



# The Effects of Age Structure on Economic Growth: An Application of Probabilistic Forecasting to India

## **Abstract**

During recent years there has been an increasing awareness of the explanatory power of demographic variables in economic growth regressions. We estimate a new model of the effects of age structure change on economic growth. We use the new model and recent probabilistic demographic projections for India to derive the uncertainty of predicted economic growth rates caused by the uncertainty in demographic developments.

## **Keywords**

Economic growth, age structure, probabilistic demographic projections, India.

## **Acknowledgements**

We are grateful for comments and suggestions by participants at the Symposium on “Demographic Change and Global Income Growth in the 21st Century” at Krusenberg Mansion January 7-9, 2004 and by participants at the session on “Population Forecasting” held at PAA Boston, April 1-3.

## **Authors**

Alexia Prskawetz, Research group leader and deputy director, Vienna Institute of Demography, Austrian Academy of Sciences, Austria.

Tomas Kögel, Lecturer, Department of Economics, Loughborough University, Loughborough, Leicestershire LE11 2AG, United Kingdom.

Warren C. Sanderson, Professor of Economics, Departments of Economics and History, State University of New York at Stony Brook, US and Visiting Professor, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Sergei Scherbov, Research group leader, Vienna Institute of Demography, Austrian Academy of Sciences, Austria.

# 1 Introduction

During recent years there has been an increasing awareness of a direct influence of population age structure on the macroeconomy. The theoretical foundations of the reduced form models applied in econometric studies are: (1) the life cycle model of savings and investment and (2) age specific variations in labor productivity. Recently, Bloom et al. (2003b) argues that a falling youth dependency ratio (the population below working age divided by the population of working age) contributed to the economic growth miracle in East Asia (see also, among others, Mason (2001)). More general, recent evidence suggests that falling youth dependency ratios in developing countries can create an opportunity for economic growth. Bloom et al. (2001), Bloom & Williamson (1998), Crenshaw et al. (1997), Kelley & Schmidt (1995), Kelley & Schmidt (2001) show empirical support for this claim by finding a positive and significant effect of declining youth dependency ratios on economic growth in cross-country regressions applied to developing and developed countries. Significant age structure effects have also been found for economic growth, inflation and savings in OECD countries (Lindh (1999), Lindh & Malmberg (1999), Malmberg (1994)).

Except in the work by Bloom & Williamson (1998) and Lindh (1999) the implications of this evidence have not yet been discussed. Lindh (1999) presents evidence that statistical models of inflation and GDP growth explained mainly by the share of 5-year age groups perform well in out-of-sample forecasting on a horizon of 3-5 years ahead during the 1990s. A different methodology is used in Bloom & Williamson (1998). The authors use the estimated coefficients on population growth rates and growth rates of the economically active population (15-64 old age group) — as estimated by a cross-country regression of economic growth applied to 78 developing and developed countries — to assess the past (1965-1990) and forecast the future (1990-2025) contribution of demographic change on economic growth. Their results indicate “that population dynamics can explain between 1.37 and 1.87 percentage points of growth in GDP per capita in East Asia or as much as one-third of the miracle ...” (p.441, where miracle refers to growth rate of GDP per capita of 6% observed over the same time period in East Asia). For their forecasts the authors combine the estimated coefficients on the population variable with the medium variant of UN population projections. Their results indicate that “in East Asia, the growth in GDP per capita attributable to demographic influences is projected to be negative between 1990 and 2025, ..... a loss of 0.14 to 0.44 percentage point up to 2025”. On the other hand countries in South Asia are projected to gain from their demographic changes in the future.

We follow Bloom & Williamson (1998) here and forecast economic growth in developing countries by combining the estimated coefficients of a regression analysis of economic growth with population projections. The contribution of this paper lies in our use of probabilistic population projections instead of deterministic ones.

We use India as a case study for our approach. India is a country for which probabilistic population projections are available (Lutz and Scherbov 2004). In addition, India is an ideal country for our purpose since it has not yet had the dramatic fall in youth dependency (which we observe for many of the Asian countries) and which has not yet had a very dramatic increase in the elderly dependency ratio (Dyson, Cassen and Visaria (2004)). Recently a survey of India in *The Economist* (2004, p.13) argues similarly: ‘The most fundamental long-term reason for optimism is demographic.’ This paper provides a partial test of this hypothesis.

The structure of the paper is as follows. In the next section we discuss the role of demographic variables in explaining economic growth. In section 3 we show the results of a cross-country regression of economic growth applied to developing and developed countries.

Section 4 reviews recent applications of probabilistic demographic projections in economics. Section 5 combines the probabilistic demographic forecasts for India with the estimates obtained from the growth regression. The final section concludes.

## 2 The role of demographic variables in explaining economic growth

According to the neoclassical growth model population growth reduces economic growth due to capital dilution. However, various studies using cross-country data found an insignificant effect of population growth on economic growth. Recently, Bloom & Williamson (1998) challenged this result (see also Bloom et al. (2001)). They confirm that population growth has no effect on economic growth in growth equations with the growth rate of the total population as the only demographic variable. However, they show that demography matters for economic growth, once one considers changes in the age structure of the population, that is, once one drops the implicit assumption of a constant age composition of the population. More specifically, the authors regress the growth rate of GDP per capita on the growth rate of the working age population and the growth rate of the total population (and various other control variables). The results show a positive and significant effect on the growth of GDP per capita due to the growth of the working age population, and an opposite, negative and significant effect from the growth rate of the total population.

Since World War II developing countries have been undergoing a demographic transition at varying rates and times (Lee (2003)). During the standard demographic transition scenario infant mortality declines and fertility falls with a lag only after the mortality decline has begun. As a consequence, a demographic transition leads first to a demographic “burden” because population growth is faster than the growth of the working age population. Later, as fertility declines, the demographic transition leads to a demographic “dividend” because the growth of the working age population is faster than the growth of the total population (cf. Bloom et al. (2003b)).

Bloom & Williamson (1998) use the quantitative results of their cross-country regressions to calculate the contribution of age structure changes to East Asian economic growth (see the discussion in the introduction). According to their view, the demographic “dividend” leads to opportunities for growth of output per capita for two reasons. First, there is an *accounting effect* because a rising ratio of the working age population to the total population increases the ratio of “producers” to “consumers”. Obviously this contributes positively to the growth of output per capita. Second, there might also exist *behavioral effects* on the growth of output per capita. As Bloom & Williamson (1998) stresses, a rising growth rate of the working age population leads, on the one hand, to capital dilution, that is, a reduction of the ratio of capital to the working age population. On the other hand, a rising ratio of the working age population to the total population implies a falling dependency ratio. Bloom et al. (2003a), using aggregate data of developing and developed countries, shows that a falling dependency ratio increases aggregate savings (see, among others Kelley & Schmidt (1996)).

