2006

Income inequality and economic growth

Nor Azam Abdul Razak
Louisiana State University and Agricultural and Mechanical College

INCOME INEQUALITY AND ECONOMIC GROWTH

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in
The Department of Economics

by
Nor Azam Abdul Razak
B.S., Management, Indiana University, 1992
M.A., Economics, University of Tsukuba, 1997
M.S., Economics, Louisiana State University, 2003
December, 2006

Dedication

To Mom, Dad, Nafishah, Nor Syuhada, Kamarul Izzat, Nor Amira, Nor Aleeya, Mohd Nazri, Rafidah, and Roslina.

Acknowledgments

First and foremost, I would like to express my heartfelt gratitude to my major professor, Dr. Chris Papageorgiou, for meticulously guiding me through this arduous academic journey. I would also like to express my utmost appreciation to the distinguished members of my dissertation committee—Dr. Douglas McMillin, Dr. Areendam Chanda, Dr. Faik Koray, and Dr. Wonik Kim—for their invaluable comments and suggestions. I am deeply indebted to Dr. Carter Hill, Dr. Tibor Besedes, and Dr. Sudipta Sarangi for sharing their wisdom with me. I am very grateful to the following individuals: Dr. Gary Barrett and Dr. Stephen Donald for making their Gauss programs available, Dr. Jeff Racine for providing with data on residual growth rates of per capita income, and Dr. Jong-Wha Lee for providing me with the unpublished data on the length of education across countries. Without guidance and assistance from these individuals, this dissertation would not have been completed.

In this lengthy education process, I have also benefited from the moral support of my family members and colleagues. On the family side, I truly appreciate the support of my mother, my wife, my brother, and my sisters. I am especially grateful to my mother, wife, and children for consoling me whenever I am stressed out. On the colleague side, I acknowledge the support of Mohd, Burak, Ozlem, Kelly, Anca, Pavlo, Beatrice, Kevin, Subaran, Mohd Faisol, Dr. Sobri, Dr. Mohd Dan, Dr. Mohd Zaini, and Dr. Juzhar.

Table of Contents

Abstract

A central issue in the growth literature is whether initial conditions matter for income disparity among nations. If they don't, then countries will converge to a single regime once the structural features of the economies are controlled for. If they do, then countries will converge to multiple regimes even if the structural features are controlled for. This dissertation is designed to investigate whether the world is characterized by a single or multiple regimes.

The first paper investigates whether the predictions of a particular multiple-regime model due to Galor and Zeira (1993) are borne out by the data. The baseline analysis is carried out with a sample of 46 countries (for which data are available) during the period 1970-2000. The analysis produces results that are consistent with the model. These results are broadly robust to different model specifications, sample periods, and permutations of alternative control variables. We take these results as evidence in support of the multiple-regime models.

The second paper examines whether the conclusions of another multiple-regime model due to Kremer and Chen (2002) are borne out by the data. The baseline analysis is also carried out with a sample of 46 countries during the period 1970-2000. The analysis produces results that are consistent with the model. These results are broadly robust across different human capital differential variables, different sample sizes, and additional control variables. We take these results as evidence in support of the multiple-regime models.

The third paper revisits the convergence hypothesis test using a new stochastic dominance method. The baseline analysis is carried out with a sample of 100 countries during the period 1960-2000. Together with the robustness check, the analysis yields results that are consistent with the club convergence hypothesis. We take these results as evidence in support of the multiple-regime models.

Chapter 1 Introduction

Economists have long been concerned about the issue of income disparity among nations. This concern is based on whether income disparity is expected to be temporary or permanent; see Durlauf, Johnson, and Temple (2005). If the disparity is expected to be temporary, then the market system may be left alone to run its course. If the disparity is expected to be permanent, however, then economic policy may be called upon to intervene. Hence, an important question is whether income disparity is likely to be temporary or permanent.

The theory of economic growth answers this question by advancing three competing hypotheses referred to as the absolute, conditional, and club convergence hypotheses; see Galor (1996) for a lucid discussion of these hypotheses. According to the absolute convergence hypothesis, regardless of their initial conditions, countries will eventually converge to each other in terms of per capita income. If this hypothesis is borne out in the real world, then income disparity is temporary. According to the conditional convergence hypothesis, regardless of their initial conditions, countries will eventually converge to each other only if they have similar structural features such as technologies, saving rates, population growth rates. According to the club convergence hypothesis, depending on their initial conditions, countries might diverge from each other in the long run even if they have similar structural features. If the hypothesis of either conditional or club convergence holds in the real world, then income disparity might be permanent. However, economic policies should be targeted toward structural features in the former and initial conditions in the latter.

Of the three hypotheses, the first one is neither predicted by any growth models nor substantiated by any empirical evidence. Thus, the first hypothesis can be readily dismissed. Of the latter two, the former is predicted by the neoclassical model such as the one developed by

Solow (1956) while the latter is predicted by the so-called multiple-regime models such as the one developed by Azariadis and Drazen (1990). Of these two models, the bulk of empirical evidence appears to lend support to the latter (see Durlauf and Johnson (1995) and Hansen (2000), among others). Nevertheless, the regression-based models on which this evidence rests have come under attack under statistical grounds. This attack has cast doubt on the findings of the convergence hypothesis. Despite this attack, some researchers have responded to its empirical support by developing alternative growth models that are characterized by multiple steady states.

This dissertation is organized as follows. In Chapter 2, we test whether the predictions of a particular multiple-regime model, developed by Galor and Zeira (1993), are supported by the international data. In Chapter 3, we test whether the predictions of another multiple-regime model, developed by Kremer and Chen (2002), are supported by the international data. In Chapter 4, we revisit the convergence hypothesis using a novel method called stochastic dominance. In Chapter 5, we offer some concluding observations.

Chapter 2, *Empirical Analysis of the Galor-Zeira Model*, investigates whether the theoretical predictions of the Galor-Zeira model are borne out by the experience of modern economies. According to this model, the initial distribution of income is the initial condition which determines whether a given economy converges to a low- or high-income group. This conjecture depends on the assumption that individuals vary in the amount of wealth inherited from their parents. Those who inherit a sufficient amount of wealth may invest in education and become skilled workers. Those who inherit an insufficient amount of wealth have two options: they may borrow to finance their education and become skilled workers during the second period of their lives or choose not to acquire education and remain as unskilled workers for the rest of their lives. If they borrow, they need to repay the debt plus its interest payments. Due to the assumed imperfect credit markets, banks charge a higher interest rate than the one when credit markets are perfect. This higher-than-normal interest rate deters some poor individuals from investing in education (i.e., they are credit-constrained).

It follows then that a) there exists a threshold level of wealth below (above) which individuals choose to become unskilled (skilled) workers, and b) if the population of an economy is concentrated below (above) this threshold level, then the economy converges to a low-income (high-income) group. Together, both of these statements suggest that income inequality has an adverse impact on the long-run income level.

Based on the theoretical predictions of the Galor-Zeira model, we construct an appropriate specification for a regression model. Since income inequality affects the income level through human capital investment, we employ an instrumental variable (IV) model in which the level of income per capita is regressed on human capital investment, where the latter is instrumented by a measure of income inequality. Following the literature, we adopt a Gini index variable compiled by Deininger and Squire (1996) as a proxy for income inequality. In the second-stage regression, we also include other control variables as implied by the Solow (1956) model.

In the baseline estimation, we obtain results that lend support to the Galor-Ziera model. That is, the coefficient on the income inequality measure is negative and significant and the coefficient of human capital investment is positive and significant. To control for idiosyncratic factors that might vary across regions, we add dummy variables for Latin American and Asian countries. We find that the results in the baseline estimation continue to hold. Therefore, we conclude that our analysis yields evidence in support of the Galor-Ziera model.

To determine whether these baseline results are robust to alternative model specifications, sample sizes, etc., we proceed with a series of robustness checks. First, we replace the dependent variable, the level of income per capita, with the growth rate of income per capita. Second, we replace the stock measure of human capital that we employ with a flow measure of human capital. Third, we alter the sample size by curtailing the sample period. Next, we expand the sample size by utilizing the panel data, and we control for individual country effects by estimating the fixed effects (FE) model, and for measurement error by estimating the between effects (BE) model. Additional robustness tests are also considered. In most of these exercises, we find that the results are broadly consistent with those in the baseline estimation. We take all of these results as evidence in favor of the Galor-Zeira model.

Chapter 3, *Empirical Analysis of the Kremer-Chen Model*, investigates whether the theoretical predictions of the Kremer-Chen model are borne out in the real world. Like the Galor-Zeira model, this model also conjectures that the initial distribution of income determines whether a particular economy converges to a low- or high-income group. Unlike the Galor-Zeira model, this conjecture depends on the assumption that individuals vary in the opportunity cost of rearing children. For poor (rich) households, the opportunity cost childrearing is low (high); thus, they end up having many (few) children. In addition to the opportunity cost, childrearing entails direct costs, one component of which is education. For poor (rich) households, the cost of education as a fraction of the total cost of childrearing is substantial (immaterial); therefore, they end up investing less (more) in the education of their children. Hence, poor people end up with many yet uneducated children while rich households end up with few yet educated children (there is a trade-off between the quantity and quality of children).

It follows then that a) there exists a threshold level of wealth below (above) which households invest more in the quantity (quality) of their children, and b) if the population of an economy is concentrated below (above) this threshold level, then the economy converges to a low-education (high-education) group. Taken together, both statements imply that income inequality has a negative impact on the long-run income level.

We begin by specifying a regression model that is consistent with the theoretical predictions of the Kremer-Chen model. Since income inequality affects the income level through human capital differential, we adopt an IV model with the level of income per capita is regressed on human capital differential, where the latter is instrumented by the Gini index which serves as a proxy for income inequality. Like Chapter 2, other control variables as implied by the Solow (1956) model are also included.

It turns out that there are a number of alternative proxies for the human capital differential variable. Experimenting with all of these proxies, we find that most results are broadly consistent with the prediction of the Kremer-Chen model. That is, the coefficient of the Gini index is negative and significant and the coefficient of human capital differential is positive and significant. As before, we add dummy variables for Latin American and Asian countries to control for idiosyncratic factors that might vary across regions. The estimation results remain intact even with the inclusion of these regional dummies. We take all of these results as evidence in support of the Kremer-Chen model. To limit the scope of our analysis, we choose the best proxy for the robustness analysis.

As in Chapter 2, we proceed with a series of robustness checks. First, we improve human capital data by weighting the original data with data on the return to schooling and the duration of schooling. Second, we expand the sample size by employing the panel data. Next, we conduct other sensitivity checks such as controlling for individual country effects (by estimating the FE model) and measurement error (by estimating the BE model). In many cases considered, we find that the results are broadly consistent with those in the baseline estimation. We take all of these results as evidence in favor of the Kremer-Chen model.

Chapter 4, *A Reexamination of the Convergence Hypothesis*, revisits the convergence hypothesis using a novel method called stochastic dominance. An advantage of the stochastic dominance method over regression-based convergence analysis is that it considers the entire growth distribution. In brief, this method works as follows: given the growth distribution of, say, Africa and OECD, then convergence is said to occur if the growth distribution of Africa fails to stochastically dominate that of OECD and vice versa. If the growth distribution of Africa stochastically dominates that of OECD or vice versa, then they diverge from each other.

We begin the analysis by utilizing the data on actual growth rates of per capita income for 100 countries during the period 1960-2000. This 100-country sample is then partitioned into four regional groups: Africa, Latin America, Asia, and OECD countries. Next, we break the sample period into two: 1961-1980 and 1981-2000. In interregional analysis, we find the following. First, convergence occurs between a) Asia and OECD, and b) Latin America and the rest of the world (ROW) only. Second, the evidence is mixed for Africa and Latin America. Finally, divergence occurs between any other pairs of regions. In intertemporal analysis, we find that a) divergence occurs for Latin America, and b) convergence occurs for each of the remaining regions plus the world as a whole.

The baseline analysis is carried out without conditioning on the structural features of the economies. Hence, its results can merely be interpreted in terms of absolute versus club convergence. Building on the work of Maasoumi, Racine, and Stengos (2006), we repeat the analysis with the data on residual growth rates of per capita income. Since residual growth rates are obtained by conditioning actual growth rates on structural features of the economies, the analysis can now be construed in terms of conditional versus club convergence. In interregional analysis, we find evidence of convergence in most cases considered. In intertemporal analysis, we find that a) convergence occurs for Africa only, b) the evidence is mixed for Latin America, and c) divergence occurs for each of the remaining regions plus the entire world. We take these results as evidence in favor of club convergence.

