11-24-2008

The Evolution of the Conditional Joint Distribution of Life Expectancy and Per Capita Income Growth

Thanasis Stengos
University of Guelph

Brennan S. Thompson
Ryerson University

Ximing Wu
Texas A&M University

Follow this and additional works at: <http://digitalcommons.ryerson.ca/economics>

Part of the [Economics Commons](#)

Recommended Citation

Stengos, Thanasis; Thompson, Brennan S.; and Wu, Ximing, "The Evolution of the Conditional Joint Distribution of Life Expectancy and Per Capita Income Growth" (2008). *Economics Publications and Research*. Paper 44.
<http://digitalcommons.ryerson.ca/economics/44>

This Working Paper is brought to you for free and open access by the Economics at Digital Commons @ Ryerson. It has been accepted for inclusion in Economics Publications and Research by an authorized administrator of Digital Commons @ Ryerson. For more information, please contact bcameron@ryerson.ca.

The Evolution of the Conditional Joint Distribution of Life Expectancy and Per Capita Income Growth*

Thanasis Stengos[†], Brennan S. Thompson[‡], Ximing Wu[§]

November 24, 2008

*The authors wish to thank the participants of the 7th Annual Advances in Econometrics Conference, Nov. 14-16, Baton Rouge, LA. for their useful comments and questions. In particular, the authors are indebted to Jeff Racine for his insightful suggestions.

[†]Department of Economics, University of Guelph, Guelph, ON N1G 2W1; tstengos@uoguelph.ca

[‡]Department of Economics, Ryerson University, Toronto, ON M5B 2K3; brennan@ryerson.ca

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

Abstract

In this paper we investigate the joint conditional distribution of health (life expectancy) and income growth and its evolution over time. The conditional distributions of these two variables are obtained by applying nonparametric methods to a bivariate nonparametric regression system of equations. Analyzing the distributions of the nonparametric fitted values from these models we find strong evidence of movement over time and strong evidence of first-order stochastic dominance of the earlier years over the later ones. We also find strong evidence of second-order stochastic dominance by non-OECD countries over OECD countries in each period. Our results complement the findings of Wu, Savvides and Stengos (2008) who explored the unconditional behaviour of these joint distributions over time.

JEL Code: 01; 015; 057

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

1 Introduction

Even though the concept of human development is a very broad concept, it certainly would include health and the standard of living as two of its fundamental components. The Human Development Report, first published in 1990, includes the United Nations Development Programme report of a composite index for each member country's average achievements. This index, the Human Development Index (HDI), covers three basic dimensions of human development: health, education and standard of living.

An important question for policy-makers is how to improve health, especially in developing countries. Many researchers, see Caldwell (1986) and Musgrove (1996) argue that development should focus on income growth, since higher incomes indirectly lead to health improvements. Others, Anand and Ravallion (1993) and Bidani and Ravallion (1997) take the stand that income growth alone is not enough as people's ability to function and perform in their economic tasks is affected by their health status and not the other way around. We intend to contribute to this debate by looking at the evolution of per capita income and health as measured by life expectancy over time for a number of countries over a thirty year period.

According to recent World Bank data, over the last forty years, the world's real GDP has increased by more than 100 percent although there exist important differences among individual country experiences. For the richest country quartile this increase is more than 150 percent, whereas for the poor-

est quartile this number was only 50 percent. Extreme poverty (the share of population living on less than \$1 per day) in developing countries has fallen by about 20 percent over the last 10 years alone, especially in East and South Asia where the accelerating growth of China and India has propelled these regions to be well within the target of the Millennium Development Goals to reduce in half the fraction of people below the cutoff of \$1 per day by 2015. Between 1960 and 2000 average life expectancy has increased by 15 years and infant mortality has fallen by more than 50 percent around the world, giving hope that the Millennium Development Goal of reducing infant and child mortality rates to one third of their 1990 levels would be met.

The rapid health improvements over the last forty years raise the question of the driving forces behind this trend. Most of the empirical studies, see Musgrove (1996) and Filmer and Pritchett (1999) for example assume that health improvements are the by-product of higher income as countries with higher income devote more resources for their health services, something that would translate into improved health status for their population.