Mathematically the accounting effect and the behavioral effects can be decomposed from the following identity that results from taking the natural logarithm of output per capita and differentiating the resulting expression with respect to time. This yields

$$\hat{y} = \hat{\tilde{y}} + \hat{l} \quad (1)$$

where  $y = Y/N$ ,  $\tilde{y} = Y/L$ ,  $l = L/N$  and  $Y$  denotes total output. Further,  $N$  denotes the population and  $L$  denotes the working age population and a “hat” on top of a variable

indicates the growth rate. (A comprehensive review of the recent economic-demographic modelling and the role of accounting effects can be found in Kelley & Schmidt (2003)). The growth rate of output per working age person,  $\hat{y}$ , contains the behavioral effects, while the term  $\hat{l}$ , which equals the difference between the growth rate of the working age population and the growth rate of the total population, is the accounting effect.

While Bloom & Williamson (1998) focuses on output per capita as the dependent variable, the recent literature argues that an understanding of international differences in output per worker is needed, since only workers contribute to production. Since forecasts of labor force participation rates are difficult and lie outside the scope of the paper, changes of output per worker are approximated here by changes in output per working age person. Most importantly, as can be seen from (1), the growth rate of output per capita contains behavioral effects, as well as, an accounting effect from age structure changes. Hence, from regressions with the growth rate of output per capita as the dependent variable it is rather complicated to assess behavioral effects (see Bloom et al. (2001) for an assessment of behavioral effects from such regressions). In contrast, the growth rate of output per working age person is free of an accounting effect. Hence, all age structure effects in regressions with the growth rate of output per working age person as the dependent variable are behavioral effects. In addition, the accounting effect can be calculated directly from calculating the difference between the growth rates of the working age population and the growth rate of the total population, with no need of regressions. For this reason, we estimated behavioral effects by regressing the growth of output per working age person on its determinants and have calculated the accounting effect directly without the use of regression exercises (see some regressions in Kelley & Schmidt (2003) for a similar estimation strategy and Kögel (2003) for the estimation of behavioral effects using regressions on the growth of output per worker).

### 3 Results of economic growth regressions

Economic growth regressions suffer from estimation biases due to measurement errors and omitted variables such as the initial level of technology. Biases from omitted variables can be avoided by applying the fixed effects estimator or the Arellano-Bond estimator to pooled time-series and cross-country data. However, as Hauk and Wacziarg (2004) stress, the fixed effects estimator or the Arellano-Bond estimator exacerbate measurement error biases when the right-hand side variables are more time persistent than the measurement errors.<sup>1</sup> Using Monte Carlo simulations, the authors find that using the OLS estimator applied to a single cross-section of variables averaged over some decades contains fewer biases than the fixed effects estimator or the Arellano-Bond estimator. For this reason, we used regressions applied to 97 developing and developed countries averaged over the time period between 1960 and 1995. (The definition and data sources of the variables in the regressions are shown in Appendix A, while a list of countries included in the analysis is shown in Appendix B).

Based on the conditional convergence hypothesis, we included the natural logarithm of output per working age person in 1960 as an independent variable affecting average annual growth of output per working age person.<sup>2</sup> If negative and significant, then the coefficient of

---

<sup>1</sup>The problem of time persistent independent variables and measurement errors is particularly relevant for our study where we choose slow moving demographic variables and variables that measure social infrastructure as independent variables. While the former set of variables is time persistent, the latter variable is often flawed by measurement errors. Since fixed effects models aim to capture the time variation in independent variables, application of this method to our data would imply that we might capture measurement errors (which are more time varying) of social infrastructure instead of the demographic effects and the institutional context as measured by the social infrastructure variable.

<sup>2</sup>Henceforth, the average annual growth rate of a variable  $x$  in country  $i$  is defined as  $(1/35)(\ln x_{i,1995} -$

this variable indicates the tendency of convergence of output per working age person to a value that is in the steady state the same for all countries, provided all countries have the same value for their control variables.

We initially followed Bloom & Williamson (1998) and Hall and Jones (1999) in controlling for social infrastructure with indicators of the quality of institutions and openness to trade. However, we could not find strong instrumental variables for those indicators. Therefore, we used somewhat different indicators for which we found stronger instruments. We followed Rodrik et al. (2002) and used the indicator of rule of law in Kaufman et al. (2003), abbreviated RULE. The updated indicator is available for 1996-2002. Since our sample ends in 1995, we used the indicator for 1996. RULE is a composite indicator that measures protection afforded to property rights and the strength of the rule of law.<sup>3</sup> Instead of using the Sachs and Warner (1995) index of the average years of openness to trade, we again followed Rodrik et al. (2002) in using the natural logarithm of the ratio of the sum of exports and imports to output in current international dollars, henceforth abbreviated lnCOPEN, as our measure of openness to trade. We followed Hall & Jones (1999) in calculating a composite indicator for social infrastructure, abbreviated SOCINF, as the average value of the indicators for quality of institutions (in our case RULE) and for openness to trade (in our case lnCOPEN). As Hall & Jones (1999) stresses, doing so imposes the restriction that the coefficients for RULE and lnCOPEN are the same, which we were able to confirm empirically (results not shown). In our view, including the composite of the two variables solves the problem of strong multicollinearity between them. This multicollinearity shows that countries that have high quality institutions are also more open to trade.

We also included the level of the youth dependency ratio as measured by the population younger than 15 divided by the population aged fifteen to sixty-four. The motivation for including the level of age structure and not its growth rate is the fact that, at least in the transition to a steady state, growth of output per working age person is affected by the level of savings. In turn, Bloom et al. (2003a) have empirically confirmed that the level of savings is affected by the level of age structure variables and not their growth rates (Kelley & Schmidt (2003) and Kögel (2003) independently suggest including only the level of youth dependency ratio for this reason). Nevertheless, in the analysis that follows we also investigated empirically, whether our hypothesis that growth of youth dependency ratio does not matter for growth of output per working age person finds empirical support. We included only the youth dependency ratio in the regression equation, because the elderly dependency ratio turned out to be insignificant for growth of output per working age person.

The first column in Table 1 shows the results of an OLS regression of the (annual average) growth rate of output per working age person on the natural logarithm of initial output per working age person, SOCINF and the natural logarithm of the youth dependency ratio.<sup>4</sup> As expected, the results show a negative and significant effect of the youth dependency ratio. In addition, the coefficient of output per working age person is negative and significant, while the coefficient of social infrastructure is positive and significant.

In the second column in Table 1 we included the (annual average) growth rates of the working age population and of the total population as further independent variables. Both new variables are significant. (When the growth rate of the youth dependency ratio is included

---

ln  $x_{i,1960}$ ), where for example  $x_{i,1995}$  denotes the value of  $x$  in country  $i$  in the year 1995.

<sup>3</sup>Using RULE helped identification of the partial economic growth effects of social infrastructure in comparison with age structure, because the best instruments that we could find for RULE are less strongly correlated with age structure than the best instruments that we could find for the preferred indicator used in Bloom & Williamson (1998) and Hall & Jones (1999).