Together, these three chapters yield results that are consistent with multiple-regime models. As mentioned earlier, these models are consistent with the idea of sustained income disparity which is due to initial conditions. Accordingly, economic policies need to be focused on altering these initial conditions in order to reverse the trend.

Chapter 2 Empirical Analysis of the Galor-Zeira Model

2.1 Introduction

The idea that income distribution affects economic growth dates back to at least as early as Kaldor (1957). According to Kaldor, income inequality is good for growth because concentrated wealth in the hands of a few permits greater savings, which are conducive for investment. Since the prediction of this model is often at odds with empirical evidence, this view has recently been challenged by a handful of economists. In the 1990s, at least three alternative theoretical models were developed to explain this phenomenon. According to the first model, known as the political-economy model, income inequality is bad for growth because average citizens would push the government for more extensive redistributive policies, which are detrimental for investment and growth. According to the second model, known as the sociopolitical instability model, income inequality is bad for growth because it might create social tension which is harmful for investment. According to the third model, known as the credit constraint model, income inequality is bad for growth because it restricts the number of people who have access to costly education. (An extensive overview of the three models is offered by Benabou (1996).)

The political-economy model was developed and tested by Alesina and Rodrik (1994) and Persson and Tabellini (1994). They found evidence consistent with their theoretical prediction: income inequality has a negative impact on economic growth. Their empirical analysis was followed by George Clarke (1995) who used alternative measures of income inequality. He showed that the basic result of a negative inequality-growth relationship is robust across alternative measures of income inequality. Next, Alesina and Perotti (1996) conducted an empirical analysis of the sociopolitical instability model. They found evidence in support of the model: income inequality tends to raise social instability which, in turn, adversely affects investment and growth.

Then, Roberto Perotti (1996) tested all of the inequality-growth models using a carefully specified reduced-form model and a variety of structural models. He concluded that the result is robust in a reduced-form model but mixed in structural models. However, the test of the credit constraint model is supported by the data.

Next, Deininger and Squire (1998) revisited the analysis with an improved measure of income inequality that they carefully compiled earlier (Deininger and Squire (1996)). They found that the basic result is not robust with their improved inequality data. Then, Li and Zou (1998) generalized the Alesina-Rodrik specification to a panel data of countries. They found that the basic negative inequality-growth result is overturned. That is, income inequality has a positive impact on economic growth. This reversed result was also found by Forbes (2000), who generalized Perotti's reduced-form specification to the panel data. However, this new result collapsed in Barro (2000), who deviated from the practice of adopting a parsimonious specification by including the inequality data that Deininger and Squire (1996) regard as low quality. In this instance, Barro showed that income inequality has no effect on growth. Finally, Sylwester (2000) restored the basic result when he tested the inequality-education-growth link with cross-country data.

A review of these empirical studies reveals that most of them are based on reduced-form models of the inequality-growth relationship. The problem with this approach is that any empirical evidence obtained cannot be directly attributed to any one of the structural models. In some cases, researchers try to patch up this deficiency by testing a structural model as well. However, most researchers do so for the political-economy model, one study is exclusively concerned with the sociopolitical instability model (Alesina and Perotti, 1996), one is devoted to an unspecified model which is similar to the credit constraint model (Sylwester, 2000), and one briefly tests the credit constraint model (Perotti, 1996).

The credit constraint model was developed by Galor and Zeira (1993). Since its introduction, this model has occupied a central place in the growth literature due to its theoretical implications on the phenomenon of persistent income inequality. Some economists argue that this phenomenon is due to structural features of the economy. If this is the case, then different economies will converge to a single regime once these structural features are controlled for. Other economists claim that this phenomenon is due to initial conditions of the economy. If so, then different economies will converge to multiple regimes even if these structural features are held constant.

The model developed by Galor and Zeira (1993) contributes to this single- vs. multiple-regime debate by demonstrating that a particular initial condition (the initial distribution of income), determines whether an economy will converge to a low- or high-income regime. Accordingly, an empirical test of this model may shed some light on this debate. Despite its prominent role in the literature, the Galor-Zeira model has not been comprehensively tested. Therefore, this paper attempts to fill this gap. In particular, we conduct a cross-country empirical analysis for 46 countries during the 1970-2000 period based on the structural model of the inequality-education-growth relationship developed by Galor and Zeira (1993). In the baseline analysis, we find that income inequality has a negative impact on human capital investment, which in turn has a positive impact on the long-run per capita income. This result is quite robust across different model specifications, estimation methods, and various permutations of variables. We interpret these results as evidence in favor of the Galor-Zeira model.

We begin our analysis in Section 2.2 by discussing the model developed by Galor and Zeira (1993), with particular attention to their testable implications. In Section 2.3, we specify our empirical model and then discuss the data used in this paper. In Section 2.4, we report the results of our basic and robustness analyses. In Section 2.5, we discuss and interpret our main results. In Section 2.6, we highlight the major findings of this paper and conclude.

2.2 The Galor-Zeira Model

Galor and Zeira (1993) introduce an overlapping-generation model of the economy with altruism. Here we merely provide a sketch of the model. An economy consists of individuals who live for two periods. During the first period, they may choose to work or invest in human capital; during the second period, they simply work. If they invested in human capital during the first period, they would work as skilled workers in the second period and receive high wages; otherwise, they would work as unskilled workers in both periods and receive low wages. The work-study decision in the first period depends partly on the amount of wealth they inherit from their parents. Assuming that this inheritance varies from one person to another, then those with greater inheritance stand a better chance of acquiring education. If one's inheritance is not sufficient, then one can still invest in human capital by borrowing. However, due to assumed imperfect credit markets, some individuals are credit-constrained. That is, there are individuals who cannot afford to acquire education because their inheritance falls short of a certain minimum amount, and they are denied educational loans. To formalize the model, let the utility function of the representative individual be expressed as

$$V = \alpha \log c + (1-\alpha) \log b, \qquad \alpha \in (0,1), \tag{2.1}$$

where V is the level of utility, c is the amount of consumption of the individual, b is the amount of the bequest the individual makes to his children, α is the elasticity of utility with respect to consumption, and $(1-\alpha)$ is the elasticity of utility with respect to bequest. (Since bequest, b, is made in the second period, Eq.(2.1) refers to the utility function for the second period.) The argument b that enters the utility function suggests that individuals derive some amount of utility from sharing their wealth with their offspring. This bequest then becomes an inheritance of their children.

Let x denote the amount of individual inheritance, which varies across individuals. Let h denote the fixed amount of human capital investment needed for individuals to become skilled workers. Regardless of the relative magnitude of x and h, individuals may or may not choose to invest in education. If they choose not to invest in education, then they will become unskilled workers during both periods of their lives, and receive the wage rate of w_u. In addition, they may lend money to others, at the rate of r. As such, their budget constraint for the second period is

$$c = (x + w_u)(1 + r) + w_u - b. \tag{2.2}$$

If they choose to invest in education, then they will become skilled workers during the second period, and receive the wage rate of w_s. However, the funding for the education expenditures depends on the relative size of x and h: a) if $x < h$, then the expenses will be funded through borrowing, at the rate of i, b) if $x \geq h$, then the expenses will be funded through inheritance. As such, the respective budget constraints for the second period are

$$c = (x - h)(1 + i) + w_s - b, \tag{2.3}$$

$$c = (x - h)(1 + r) + w_s - b. \tag{2.4}$$

If we maximize Eq.(2.1) with respect to b subject to the appropriate budget constraint, we will obtain the optimal bequest for various individuals. With some plausible assumptions and algebra, Galor and Zeira (1993) show that the optimal amount of inheritance is given by a first-order nonlinear difference equation in (x_{t+1}, x_t) space:

$$x_{t+1} = (1 - \alpha) \left\{ \begin{array}{c} [(x_t + w_u)(1 + r) + w_u] \text{ if } x_t < f \\ [(x_t - h)(1 + i) + w_s] \text{ if } f \leq x_t < h \\ [(x_t - h)(1 + r) + w_s] \text{ if } x_t > h \end{array} \right\}, \tag{2.5}$$

where x_{t+1} is the amount of inheritance at time $t+1$, x_t is the amount of inheritance at time t, and the remaining variables and parameters are as defined before (see Appendix A for the derivation of Eq.(2.5)).

Eq.(2.5) can be alternatively represented by a curve with two kinks as depicted in Figure 2.1. The figure shows that individuals are divided into three groups: those with inheritance less than f, those with inheritance between f and h, and those with inheritance more than h. Individuals whose inheritance is less than f are those who work during both periods of their lives as unskilled workers. Individuals whose inheritance is between f and h are those who acquire education during the first period through borrowings and work during the second period as skilled workers. Individuals whose inheritance is more than h are those who acquire education during the first period through inheritance and work during the second period as skilled workers.

Even though there are three groups of individuals to begin with, it turns out that the population is finally partitioned into two groups separated by point g, an unstable equilibrium point. That is, those individuals who receive inheritance less than g will end up in the poor group, x_{poor}, and those who receive inheritance more than g will end up in the rich group, x_{rich}, in the long-run. The reason for this dynamic evolution is that a minimum amount of inheritance is needed before subsequent generations can provide enough bequests for their offspring as well.

It can be inferred from Figure 2.1 that the long-run levels of income are positively related to the initial number of individuals who inherit more than g. To illustrate, consider an economy characterized by three different scenarios. First, one-half of the population is concentrated around f and the remaining one-half around h. Second, one-third of the population is concentrated around f and the remaining two-third around h. Third, two-third of the population is concentrated around f and the remaining one-third around h. In all cases, the fraction of population that lives around f will move to x_{poor} and the fraction of population that lives around h will move to x_{rich}.

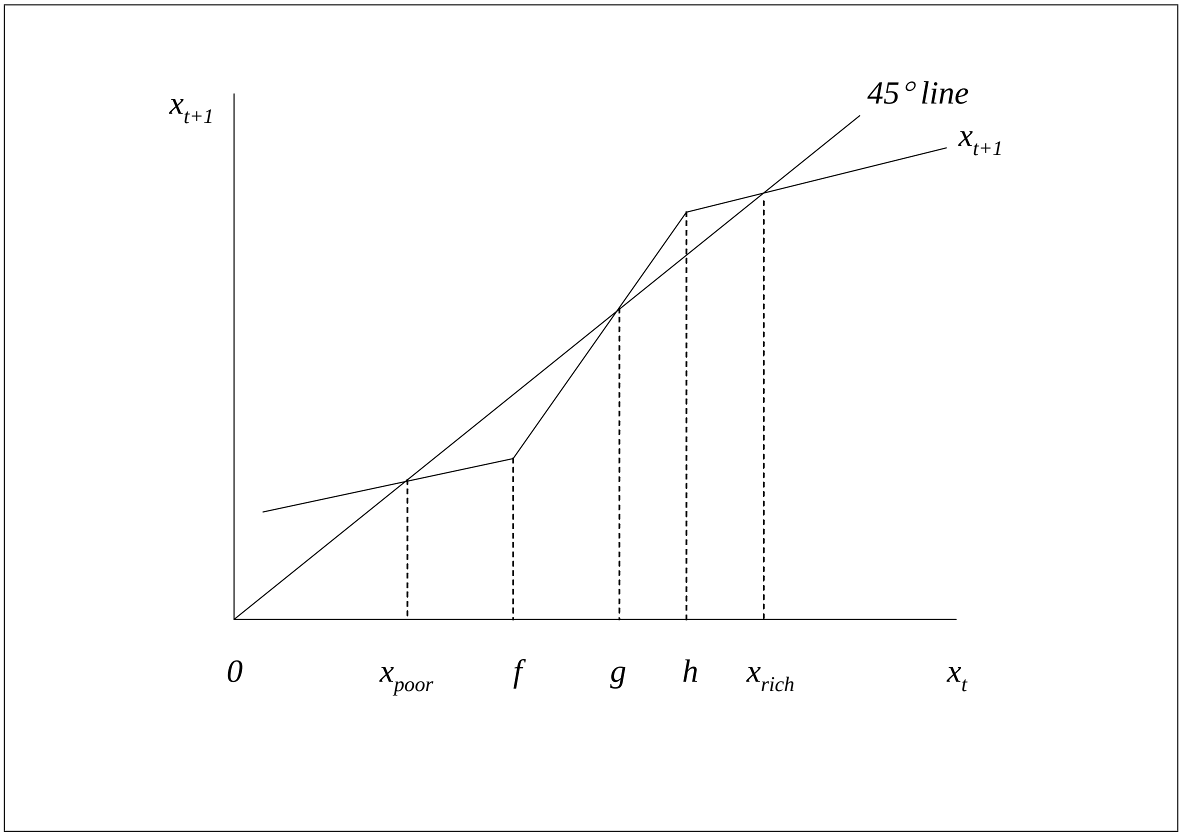

Figure 2.1: The Dynamics of Galor-Zeira Model

With reasonable values of income at x_{poor}, f, g, h, and x_{rich} (say, 10, 30, 50, 70, and 90, respectively), we can deduce the following: income tends to remain unchanged in the first scenario, income tends to rise in the second scenario, and income tends to fall in the third scenario. Thus, the larger the fraction of people who inherit more than g, the higher the long-run income tends to be. If we let g be the threshold that separates a poor from a non-poor economy, then we obtain the following conclusions: 1) An initially poor economy will end up poor in the long run, 2) An initially non-poor economy with wealth distributed among many will end up rich, and 3) An initially non-poor economy with wealth distributed among few will end up poor.

