributes below the poverty line, while the latter includes those with at least one of their attributes below the poverty line.

In this study, let (z_x, z_y) be the poverty line for health and income distribution. The lower and upper bound of poverty rate is defined as

$$P_0^l = \frac{1}{N} \sum_{i=1}^N I(X_i < z_x \text{ and } Y_i < z_y),$$

$$P_0^u = \frac{1}{N} \sum_{i=1}^N I(X_i < z_x \text{ or } Y_i < z_y).$$

Note that when $\alpha \neq 0$, P_α is generally not well defined for multivariate variables except for the lower bound case. This is because $\left(\frac{z_j - X_{ij}}{z_j}\right)^\alpha$, $1 \leq j \leq d$, can take negative values except for the lower bound case where $X_{ij} < z_j$ for all j . Hence, the discussion of the poverty gap (P_1) is restricted to the marginal distributions in this study.

- Stochastic dominance: See Atkinson (1987) for multivariate stochastic dominance and its relationship to the lower and upper bounds of the poverty rate.

Table 1: Welfare indices of income distribution

| | 1970 | 1980 | 1990 | 2000 |
|---------------------------------------|---------|---------|---------|---------|
| Population (millions) | 3,507 | 4,209 | 5,007 | 5,763 |
| Mean | \$4,191 | \$5,119 | \$6,038 | \$7,282 |
| Median | \$639 | \$721 | \$1,044 | \$2,780 |
| Inequality index | | | | |
| Gini | 0.681 | 0.691 | 0.687 | 0.643 |
| Mean log deviation | 0.950 | 0.994 | 0.952 | 0.824 |
| Theil's Entropy | 0.889 | 0.920 | 0.936 | 0.846 |
| Poverty index (poverty line: \$1/day) | | | | |
| Poverty rate (%) | 19.20 | 14.91 | 8.31 | 5.69 |
| Poverty count (millions) | 673 | 628 | 416 | 328 |
| Poverty gap (%) | 5.41 | 3.86 | 1.95 | 1.78 |
| Poverty index (poverty line: \$2/day) | | | | |
| Poverty rate (%) | 33.47 | 28.66 | 18.93 | 11.05 |
| Poverty count (millions) | 1,174 | 1,206 | 948 | 637 |
| Poverty gap (%) | 12.28 | 9.73 | 5.71 | 3.90 |

Table 2: Welfare indices of life expectancy distribution

| | 1970 | 1980 | 1990 | 2000 |
|---|-------|-------|-------|-------|
| Population (millions) | 3,507 | 4,209 | 5,007 | 5,763 |
| Mean | 46.8 | 50.8 | 54.3 | 57.3 |
| Median | 45 | 51 | 56 | 59 |
| Inequality index | | | | |
| Gini index | 0.167 | 0.153 | 0.136 | 0.122 |
| Mean log deviation | 0.045 | 0.040 | 0.033 | 0.027 |
| Theil's Entropy | 0.043 | 0.037 | 0.030 | 0.025 |
| Poverty index (poverty line: expectancy = 40 years) | | | | |
| Poverty rate (%) | 37.32 | 26.28 | 17.94 | 12.28 |
| Poverty count (millions) | 1,309 | 1,106 | 898 | 708 |
| Poverty gap (%) | 6.97 | 4.53 | 2.79 | 1.58 |