<sup>4</sup>All regressions in Table 1 contain White-Huber heteroscedasticity-consistent p-values, reported in parentheses below the coefficient estimates.

instead it also has a statistically significant coefficient - not shown). In the third model (column (3) in Table 1) we further add the natural logarithm of the average life expectancy at birth between 1960-65 and 1990-95. This can be justified with the finding of Bloom et al. (2003a) that rising life expectancy boosts aggregate savings. Life expectancy turns out to be positive and significant, as expected. In addition, the growth rate of the working-age population and the growth rate of the total population become insignificant when we control for life expectancy (see also Kelley & Schmidt (2003)). The fact that the p-value of the youth dependency ratio falls to seven percent when we control for life expectancy is very likely due to some multicollinearity between the youth dependency ratio and life expectancy. In the fourth column in Table 1 we exclude the growth rates of the working age population and the total population but keep life expectancy. The fact that the value of the coefficient of the youth dependency ratio falls from -2.41 in Model 1 (column 1) to -1.08 in Model 4 (column 4) is an indication that the youth dependency ratio in Model 1 partly captures the effect of life expectancy.

[Table 1 about here.]

Recent literature (and most prominently Hall & Jones (1999)) argues that indicators for quality of institutions contain measurement errors. Moreover, because of feedback effects from the level of output per working-age person or its growth rate, the indicators for quality of institutions and openness for trade might be endogenous. For this reason, Hall & Jones (1999) suggest instrumenting the indicator for social infrastructure. We follow the authors to use the predicted trade share of Frankel and Romer (1999), abbreviated *lnFrankRom*, as an instrument for our indicator of openness in *SOCINF*. As second instrument for *SOCINF* we choose English legal origin, abbreviated *EngLegal*, as taken from La Porta et al. (1999).

To allow for the possibility that the youth dependency ratio is endogenous, we applied IV estimation with the youth dependency ratio being instrumented, as well. For the set of instruments for the youth dependency ratio we chose the following four variables for ecozones, as taken from the data appendix of Gallup et al. (1999): *ztropics* (percent of land in a tropical zone), *zsubtropics* (percent of land in a subtropical zone), *zdestmp* (percent of land in a temperate desert zone) and *zdestrp* (percent of land in a tropical or subtropical desert zone).

The first column in Table 2 shows the IV regression results. (In Table 2 we did not correct for heteroscedasticity, because tests for heteroscedasticity (White test and Breusch-Pagan test) could not reject homoscedasticity). Again, the coefficient of initial output per working-age person is negative and significant, while *SOCINF* is positive and significant. Most importantly, the coefficient of the youth dependency ratio remains negative and significant. Because of space limitation we refrain from showing the first-stage regression results.<sup>5</sup>

To show the consequences of applying IV estimation, we proceed with a comparison of the regression results of IV estimation (column 1, Table 2) with the OLS regression results (column 1, Table 1). One can see that IV estimation leads to a larger coefficient for the youth dependency ratio. For *SOCINF* the magnitude of the coefficient falls and the magnitude of the coefficient of initial output per working-age person is almost unaffected.

Regression (1) does not contain life expectancy at birth. It is very likely that life expectancy is affected by economic growth and therefore endogenous, as well. For this reason,

---

<sup>5</sup>We applied for the IV estimation the Fuller's method of modified limited-information maximum likelihood (LIML). Stock et al. (2002) and Hahn and Hausman (2003) suggest this estimation method when instruments are weak. Among other possible reasons, instruments can be weak when the instruments have strong explanatory power for both endogenous variables.

we also applied IV estimation to the ln of average life expectancy at birth. As an instrument for life expectancy we chose the variable ln of one plus malaria ecology, abbreviated ME. ME is an ecologically-based variable that is predictive of the extent of malaria transmission and is taken from Kiszewski et al. (2004). The second column in Table 2 shows the regression results where we applied IV estimation to SOCINF, ln of youth dependency ratio and ln of life expectancy. The coefficients of all right-hand side variables have the expected sign and are significant. Compared to the OLS regressions in column 4 in Table 1 the coefficients on the youth dependency ratio and on life expectancy are smaller indicating a feedback effect from economic growth to life expectancy and the youth dependency ratio. Furthermore, a comparison across model 1 and model 2 in Table 2 shows that the coefficient on the youth dependency ratio declines by almost fifty per cent once we control for life expectancy pointing to a high correlation between the variables (indeed, the correlation coefficient between the ln of the youth dependency ratio and the ln of life expectancy has in our sample the value -0.72).

[Table 2 about here.]

## 4 Applications of probabilistic demographic projections in economics

One common approach to quantifying the uncertainty in future outcomes of economic and demographic variables is the use of scenarios. It starts by constructing alternative trajectories (commonly referred to as low, medium and high variants) for the various inputs into the forecasting model. The next step is then to 'bundle combinations of these trajectories together to calculate high, medium and low projection "scenarios". The way the bundling is done depends on the purpose of the projections, and has an important influence on the results.' (Lee & Edwards (2001, p.6)).

For instance consider the population projections by United Nations (2003) (also used by Bloom & Williamson (1998)). The UN provides three main projections, labelled as low, medium and high variants. The low, medium and high scenarios only differ with respect to their fertility assumptions (see United Nations (2003) for further details). The assumptions are set in order to highlight the effects that different fertility paths have on demographic outcomes. According to Lee & Edwards (2001, p.6) not only does the scenario based approach not assign any probability to its alternative variants but the approach is seriously flawed with respect to several of its assumptions.

During the last decade several authors have developed stochastic population forecasts (see Lee (1998) for a review of probabilistic approaches to population forecasting). While alternative approaches differ with respect to the methodology that generates the input vectors for the projection model (e.g. time series modelling versus expert opinions are used to yield future time paths of fertility and mortality), the ultimate aim of those methods is the same: to derive prediction intervals for future realizations of population numbers or more general population functionals such as the dependency ratios.

In recent years several authors (Lee & Edwards (2001), Tuljapurkar (1992)) have also started to apply stochastic population forecasts in economics and environmental studies (O'Neill (2003)). E.g. Lee and Edwards (2001) apply stochastic population forecasts to predict the fiscal impact of population ageing in the US and to assess its uncertainty. Specific applications include stochastic projections for social security, stochastic projections for

medicare costs and stochastic projections of the balance in the Social Security Trust fund (all applied for the USA). Stochastic simulations are applied to generate the stochastic forecasts for the quantity of interest. For example to forecast the balance of the Social Security Trust Fund one combines the stochastic population forecasts with forecasts of economic variables (like productivity changes and changes in the interest rate) and applies those to the mathematical expression which yields the social security balance each time period. Each realization of a demographic forecast and an economic forecast will yield one stochastic simulation trajectory. After generating at least one thousand of those stochastic trajectories we can apply a frequency distribution to the outcome variable of interest (the social security balance) to estimate the probability distribution of the forecast. We shall proceed in a similar fashion in section 5 where we combine the econometric results from the previous section with stochastic forecasts of the population.