2.3 Model Specification and Data

We begin by specifying our model based on the description of the Galor-Zeira model. Next, we compare our specification with the related ones employed in the literature. Finally, we describe the data used in this paper based on our model specification.

2.3.1 Model Specification

The above conclusions constitute a testable implication of the Galor-Zeira model. That is, an initially non-poor country has a better chance to enrich itself in the future if its income is more equally distributed among its citizens. Thus, a priori we would expect that higher income inequality will have a negative impact on a country's income in the long-run. Since this implication is not applicable to an initially poor country, we can control for this effect by introducing a dummy variable for poor countries. Since income inequality affects future income through education, we estimate a structural model which consists of two equations. In the first equation, income is a function of education and other explanatory variables in the Solow growth regression. In the second equation, education is a function of income inequality and a dummy variable for poor countries. In particular, we estimate the following structural model:

$$\begin{aligned} \mathit{Income} &= \alpha_1 + \alpha_2.\mathit{Educ} + \alpha_3.\mathit{Invest} + \alpha_4.(n+g+\delta) + u, && (2.6) \\ \mathit{Educ} &= \beta_1 + \beta_2.\mathit{Gini} + \beta_3.\mathit{Poor} + v, && (2.7) \end{aligned}$$

where Income is the level of long-run income per capita, Educ is the amount of human capital investment, Invest is the amount of physical capital investment, $(n + g + \delta)$ is the sum of the rates of population growth, technological progress, and capital depreciation, Gini is the Gini index which measures the degree of inequality, Poor is a dummy variable equal to 1 for an initially poor country and 0 otherwise, and u and v are the error terms. A priori, we expect the coefficients of Gini, Poor, and $(n + g + \delta)$ to be negative and those of Educ and Invest to be positive.

2.3.2 Related Model Specifications

At this point, we may compare our structural specification with two closely related ones, namely, those of Perotti (1996) and Sylwester (2000). Perotti (1996) employs the following structural model:

$$\begin{aligned} \textit{Growth} &= \alpha_1 + \alpha_2.\textit{Educflow} + x'\delta + u, && (2.8) \\ \textit{Educflow} &= \beta_1 + \beta_2.\textit{Mid} + \beta_3.\textit{Educstockf} + \beta_4.\textit{Educstockm} + v, && (2.9) \end{aligned}$$

where Growth is the growth rate of per capita income for the period 1960-1985, Educflow is the flow of human capital, x is a vector of control variables (which includes initial income per capita and PPP investment deflator), Mid is the income share of the third and fourth quintiles of population which measures income equality (as opposed to income inequality), Educstockf is the stock of female human capital, and Educstockm is the stock of male human capital.

There are a few notable differences between Perotti's and our structural model. First, Perotti's dependent variable in the first equation is Growth while ours is Income. We use Income because that is what is implied by the Galor-Zeira model; Perotti uses Growth because that is a standard practice in the growth empirics. This should not be a problem, however, because we can always transform our level regression into the growth regression. Second, Perotti discriminates between two measures of human capital—stock and flow—and he treats the flow measure as endogenous and the stock measure as exogenous. Third, Perotti includes a PPP investment deflator in order to account for market distortion. However, this variable is not an important determinant of growth. Finally, Perotti does not include Invest and Poor. The omission of the former follows from his reduced-form model which tries to accommodate other theoretical models. Nevertheless, this variable is an important determinant of growth. The exclusion of Poor is unfortunate since this is implied by the Galor-Ziera model.

Sylwester (2000) employs the following structural model:

$$\begin{aligned} \textit{Growth} &= \alpha_1 + \alpha_2.\textit{Educ\$} + x'\delta + u, && (2.10) \\ \textit{Educ\$} &= \beta_1 + \beta_2.\textit{Gini} + \beta_3.\textit{Democracy} + \beta_4.\textit{Pop} + v, && (2.11) \end{aligned}$$

where Educ$ is the amount of educational expenditures, x is a vector of control variables (which includes the lagged value of Educ$, the stock of human capital, and initial income per capita), Democracy is a dummy variable equal to one for a democratic country and zero otherwise, and Pop is the growth rate of population; other variables are as defined before.

To begin with, Sylwester does not base his specification on a theoretical model. His main concern is to determine whether income inequality affects growth through education. It turns out that his specification is consistent with the credit constraint model. There are several differences, though. First, Sylwester uses a distinctive measure of human capital, educational expenditures. This measure can be thought of as another proxy for the flow of human capital. Second, he employs both the stock and flow of human capital in Eq. (2.10). Third, he also includes the lagged value of educational expenditures in both of his equations. Finally, he adds Democracy, a variable which is implied by the political-economy model but not by the Galor-Zeira model.

2.3.3 Basic Data

On the basis of the preceding discussion, we believe that our specifications in Eqs. (2.6) and (2.7) are appropriate. We proceed by collecting the cross-country data for all of the variables identified in those equations from various sources. It turns out that the Gini data imposes substantial restrictions on the number of available observations. If we wish to use this data for as early as 1960, then we end up with as few as 14 observations. The number of available observations rises as we adjust the beginning period upward: 27 if we begin from 1965, 41 if 1970, 52 if 1975, and 62 if 1980.

To have as many observations as possible while having data for a relatively long period of time, we relax the time classification for the inequality data. That is, data that ranges between 1960 and 1965 is treated as the 1960 data, data that ranges between 1970 and 1975 is treated as the 1970 data, and data that ranges between 1980 and 1985 is treated as the 1980 data. With this slight relaxation of classification, we have the following: 75 observations if we begin from 1980, 56 if

1970, 29 if 1960, etc. We settle for data that begins from 1970; hence, we have 56 observations. When we match these data with the data on other variables, we lose another 10 observations. Thus, we end up with 46 observations.

Given this restriction, we collect the necessary data for 46 countries during the period 1970-2000. Based on Eqs. (2.6) and (2.7), we need data for the following six variables:

- Inc2000: This variable is defined as the log of the real GDP per capita in 2000 and is taken from the Penn World Table version 6.1 (PWT6.1).

- Invest: This variable is defined as the log of the annual average of the ratio of real investment to GDP during the period 1970-2000. This variable is taken from PWT6.1.

- $(n + g + \delta)$: This variable is defined as the log of the sum of the rates of population growth (n), technological progress (g), and capital depreciation (δ). The population growth rate data (n), taken from PWT6.1, is defined as the annual average of the population growth rate during 1970-2000. We follow the literature by setting $g + \delta = 0.05$.

- Educ: This variable is defined as the log of average years of schooling for population over 25 years old during the period 1970-2000.[1,2] This measure is taken from Barro and Lee (2001).

- Gini: This variable, which measures the degree of income inequality, is defined as the log of the Gini index in 1970 or its closest neighboring period but cannot exceed 1975. *Gini* is taken from Deininger and Squire (1996), who make the necessary efforts to compile high-quality income distribution data. In particular, they impose three stringent quality criteria before the data can be accepted. First, data must be based on household surveys (not from national accounts that make some assumptions about patterns of income inequality). Second, data must be based on comprehensive coverage of population (not based on some segments of population only). Third, data must be based on comprehensive coverage of income sources (not based on wage incomes only but also nonwage incomes).[3]

- Poor: This variable is defined as a dummy variable, which is equal to 1 for any countries that are classified by the World Bank as low-income countries in 1970 (and 0 otherwise) based on their income range. Since the data for 1970 is not available, we use the data for 1972. This data is taken the World Tables 1976, published by the World Bank.

[1] We could have defined *Educ* as the log of annual average of the educational attainment for population over 15 years old. However, this measure of educational attainment may potentially create bias since part of the population, ages 15–24, do not have a chance to complete tertiary education.

[2] It has been suggested that the education variable be measured in 2000 (i.e., *Educ2000*) instead of as the average of 1970-2000. However, doing so makes us prone to simultaneity bias between *Educ2000* and *Inc2000*.

[3] It has been suggested that we consider another source of Gini data from Texas income inequality data.

2.4 Empirical Analysis

Using cross-country data for 46 countries during the period 1970-2000, we conduct an empirical analysis of the Galor-Zeira model based on Eqs. (2.6) and (2.7). In particular, we estimate Eq. (2.6) by the instrumental variable (IV) method, where Educ is instrumented by Gini and Poor. Hence, Eq. (2.7) corresponds to the first-stage regression and Eq.(2.6) the second-stage regression.[4]

2.4.1 Basic Analysis

We begin by running the first-stage regression corresponding to Eq. (2.7) and present the estimation results in Table 2.1. Column (1b) shows that the coefficients of Gini and Poor are individually significant at the 1% level. Since both coefficients are also jointly significant at the 5% level, we proceed with the second-stage regression and present the results in Column (1a). (In any analysis, the second-stage regression is conducted only if Gini and Poor are jointly significant.) We observe that the coefficient of Educ enters with the expected sign and significant at the 1% level of significance. While the coefficient of Invest is of the anticipated sign, it is insignificant. Finally, the coefficient of $(n + g + \delta)$ enters with the wrong sign and is insignificant. Since the coefficients of key variables—Gini, Poor, and Educ—enter with the correct signs and significant, we take these results as evidence in favor of the Galor-Zeira model.

The empirical literature on the inequality-growth relationship usually adds three regional dummy variables—the Latin American countries, the Asian countries, and the African countries—in order to control for institutional and cultural factors that might differ across regions. Since there are only two African countries in our 46-country sample, we add two regional dummies only, Latin and Asia, to our second-stage regression.[5] In the first-stage regression, Column (2b), we

4 Since *Invest* and $(n + g + \delta)$ are assumed to be exogenous, their coefficients will enter the first-stage regression as well to ensure that *Educ* is estimated with the optimal set of instruments; see Chapter 5 of Wooldridge (2002). However, these exogenous variables have little meaning in the first-stage regression. Hence, their coefficients will be suppressed from the first-stage regression results.

5 Like *Invest* and $(n + g + \delta)$, these regional dummies are assumed to be exogenous. Hence, they will enter the first-stage regression too. However, since they have little meaning there, their estimates will be suppressed.

see that the coefficients of Gini and Poor continue to be significant and their magnitudes remain robust.[6] In the second-stage regression, Column (2a), we observe that the coefficients of regional dummies are significant, the coefficient of Educ continues to be significant and its magnitude remains robust, and the coefficient of $(n + g + \delta)$ continues to be insignificant and enters with the wrong sign. However, the coefficient of Invest becomes significant now.

Table 2.1: Preliminary Estimation

Dep.Variable	*Inc2000*	*Educ*	*Inc2000*	*Educ*
	(1a)	(1b)	(2a)	(2b)
Constant	5.526^{**} (2.51)	2.627 (1.36)	7.750^{***} (3.39)	2.706 (1.26)
Gini	–	-0.839^{***} (−3.51)	–	-0.999^{***} (−3.57)
Poor	–	-0.872^{***} (−4.84)	–	-0.846^{***} (−4.30)
Educ	1.774^{***} (5.40)	–	1.610^{***} (5.64)	–
Invest	0.260 (1.44)	–	0.337^{**} (2.04)	–
$(n + g + \delta)$	0.039 (0.04)	–	0.726 (0.81)	–
Latin	–	–	-0.591^{***} (−2.95)	–
Asia	–	–	-0.370^{*} (−1.75)	–
Adj.R^2	0.68	0.63	0.75	0.63
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

We repeat our preliminary estimation without African countries. As shown in Table 2.2, the coefficients of Gini, Poor, Educ, and Invest enter with the correct signs and significant in both basic and augmented specifications. However, the coefficient of $(n + g + \delta)$ enters with the wrong sign and insignificant in both specifications. Finally, the coefficients of regional dummies enter with the negative signs (but only the coefficient of Latin is significant). Compared to the corresponding coefficients in Table 2.1, these results are qualitatively similar with one major

[6] The magnitude of a coefficient is said to be robust if it differs from the corresponding coefficient in the basic specification within one standard deviation. In our sample, one standard deviation of *Gini*, *Poor*, *Educ*, *Invest*, $(n + g + \delta)$, *Latin*, and *Asia* is 0.25, 0.31, 0.51, 0.52, 0.14, 0.46, and 0.47, respectively.

exception: the coefficient of Invest becomes significant in both specifications. This result suggests that the presence of African countries drives the previous results. Therefore, we will include a dummy variable for African countries, Africa, in each basic specification in subsequent exercises.