One of the earlier benchmark studies of the income-health relationship is Preston (1975) who compared different countries' life expectancy and per capita income for different benchmark years (1900, 1930 and 1960) and proposed the "Preston curve" a non-linear and concave empirical relationship between the two. The concave Preston curve has provided the rationale for much of the empirical work that has followed. However, simple health-per capita income relationships may suffer from endogeneity, especially when it

comes to countries on the flat portion of the Preston curve, where health has reached such an advanced stage where additional improvements coming from income growth cannot be attained. In that case it would be the reverse impact from health to income that would be important. Papers, such as Pritchett and Summers (1996) address this issue by relying on an Instrumental Variable (IV) methodology. However, the difficulty here is the choice of instruments as many of those chosen as instruments may not be appropriate or may be weak, e.g. the investment ratio (ratio of investment to GDP) will itself be endogenous in a health type production function.

In a recent paper Maasoumi, Racine and Stengos (2007) (MRS hereafter) examined the entire distribution of income growth rates, as well as the distributions of parametrically and nonparametrically fitted and residual growth rates relative to a space of popular conditioning variables in this literature. In that respect they were able to compare convergence in distribution and “conditional convergence” as they introduced some entropy measures of distance between distributions to statistically examine the question of convergence or divergence. This approach can be viewed as alternative quantifications within a framework of distributional dynamics discussed in Quah (1993, 1997). Quah focused on the distribution of per capita incomes (and relative incomes) by introducing a measure of “transition probabilities”, the stochastic kernel, to analyze their evolution. The MRS paper’s focus on significant features of the probability laws that generate growth rates goes beyond both the standard “ β -convergence” and “ σ -convergence” in the lit-

erature, see Barro and Sala-i-Martin (2004). The former concept refers to the possible equality of a single coefficient of a variable in the conditional mean of a distribution of growth rates. The latter, while being derivative of a commonplace notion of “goodness of fit,” also is in reference to the mere fit of a conditional mean regression, and is plagued with additional problems when facing nonlinear, non-Gaussian, or multimodal distributions commonly observed for growth and income distributions. As has been pointed out by Durlauf and Quah (1999), the dominant focus in these studies is on certain aspects of estimated conditional means, such as the sign or significance of the coefficient of initial incomes, how it might change if other conditioning variables are included, or with other functional forms for the production function or regressions. All of the above studies rely on “correlation” criteria to assess goodness of fit and to evaluate “convergence”.

In the first study to use a bivariate framework, Wu, Savvides and Stengos (2008) (WSS hereafter) investigate the unconditional evolution of income per capita and life expectancy using a maximum entropy density estimator. They consider income and life expectancy jointly and estimate their unconditional bivariate distribution for 137 countries for the years 1970, 1980, 1990 and 2000. Their main conclusion is that the world joint distribution has evolved from a bimodal into a unimodal one, that the evolution of the health distribution has preceded that of income and that global inequality and poverty has decreased over time. They also find that global inequality and poverty would be substantially underestimated if the dependence between the income

and health distributions is ignored.

In this paper we extend the work of WSS by estimating the joint conditional distribution of health (life expectancy) and income growth and we examine its evolution over time. The conditional distributions of these two variables is obtained by applying nonparametric methods to a bivariate nonparametric system of equations. This generalizes the MRS approach to a multidimensional context. Using a similar data set as WSS, we extend their analysis to go beyond unconditional distributions. As in the MRS univariate framework we will be examining conditional distributions by looking into a bivariate system of per capita income growth and life expectancy growth equations. We will then analyze the distributions of parametrically and nonparametrically fitted values and residuals from these models using a bivariate growth framework relative to the standard conditioning variables that are employed in the literature. The resulting analysis produces “fitted values” of growth rates and life expectancy as well as “residual growth rates and life expectancy” which will be used for to look at the question of “conditional” convergence in a bivariate context. Note that in contrast with the WSS study which was conducted for the unconditional joint distribution of per capita income and life expectancy in levels, our approach will be based on analyzing the conditional joint distribution of growth rates, which provides new insight into the driving forces of their joint evolution over time.