Let us try to translate the meaning of the probabilistic forecasts in terms of our research question, long run economic growth forecasts. Demographic change is only one among a host of factors influencing economic growth. Since we concentrate only on this one factor, we cannot forecast the unconditional distribution of India's future rate of economic growth. We can only report on the effects of demographic uncertainty on economic growth, conditional on pre-specified values of SOCINF, our variables that summarizes the effects of all non-demographic factors on economic growth. Conditional forecasted distributions of economic growth rates can be used to answer questions such as 'What is the probability that India in 2035 will surpass the growth rate in output per worker that has been achieved in 2000'.<sup>6</sup> We can also consider how the conditional distribution of economic growth rates changes when SOCINF changes. For example, below we study what would happen to economic growth rates if India, by 2035, would be able to achieve the same level of SOCINF as observed in Singapore in 2000 (the country with the highest value of SOCINF in 2000)? Clearly, this sort of analysis is only a first step, but nevertheless, it may still help in the policy making process.

## 5 Probabilistic demographic and long run economic growth projections for India

To provide some context, we compare the past and projected youth dependency ratio of India (according to the UN medium variant) with the age structure observed in Japan and Niger, two countries at different stages of their development as compared to India. Figure 1 shows the youth dependency ratios (the population of age 0-15 divided by the economically active population of age 15-64) for these three countries over the period 1960 to 2050.<sup>7</sup> Clearly, India's youth dependency ratio has been falling since 1970 and the UN forecasts that it will fall until 2045. In contrast, Niger (one of the poorest countries of the world) has not yet been experiencing a fall of the youth dependency ratio while Japan had its demographic transition much earlier than India. It is important to note that the old age dependency ratio (the number of people of age 65 and older divided by the number of people 15 to 64) in India rises only modestly from 2000 to 2035 (compare United Nations (2003)) since our regression specification does not include the old age dependency ratio as an independent variable.

[Figure 1 about here.]

---

<sup>6</sup>Since our estimation for the average annual growth rate of per working age person is based on a 35 year interval we choose the same time horizon for our predictions.

<sup>7</sup>We add a vertical line at the year 2000 since this will constitute the starting period for our probabilistic projections.

In Figure 2 we plot percentiles of the probability distribution of the youth dependency ratio based on Lutz and Scherbov (2004). During the first two decades the uncertainty in the number of births and the uncertainty in the number of surviving people in the working age population mainly determine the probability distribution. As the uncertain births start to enter the denominator (working age population) the uncertainty in the youth dependency ratio increases further. An additional broadening of the probability distribution results from the uncertain fertility applied to uncertain numbers of young women in fertile ages. By 2050 there is a 2.5% chance that the youth dependency ratio will have fallen by 82% or more while there is an equal chance that it will have fallen by 34% or less. Put differently, there is a 95% chance that the youth dependency ratio will be between 0.10 and 0.36 in 2050 compared to a value of 0.54 in 2000. The evolution of the probabilistic youth dependency ratios follows a trajectory that is similar to the UN ratios in Figure 1, but because of differences in assumptions, the probabilistic forecasts of youth dependency ratios tend to be somewhat lower than the UN forecasts.

[Figure 2 about here.]

In Figure 3 we plot the probabilistic projections for life expectancy at birth. According to these projections the uncertainty in the projected life expectancy is substantial. With a 95% chance the life expectancy is predicted to be between 63 and 78 years in 2050 and this range is reduced to between 68 and 74 years if we consider the 60% prediction interval. The median value of life expectancy in 2050 is 71 years and therefore by 10 years above the value currently observed in India (61 years in 2000).

[Figure 3 about here.]

In Figure 4 we plot the probabilistic projections for the accounting effect (the difference between the growth rate of the working age population and the total population). Our probabilistic projections indicate that the uncertainty about the future path of the accounting effect is not negligible. In fact, there is an equal chance of 2.5% that the accounting effect will be positive and exceed .4 or be negative and fall below  $-1$  in the time period 2045-2050. Note, however that the probability that the accounting will become negative before 2025 is very small. This is because of the falling youth dependency ratio which implies that the ratio of consumers to producers is falling during the next decades while the old age dependency ratio does not change much. At the time when the smaller cohorts of children start entering the labour force (around 2020 and 2030) the ratio of consumers to producers will increase again and reduce the accounting effect.

[Figure 4 about here.]

As a next step we combine the probabilistic demographic projections of India with the regression results from model (2) in Table 2 to predict an annual average growth rate between 2000 and 2035. More specifically we draw one realization of the stochastic population projections in 2035 and solve the following equation

$$-8.69 + recal - 1.56 \ln y_{2000} - 1.60 \ln YD + 0.83 \text{ SOCINF} + 5.14 \ln e_0$$

where  $y$  indicates the output per working age person,  $YD$  the youth dependency ratio and  $e_0$  the life expectancy at birth with  $\ln YD = 0.5[\ln YD(2000) + \ln YD(2035)]$ ,  $\ln e_0 = 0.5[\ln e_0(2000) + \ln e_0(2035)]$  and SOCINF is set equal to the value of SOCINF observed in India in 2000. The constant *recal* (for recalibration) is the difference of the

predicted and observed annual growth rate between 1995 and 2000. It essentially updates the value of the constant SOCINF to the average value for 1995-2000. Note, however that for the predicted value of average annual growth rates between 2000 and 2035 we keep the level of the SOCINF constant and therefore will forecast a growth rate that is likely to be too low. We assemble 1,000 stochastic simulations of the growth rate of output per working age person and summarize the resulting probability distribution in Table 3, column (1). Note, that we assume uncertainty only in the demographic variables and ignore any uncertainty that is caused by the estimation procedure.

[Table 3 about here.]

Compared to the past experience of annual economic growth rates of 4.13 per cent between 1995 and 2000 our results indicate that there is about the same likelihood that the annual growth rate of output per working age person between 2000 and 2035 will be either below or above the growth rate of output per working age person reached in 1995–2000. Recalling that our structural equation used to forecast economic growth only includes three time varying regressors — the lagged dependent variable, the youth dependency ratio and life expectancy at birth — these results lend themselves to another interesting interpretation. In the median path, the behavioral effect (represented by the youth dependency ratio and the life expectancy at birth) will almost exactly counterbalance the negative convergence term and the growth rate of output per working age person will remain close to 4.13 per cent, the value attained at the beginning of our forecasting time period. On the other hand, projected growth rates that are lower (higher) than 4.13 per cent imply that the convergence term dominates (is dominated by) the behavioral effect.

In Table 3, column (1') we add to the uncertainty of the youth dependency ratio the uncertainty of the accounting effect to arrive at predictions of output per capita. Our results indicate that for each reported percentile the annual growth rate of per capita output between 2000 and 2035 will be about 0.4 percentage point larger compared to the annual growth rate of output per working age person. This result can be explained by the fact that according to our probabilistic predictions (Figure 3), the accounting effect will most likely stay positive until 2020. More specifically, our results indicate that there is a 95 per cent chance that the predicted annual per capita output growth rate between 2000 and 2035 will be between 4.2 and 5.0 per cent (compared with a value of 4.5 per cent observed between 1995 and 2000).