Table 2.2: Preliminary Estimation without African Countries

Dep.Variable	*Inc2000* (1a)	*Educ* (1b)	*Inc2000* (2a)	*Educ* (2b)
Constant	5.244^{**} (2.40)	2.228 (1.20)	7.122^{***} (3.12)	2.578 (1.23)
Gini	–	-0.718^{***} (-3.08)	–	-0.779^{**} (-2.64)
Poor	–	-0.809^{***} (-4.52)	–	-0.781^{***} (-3.91)
Educ	1.878^{***} (4.89)	–	1.677^{***} (4.49)	–
Invest	0.342^{*} (1.87)	–	0.322^{*} (1.86)	–
$(n+g+\delta)$	0.101 (0.11)	–	0.550 (0.60)	–
Latin	–	–	-0.504^{**} (-2.22)	–
Asia	–	–	-0.268 (-1.09)	–
Adj.R^2	0.63	0.57	0.69	0.55
Obs.	44	44	44	44

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 2.3 shows the results of our baseline estimation. In Column (1b), we see that the coefficients of Gini and Poor enter with the expected signs and significant. In Column (1a), we see that the coefficients of Educ and Invest enter with the correct signs and significant, the coefficient of $(n + g + \delta)$ enters with the incorrect sign and insignificant, and the coeffiient of Africa enters with the positive sign and significant. This last result mirrors the previous finding that African countries matter. We proceed by including all regional dummies and report the results in Columns (2a) and (2b). The coefficients of Gini, Poor, Educ, and Invest enter with the correct signs and significant, and their magnitudes are robust to those in the basic specification. However, the coefficient of Africa becomes insignificant due to the presence of other regional dummies. Since the coefficients of key variables remain intact, we conclude that these findings lend support to the Galor-Zeira model.

Table 2.3: Baseline Estimation

Dep.Variable	*Inc2000* (1a)	*Educ* (1b)	*Inc2000* (2a)	*Educ* (2b)
Constant	5.437^{**} (2.50)	2.164 (1.20)	7.314^{***} (3.15)	2.500 (1.23)
Gini	–	-0.710^{***} (−3.15)	–	-0.766^{***} (−2.70)
Poor	–	-0.799^{***} (−4.74)	–	-0.770^{***} (−4.06)
Educ	1.940^{***} (5.23)	–	1.776^{***} (4.85)	–
Invest	0.368^{**} (2.09)	–	0.359^{**} (2.12)	–
$(n+g+\delta)$	0.238 (0.26)	–	0.725 (0.79)	–
Africa	1.064^{*} (1.99)	–	0.578 (1.01)	–
Latin	–	–	-0.485^{**} (−2.10)	–
Asia	–	–	−0.251 (1.00)	–
Adj.R^2	0.68	0.68	0.72	0.67
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

2.4.2 Robustness Analysis

Previous work in this area has employed a different measure of income inequality such as income share or a Gini index from other sources than Deininger and Squire (1996). To cite a few examples, Persson and Tabellini (1994) use the income share data taken from Paukert (1973), Alesina and Rodrik (1994) utilize the Gini data from Fields (1989), and Perotti (1996) employs the income share data from Lecaillon et al. (1984). Therefore, it could be argued that our benchmark results in Table 2.1 are driven by either a specific source or measure of income inequality that we employ.

As far the specific data source is concerned, we could argue that data taken from other sources are of inferior quality because they do not meet the quality standard imposed by Deininger and Squire (1996). As far as the specific measure of income inequality data is concerned, we entertain this objection by replacing our measure of inequality, the Gini index, with the income share measure. Following Persson and Tabellini (1994) and Perotti (1996), we utilize data on the income share of the third quintile of population, Mid. However, we take this data from Deininger and Squire. Furthermore, unlike Gini, Mid is a measure of income equality: hence, we expect the coefficient of Mid to be positive in the first-stage regression.

With this change, we repeat our estimation and report the results in Table 2.4. (Note that the number of observations has reduced to 39 with this alternative measure of income inequality.) In Columns (1a) and (1b), we observe that the coefficients of Mid, Poor, and Educ enter with the anticipated signs and significant. It is interesting to note that the coefficients of Invest and $(n + g + \delta)$ enter with the correct signs too although they remain insignificant. In Columns (2a) and (2b), where regional dummies are included, we see that the coefficients of Mid, Poor, Educ, and Invest continue to be of the correct signs and significant (except for Invest). However, the coefficient of $(n + g + \delta)$ enters with the wrong sign now although it continues to be insignificant. As before, the presence of other regional dummies renders the coefficient of Africa insignificant. Compared to the corresponding coefficients in Columns (1a) and (1b), these coefficients are fairly robust except for Mid. Compared to the corresponding coefficients in Table 2.3 (excluding Mid), we see that there is a fairly huge change in a) the significance of Invest and Latin in Column (2a), b) the sign of $(n + g + \delta)$ in Column (1a), and c) the magnitude of $(n + g + \delta)$ in Column (2a). However, there is little change in the coefficients of interest. Therefore, we take these results as evidence in favor of the Galor-Zeira model.

Previous empirical studies usually employ the growth rate of per capita income as the dependent variable. When this is the case, it is instructive to add a measure of initial income per capita, Inc70 in our case, as another explanatory variable. Hence, it could be argued that our results are driven by these two differences. To accommodate this argument, we reestimate our model in the growth regression form. Note, however, that Poor is highly correlated with Inc70, with a simple correlation coefficient of 0.76. This is hardly surprising since the dummy variable Poor is derived from the continuous variable Inc70. In order to avoid the problem of omitted variable bias, however, we include both of them anyway.

Table 2.4: Estimation with *Mid* in Lieu of *Gini*

Dep.Variable	*Inc2000* (1a)	*Educ* (1b)	*Inc2000* (2a)	*Educ* (2b)
Constant	5.078^{**} (2.41)	-0.702 (-0.62)	7.121^{***} (3.13)	-0.533 (-0.41)
Mid	–	7.460^{***} (3.40)	–	6.059^{**} (2.48)
Poor	–	-0.824^{***} (-4.98)	–	-0.932^{***} (-4.87)
Educ	1.851^{***} (5.62)	–	1.774^{***} (5.45)	–
Invest	0.233 (1.34)	–	0.243 (1.43)	–
$(n+g+\delta)$	-0.073 (-0.08)	–	0.553 (0.62)	–
Africa	0.911^{*} (1.86)	–	0.543 (1.00)	–
Latin	–	–	-0.386 (-1.63)	–
Asia	–	–	-0.316 (-1.29)	–
Adj.R^2	0.74	0.72	0.76	0.72
Obs.	39	39	39	39

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

With these modifications, we repeat our estimation and document the results in Table 2.5. Columns (1a) and (1b) show that the coefficients of Gini, Poor, Educ, and Invest enter with the expected signs and significant. It is interesting to note that the coefficient of Inc70 is of the correct negative sign and significant; this last result is usually taken as evidence of conditional convergence. Compared to the corresponding coefficients in the baseline estimation (Table

2.3), we observe a fairly appreciable change in the magnitude of Poor in Column (1b), and the significance of Africa in Column (1a).

Table 2.5: Estimation with *Growth* in Lieu of *Inc2000*

Dep.Variable	*Growth* (1a)	*Educ* (1b)	*Growth* (2a)	*Educ* (2b)
Constant	4.477^{*} (1.73)	0.799 (0.48)	3.898 (1.16)	-0.696 (-0.36)
Gini	–	-0.558^{**} (-2.67)	–	-0.602^{**} (-2.42)
Poor	–	-0.473^{**} (-2.58)	–	-0.453^{**} (-2.46)
Educ	1.640^{***} (2.75)	–	1.312^{**} (2.49)	–
Invest	0.363^{**} (2.27)	–	0.314^{**} (2.22)	–
$(n+g+\delta)$	0.344 (0.42)	–	0.328 (0.38)	–
Inc70	-0.790^{***} (-3.43)	–	-0.637^{**} (-2.47)	–
Africa	0.879 (1.48)	–	0.518 (1.07)	–
Latin	–	–	-0.304 (-1.36)	–
Asia	–	–	0.095 (0.32)	–
Adj.R^2	0.14	0.74	0.35	0.75
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

As before, we add regional dummies to the basic specification. In the second-stage regression, Column (2a), the coefficient of Latin is negative but insignificant, the coefficients of Asia and Africa are positive yet insignificant, and the coefficients of other variables remain fairly robust to the corresponding coefficients in Column (1a). In the first-stage regression, Column (2b), the coefficients of Gini and Poor remain robust with respect to their counterparts in Column (1b). Compared to the corresponding coefficients in Table 2.3, we see that there is a fairly huge change in the magnitude of Poor and $(n+g+\delta)$ Columns (2b) and (2a), respectively. Note that the coefficient of Latin has become insignificant and the sign of Asia has reversed. Nevertheless, since changes in the dependent and explanatory variables do not alter our central results, we interpret these findings as further evidence in favor of the Galor-Zeira model.

Perotti (1996) argues that, in the credit imperfection model such as Galor and Zeira (1993), the human capital variable should be measured in terms of a flow rather than a stock because that is what is implied by the model. It follows then that the average-years-of-schooling variable that we have employed up to this point should be replaced by the school-enrollment-ratio variable. By definition, however, the flow and stock of human capital should be correlated with each other. This means that if income inequality affects the flow of human capital, then it should also affect the stock of human capital as well. In fact, for the sample period of 1970-2000 that we utilize, we find that the correlation between them is 0.87.

To verify our claim, we substitute the school enrollment ratio, Educf, for the average years of schooling, Educ, that we employed earlier. Educf is taken from Bernanke and Gurkaynak (2001) and is defined as the average secondary school enrollment ratio for the period 1970-1995. As in MRW (1992), the secondary school enrolment is defined as the product of the fraction of secondary-school-age population in school and the fraction of working-age-population that is of secondary school age. Even though Educf is available for more than 90 countries, the number of observations is reduced to 41 when this variable is matched with the Gini variable.

With this slight change, we repeat our estimation and report the results in Table 2.6. In Columns (1a) and (1b), we observe that the coefficients of Poor and Educf enter with the correct signs and significant. While the coefficients of Gini, Invest, and $(n + g + \delta)$ enter with the expected signs, they are insignificant. Compared to the corresponding coefficients in Table 2.3, we see that there is a substantial change in a) the significance of Gini and Invest in Columns (1b) and (1a), respectively, b) the magnitude and sign of $(n + g + \delta)$ in Column (1a), and c) the magnitude and significance of Africa in Column (1a).

When we add regional dummies to the second-stage regression in Column (2a), their coefficients are insignificant except for Latin, and the coefficient of $(n + g + \delta)$ changes its sign. However, the coefficients of Educf and Invest remain robust. In the first-stage regression, Column (2b), the coefficients of Gini and Poor remain robust with respect to their counterparts in Column (1b). Compared to the corresponding coefficients in Table 2.3, we find that there is a huge change in a) the magnitude and significance of Gini in Column (2b), b) the significance of Invest in Column (2a), and c) the magnitude of Educf, $(n + g + \delta)$, and Africa in Column (2a). Since Gini is insignificant in both specifications, we take these findings as evidence against the Galor-Zeira model.