Table 3: Welfare indices of joint distribution

| | 1970 | 1980 | 1990 | 2000 |
|--|--------------|--------------|--------------|--------------|
| Population (millions) | 3,507 | 4,209 | 5,007 | 5,763 |
| Mean (dollar, year) | (\$4,191;47) | (\$5,119;51) | (\$6,038;54) | (\$7,282;57) |
| Correlation | 0.673 | 0.624 | 0.565 | 0.513 |
| Inequality index | | | | |
| Gini | 0.447 | 0.450 | 0.431 | 0.395 |
| Mean log deviation | 0.329 | 0.333 | 0.305 | 0.263 |
| Theil's entropy | 0.337 | 0.342 | 0.315 | 0.259 |
| Poverty index (poverty line: \$1/day; expectancy = 40 years) | | | | |
| Poverty rate lower bound (%) | 11.94 | 7.11 | 5.36 | 4.68 |
| Poverty rate upper bound (%) | 40.37 | 30.91 | 19.19 | 13.04 |
| Poverty count lower bound (millions) | 419 | 299 | 268 | 270 |
| Poverty count upper bound (millions) | 1,416 | 1,301 | 961 | 751 |
| Poverty index (poverty line: \$2/day; expectancy = 40 years) | | | | |
| Poverty rate lower bound (%) | 18.98 | 12.12 | 9.41 | 7.75 |
| Poverty rate upper bound (%) | 47.29 | 39.68 | 25.52 | 15.52 |
| Poverty count lower bound (millions) | 666 | 510 | 471 | 447 |
| Poverty count upper bound (millions) | 1,659 | 1,670 | 1,278 | 894 |

Table 4: Decomposition of global inequality into between- and within-country components

| | | 1970 | | 1980 | | 1990 | | 2000 | | |
|------------------------------|---------|-------|-------|-------|-------|-------|-------|-------|-------|--|
| Income distribution | | | | | | | | | | |
| Mean log deviation | Overall | 0.950 | | 0.994 | | 0.952 | | 0.824 | | |
| | Between | 0.641 | (67%) | 0.670 | (67%) | 0.609 | (64%) | 0.527 | (64%) | |
| | Within | 0.309 | (33%) | 0.324 | (33%) | 0.343 | (36%) | 0.298 | (36%) | |
| Theil's entropy | Overall | 0.889 | | 0.920 | | 0.936 | | 0.846 | | |
| | Between | 0.579 | (65%) | 0.605 | (66%) | 0.589 | (63%) | 0.520 | (61%) | |
| | Within | 0.310 | (35%) | 0.315 | (34%) | 0.347 | (37%) | 0.326 | (39%) | |
| Life expectancy distribution | | | | | | | | | | |
| Mean log deviation | Overall | 0.045 | | 0.040 | | 0.033 | | 0.027 | | |
| | Between | 0.016 | (34%) | 0.011 | (28%) | 0.009 | (27%) | 0.008 | (30%) | |
| | Within | 0.030 | (66%) | 0.029 | (72%) | 0.024 | (73%) | 0.019 | (70%) | |
| Theil's entropy | Overall | 0.043 | | 0.037 | | 0.030 | | 0.025 | | |
| | Between | 0.016 | (37%) | 0.011 | (31%) | 0.009 | (29%) | 0.008 | (31%) | |
| | Within | 0.027 | (63%) | 0.026 | (69%) | 0.021 | (71%) | 0.017 | (69%) | |
| Joint distribution | | | | | | | | | | |
| Mean log deviation | Overall | 0.329 | | 0.333 | | 0.305 | | 0.263 | | |
| | Between | 0.215 | (65%) | 0.218 | (65%) | 0.192 | (63%) | 0.166 | (63%) | |
| | Within | 0.114 | (35%) | 0.116 | (35%) | 0.113 | (37%) | 0.097 | (37%) | |
| Theil's entropy | Overall | 0.337 | | 0.342 | | 0.315 | | 0.260 | | |
| | Between | 0.224 | (66%) | 0.227 | (66%) | 0.206 | (65%) | 0.176 | (68%) | |
| | Within | 0.113 | (34%) | 0.115 | (34%) | 0.109 | (35%) | 0.084 | (32%) | |

Table 5: Welfare indices from joint distributions and under independence assumption