We also have combined the probabilistic demographic projections with the regression results in Table 1, column (4) to quantify the effect on predicted economic growth rates if we ignore the endogeneity of demographic variables in the cross country regression. The resulting percentiles are reported in the second column in Table 3. Predicted average annual growth rates between 2000 and 2035 ignoring endogeneity are lower than when the effects of endogeneity are controlled. Ignoring the endogeneity of the demographic variables implies that the coefficient on the life expectancy is upwards biased while the coefficient on the youth dependency ratio is downwards biased. This can be explained by the fact that these variables are multicollinear and the endogeneity of life expectancy is greater than the endogeneity of the youth dependency ratio. Hence, since the predicted gain in life expectancy is lower than the predicted decrease in the youth dependency ratio, the predicted growth rate of output per working age person is too low if we apply OLS estimates. Hence, applying instrumental variable estimation turns out to be essential if we make predictions based on past effects of demographic variables.

We may compare the conditional distribution of growth rates in the first column in Table 3 to the effect of a change in the variable SOCINF. If between 2000 and 2035 India's social

infrastructure increased from its current level to the current level of Singapore (the country with the highest social infrastructure in our world sample in 2000) the growth gain would be about 1.18 per cent per year for the median value of the conditional distribution (cf. column (3) in Table 3). Cumulated over 35 years this would imply a difference in the level of output per working age person of almost 50 per cent between 2000 and 2035. More accurately, while output per working age person would be predicted to increase by about 3.6 times the value obtained in 2000 if the value of SOCINF stays constant at the value observed in India in 2000, output per working age person would be predicted to increase by about 5.4 times if SOCINF would approach the value of Singapore by 2035.

Another interesting exercise is to see how long it would take India to reach the value of output per working age person observed in 2000 in UK for example given the predicted conditional growth rates of output per working age person in Table 3. The results are summarized in Table 4 for growth rates based on column (1) and column (3) in Table 3. Based on the median forecast, with the 1995-2000 level of SOCINF it would take 51 years for India to reach the level of output per working age person experienced by the UK in the year 2000. If India has Singapore's value of SOCINF in 2035 it would take 41 years to achieve that level of output per working age person.

Given that today's value of the social infrastructure is unchanged a difference of the annual growth rate of about 0.72 percentage points (as observed between the 25th and 97.5th percentile in column (1) in Table 3) would result in a difference of about 8.8 years that it takes India to reach the level of the output per working age person of the UK. If however, India would approach the value of SOCINF that currently prevails in Singapore, the time it takes India to reach today's level of output per working age person of the UK is reduced markedly. For the 25th percentile the difference among column (1) and column (2) in Table 4 is about 9 years. This difference increases to 12.2 years for the 97.5th percentile. These figures indicate that the uncertainty of predicted growth rates of output per working age person due to uncertainty in demographic indicators is of about the same order as uncertainty due to limited knowledge of the future path of the social infrastructure variable.

[Table 4 about here.]

## 6 Conclusion

We draw on recent literature that has shown that age structure may have an important effect on economic growth. These effects work through a pure accounting effect (a change in the ratio of producing to consuming people in an economy) and a behavioral effect which relates to the change in output per working age person.

To assess the behavioral effect we applied a new cross-section estimation approach to data of 97 developing and developed countries from 1965 to 1990. Our results indicate a negative and significant effect of the youth dependency ratio and the output per working age person in the base year and a positive and significant effect of social infrastructure. We have applied instrumental variable methods to control for the endogeneity of demographic variables and measurement errors in variables that capture the social infrastructure.

In a second step we combined the estimated coefficients of the regression on economic growth with probabilistic population projections of India to qualify the growth potential embodied in India's changing age structure. If one considers only demographic uncertainty we find that there is a 95 per cent chance that the predicted growth rate of output per

working age person in 2035 will be between 3.8 and 4.6 percent compared to a value of 4.1 per cent in 1995-2000.

Changing India's social infrastructure variable from its value in 1995-2000 to the level currently observed in Singapore increases the average rate of economic growth of output per working age person by about 1 percentage point. La Porta et al. (1999) argues that government performance improves with per capita income. Since in this case there would be bi-directional causality, it would be infeasible to quantify a feedback effect from economic growth to better social infrastructure. Nevertheless, keeping in mind the possibility of improving social infrastructure, India's future economic growth might be faster than our predictions that consider only the change in the age structure.

## References

- Alho, J. (1990) Stochastic methods in population forecasting, *International Journal of Forecasting*, 6(4), 521–530.
- Bloom, D.D., Canning, D. & Malaney, P. (2001). Demographic change and economic growth in Asia, *Population and Development Review*, 26, supp., 257–290.
- Bloom, D.D., Canning, D. & Graham, B. (2003a). Longevity and life cycle savings. *Scandinavian Journal of Economics*, 105, 319–338.
- Bloom, D.E., Canning, D. & Sevilla, J. (2003b). *The Demographic Dividend: A New Perspective on the Economic Consequences of Population Change*, Rand.
- Bloom, D.E. & Williamson, J.G. (1998). Demographic transitions and economic miracles in emerging Asia, *World Bank Economic Review*, 12, 419–455.
- Crenshaw, E.M., Ameen, A.Z. & Christenson, M. (1997). Population dynamics and economic development: age-specific population growth rates and economic growth in developing countries, 1965 to 1990, *American Sociological Review*, 62, 974–984.
- Dyson, T. & Cassen, R. & Visaria, L. (2004). *Twenty-first century India: Population, Economy, Human Development, And the Environment*, Oxford University Press.
- ESRI (1996). ArcWorld Supplement Data Reference and Users Guide, Environmental Systems Research Institute, Redlands, CA.
- Frankel, J.A. & Romer D. (1999). Does trade cause growth? *American Economic Review*, 89, 379–399.
- Gallup, J. L. & Sachs, J.D. with Mellinger, A. (1999). Geography and economic development, CID Working Paper No. 1, Harvard, MA.
- Hahn, J. & Hausman, J. (2003). Weak instruments: diagnostics and cures in empirical econometrics, *American Economic Review, Papers and Proceedings* 93, 118–125.
- Hall, R. & Jones, C.I. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, 114, 83–116.
- Hauk W. R. & Wacziarg, R. (2004). A Monte Carlo study of growth regressions, National Bureau of Economic Research Working Paper No. T0296, Cambridge, MA.
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2003). Governance Matters III: Governance Indicators for 1996–2002, World Bank Policy Research Working Paper 3106.
- Kelley, A.C. & Schmidt, R.M. (1995) Aggregate population and economic growth correlations: the role of the components of demographic change, *Demography*, 32(4), 543–555.
- Kelley, A.C. & Schmidt, R.M. (1996). Savings, dependency and development, *Journal of Population Economics*, 9, 365–386.
- Kelley, A.C. & Schmidt, R.M. (2001) Economic and demographic change: a synthesis of models, findings, and perspectives, in: N. Birdsall, A.C. Kelley & S. Sinding (eds.) *Population Matters: Demographic Change, Economic Growth, and Poverty in the Developing World*, New York, Oxford University Press, 67–105.