Table 2.6: Estimation with *Educf* in Lieu of *Educ*

Dep.Variable	*Inc2000* (1a)	*Educf* (1b)	*Inc2000* (2a)	*Educf* (2b)
Constant	11.962^{**} (2.22)	-3.927^{*} (−1.73)	17.553^{***} (2.77)	-4.533^{*} (−1.79)
Gini	–	−0.270 (−0.95)	–	−0.293 (−0.86)
Poor	–	-0.624^{***} (−3.48)	–	-0.666^{***} (−3.24)
Educf	2.360^{***} (3.41)	–	2.410^{***} (3.34)	–
Invest	0.165 (0.64)	–	0.159 (0.59)	–
$(n+g+\delta)$	−0.927 (−0.68)	–	0.934 (0.54)	–
Africa	0.541 (0.81)	–	0.040 (0.05)	–
Latin	–	–	-0.770^{*} (−1.90)	–
Asia	–	–	−0.440 (−1.14)	–
Adj.R^2	0.46	0.52	0.47	0.50
Obs.	41	41	41	41

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

One may argue that our benchmark results are subject to sample selection bias since they are based on a sample of 46 countries only. After all, there is a severe under representation of African countries (only two countries are included). As discussed before, this problem is due to the unavailability of the Gini data across a large sample of countries. Barro (2000) gets around this problem by including some Gini data dubbed as low quality by Deininger and Squire (1996) at the risk of exacerbating measurement error. In particular, he adds the Gini data with dubious documentation as long as the data meet the three criteria set by Deininger and Squire. In contrast, Deininger and Squire filter out these dubious data before applying the three criteria.

As a compromise between sample selection bias and measurement error, we curtail the sample period to 1980–2000 and end up with 61 countries. We replicate our benchmark estimation in Table 2.3 with this new sample and present the results in Table 2.7. In Columns (1a) and (1b), we see that the coefficients of Poor80, Educ, and Invest enter with the expected signs and significant. Although the coefficients of Gini80 and $(n+g+\delta)$ have the expected signs, they are insignificant. Compared to the corresponding coefficients in Table 2.3, we find that there is a noticeable change in a) the magnitude and significance of Gini and Africa in Columns (1b) and (1a), respectively, and b) the magnitude and sign of $(n+g+\delta)$ in Column (1a).

When we add regional dummies to the second-stage regression in Column (2a), their coefficients are negative (except for Africa) and insignificant while the coefficients of other variables remain robust with respect to their counterparts in Column (1a). In the first-stage regression, Column (2b), the coefficients of Gini80 and Poor80 remain robust with respect to the corresponding coefficients in Column (1b). Compared to the corresponding coefficients in Table 2.3, we find that there is an appreciable change in a) the magnitude and significance of Gini in Column (2b), b) the significance of Invest and Latin in Column (2a), and c) the magnitude and sign of $(n+g+\delta)$ in Column (2a). We take the insignificance of Gini80 as evidence against the Galor-Zeira model.

One may justifiably argue that these different results might be driven by a different sample size (46 versus 61) rather than a different sample period. To control for this effect, we repeat our estimation in Table 2.7 by reducing the sample size to 46. It turns out that controlling for the sample size does change the results materially (see Table B1 in Appendix B). First, the coefficient of Gini80 becomes significant in Columns (1b) and (2b). Second, the coefficient of Invest becomes significant in Columns (1a) and (2a). Third, the magnitude of $(n + g + \delta)$ falls in Column (1a) and its sign changes in Column (2a). We conclude that differences in the magnitude and significance of coefficients in Tables 2.3 and 2.7 are driven by the difference in sample size.

Table 2.7: Estimation with 1980-2000 Sample

Dep.Variable	*Inc2000* (1a)	*Educ* (1b)	*Inc2000* (2a)	*Educ* (2b)
Constant	3.187 (1.63)	-0.271 (-0.16)	3.711^{*} (1.68)	-0.143 (-0.08)
Gini80	–	-0.370 (-1.62)	–	-0.322 (-1.21)
Poor80	–	-0.649^{***} (-4.70)	–	-0.612^{***} (-3.58)
Educ	1.557^{***} (4.13)	–	1.458^{***} (3.05)	–
Invest	0.507^{*} (1.78)	–	0.560 (1.48)	–
$(n + g + \delta)$	-0.620 (-0.83)	–	-0.469 (-0.60)	–
Africa	0.324 (1.22)	–	0.190 (0.55)	–
Latin	–	–	-0.111 (-0.50)	–
Asia	–	–	-0.155 (-0.58)	–
Adj.R^2	0.73	0.69	0.74	0.68
Obs.	61	61	61	61

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Hitherto the results seem to indicate that a larger sample size in our baseline estimation is unfavorable to the Galor-Zeira model. To pursue this line of thought further, we expand the sample size substantially by replacing our cross-section data with panel data. Following the literature, we utilize the five-year interval panel data, where Gini and Poor are measured at 1970, 1975, ..., 1995, Invest and $(n + g + \delta)$ are measured as averages of 1971-1975, 1976-1980, ..., 1996-2000, and Income and Educ are measured at 1975, 1980, ..., 2000. Including only those data for which there are at least two consecutive observations, we end up with an unbalanced panel of 53 countries and 226 observations.

With this expanded sample size, we reestimate our model by the pooled IV method. As before, we start with the first-stage regression without Latin and Asia. As shown in Column (1b) of Table 2.8(a), the coefficients of Gini and Poor enter with the correct signs and are significant. Since they are jointly significant, we proceed with the second-stage regression. In Column (1a), we see that the coefficients of Educ and Invest enter with the correct signs and significant. However, the coefficient of $(n + g + \delta)$ enters with the wrong sign and is insignificant. As before, we compare these coefficients to their counterparts in Table 2.3. Except for changes in the magnitude of Poor and $(n + g + \delta)$, the estimates are fairly robust.

Table 2.8(a): Estimation with Panel Data

Dep.Variable	*Inc2000* (1a)	*Educ* (1b)	*Inc2000* (2a)	*Educ* (2b)
Constant	5.044^{***} (3.84)	0.496 (0.47)	5.704^{***} (5.15)	0.757 (0.68)
Gini	–	-0.508^{***} (−3.65)	–	-0.490^{***} (−2.90)
Poor	–	-0.494^{***} (−5.77)	–	-0.451^{***} (−4.51)
Educ	1.888^{***} (6.66)	–	1.388^{***} (5.27)	–
Invest	0.598^{***} (4.18)	–	0.789^{***} (6.37)	–
$(n + g + \delta)$	0.449 (0.81)	–	0.489 (1.09)	–
Africa	0.860^{***} (4.25)	–	0.391^{*} (1.85)	–
Latin	–	–	−0.186 (−1.61)	–
Asia	–	–	-0.499^{***} (−4.25)	–
Adj.R^2	0.59	0.56	0.75	0.56
Obs.	226	226	226	226

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Now we proceed with the first-stage regression with regional dummies. As shown in Column (2b), the results on Gini and Poor are highly similar to those in Column (1b). In the second-stage regression in Column (2a), the coefficients of Educ and Invest continue to be significant, enter with the correct signs, and the change in their magnitudes is within one standard deviation. Compared to the corresponding coefficients in Table 2.3, we see that there is a significant change in a) the magnitude of Gini, Poor, and $(n + g + \delta)$, and b) the significance of all regional dummies. Once again, the difference in our results are driven by the difference in sample size. As a matter of fact, since the results in the baseline estimation appear to be restored, we reject the idea that a larger sample size is unfavorable to the Galor-Zeira model. In this case, since the results on key variables remain intact, we take these findings as further support for the Galor-Zeira model.

Kristin Forbes (2000) argues that the negative impact of income inequality on growth could be attributed to omitted variable bias due to factors that are not adequately captured by regional dummy variables. To overcome this deficiency, she estimates the growth regression using the fixed-effect (FE) method of panel data. She finds that inequality actually has a positive effect on growth, thereby discrediting the inequality-growth model in general (similar results are obtained by Li and Zou (1998)). However, these findings are based on the estimation of a reduced-form equation. It would be interesting to see whether the use of a FE method might reject a structural equation as well, which is consistent with the Galor-Zeira model.

We begin by conducting the first-stage regression of the FE model. Since the FE model controls for any time-invariant country-specific characteristics, we drop any dummy variables that we have used so far (Poor, Latin, Asia, and Africa) since they will be captured by individual country effects. Since Poor is dropped out, only Gini appears in the first-stage regression. As shown in Column (1b) of Table 2.8(b), the coefficient of Gini is insignificant although it enters with the correct sign. The insignificance of Gini precludes us from conducting the second-stage regression, and therefore prevents us from rejecting Forbes' findings.[7]

[7] In practice, we could have proceeded by estimating the first-stage regression of a random effect (RE) model. In principle, however, the use of a RE method is unwarranted for two reasons. First, the use of a panel IV estimation based on a RE method is inappropriate if *Educ* is correlated with individual country effects. Second, the unwarranted

Table 2.8(b): Estimation with Panel Data

Dep.Variable	*Educ* (1b)	*Inc2000* (2a)	*Educ* (2b)
Constant	-1.722^{**} (-2.15)	4.925^{**} (2.36)	3.521 (1.16)
Gini	-0.126 (-0.76)	–	-0.964^{**} (-2.18)
Poor	–	–	-0.322 (-1.41)
Educ	–	0.837^{*} (1.75)	–
Invest	–	1.213^{***} (5.35)	–
$(n+g+\delta)$	–	0.277 (0.32)	–
Africa	–	0.249 (0.73)	–
Latin	–	-0.329 (-1.56)	–
Asia	–	-0.678^{***} (-3.10)	–
Adj.R^2	0.19	0.81	0.57
Obs.	226	226	226

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Hauk and Wacziarg (2004) conduct a Monte Carlo study of the augmented Solow growth model using a few alternative estimation techniques. They find that the between effect (BE) estimator outperforms other estimators in terms of the degree of bias of the estimated coefficients. Based on this consideration, we reestimate the first-stage regression using a BE method. As shown in Column (2b) of Table 2.8(b), the coefficients of Gini and Poor enter with the correct signs; however, only the coefficient of Gini is significant. Since both coefficients are jointly significant, we proceed with the second-stage regression. The results in Column (2a) indicate that the coefficients of Educ and Invest enter with the anticipated signs and significant. We take these results as evidence in favor of the Galor-Zeira model.

The fact that our dependent variable, Inc2000, is measured in 2000, while some of our explanatory variables (Invest and n) are measured as averages over the period 1970-2000 may

use of a panel IV estimation based on a FE method precludes us from evaluating the appropriateness of a RE method.

make us susceptible to simultaneity bias (i.e., the direction of causality may run from these variables to Inc2000 instead).[8] In the growth empirics, this endogeneity issue is usually taken care of by instrumenting the relevant regressors (Invest and n in our context) with their lagged values (see Barro and Sala-i-Martin (2004)). If the baseline results are sensitive to these changes, this is taken as evidence that Invest and n are endogenous.

Before we do that, however, it is imperative that we test the endogeneity of Invest and n using a Hausman test. First, we estimate the second-stage regression with and without instrumenting Invest and n with their respective lagged values, which are measured as averages over the period 1965-1995. Second, we test whether the difference between estimates obtained from the regression with and without instrumentating Invest and n is statistically significant. (Note that Educ is always instrumented by Gini and Poor by theoretical implication.) Unfortunately, the Hausman test fails to deliver any results in the basic specification because the variance-covariance matrix is not positive definite. In the augmented specification, the test fails to deliver useful results because the test statistic is negative.

As an alternative to the Hausman test, we adopt an auxiliary regression approach.[9] This method can be summarized in the following steps. First, we run the first-stage regression for each Educ, Invest, and $(n + g + \delta)$. (Note that the first-stage regression needs to conducted for Educ as well.) Second, we extract residuals obtained from each first-stage regression. Third, we run the second-stage regression with the inclusion of these residuals using the method of ordinary least squares (OLS). Finally, we test whether the estimated coefficients from the residuals are jointly significant. Performing all of these steps, we find that the estimates are jointly significant. We take this result as evidence that Educ, Invest, and $(n + g + \delta)$ are endogenous.

---

8 Although *Educ* data is also an average of the period 1970-2000, this should not pose any simultaneity problem because it is instrumented by *Gini* and *Poor* that are measured at the beginning of the period of analysis.

9 I am grateful to Dr. Carter Hill for suggesting this method as an alternative to the Hausman test; see Baum, Schaffer, and Stillman (2003) for reference.

Given the above results, we repeat our baseline estimation by instrumenting Invest and n with their respective lagged values and report the estimation results in Table 2.9. In Columns (1a) and (1b), we find that the coefficients of all variables enter with the anticipated signs and, except for $(n + g + \delta)$, are significant. These results continue to hold in Columns (2a) and (2b), where we add all regional dummies to the basic specification. Compared to the corresponding coefficients in Table 2.3, we see that the results are fairly robust. (The only sensitive coefficients are Educ in Column (2a) and $(n + g + \delta)$ in Columns (1a) and (2a).) We take these results as evidence that our baseline results are not affected much by the endogeneity of Invest and n.