The paper is organized as follows. In the next section we discuss the data used. We then proceed to discuss the empirical methodology and results of

both the parametric and nonparametric approaches that we pursue. Finally we conclude.

2 Data

To estimate the global joint distribution of income and life expectancy, we collected data on 124 countries to construct 10 year averages for the 1970's, the 1980's and the 1990's for a total number of observations of 372. These countries account for approximately 80 percent of global population. Below we describe in more detail the data that we use and their source. Similar data have been used by WSS.

Data on income per capita are in PPP dollars from the *Penn World Tables 6.2* and they are used to construct the real per capita GDP growth. This data base provides estimates in 2000 international prices for most countries beginning in 1950 until 2004.

For each country in our sample, the income information is reported in the form of interval summary statistics. 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). Since we construct an average over a 10 year period we do not need to have the same number of intervals to be the same in each year. For 1970, 1980, and 1990, we used income interval data from Bourguignon and Morrisson (2002). We construct an average income observation

for each country for each 10 year period. An alternative source of income data 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.¹ For 2000, Bourguignon/Morrisson do not provide data and we used income interval data from the WDI.² These data are based on household surveys of income (in some cases consumption) from government statistical agencies and World Bank country departments.

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). This is the most comprehensive collection of demographic statistics. For each of the 124 countries, it provides data on the number of persons in each age group for each of the four years (1970, 1980, 1990 and 2000).

¹Bourguignon 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 very few individual countries but in most cases for geographic groups of countries (see their study for group definitions). Our study is based on country-level data. Therefore, where individual-country interval data were unavailable we used the corresponding geographic-group data.

²Income 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 interval 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.

The U.N. Population Division begun compiling estimates of life expectancy at 5-year intervals in 1950. For each country we constructed average life expectancy over the relevant 10 year period. For more details about the data construction, see the WSS study.

3 Empirical Results

In this paper, we use both parametric and nonparametric techniques to estimate a bivariate system of equations that describe real per capita growth and life expectancy growth. The framework of analysis is an extension of the MRS framework to account for the simultaneous evolution of per capita income and life expectancy. We proceed by first estimating a bivariate system of equations parametrically and then we continue with the nonparametric analysis.

3.1 Parametric Results

We first consider a bivariate parametric system of seemingly unrelated regressions (SUR) to model the growth path of per capita income and life expectancy. The dependent variables are $Y = (Y_1, Y_2)$, where Y_1 is real GDP per capita growth and Y_2 is life expectancy growth. For each country-year, the list of independent variables is given by $X = (X_1, X_2, \dots, X_7)$, where X_1 is a dummy variable indicating OECD status, X_2 is a dummy variable for the 1980s, X_3 is a dummy for the 1990s, X_4 is the log of population growth

plus 0.05 to account for a constant rate of technical change of 0.02 and a depreciation rate of 0.03, X_5 is the log of investment share of GDP, X_6 is the log of real GDP at the start of the period and X_7 is the log of life expectancy at the start of period. The last two variables capture initial conditions and their effect on the transition to a steady state. The specification of the equation describing the evolution of per capita income is a standard growth regression of an extended Solow type model; the evolution of life expectancy is modeled in a symmetric way.

We begin by estimating a simple benchmark bivariate parametric regression model that is standard for the bivariate extension of the standard workhorse model of the empirical literature

$$Y_1 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \epsilon_1 \quad (1)$$

$$Y_2 = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \gamma_4 X_4 + \gamma_5 X_5 + \gamma_6 X_6 + \gamma_7 X_7 + \epsilon_2$$

We estimate the above system of equations as a seemingly unrelated regression (SUR). However, since the right hand side variables are identical in the two equations, GLS is identical to estimating each equation separately by OLS. Note that, in each equation, both GDP per capita and life expectancy enter in lagged (i.e., initial) values to guard against endogeneity.