- Kelley, A.C. & Schmidt, R.M. (2003) Evolution of recent economic-demographic modelling: a synthesis, Working Paper, Duke University.
- Kiszewski, A., Mellinger, A., Malaney, P., Spielman, A., Ehrlich, S. & Sachs, J.D. (2004). A global index of the stability of malaria transmission based on the intrinsic properties of anopheline mosquito vectors, *American Journal of Tropical Medicine and Hygiene*, forthcoming.
- Kögel, T. (2003). Youth dependency and total factor productivity, *Journal of Development Economics*, forthcoming.
- La Porta, R., Lopez-de-Silanes, F., Shleifer, A. & Vishny, R. (1999). The quality of government, *Journal of Law, Economics & Organization* 15, 222–279.
- Lee, R. & Edwards, R. (2001). The fiscal impact of population aging in the US: assessing the uncertainties, Department of Demography, University of California, Berkeley.
- Lee, R.D. (1998). Probabilistic approaches to population forecasting, in: Lutz, W., Vaupel, J.W. and Ahlburg, D.A. (eds.) *Frontiers of Population Forecasting*, A supplement to Vol. 24, Population and Development Review, 156–190
- Lee, R.D. (2003). The Demographic Transition: Three Centuries of Fundamental Change, *Journal of Economic Perspectives*, Vol. 17, No.4, 167–190
- Leemans, R. (1990). Possible changes in natural vegetation patterns due to a global warming.” IIASA Working Paper WP–90-08. International Institute of Applied Systems Analysis, Laxenburg, Austria.
- Mason, A. (ed.) , (2001). *Population Change and Economic Development in East Asia: Challenges Met, Opportunities Seized*, Stanford, Stanford University Press.
- Lindh, T. (1999). Medium-term forecasts of potential GDP and Inflation using age structure information, Department of Economics, Uppsala University, Sweden.
- Lindh, T. & Malmberg, B. (1999). Age structure effects and growth in OECD, 1950-90, *Journal of Population Economics*, 12(3), 431–449.
- Lutz, W. & Scherbov, S. (2004). Probabilistic population projections for India with explicit consideration of the education-fertility link, *International Statistical Review*, Volume 74, no. 1, 81–92.
- Malmberg, B. (1994). Age structure effects on economic growth: Swedish evidence, *Scandinavian Economic History Review*, 42(3), 279–295.
- O’Neill, B. C. (2003). Conditional probabilistic population projections: an application to climate change, IIASA, Interims Report-03-051, International Institute of Applied Systems Analysis, Laxenburg, Austria.
- Rodrik, D., Subramanian, A., & Trebbi, F. (2002). Institutions rule: the primacy of institutions over geography and integration in economic development. Center for International Development (CID) at Harvard University, CID-Working Papers No.97.
- Sachs, J.D. & Warner, A.M. (1995) Economic reform and the process of global integration, *Brooking Papers of Economic Activity*, 1-118.
- Stock, J.H., Wright, J.H, & Yogo, M. (2002) A survey of weak instruments and weak identification in general method of moments, *Journal of Business and Economic Statistics*, 20, 518–529.

The Economist (2004) India's shining hopes: A survey of India, February 21st.

Tuljapurkar, S. (1992) Stochastic population forecasts and their uses, *International Journal of Forecasting*, 385–391.

United Nations (2003). World Population Prospects: The 2002 Revision, <http://esa.un.org/unpp>.

## **Appendix A: the data of the regressions in section 3**

**Output per working-age person** (in constant international dollars of 1996 according to the Laspeyres method)

*Calculation:* Multiplication of output per capita in constant international dollars of 1996 according to the Laspeyres method with total population and division by population of age 15-64.

*Sources:* Penn World Table 6.1 (PWT 6.1) (output per capita and total population) and United Nations (2003) (population of age 15-64).

### **Youth dependency ratio**

*Definition:* population of age 0-14 divided by population of age 15-64.

*Source:* United Nations (2003) .

### **RULE**

*Definition:* indicator of rule of law.

*Source:* Kaufmann et al. (2003).

### **LnCOPEN**

*Definition:* Ln of the ratio of the sum of exports and imports to output in current international dollars

*Source:* Penn World Table 6.1 (PWT 6.1).

### **Annual average growth of working-age population and of total population**

*Definition:* working-age population is defined as population of age 15-64.

*Source:* Penn World Table 6.1 (PWT 6.1).

### **Ln of life expectancy at birth**

*Definition:* average of Ln of life expectancy at birth in 1960-65 and 1990-95.

*Source:* Penn World Table 6.1 (PWT 6.1).

### **ztropics, zsubtropics, zdestmp, zdestrp, zdrytemp, zwater**

*Definition:* ecozones, which are: ztropics (percent of land in a tropical zone), zsubtropics (percent of land in a subtropical zone), zdestmp (percent of land in a temperate desert zone) and zdestrp (percent of land in a tropical or subtropical desert zone), zdrytemp (percent of land of a country in a dry-temperate zone) and zwater (percent of the area of a country that is lake or ocean)

*Source:* data appendix of Gallup et al. (1999) and personal information of John Gallup. The percent of land of a country in the aforementioned zones was calculated by combining information of country boundaries from ESRI (1996) with the thirty-seven Holdridge Life Zones from Leemans (1990). Next, the resulting thirty-seven variables were aggregated into the aforementioned broader categories.

### **LnFrankRom**

*Definition:* predicted trade share of an economy.

*Source:* Frankel and Romer (1999) (available at Chad Jones' data archive at <http://elsa.berkeley.edu/~chad/datasets.html>)

### **EngLegal**

*Definition:* dummy variable which takes the value one, if a country's legal origin is British (and takes otherwise the value zero).

*Source:* data appendix of LaPorta et al. (1999).

**LnME**

*Definition:* ln of an ecologically-based variable that is predictive of the extent of malaria transmission.

*Source:* Kiszewski et al. (2004).