Table 2.9: Estimation with Instrumented *Invest* and n

Dep.Variable	*Inc2000*	*Educ*	*Inc2000*	*Educ*
	(1a)	(1b)	(2a)	(2b)
Constant	2.681 (1.06)	1.336 (0.70)	4.844^{*} (1.90)	2.007 (0.96)
Gini	–	-0.622^{**} (-2.61)	–	-0.754^{**} (-2.67)
Poor	–	-0.684^{***} (-3.61)	–	-0.586^{**} (-2.69)
Educ	1.498^{***} (3.27)	–	1.278^{**} (2.67)	–
Invest	0.765^{**} (2.65)	–	0.802^{**} (2.66)	–
$(n + g + \delta)$	-0.624 (-0.63)	–	-0.030 (-0.03)	–
Africa	0.998^{*} (1.92)	–	0.448 (0.77)	–
Latin	–	–	-0.463^{*} (-1.99)	–
Asia	–	–	-0.355 (-1.31)	–
Adj.R^2	0.70	0.68	0.73	0.67
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

According to the Galor-Zeira model, the impact of inequality on long-run income per capita works through a sequence of financial constraints and education channel. That is, income inequality imposes financial constraints for people to invest in education, and education in turn affects long-run income. If so, then it stands to reason that the impact can be undermined if the financial constraints can somehow be relaxed. One way of doing this is for the government to provide an education subsidy for the poor. A priori, we expect that the presence of this subsidy would yield one of the following: either inequality or the subsidy itself becomes less significant and inequality declines in magnitude.

To test this idea, the value of government expenditures on education, Goveduc, is used as a proxy for the education subsidy variable.[10] Goveduc is admittedly an imperfect proxy for an education subsidy variable because it is available to poor and non-poor alike. In the absence of a better alternative, this variable is used anyway. We begin by entering Goveduc into, and excluding Gini from, the first-stage regression (i.e., Eq.(2.7)). In Column 1 of Table 2.10, the coefficient of Goveduc enters with the correct positive sign and is significant. In Column 2, where both Goveduc and Gini are included, the coefficient of the former falls in magnitude and significance while that of the latter is negative and significant. In both specifications, the coefficient of Educ is of the expected sign and significant (note also that the coefficient of Gini is robust to its counterpart in Table 2.3). We take these results as evidence in support of the Galor-Zeira model.

According to the Galor-Zeira model, the impact of income distribution on long-run income per capita is not applicable to an initially poor society because such a society will still be financially constrained even if income is evenly distributed. If this is the case, then it follows that there exists a threshold level of initial income above which inequality matters and below which it does not. This can be done by relaxing the definition of poor countries, Poor, to include middle-income countries as well. Renaming our previous measure of Poor as Poor1, we include two additional categories of poor countries: a) Poor2, which is defined as the sum of Poor1 and lower-middle-income countries and b) Poor3, which is the sum of Poor2 and upper-middle-income countries.[11] A priori,

[10] More specifically, *Goveduc* is measured as the average government expenditures as a fraction of GDP for 1970-1984. The data is taken from Barro and Lee (1994).

[11] The classification of countries into lower- and upper-middle-income countries is made by the World Bank based on the income range of these countries.

we expect that the magnitude and significance of Gini would decline as we vary the definition of Poor from Poor1 to Poor2, and from Poor2 to Poor3 because this entails the relaxation of financial constraints. As shown in Columns (3)–(5) of Table 2.10, the estimation results strongly support this idea. We take these results as further support of the Galor-Zeira model.

Table 2.10: Estimation with Government Education Expenditures, Alternative Measures of Poor, and Interactive Term

Specification	*Goveduc* (1)	*Goveduc* (2)	*Poor1* (3)	*Poor2* (4)	*Poor3* (5)	*Gini×Poor* (6)
Gini	–	-0.799^{**} (−2.39)	-0.999^{***} (−3.57)	-0.860^{**} (−2.68)	-0.657^{*} (1.87)	-0.932^{***} (−3.03)
Goveduc	8.695^{**} (2.31)	6.314^{*} (1.72)	–	–	–	–
Gini×Poor	–	–	–	–	–	-0.457 (−0.56)
Educ	1.027^{***} (2.83)	0.988^{***} (3.35)	1.610^{***} (5.64)	1.905^{***} (4.62)	1.217^{***} (3.27)	1.627^{***} (5.71)
Obs.	45	45	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS. The coefficient of Educ comes from the second-stage regression, where Inc2000 is the dependent variable. The coefficients of other variables come from the first-stage regression, where Educ is the dependent variable. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Regional dummies are always included.

Another way to test the above hypothesis is by including an interactive term between Gini and Poor in the first-stage regression, Gini×Poor. If the interactive term is significant, this implies that the impact of Gini on Educ is different between poor and non-poor countries. Adding this interactive term, we reestimate the first-stage regression corresponding to Column (2b) in Table 2.3. As shown in Column (6) of Table 2.10, however, the interactive term is insignificant. It should be noted, however, that this poor result is driven by the high correlation between Poor and Gini×Poor (their correlation coefficient is 0.99) because the coefficient of Poor enters with the incorrect sign and insignificant. Again, the coefficients of Gini and Educ are robust to their counterparts in Table 2.3.

As an additional robustness check, we test whether the coefficients of Gini and Educ are sensitive to the inclusion of other additional control variables. In practice, this is hard to implement because different papers employ different control variables in their robustness tests. We employ three criteria in choosing these additional control variables. That is, the variables must not be implied by other theoretical models in this area. Second, the variables should be sensible. Finally, the variables should be frequently employed. (The most frequent ones are regional dummies; for this reason, they are always included in each of our augmented specifications.)

Applying these criteria, we obtain six control variables: government consumption expenditures (Govcons), government consumption expenditures net of education spending (Govconse), government expenditures on defense (Govdef), the price of investment (PPI), life expectancy (Life), and the fertility rate (Fert). Each of these control variables is added one at a time in our estimation.[12] As shown in Table 2.11, the coefficient of Gini is robust to the inclusion of these additional variables when it is compared against its counterpart in Table 2.3. However, this is not the case with Educ, whose magnitude deviates substantially from its counterpart in Table 2.3. Note also that the magnitudes of Gini and Educ decline considerably (and Educ becomes insignificant) when Life and Fert are added to the regression; see Columns (5) and (6). These results suggest that Gini may affect Inc2000 through Life and Fert channels, too. Since the coefficients of Gini and Educ enter with the correct signs and significant, we take these results as further support for the Galor-Zeira model.

Previous empirical studies that analyze the inequality-growth relationship express the Gini variable in levels (as opposed to logs). To test whether our results are robust to this variable definition, we reestimate our baseline estimation in Table 2.3 with the Gini index expressed in levels, Ginilev. As shown in Table B2, the coefficients of all variables retain their signs and significance. Hence, our baseline estimation is robust to the way Gini is expressed.

It has been suggested that a measure of the wage premium between skilled and unskilled workers, defined as the ratio of skilled-worker wages to unskilled-worker wages, be included in

[12] Each of these variables is measured as the average for 1970-1984 or 1970-1985, depending on the availability of data. All data are taken from Barro and Lee (1994).

the first-stage regression in order to control for the incentive in acquiring education. We take this variable from Caselli and Coleman (2006), who construct the variable based on some functional relationship between wages and human capital data. A priori, we expect the coefficient of this new variable, Prem, to be positive. Adding Prem to the specification results in the number of observations drops to 33.

Table 2.11: Estimation with Alternative Control Variables

Specification	*Govcons* (1)	*Govconse* (2)	*Govdef* (3)	*PPI* (4)	*Life* (5)	*Fert* (6)
Gini	-0.951^{***} (-2.84)	-0.943^{***} (-2.82)	-0.906^{***} (-2.81)	-0.835^{**} (-2.68)	-0.601^{**} (-2.22)	-0.618^{**} (-2.09)
Educ	1.159^{***} (3.40)	1.126^{***} (3.40)	1.125^{***} (3.86)	0.978^{***} (3.06)	0.700 (1.39)	0.397 (0.98)
Control	-2.294 (-1.32)	-2.659 (-1.20)	1.077 (0.30)	0.507^{**} (2.52)	0.333 (1.02)	-0.411^{***} (-2.97)
Obs.	45	45	45	46	46	43

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS. The coefficient of Gini comes from the first-stage regression, where Educ is the dependent variable while the coefficients of Educ and additional control variables from the second-stage regression, where Inc2000 is the dependent variable. t-statistics are given in parentheses. t-statistics in Columns (3) and (5) are based on White's heteroskedasticity-consistent standard error. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. Regional dummies are always included.

We repeat our baseline estimation and show our results in Table B3. In the two versions of the first-stage regression (with Africa and with all regional dummies), we find the following: a) the coefficients of Gini and Poor enter with the correct signs (but the coefficient of Gini is insignificant in Column (2b)), and b) the coefficient of Prem enters with the wrong sign in both columns but is significant in Column (1b) only. In both cases, however, these variables are jointly significant; so we proceed with the second-stage regression. In Columns (1a) and (2a), we find that a) the coefficients of Educ and Invest enter with the expected signs and significant, and b) the coefficient of $(n + g + \delta)$ is insignificant. Since the results on key variables remain intact, we interpret these findings in favor of the Galor-Zeira model .

Manuelli and Seshadri (2005) argue that the average-years-of-education variable that is usually employed in the growth empirics (and in this paper too) merely captures the quantity of human capital only. To capture the quality of education as well, they suggest the inclusion of education expenditures as a fraction of GDP as a proxy. Since this variable varies considerably across countries, it is imperative that this variable be incorporated into our analysis. The inclusion of this variable is left for future work.

Finally, one may argue that, given our small sample size, the asymptotic inference of IV estimation should be complemented with the inference based on bootstrap standard errors. As Mackinnon (2002) puts it, however, bootstrap standard errors perform poorly in models with endogenous regressors.

In addition to estimating a structural model, we also estimate a reduced-form version of the Galor-Zeira model where Eq.(2.7) is substituted into Eq.(2.6). The results of this reduced-form model, which are reported in Columns (1) and (2) of Table B4, indicate that inequality exerts a negative influence on long-run income per capita. We take these findings as further evidence, albeit indirect, supporting the Galor-Zeira model.

2.5 Discussion

As implied by Eqs. (2.6) and (2.7), income inequality affects income per capita through education. Since each of these variables is measured in logs, the connection between any two variables is in the form of an elasticity. Consider Columns (1a) and (1b) in Table 2.3 again. We observe that the coefficients of Gini and Educ are –0.71 and 1.94, respectively. This means that a 10% increase in inequality would decrease educational investment by 7.1% and a 10% increase in education would raise income per capita by 19.4%. Taken together, both coefficients imply that a 10% increase in inequality would reduce the level of income per capita by 13.77% (i.e., $19.4\% \times (-7.1\%/10\%) = -13.77\%$).

The above estimate is based on the impact of inequality on income through the education channel. To measure the net effect of inequality on income, we may resort to a reduced-form estimate described in Table B4. In Column (1) of Table B4, the counterpart of Columns (1a) and

(1b) in Table 2.3, we see that the coefficient of Gini is –1.462. This means that a 10% increase in inequality would decrease income per capita by 14.62%. Hence, the magnitudes of both estimates are quite similar.

Now we compare our finding on the impact of income inequality on income with the findings in previous studies. In order to make our quantitative effect of inequality on income comparable to that in previous work, we need to make a few adjustments. First, we replace our dependent variable, Inc2000, with Growth. Second, we express Gini in levels (instead of logs). Finally, we assess how much a one-standard-deviation change in Gini affects Growth. With these modifications, we reestimate our reduced-form model in Column (1) of Table B4. We find that the coefficient of Gini is –0.03 and significant (see Column (3) in Table B4). This result implies that an increase in income inequality by one standard deviation (i.e., 10 in our sample) is expected to decrease economic growth by 30%.