The parameter estimates for specification (1) are given in Tables 1 and 2, and are in line with results from the extensive univariate growth literature. For the per capita income growth regression, we find investment having a

positive effect on growth, while population growth seems to have a negative effect. Initial GDP has a negative effect on growth (although not statistically significant) suggesting the presence of (statistically weak) conditional or β -convergence. The initial life expectancy variable also turns out to be statistically insignificant. In the context of an income growth regression life expectancy stands for a proxy for human capital and as such the latter often does not appear significant in parametric specifications, especially with panel data, see Savvides and Stengos (2008).

In the life expectancy growth equation, investment is also positive and significant, while population growth is positive but not highly significant. The initial life expectancy variable has a strongly negative effect which would seem to imply β -convergence in health outcomes. Initial GDP has an significant effect. I

Despite its use in the literature, there is evidence that the above parametric linear specification (1) is inadequate and misspecified, especially when it comes to describing the effect of initial conditions on the growth process. Following the per capita income growth literature we allow the initial condition variables X_6 and X_7 to enter as third degree polynomials, see Liu and

Stengos (1999), i.e.,

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \\
&\quad \beta_7 X_6^2 + \beta_8 X_6^3 + \beta_9 X_7 + \beta_{10} X_7^2 + \beta_{11} X_7^3 + \epsilon_1 \\
Y_2 &= \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \gamma_3 X_3 + \gamma_4 X_4 + \gamma_5 X_5 + \gamma_6 X_6 + \\
&\quad \gamma_7 X_6^2 + \gamma_8 X_6^3 + \gamma_9 X_7 + \gamma_{10} X_7^2 + \gamma_{11} X_7^3 + \epsilon_2
\end{aligned} \tag{2}$$

The results from the above parametric SUR system are given in Tables 3 and 4. These results are in line with results from the simple parametric specification (1) discussed above. Investment is found to positively affect both per capita GDP and life expectancy growth. Population growth has a negative effect on GDP per capita growth, but an insignificant effect on life expectancy growth. Interestingly, in both of the equations, none of the polynomial terms for either initial GDP per capita or initial life expectancy appear to be significant, which may suggest overparameterization.

We next test these parametric specifications against some unknown non-parametric alternative. If we denote the parametric model given by the above system of equations as $m_g(x_i, \beta)$, $g = 1, 2$ and the true but unknown regression functions by $E_g(y_{gi}|x_i)$, $g = 1, 2$, then a test for correct specification is a test of the hypothesis $H_0 : E_g(y_{gi}|x_i) = m_g(x_i, \beta)$, $g = 1, 2$ almost everywhere versus the alternative $H_1 : E_g(y_{gi}|x_i) \neq m_g(x_i, \beta)$, $g = 1, 2$ on a set of positive measure. That is equivalent to testing that $E_g(\epsilon_{gi}|x_i) = 0$ almost everywhere, where $\epsilon_{gi} = y_{gi} - m_g(x_i, \beta)$. This implies that for an incorrect

specification, $E_g(\epsilon_{gi}|x_i) \neq 0$ on a set of positive measure.

To avoid problems arising from the presence of random denominator in the nonparametric estimator of the regression functions $E_g(y_{gi}|x_i)$, the test employs a density weighted estimator of the regression function. To test whether $E_g(\epsilon_{gi}|x_i) = 0$ holds over the entire support of the regression function, we use the statistic $J = E_g\{[E_g(\epsilon_{gi}|x_i)]^2 f(x_i)\}$ where $f(x_i)$ denotes the density weighting function. Note that $J = 0$ if and only H_0 is true. The sample analogue of J , J_n is obtained by replacing ϵ_{gi} with the residuals from the parametric model and both $E_g(\epsilon_{gi}|x_i)$ and $f(x_i)$ by their respective kernel estimates. The null distribution of the statistic is obtained via bootstrapping (see Hsiao et al., 2008, for details).

For specification (1), we are able to reject the null of correct specification at the 5% and 1% levels for the first and second equations, respectively (the test statistics, J_n , are 0.6919 and 4.411, with bootstrap p -values of 0.0276 and 0.0025, respectively). Similarly, for (2), we are able to reject at the 5% and 0.1% levels for the first and second equations, respectively (the test statistics, J_n , are 0.3658 and 2.1892, with bootstrap p -values of 0.0401 and 2.22e-16, respectively). We use 399 bootstrap replications throughout the paper.