| | | 1970 | | 1980 | | 1990 | | 2000 | |
|--|-----|-------|--------|-------|--------|-------|--------|-------|--------|
| Gini | (a) | 0.447 | | 0.450 | | 0.431 | | 0.393 | |
| | (b) | 0.393 | (88%) | 0.401 | (89%) | 0.387 | (90%) | 0.354 | (90%) |
| Mean log deviation | (a) | 0.329 | | 0.333 | | 0.305 | | 0.263 | |
| | (b) | 0.247 | (75%) | 0.258 | (77%) | 0.239 | (78%) | 0.207 | (79%) |
| Theil's entropy | (a) | 0.337 | | 0.342 | | 0.315 | | 0.259 | |
| | (b) | 0.259 | (77%) | 0.270 | (79%) | 0.255 | (81%) | 0.212 | (82%) |
| Poverty index (poverty line: \$1/day; expectancy = 40 years) | | | | | | | | | |
| Lower bound | (a) | 0.119 | | 0.071 | | 0.054 | | 0.047 | |
| | (b) | 0.070 | (58%) | 0.038 | (54%) | 0.014 | (27%) | 0.007 | (14%) |
| Upper bound | (a) | 0.404 | | 0.309 | | 0.192 | | 0.130 | |
| | (b) | 0.485 | (120%) | 0.366 | (118%) | 0.241 | (126%) | 0.167 | (128%) |
| Poverty index (poverty line: \$2/day; expectancy = 40 years) | | | | | | | | | |
| Lower bound | (a) | 0.190 | | 0.121 | | 0.094 | | 0.078 | |
| | (b) | 0.121 | (64%) | 0.073 | (60%) | 0.033 | (35%) | 0.013 | (17%) |
| Upper bound | (a) | 0.473 | | 0.397 | | 0.255 | | 0.155 | |
| | (b) | 0.575 | (122%) | 0.467 | (118%) | 0.329 | (129%) | 0.214 | (138%) |

Note: For each index, row (a) is computed used the estimated joint densities; row (b) is computed using the products of two marginal densities, under the (false) assumption of independence between income and health distributions.

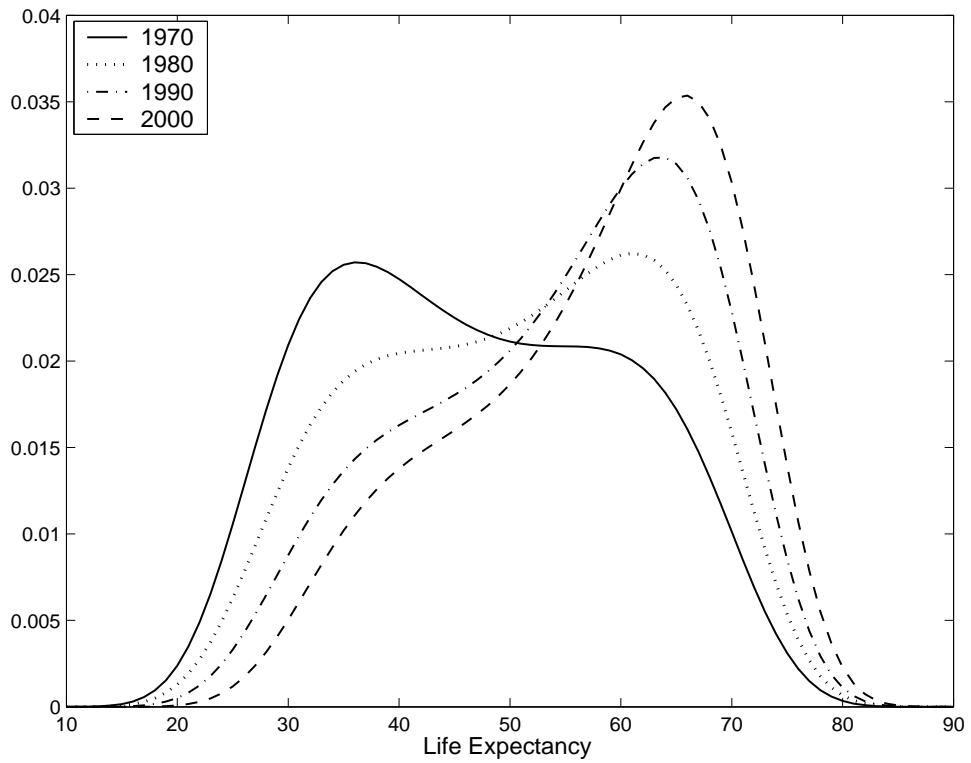
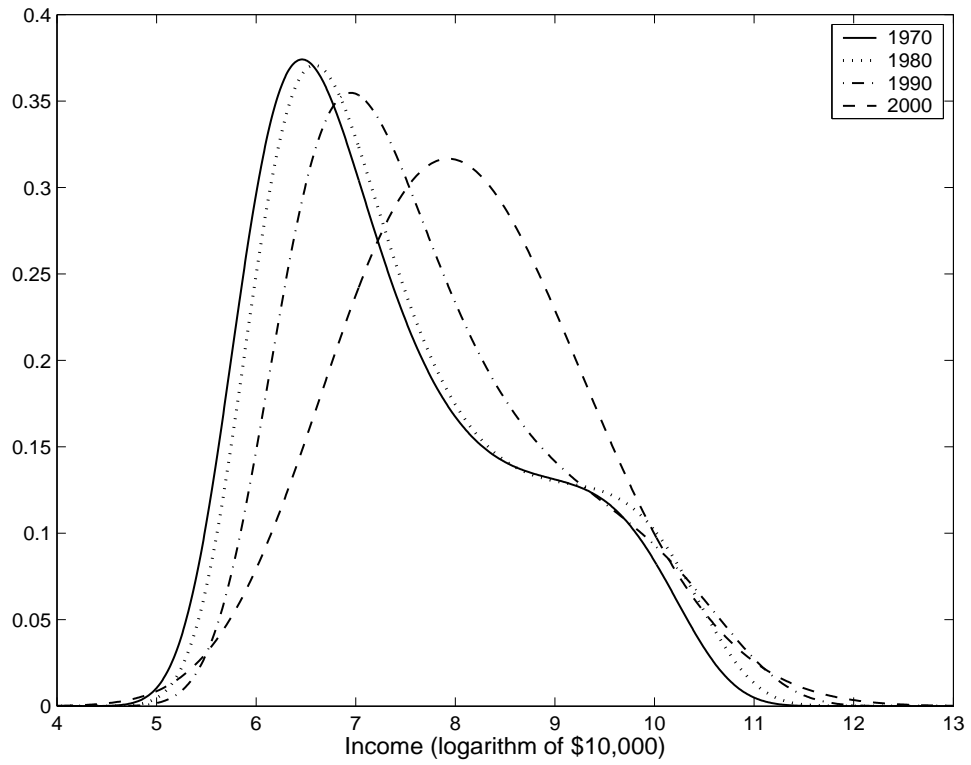