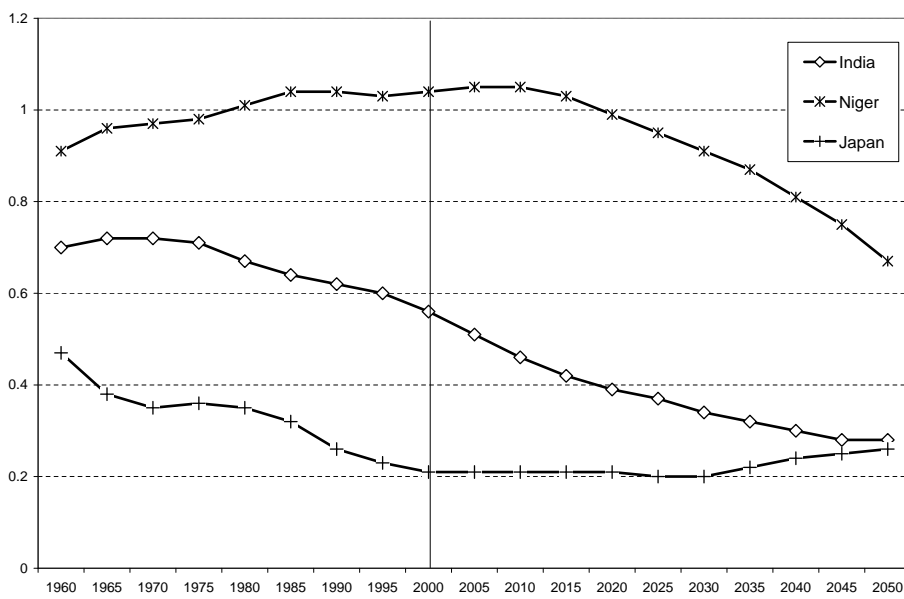
## Appendix B: list of countries included in data set of the regressions in section 3

Angola	Egypt	Kenya	Portugal
Argentina	Spain	Korea, Rep.	Paraguay
Australia	Ethiopia	Sri Lanka	Romania
Austria	Finland	Lesotho	Rwanda
Burundi	France	Morocco	Senegal
Belgium	Gabon	Madagascar	Singapore
Benin	United Kingdom	Mexico	El Salvador
Burkina Faso	Ghana	Mali	Sweden
Bangladesh	Guinea	Mozambique	Syria
Bolivia	Gambia, The	Mauritania	Chad
Brazil	Guinea-Bissau	Mauritius	Togo
Botswana	Greece	Malawi	Thailand
Central African Rep.	Guatemala	Malaysia	Trinidad & Tobago
Canada	Hong Kong	Namibia	Turkey
Switzerland	Honduras	Niger	Uganda
Chile	Haiti	Nigeria	Uruguay
China	Indonesia	Nicaragua	USA
Cote d'Ivoire	India	Netherlands	Venezuela
Cameroon	Ireland	Norway	South Africa
Congo, Rep.	Iran	New Zealand	Congo, Dem. Rep.
Colombia	Israel	Pakistan	Zambia
Costa Rica	Italy	Panama	Zimbabwe
Denmark	Jamaica	Peru	
Dominican Rep.	Jordan	Philippines	
Ecuador	Japan	Papua New Guinea	

# List of Figures

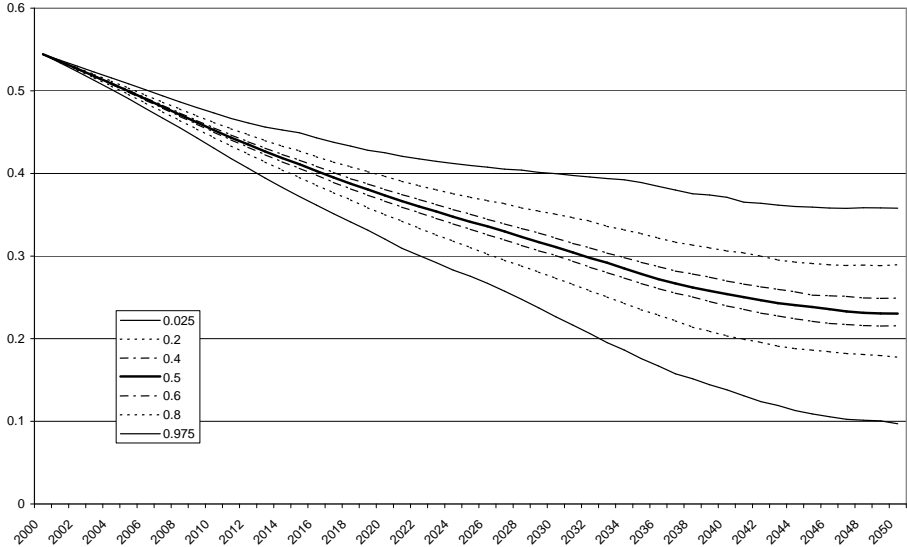
- 1 Youth dependency ratios; selected countries . . . . . 20
- 2 Probabilistic forecasts for the youth dependency ratio in India . . . . . 21
- 3 Probabilistic forecasts for the life expectancy at birth . . . . . 22
- 4 Probabilistic forecasts for the accounting effect . . . . . 23

Figure 1: Youth dependency ratios; selected countries



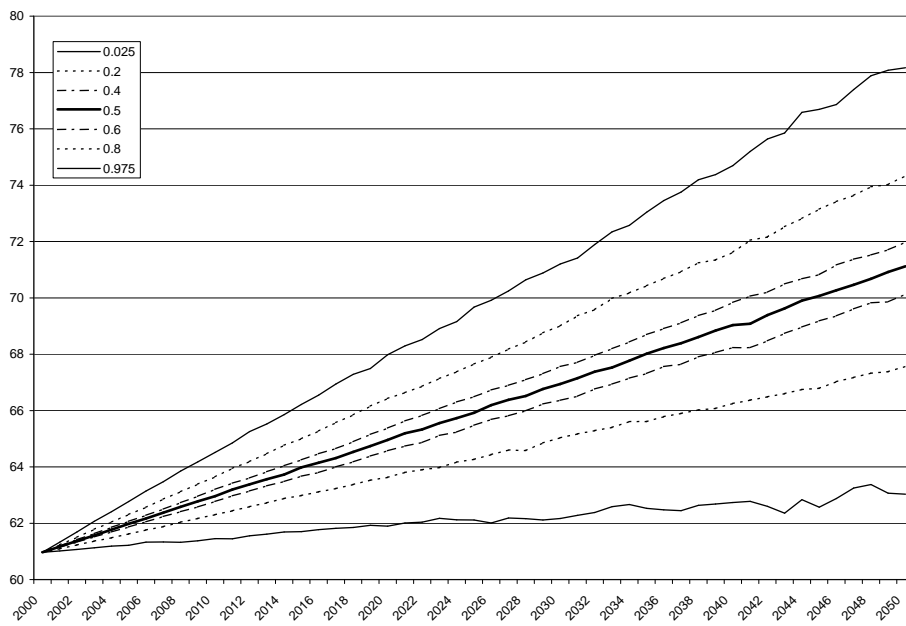
Source: United Nations (2003)