3.2 Nonparametric Results

Next, we use local linear estimation to estimate nonparametric regression models

$$Y_1 = g_1(X) + \epsilon_1$$

$$Y_2 = g_2(X) + \epsilon_2.$$

We use least squares cross-validation techniques to obtain the appropriate bandwidths for the discrete and continuous regressors, see Racine and Li (2004). This approach allows for interactions among all variables and also allows for nonlinearities in and among variables. The method has the additional feature that if there is a linear relationship in a variable then the cross validated smoothing parameter will automatically detect this. We estimate these equations separately using local linear estimation. A second-order Gaussian kernel is used for the continuous variables, while the Aitchison and Aitken kernel is used for the unordered categorical variable (OECD status) and the Wang and Van Ryzin kernel is used for the ordered categorical variable (decade). For details, see Racine and Li (2004).

In Figures 1-4, we summarize the nonparametric results using partial regression plots. These plots simply present the estimated multivariate regression function through a series of bivariate plots in which the regressors not appearing on the horizontal axis of a given plot have been held constant at their respective (within group and decade) medians. For example, in the

upper-left plot in Figure 1, we plot the estimated level of GDP per capita growth conditioned on population growth for just OECD members in the 1970s holding all the other conditioning variables at their respective median levels for OECD members in the 1970s. In this way we are able to visualize the multivariate regression surface via a series of two-dimensional plots.

The level of investment appears to have a (linearly) positive and stable effect across decade and country group for both equations. Population growth appears to be unrelated to the dependent variables except in the 1980s, where it is slightly negative for the GDP per capita growth equation and slightly positive for the life expectancy growth equation (for both OECD and non-OECD members). For the GDP per capita growth equation, initial GDP appears to have a slightly negative effect in the 1970s, but little effect in either the 1980s or the 1990s (for both OECD and non-OECD members). For OECD members, initial life expectancy seems to have a negative effect on GDP per capita growth in the 1980s, but little effect in the other decades. However, for non-OECD members, the effect of initial life expectancy on GDP per capita growth is mixed: The effect seems to be positive in the 1970s, negative in the 1980s, and non-existing in the 1990s. For the life expectancy growth equation, initial GDP appears to have a slight negative effect in all decades and groups. However, initial life expectancy appears to have a generally negative, but non-linear effect in all decades and groups.

To further examine how the joint distribution of per capita GDP and life expectancy growth rates differ between groups and over time, we use

the notion of stochastic dominance, which is defined as follows. We say distribution G stochastically dominates distribution F at first-order if

$$F(x_1, x_2) \geq G(x_1, x_2),$$

for all (x_1, x_2) .

More generally, we can say that distribution F dominates distribution G stochastically at order s (an integer) if

$$D_F^s(x_1, x_2) \leq D_G^s(x_1, x_2)$$

for all (x_1, x_2) , where $D_F^1(x_1, x_2) = F(x_1, x_2)$ and $D_F^s(x_1, x_2)$ is defined recursively as

$$D_F^s(x_1, x_2) = \int_0^{x_1} \int_0^{x_2} D_F^{s-1}(u_1, u_2) du_1 du_2, \quad s \geq 2.$$

D_G^1 and D_G^s are defined analogously. In what follows, we will denote this relation by $F \succeq_s G$.

To empirically test such a relationship, we use the approach of McCaig and Yatchew (2007). To test the null hypothesis that $F \succeq_s G$, these authors introduce the test statistic

$$T_{F,G} = \left\{ \int \int [\psi^s(x_1, x_2)]^2 dv_1 dv_2 \right\}^{1/2},$$

where $\psi^s(x_1, x_2) = \max\{D_F^s(x_1, x_2) - D_B^s(x_1, x_2), 0\}$. Of course, when the null is true, T is equal to zero.

In practice, this test involves estimating T and testing whether it is statistically different from zero. This process will involve estimating $\psi^s(x_1, x_2)$ over a set of grid points on the common support of the two distributions under consideration. The p -value of this test statistic is obtained via bootstrapping (see McCaig and Yatchew, 2007, for details).