Figure 1: Marginal distributions: upper, income; lower, life expectancy

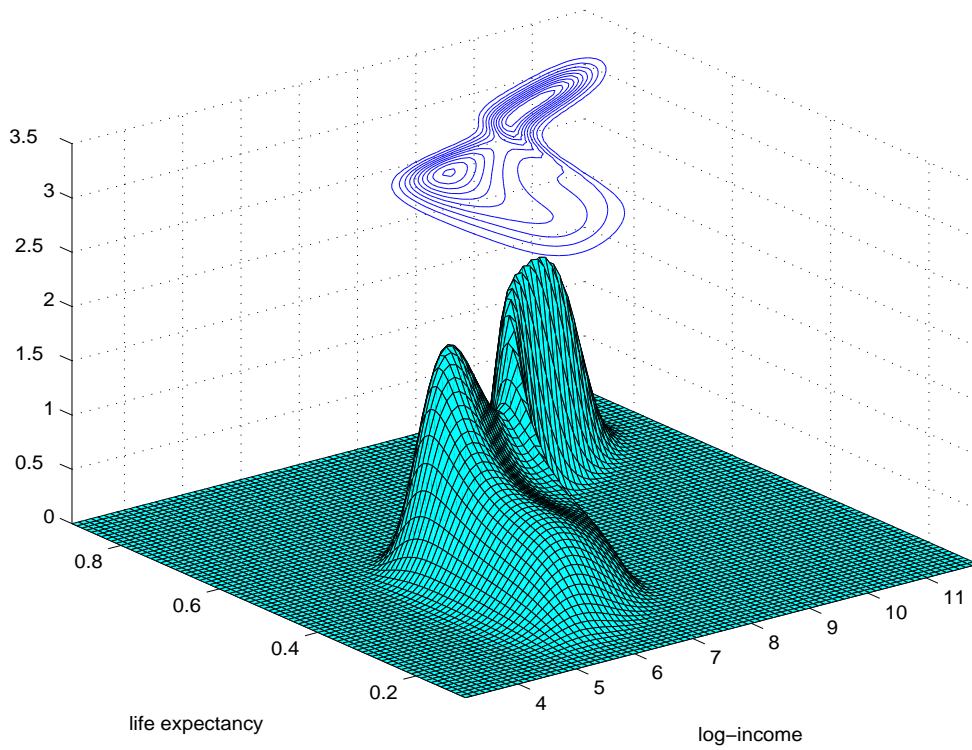
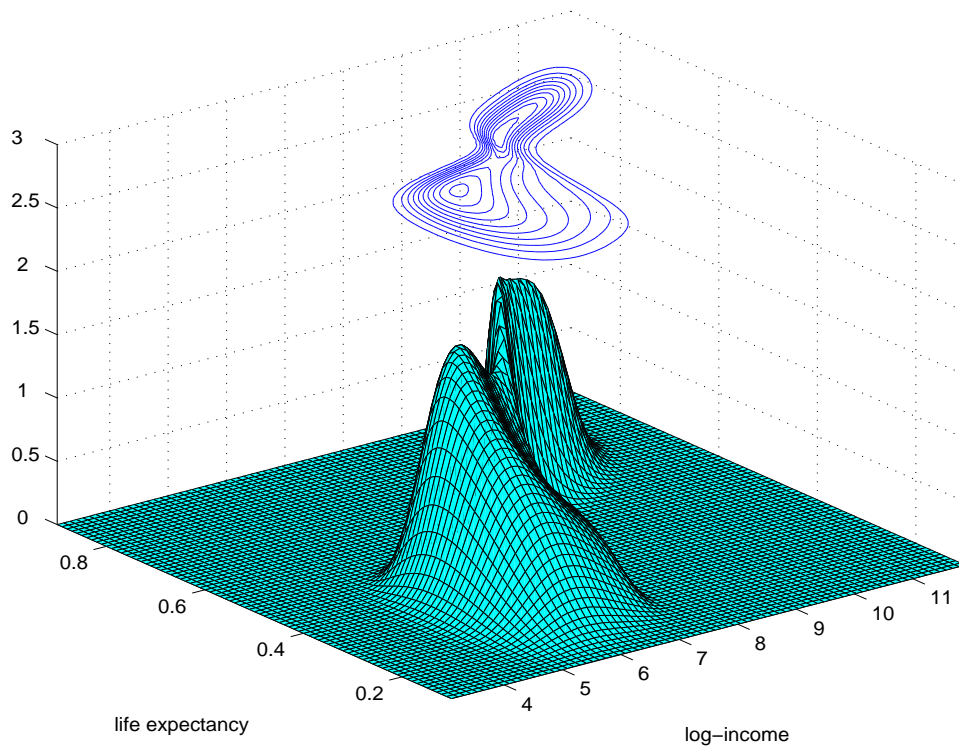


Figure 2: Joint distributions: upper 1970, lower 1980 (life expectancy in 100 years and income in logarithm of \$10,000)