Persson and Tabellini (1994) employ income share of the third quintile as a measure of income equality. In Column (1) of their Table 5, they report that the coefficient of this variable is 0.189 and significant. This implies that an increase in income equality by one standard deviation (i.e., 3 in their sample) is expected to increase growth by 0.57%. Perotti (1996) employs income share of the third and fourth quintiles as a measure of income equality. In Column (1) of his Table 4, he reports that the coefficient of this variable is 0.118 and significant. This suggests that a rise in income equality by one standard deviation (i.e., 0.16 in his sample) is expected to increase growth by 0.58%. Alesina and Rodrik (1994) employ the Gini coefficient in 1960 (measured in levels) as a measure of income inequality. In Column (6) of their Table 1, they report that the coefficient of Gini is –5.23. This implies that a rise in Gini by one standard deviation (i.e., 0.16 in their sample) is expected to decrease growth by almost 0.84%. Sylwester (2000) employs the Gini coefficient in 1970 (also measured in levels) as a measure of income inequality. In Column (1) of his Table

1, he reports that the coefficient of Gini is –0.0007. This suggests that a rise in inequality by one standard deviation (i.e., 10 in his sample) is expected to decrease growth by 0.7%. Finally, Forbes (2000) employs lagged values of Gini (also measured in levels) as a measure of income inequality. In Column (4) of her Table 3, she reports that the coefficient of Gini is 0.0013. This implies that a rise in inequality by one standard deviation (i.e., 10 in her sample) is expected to increase growth by 1.3%. Taken together, these findings indicate that the impact of income inequality on economic growth takes a very close range of values, from as low as 0.57% (Persson and Tabellini's findings) to as high as 0.84% (Alesina and Rodrik's finding). Therefore, our estimate is too high.

2.6 Conclusion

In this paper, we conduct an empirical analysis to test the implications of the Galor-Zeira model based on a cross-section of 46 countries during the period 1970–2000. Our baseline analysis in Table 2.3 yields the results that income inequality exerts a negative impact on long-run per capita income, and it does so through the education channel (which has a positive impact on income). These findings continue to hold when we a) replace our dependent variable, Inc2000, with a typical variable in the growth regression, Growth, b) expand the sample size by using panel data and employ the pooled IV method, c) employ the BE method, d) add government education expenditures variable to the specification, e) vary the threshold of poor countries from low-income countries to upper-middle-income countries, f) employ alternative control variables, g) express the Gini variable in levels, and h) add wage premium variable in the baseline estimation. However, we obtain unfavorable results when we a) expand the sample size by curtailing the period of analysis to 1980–2000, b) replace a stock measure of human capital variable by a flow measure, c) employ the FE method in panel data estimation, and d) include the interactive term between Gini and Poor. Overall, the balance of evidence appears to be broadly consistent with the theoretical implications of the Galor-Zeira model.

Nevertheless, it is important to acknowledge the limitations of our study. First, our findings are based on a cross-country data that excludes many African countries. Since most of them are poor-income countries, our results may be driven by this omission. Second, the Galor-Zeira model conjectures that the link between inequality and education is bridged by financial or credit constraints. However, we have not been able to test this link due to the unavailability of data on credit constraints. It has been argued that financial constraints are likely to be reduced as countries develop their financial institutions. If so, measures of the level of financial development such as the ones considered by Levine (1997) could be used as proxies for the lack of credit constraints. The incorporation of these measures is left for future work.

Chapter 3 Empirical Analysis of the Kremer-Chen Model

3.1 Introduction

In the 1990s, there has been renewed interest in the relationship between income inequality and economic growth. Although there is a consensus among economists that there is an adverse impact of inequality on growth, there is a disagreement over the channel in which inequality affects growth. Some economists conjecture that inequality and growth are linked by the interaction between redistributive fiscal policy and physical capital investment; this is known as the political economy model. Other economists postulate that they are linked by the interaction between socio-political instability and physical capital investment; this is called the socio-political instability model. Still others argue that they are linked by the interaction between credit constraints and human capital investment; this is known as the credit constraint model. Recently, a handful of economists have conjectured yet another channel: inequality and growth are linked by the interaction between fertility and human capital investment. According to this so-called endogenous fertility model, income inequality is bad for growth because a fertility differential between poor and rich households rises with a rise in inequality. The increase in this fertility differential, in turn, results in a lower stock of human capital and income per capita.

The theoretical literature which examines the impact of income inequality on economic growth through a fertility differential includes Dahan and Tsiddon (1998), Kremer and Chen (2002), De La Croix and Doepke (2003), and Moav (2005). In Dahan and Tsiddon (1998) and De La Croix and Doepke (2003), the impact of a fertility differential on growth is transitory (i.e., the economy ends up with a single steady state in the long run.) In Kremer and Chen (2002) and Moav (2005), however, the impact of a fertility differential on growth is permanent (i.e., the economy ends up with multiple steady states in the long run). Given the burgeoning evidence on multiple steady states in the literature [see Durlauf, Johnson, and Temple (2005)], the models developed by Kremer and Chen (2002) and Moav (2005) are more realistic. Between these two, there is little to choose from. In Kremer and Chen (2002), a fertility differential is achieved through the notion that rich (as opposed to poor) households are time constrained to raise many children. In Moav (2005), a fertility differential is achieved through the premise that poor (as opposed to rich) households have a comparative disadvantage in providing education to their children.

The empirical literature which investigates the inequality-fertility-growth link includes Peretto (1996), Barro (2000), Kremer and Chen (2002), and De La Croix and Doepke (2003). In Peretto (1996) and Barro (2000), however, the analysis is carried out in terms of overall fertility (as opposed to a fertility differential). This is not surprising given the fact that their studies are motivated by early endogenous fertility models that ignore a fertility differential. Kremer and Chen (2002) conduct an empirical analysis based on their own theoretical model and find evidence that inequality has a positive effect on the fertility differential between poor and rich households. Building upon Kremer and Chen's study, De La Croix and Doepke (2003) find evidence that the fertility differential has a negative impact on economic growth. Despite these encouraging results, each of these papers focuses on one part of the inequality-fertility-growth link only. That is, Kremer and Chen focus on the link between inequality and the fertility differential only while De La Croix and Doepke focus on the link between the fertility differential and growth only.

Since none of these empirical studies investigates the impact of income inequality on growth though a fertility differential as a system, this paper attempts to fill this gap. In doing so, we recognize that a high fertility differential between poor and rich households implies a low ratio of skilled to unskilled labor. Given this connection, we conduct a cross-country empirical analysis for 46 countries during the 1970-2000 period which systematically tests the relationship between inequality, the skilled-unskilled labor ratio, and long-run income per capita as implied by the

Kremer-Chen model. In the baseline analysis, we find that income inequality has a negative impact on the skilled-unskilled labor ratio, which in turn has a positive impact on per capita income in the long run. These baseline results are broadly robust across alternative model specifications and samples. We do obtain unfavorable results in some robustness checks. However, there is evidence that these poor results are due to a reduced sample size in one case, and a multicollinerity problem in another case. Given these qualifications, we conclude that our empirical analysis produces evidence that lends support to the Kremer-Chen model.

We begin our analysis in Section 3.2 by describing the model developed by Kremer and Chen (2002). In Section 3.3, we specify our empirical model and then discuss the data employed in the paper. In Section 3.4, we report and discuss our baseline and robustness analyses. In Section 3.5, we discuss how our main results relate to those in previous studies. In Section 3.6, we conclude by highlighting the major findings of our paper.

3.2 The Kremer-Chen Model

Kremer and Chen (2002) introduce a representative agent model of the economy with endogenous fertility decisions. (A shortened version of Kremer and Chen (2002) circulates under Kremer and Chen (1999).) That is, households make a conscious decision on the optimal number of children that they wish to have. This optimal decision hinges on the trade-off that households face between the quantity and quality of children that they wish to have. This trade-off arises from the total cost of raising children, which consists of direct cost (food, clothing, and education) and indirect cost (the opportunity cost of raising children). As household incomes rise, the direct cost of childrearing (as a fraction of the total cost of childrearing) becomes less important; thus, education (which is part of the direct cost) rises with income. As their incomes rise, however, the indirect or opportunity cost of childrearing becomes more prominent; hence, fertility declines with income. As a result, rich people tend to have few yet more educated children and poor people tend to have many yet less educated children. It follows, then, that the higher the fertility and education differentials are, the smaller the stock of human capital and the lower the level of per capita income in the future.

To formalize the model, Kremer and Chen make the following assumptions. First, optimal fertility is a decreasing function of wages. (This assumption is actually a result of other assumptions; see Appendix C.) This is due to the substitution effect of a wage increase: the opportunity cost of childrearing rises as household income increases. However, this inverse relationship is most likely applicable to middle-income households only. For very poor households, a rise in wages is probably going to generate an income effect as the direct cost of childrearing as a fraction the total cost falls with a rise in income. For very rich households, a further rise in wages is unlikely going to reduce fertility anymore. There must be a threshold level of income below and above which fertility is not affected by changes in wages. For these reasons, Kremer and Chen postulate that a fertility differential between poor and rich households is most likely applicable to middle-income countries.

Second, educational decisions depend on the incentives provided by wage premium; i.e, the ratio of wages earned by skilled workers to wages earned by unskilled workers, w_s/w_u. Third, children of unskilled (poor) parents face higher costs of education, c_H, than children of skilled (rich) parents, c_L. More precisely, they assume that all children of skilled parents and a fraction θ of children of unskilled parents face c_L costs of education while the remaining fraction $(1-\theta)$ of children of unskilled parents face c_H costs of education.(These varying costs of education among children of unskilled workers are essential in generating multiple steady states.) Suppose that each individual is endowed with a total working period which is normalized to unity, 1. Letting the costs of education be measured in units of time, then the total working period available to those who obtain education is equal to $(1-c_i)$ for $i = s, u$. Based on the first assumption, we deduce the following:

$$(1 - c_i)\, w_s \geq w_u. \tag{3.1}$$

That is, each individual will a) choose to acquire education if the lifetime income from obtaining education is greater than the lifetime income from not obtaining education, and b) be indifferent about acquiring education if the lifetime income from obtaining education is equal to the lifetime income from not obtaining education. Eq.(3.1) can be equivalently expressed as

$$\frac{w_s}{w_u} \geq \frac{1}{1 - c_i}. \tag{3.2}$$

That is, each individual will a) invest in education if the wage premium between skilled and unskilled workers is greater than the ratio of working period available to unskilled workers to working period available to skilled workers, and b) be indifferent about investing in education if the wage premium is equal to the ratio of the working periods. Note that c_i contains two parts, c_L and c_H. Therefore, $1/(1 - c_i)$ contains two parts too, $1/(1 - c_L)$ and $1/(1 - c_H)$. Since $c_L < c_H$, then $1/(1 - c_L) < 1/(1 - c_H)$. Accordingly, Eq.(3.2) can be expanded to be

$$\frac{1}{1 - c_L} \leq \frac{w_s}{w_u} \leq \frac{1}{1 - c_H}. \tag{3.3}$$

From Eq.(3.3), we can deduce the following: First, if $w_s/w_u = 1/(1 - c_L)$, then low-cost individuals will be indifferent about obtaining education but high-cost individuals will not obtain education. Second, if $w_s/w_u = 1/(1 - c_H)$, then low-cost individuals will obtain education but high-cost individuals will be indifferent. Third, if $w_s/w_u \in (1/(1 - c_L), 1/(1 - c_H))$, then low-cost individuals will obtain education but high-cost individuals will not.

Let L_s be the number of skilled workers, L_u the number of unskilled workers, and R the ratio of skilled to unskilled workers, $R = L_s/L_u$. Then, it can be shown that the wage premium is equal to the inverse of the skilled-unskilled worker ratio, $w_s/w_u = 1/R$ (see Appendix C). Given this relationship, the above deductions can be stated in terms of R as follows. First, if $R = 1 - c_L$, then low-cost individuals will be indifferent about obtaining education but high-cost individuals will not obtain education. Second, if $R = 1 - c_H$, then low-cost individuals will obtain education but high-cost individuals will be indifferent. Third, if R is between $(1 - c_H)$ and $(1 - c_L)$, then low-cost individuals will obtain education but high-cost individuals will not.

Now let us consider the dynamics of the model. Let n_{st} be the number of children born to each skilled worker and n_{ut} the number of children born to each unskilled worker. Hence, the number of children of skilled and unskilled workers can be expressed as the product of the number of the respective workers and the number of their children: $C_{st} = L_{st}n_{st}, C_{ut} = L_{ut}n_{ut}$.