To make such comparisons in a conditional manner, we use the fitted values from the nonparametric regressions considered above. The estimated joint density and distribution functions of these fitted values are shown in Figures 5 and 6, respectively. We separate the observations by group and decade; that is, we consider 6 unique groupings (OECD and non-OECD members for the 1970s, OECD and non-OECD members for the 1980s, and OECD and non-OECD members for the 1990s). As seen in Figure 5, the distribution of bivariate conditional growth rates has become more concentrated within each group (OECD and non-OECD members) over time. Also, it is interesting to note that the (conditional) GDP per capita growth rates tend to be higher among OECD members, but that the (conditional) life expectancy growth rates tend to be higher among non-OECD members. However, these differences appear to be diminishing over time.

We now proceed to testing for stochastic dominance of the fitted (conditional) bivariate growth rates between the two groups of countries under consideration: OECD members and non-OECD members. The values of the

test statistics and their bootstrap p -values are presented in Table 5. As can be seen, we can strongly reject the null of first-order stochastic dominance of OECD members over non-OECD members (and vice-versa) in each of the three decades under consideration. We can also strongly reject the null of second-order stochastic dominance of OECD members over non-OECD members in each of the three decades, but not vice-versa. That is, we are unable to reject the null of second-order stochastic dominance of non-OECD members over OECD members.

Next, we consider testing for first-order stochastic dominance of the same fitted values between the three decades under consideration: the 1970s, 1980s, and 1990s. The values of the test statistics and their bootstrap p -values are presented in Table 6. For both the OECD and non-OECD groups, we are unable to reject the null of first-order stochastic dominance of the 1970s over the 1980s and the 1980s over the 1990s (and, of course, the 1970s over the 1990s).

These results somewhat agree with the findings of MRS, who show that the fitted (conditional) growth rates of per capita income have "deteriorated" over time for OECD countries. However, we also want to point out that the MRS analysis is univariate and as pointed out in WSS the overall results will underestimate substantially the degree of global inequality and poverty if one ignores the dependence between the two measures of welfare. Note however, that the later analysis was conducted for the unconditional joint distribution of per capita income and life expectancy (levels), whereas here we analyze the

conditional joint distribution of growth rates. The implication is that there was a more "equal" joint distribution of growth rates in the earlier years than the later ones, not necessarily faster growth in the earlier years. Note that the interpretation of this result for growth rates is different from that for levels. For the case of the joint distribution of growth rates, the results suggest that in the earlier years "convergence" between developing and more developed countries would be more difficult to achieve since countries in these groups would be growing more or less at equal rates. It is only in the later years that a more "unequal" joint distribution of growth rates would allow for faster growing developing countries being able to catch up with slower growing developed countries. Hence, the results that we find are complementary to the ones found in WSS for levels, where the level of overall (unconditional) inequality in levels decreased over time. Overall, it seems that countries developed quite differently in the 1980s and 1990s with some jumping ahead and others falling behind. We leave it for future research to further explore the issue for subgroups of countries, such as OECD and non OECD and especially African and non African countries.

4 Conclusion

In this paper we have estimated the joint conditional distribution of health (life expectancy) and income growth and we have examined its evolution over time. The conditional distributions of these two variables is obtained

by applying nonparametric methods to a bivariate nonparametric regression system of equations. Using a similar data set as WSS, we extend their analysis to go beyond unconditional distributions. Extending the MRS univariate framework we have looked at conditional distributions of a bivariate system of per capita income growth and life expectancy growth equations. Analyzing the distributions of the nonparametric residuals from these models we establish that there is strong evidence of movement over time in the joint conditional bivariate densities of per capita growth and life expectancy. We also find strong evidence of first order stochastic dominance of the earlier years over the later ones. Our results complement the findings of WSS who explored the unconditional behaviour of these joint distribution over time.