Figure 2: Probabilistic forecasts for the youth dependency ratio in India



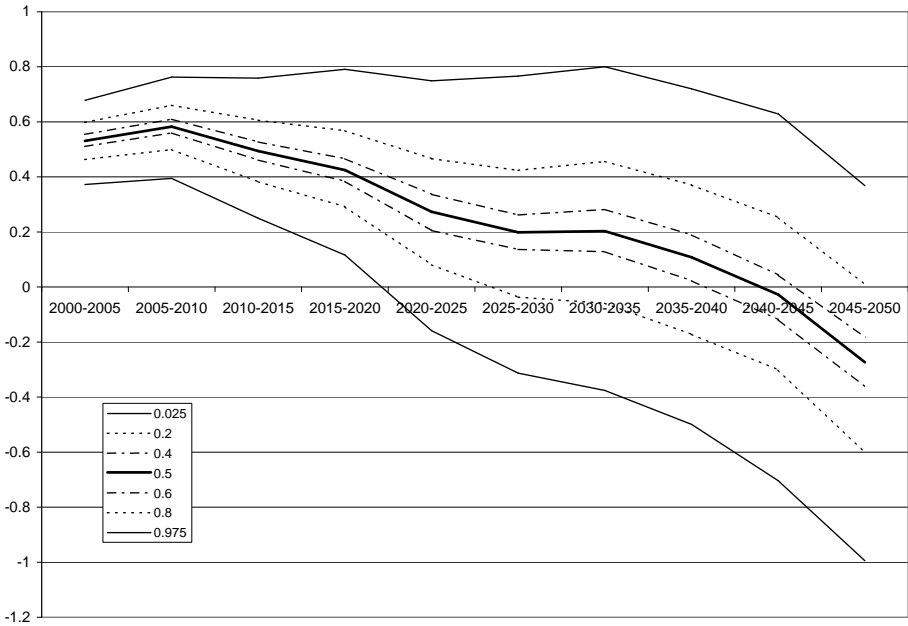
Source: United Nations (2003), own calculations based on the assumptions in Lutz and Scherbov (2004).

Figure 3: Probabilistic forecasts for the life expectancy at birth



Source: United Nations (2003), own calculations based on Lutz and Scherbov (2004)..

Figure 4: Probabilistic forecasts for the accounting effect



Source: United Nations (2003), own calculations based on Lutz and Scherbov (2004)..

# List of Tables

- 1 Explaining annual average growth of output per working age person 1960-95,  
OLS regression . . . . . 25
- 2 Explaining annual average growth of output per working age person 1960-95,  
IV estimation . . . . . 26
- 3 Predicting annual average growth of output per working age person 2000-35 . 27
- 4 Number of years it would take India to reach the value of output per working  
age person of UK in 2000 . . . . . 28

Table 1: Explaining annual average growth of output per working age person 1960-95, OLS regression

Dependent variable: Annual average growth rate of output per working-age person (x100)				
Independent variables				
Constant	6.02	7.04	-9.13	-12.79
	(0.00)	(0.00)	(0.03)	(0.00)
Ln of output per working age person 1960	-1.01	-1.14	-1.61	-1.69
	(0.00)	(0.00)	(0.00)	(0.00)
Ln of youth dependency ratio	-2.41	-2.94	-1.58	-1.08
	(0.00)	(0.01)	(0.07)	(0.01)
SOCINF	1.36	0.98	0.81	0.86
	(0.00)	(0.00)	(0.00)	(0.00)
Growth of the working-age population (ann. average)		2.43	0.86	
		(0.00)	(0.18)	
Growth of the total population (ann. average)		-2.29	-0.78	
		(0.00)	(0.27)	
Ln of life expectancy at birth			5.31	6.48
			(0.00)	(0.00)
$R^2$	0.46	0.60	0.69	0.68
Number of observations	97	97	97	97

<sup>a</sup>(White-Huber) heteroscedasticity-consistent p-values are reported in parentheses below the coefficient estimates.

Table 2: Explaining annual average growth of output per working age person 1960-95, IV estimation

Dependent variable: Annual average growth rate of output per working age person (x100)		
Independent variables	(1)	(2)
Constant	6.63 (0.00)	-8.69 (0.06)
Ln of output per working age person 1960	-1.04 (0.00)	-1.56 (0.00)
Ln of youth dependency ratio	-3.05 (0.00)	-1.60 (0.03)
SOCINF	1.04 (0.01)	0.83 (0.01)
Ln of life expectancy at birth		5.14 (0.00)
$R^2$	0.45	0.67
Number of observations	97	97

<sup>a</sup>P-values are reported in parentheses below the coefficient estimates.

<sup>b</sup>Estimation method is Fuller's modified LIML with the user-supplied Fuller parameter alpha set equal to one, as has been suggested.

<sup>c</sup>In regression (1) the instruments are ztropics, zsubtropics, zdestmp, zdestrp, lnFrankRom and EngLegal. In regression (2) the instruments are the same, except that ln(1+ME) is added as further instrument.

Table 3: Predicting annual average growth of output per working age person 2000-35

Percentiles	(1)	(2)	(3)	(1')
0.025	3.84	3.75	4.91	4.19
0.2	4.02	3.92	5.09	4.39
0.4	4.12	4.02	5.2	4.50
0.5	4.17	4.06	5.24	4.57
0.6	4.21	4.11	5.28	4.62
0.8	4.33	4.21	5.41	4.75
0.975	4.56	4.40	5.64	5.01

Table 4: Number of years it would take India to reach the value of output per working age person of UK in 2000

Percentiles	(1)	(2)
0.025	47.04	38.03
0.2	49.54	40.62
0.4	50.95	40.62
0.5	51.44	40.93
0.6	52.06	41.25
0.8	53.36	42.14
0.975	55.85	43.69

# VIENNA INSTITUTE OF DEMOGRAPHY

## *Working Papers*

Winkler-Dworak, Maria. 2003. Food Security, Fertility Differentials and Land Degradation in Sub-Saharan Africa: A Dynamic Framework. *VID Working Papers 01/2003*. Vienna Institute of Demography.

Frejka, Tomas and Jean-Paul Sardon. 2003. Fertility Austria: Past, Present and the Near Future. *VID Working Papers 02/2003*. Vienna Institute of Demography.

Sanderson, Warren C., Sergei Scherbov, Brian C. O'Neill, and Wolfgang Lutz. 2003. Conditional probabilistic population forecasting. *VID Working Papers 03/2003*. Vienna: Vienna Institute of Demography.

Fliegenschnee, Katrin, Anne Goujon and Wolfgang Lutz. 2004. Neue demographische Szenarien zur Zukunft der Evangelischen Kirche in Österreich. *VID Working Papers 01/2004*. Vienna: Vienna Institute of Demography.

Engelhardt, Henriette. 2004. Fertility Intentions and Preferences: Effects of Structural and Financial Incentives and Constraints in Austria. *VID Working Papers 02/2004*. Vienna: Vienna Institute of Demography.

Prskawetz, Alexia, Tomas Kögel, Warren C. Sanderson and Sergei Scherbov. 2004. The Effects of Age Structure on Economic Growth: An Application of Probabilistic Forecasting to India. *VID Working Papers 03/2004*. Vienna: Vienna Institute of Demography.

---