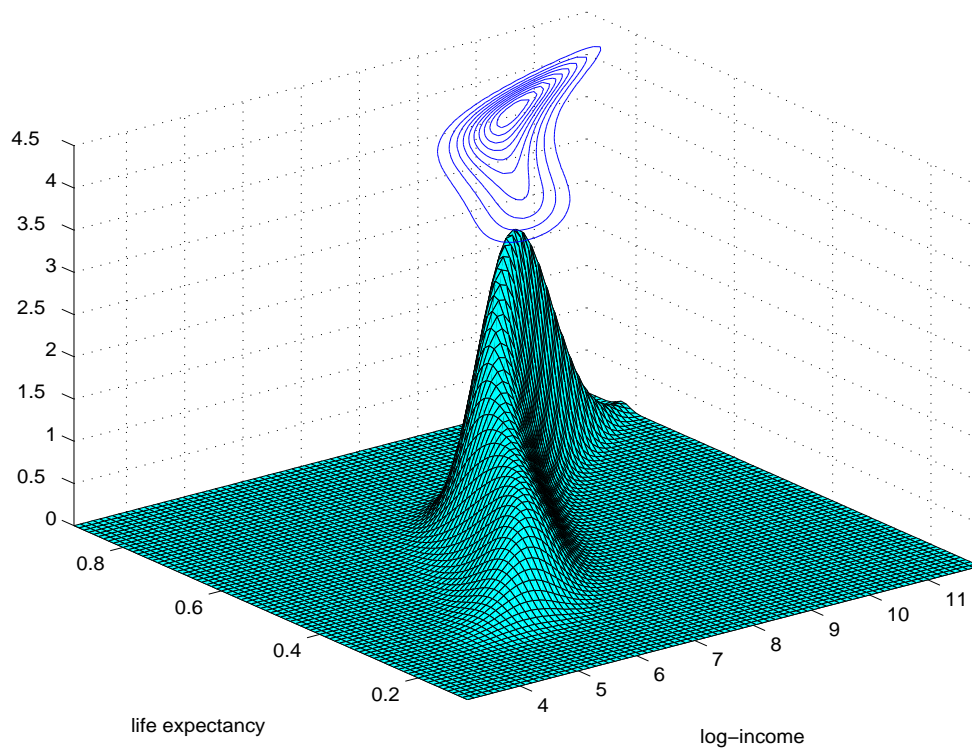
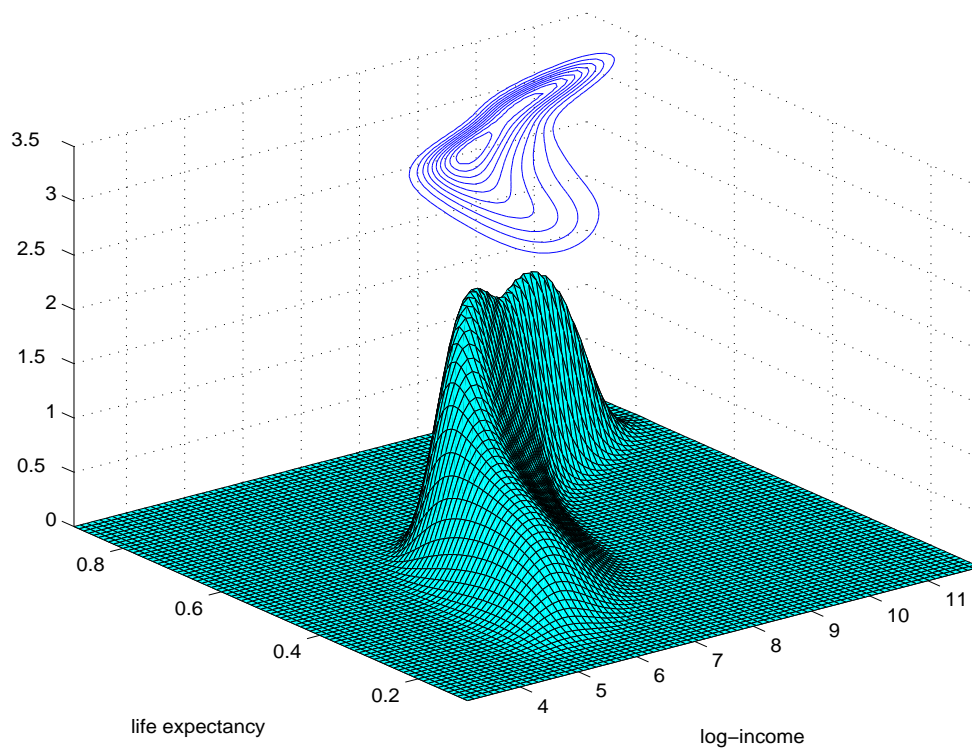


Figure 3: Joint distributions: upper 1990, lower 2000 (life expectancy in 100 years and income in logarithm of \$10,000)