References

- Anand, S. and M. Ravallion (1993). "Human development in poor countries: On the role of private incomes and public services," *Journal of Economic Perspectives* 7, 133-150.
- Barro, R., and X. Sala-i-Martin (2004). *Economic Growth*, MIT Press, Second edition.
- Bidani, B. and M. Ravallion (1997). "Decomposing social indicators using distributional data," *Journal of Econometrics* 77, 125-139.
- Caldwell, J.C. (1986). "Routes to low mortality in poor countries," *Population and Development Review* 12, 171-220.
- Durlauf, S.N. and P.A. Johnson (1995). "Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics* 10, 365-384.
- Durlauf, S.N. and D.T. Quah (1999). "The New Empirics of Economic Growth," in J.B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics I*, Elsevier, 235-308.
- Filmer, D. and L. Pritchett (1999). "The impact of public spending on health: Does money matter?" *Social Science and Medicine* 49, 1309-1323.
- Hsiao, C., Q. Li and J.S. Racine (2008). "A consistent model specification test with mixed categorical and continuous data," *Journal of Econometrics* 140, 802-826.

- Li, Q. and J.S. Racine (2003). "Nonparametric Estimation of Distributions with Categorical and Continuous Data," *Journal of Multivariate Analysis* 86, 266-292.
- Liu, Z. and T. Stengos (1999). "Nonlinearities in Cross Country Growth Regressions: A Semiparametric Approach," *Journal of Applied Econometrics* 14, 527-538.
- McCaig, B. and A. Yatchew (2007). "International welfare comparisons and nonparametric testing of multivariate stochastic dominance," *Journal of Applied Econometrics* 22, 951-969.
- Musgrove, P. (1996). "Public and private roles in health," World Bank, Discussion Paper No. 339.
- Maasoumi, E., J.S. Racine and T. Stengos (2007). "Growth and convergence: A profile of distributional dynamics and mobility," *Journal of Econometrics* 136, 483-508.
- Pritchett, L. and L. Summers (1996). "Wealthier is healthier," *Journal of Human Resources* 31, 841-868.
- Quah, D.T. (1993). "Empirical Cross-Section Dynamics in Economic Growth," *European Economic Review* 37, 426-434.
- Quah, D.T. (1997). "Empirics for Growth and Distribution: Stratification, Polarization and Convergence Clubs," *Journal of Economic Growth* 2, 27-59.

Racine, J.S. and Q. Li (2004). "Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data," *Journal of Econometrics* 119, 99-130.

Savvides, A. and T. Stengos (2008). *Human Capital and Economic Growth*, Stanford University Press.

Wu, X., A. Savvides and T. Stengos (2008). "The Global Joint Distribution of Income and Health," Department of Economics, University of Guelph, Discussion Paper No. 2008-7.

Table 1: Parameter estimates for GDP per capita growth polynomial regression

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.5256	4.2270	0.83	0.4048
oecd1	-0.7596	0.3882	-1.96	0.0511
d1980	-1.3791	0.2852	-4.84	0.0000
d1990	-1.0124	0.3088	-3.28	0.0011
pop	-2.8323	0.9554	-2.96	0.0032
inv	1.5212	0.2318	6.56	0.0000
initY	-0.0423	0.0668	-0.63	0.5265
initL	0.2787	0.9392	0.30	0.7668

Table 2: Parameter estimates for life expectancy growth linear regression

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.0159	0.5371	3.75	0.0002
oecd1	-0.1911	0.0493	-3.87	0.0001
d1980	-0.1053	0.0362	-2.91	0.0039
d1990	-0.2767	0.0392	-7.05	0.0000
pop	0.2118	0.1214	1.74	0.0819
inv	0.1556	0.0295	5.28	0.0000
initY	0.0092	0.0085	1.08	0.2805
initL	-0.5441	0.1193	-4.56	0.0000

Table 3: Parameter estimates for GDP per capita growth polynomial regression

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-973.9962	869.5863	-1.12	0.2634
oecd1	-0.6090	0.4550	-1.34	0.1816
d1980	-1.4345	0.2860	-5.02	0.0000
d1990	-1.0460	0.3183	-3.29	0.0011
pop	-3.2726	1.0003	-3.27	0.0012
inv	1.3891	0.2417	5.75	0.0000
initY	-2.4379	5.8822	-0.41	0.6788
initY ²	0.1012	0.3479	0.29	0.7712
initY ³	-0.0012	0.0068	-0.18	0.8590
initL	744.6284	676.8577	1.10	0.2720
initL ²	-184.8683	174.8208	-1.06	0.2910
initL ³	15.2578	15.0334	1.01	0.3108

Table 4: Parameter estimates for life expectancy growth polynomial regression

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-148.8768	111.0005	-1.34	0.1807
oecd1	-0.1997	0.0581	-3.44	0.0007
d1980	-0.1133	0.0365	-3.10	0.0021
d1990	-0.2873	0.0406	-7.07	0.0000
pop	0.1723	0.1277	1.35	0.1780
inv	0.1409	0.0309	4.57	0.0000
initY	-0.7304	0.7508	-0.97	0.3313
initY ²	0.0402	0.0444	0.91	0.3653
initY ³	-0.0007	0.0009	-0.83	0.4094
initL	118.3105	86.3992	1.37	0.1717
initL ²	-30.2379	22.3154	-1.36	0.1763
initL ³	2.5600	1.9190	1.33	0.1830

Table 5: Stochastic dominance tests: Between groups

	1970s	1980s	1990s
OECD $\succeq_1$ Non-OECD	4.2802	3.4902	1.8601
	0.0000	0.0000	0.0000
Non-OECD $\succeq_1$ OECD	1.0816	1.9978	2.1026
	0.0404	0.0000	0.0000
OECD $\succeq_2$ Non-OECD	27.8089	23.6515	10.0243
	0.0000	0.0000	0.0000
Non-OECD $\succeq_2$ OECD	0.0742	0.0513	1.761
	0.9495	0.9596	0.5051

Note: For each result, the first line is the value of the test statistic, while the second line is the bootstrap p -value.

Table 6: Stochastic dominance tests: Between decades

			OECD	Non-OECD
1970s $\succ_1$ 1980s			0.0000	0.0000
			0.9393	0.9899
1980s $\succ_1$ 1970s			2.0061	3.1569
			0.0000	0.0000
1980s $\succ_1$ 1990s			0.4374	0.4291
			0.4545	0.2929
1990s $\succ_1$ 1980s			1.3846	3.0810
			0.0000	0.0000
1970s $\succ_1$ 1990s			0.0000	0.0000
			0.8788	0.9899
1990s $\succ_1$ 1970s			2.9433	5.2524
			0.0000	0.000

Note: For each result, the first line is the value of the test statistic, while the second line is the bootstrap p -value.

Figure 1: GDP per capita growth nonparametric partial regression plots for OECD countries. The first, second, and third columns are for the 1970s, 1980s, and 1990s, respectively.

Figure 2: GDP per capita growth nonparametric partial regression plots for non-OECD countries. The first, second, and third columns are for the 1970s, 1980s, and 1990s, respectively.

Figure 3: Life expectancy growth nonparametric partial regression plots for OECD countries. The first, second, and third columns are for the 1970s, 1980s, and 1990s, respectively.

Figure 4: Life expectancy growth nonparametric partial regression plots for non-OECD countries. The first, second, and third columns are for the 1970s, 1980s, and 1990s, respectively.

Figure 5: Estimated density functions for the fitted values from the nonparametric regressions. The left column is for OECD countries, while the right column is for non-OECD countries. The first, second, and third rows are for the 1970s, 1980s, and 1990s, respectively. Within each plot, the lower-left axis is for the fitted values from the life expectancy growth nonparametric regression, while the lower-right axis is for the fitted values from the GDP per capita growth nonparametric regression.

Figure 6: Estimated distribution functions for the fitted values from the nonparametric regressions. The left column is for OECD countries, while the right column is for non-OECD countries. The first, second, and third rows are for the 1970s, 1980s, and 1990s, respectively. Within each plot, the lower-left axis is for the fitted values from the life expectancy growth nonparametric regression, while the lower-right axis is for the fitted values from the GDP per capita growth nonparametric regression.