Recall from the preceding discussion that unskilled workers have more children than skilled workers, $C_{st} < C_{ut}$. In order for the ratio of skilled to unskilled workers to be constant, this fertility differential has to be offset by postulating that some children of unskilled workers becoming skilled workers later on. This is accomplished by assuming that a fraction γ_t of children of unskilled workers become skilled. If we define R_t as the ratio of skilled to unskilled workers at time t (i.e., $R_t \equiv L_{st}/L_{ut}$), then the evolution of R_t can be expressed as

$$R_{t+1} = \frac{L_{s,t+1}}{L_{u,t+1}} = \frac{L_{st}n_{st} + \gamma_t L_{ut}n_{ut}}{(1 - \gamma_t)L_{ut}n_{ut}} = \frac{R_t^2 + \gamma_t}{1 - \gamma_t}. \tag{3.4}$$

Given R_{t+1}, we can deduce the following: First, if $R_{t+1} = 1 - c_L$, then the fraction γ of children of unskilled workers who become skilled will be less than or equal to the fraction θ of children of unskilled workers who incur low costs of education. This conclusion follows from the conjecture that low-cost individuals will be indifferent about obtaining education if $R = 1 - c_L$. So, some of them may actually end up not pursuing education. Second, if $R_{t+1} = 1 - c_H$, then the fraction γ of children of unskilled workers who become skilled will be more than or equal to the fraction θ of children of unskilled workers who incur low costs of education. This result follows from the postulate that high-cost individuals will be indifferent about acquiring education if $R = 1 - c_L$. So, some of them may actually end up pursuing education. Third, if $R_{t+1} \in (1 - c_H, 1 - c_L)$, then the fraction γ of children of unskilled workers who become skilled will be equal to the fraction θ of children of unskilled workers who incur low costs of education. This prediction follows from the premise that low-cost individuals will obtain education but high-cost individuals will not obtain education if $R \in (1 - c_H, 1 - c_L)$. We can summarize these conjectures as follows:

$$\begin{array}{c} \gamma_t \leq \theta \text{ if } R_{t+1} = 1 - c_L \\ \gamma_t = \theta \text{ if } R_{t+1} \in (1 - c_H, 1 - c_L) \\ \gamma_t \geq \theta \text{ if } R_{t+1} = 1 - c_H \end{array} \tag{3.5}$$

Note that, when $\gamma = \theta$, Eq.(3.4) can be rewritten as

$$R_{t+1} = \frac{R_t^2 + \theta}{1 - \theta} \equiv R_\theta. \tag{3.6}$$

Given Eqs.(3.5) and (3.6), it can be shown that (see Appendix D) R_t evolves according to

$$R_{t+1} = \left\{ \begin{array}{c} 1 - c_L \ \text{ if } R_\theta \geq 1 - c_L \\ R_\theta \text{ if } R_\theta \in (1 - c_H, 1 - c_L) \\ 1 - c_H \ \text{ if } R_\theta \leq 1 - c_H \end{array} \right\} \tag{3.7}$$

Eq.(3.7) can be alternatively represented by a nonlinear curve as depicted in Figure 3.1. The figure shows that there are two stable steady states (marked by points A and C) and one unstable steady state (marked by point B). The critical point is point B because the long-run ratio of skilled to unskilled labor, R^*, is positively related to the fraction of the initial ratio of skilled to unskilled labor, R_0, that exceeds this point: the larger the fraction of R_0 that exceeds this point, the higher the R^*, and vice versa. The higher the R^*, in turn, implies the larger stock of human capital and the higher the level of per capita income. To the extent that unskilled labor can be identified with poor households, we reach the following conclusion: an economy characterized by an initially high income inequality results in a low ratio of skilled-unskilled labor and a low standard of living.

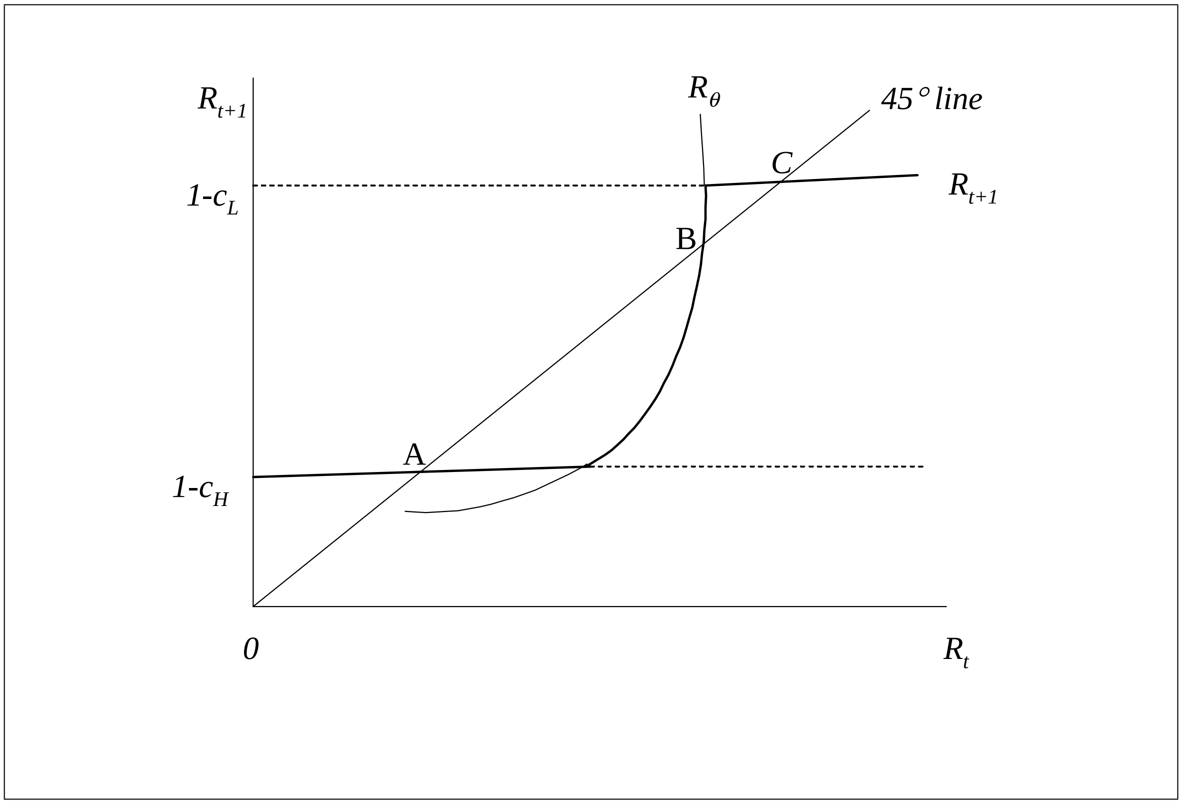

Figure 3.1: The Dynamics of Kremer-Chen Model

3.3 Model Specification and Data

We begin by specifying our model based on the description of the Kremer-Chen model. We proceed by comparing our specification with the related ones in the literature. We conclude by describing the data used in this paper based on our model specification.

3.3.1 Model Specification

Based on the conclusions of the Kremer-Chen model, we conjecture that a country has a great prospect to improve itself in the future if its income is more equally distributed among its citizens. Thus, a priori we expect that a measure of income inequality will have an adverse impact on a country's future income. Since this implication is arguably to be most likely applicable to middle-income countries, we need to qualify our analysis by introducing a dummy variable for each poor and rich country. Since income inequality affects future income through the ratio of skilled to unskilled labor, we need to estimate a structural model consisting of two equations. In the first equation, income is a function of the skilled-unskilled labor ratio and other explanatory variables in the Solow growth model. In the second equation, the ratio of skilled to unskilled labor is a function of income inequality, a dummy variable for poor countries, and a dummy variable for rich countries. Our structural model looks as follows:

$$\begin{aligned} \textit{Income} &= \alpha_1 + \alpha_2.\,(L_s/L_u) + \alpha_3.\textit{Invest} + \alpha_4.\,(n+g+\delta) + u, & (3.8) \\ L_s/L_u &= \beta_1 + \beta_2.\textit{Gini} + \beta_3.\textit{Poor} + \beta_4.\textit{Rich} + v, & (3.9) \end{aligned}$$

where Income is the level of long-run income per capita, (L_s/L_u) is the ratio of skilled to unskilled labor, Invest is the amount of physical capital investment, $(n + g + \delta)$ is the sum of the rates of population growth, technological change, and capital depreciation, Gini is the Gini index which measures the degree of income inequality, Poor is a dummy variable equal to one for an initially poor country and zero otherwise, and Rich is a dummy variable equal to one for an initially rich country and zero otherwise. A priori, we expect the coefficients of Gini, Poor, and $(n + g + \delta)$ to be negative and the coefficients of (L_s/L_u) and Invest to be positive. It is unclear, though, what sign the coefficient of *Rich* will take; hence, its coefficient can be either positive or negative.

3.3.2 Related Model Specifications

At this point, it is imperative that we compare our specification with the specification employed in previous studies. Kremer and Chen (2002) estimate the following model:

$$\textit{Fertd} = \alpha_1 + \alpha_2.\textit{Gini} + x^{'}\beta + \varepsilon, \tag{3.10}$$

where Fertd is the fertility differential, Gini is a measure of income inequality, x is a vector of control variables (such as initial income and regional as well as time dummy variables), and ε is the error term. The Fertd variable is the fitted value of the overall fertility obtained from regressing the overall fertility variable on average years of education; therefore, Fertd measures the variation in the overall fertility that is explained by educational attainment.

De La Croix and Doepke (2003) estimate the following model:

$$Growth = \alpha_1 + \alpha_2.Fertd + x^{'}\beta + \boldsymbol{\varepsilon}, \tag{3.11}$$

where Growth is the growth rate of income per capita (other variables and coefficients are as described before).

As mentioned earlier, both Eqs. (3.9) and (3.10) merely test part of the endogenous fertility model. A better approach would be to specify a model that is capable of testing the model as a system. This can be accomplished by specifying a structural model. It turns out that an earlier empirical analysis by Perotti (1996) does this. In particular, Perotti (1996) estimates the following structural model:

$$Growth = \alpha_1 + \alpha_2.Fert + x^{'}\beta + u, \tag{3.12}$$

$$Fert = \beta_1 + \beta_2.Mid + v, \tag{3.13}$$

where Fert is the overall fertility and Mid is the income share of the third and fourth quintiles of the population. Hence, Mid is a measure of income equality (as opposed to income inequality).

Despite the similarity between Perotti's specification and ours, there is one notable difference. That is, he employs a measure of the overall fertility as opposed to the fertility differential as implied by the endogenous fertility model. Instead of employing a measure of the fertility differential, we deviate further by utilizing a measure of the human capital differential (i.e., the ratio of skilled labor to unskilled labor). This measure is a perfectly acceptable alternative to the fertility differential because the fertility differential implies the human capital differential. Indeed, this measure is more appropriate because we use actual human capital differential data rather than estimated fertility data used by Kremer and Chen (2002).

Finally, it should be mentioned that our model specification is fundamentally different from the one employed by Barro (2000) in the sense that we utilize structural equations in Eqs.(3.8) and (3.9) whereas he uses a reduced-form equation of the neoclassical growth model. He finds that the impact of Gini on Growth depends on whether Fert is added to the specification: when Fert is added, Gini has zero impact on Growth; when Fert is omitted, Gini has a significant, negative impact on Growth. This difference notwithstanding, Barro's findings suggest that income inequality affects economic growth through the fertility rate channel.

3.3.3 Basic Data

On the basis of the preceding discussion, we believe that our specification in Eqs. (3.8) and (3.9) is appropriate. We proceed by collecting the cross-country data for all of the variables identified in those equations. It turns out that the Gini data imposes a severe restriction on the number of available observations. Given this restriction, we are able to collect the required data for 46 countries during the period 1970-2000 for the following variables:

- Inc2000: This variable is defined as the log of real GDP per capita in 2000. The data is taken from the Penn World Table version 6.1 (PWT6.1).

- Invest: This variable is defined as the log of the annual average of the ratio of investment to GDP during the period 1970-2000. The data is taken from PWT6.1.

- $(n + g + \delta)$: This variable is defined as the log of the sum of the rates of population growth (n), technological change (g), and capital depreciation (δ). The population growth rate (n) is defined as the annual average of the population growth rate during 1970-2000. The data is taken from PWT6.1. Following the literature, we set $g + \delta = 0.05$.

- Gini: This variable, which measures the degree of income inequality, is defined as the log of the Gini index in 1970 or its closest neighboring period as long as it does not exceed 1975. The data is taken from Deininger and Squire (1996), who make the necessary efforts to compile high-quality income inequality data. In particular, they impose three stringent quality criteria for the data to be acceptable. First, data must be based on household surveys (not from national accounts that make some assumptions about patterns of income inequality). Second, data must be based on comprehensive coverage of population (not based on some segments of population only). Third, data must be based on comprehensive coverage of income sources (not based on wage incomes only but also nonwage incomes).

- Poor: This variable is defined as a dummy variable which is equal to 1 for any countries that are classified by the World Bank as low-income countries in 1972 (and 0 otherwise); this classification is made based on the income range of these countries. There are 5 low-income countries in our sample. The data is taken from the World Tables 1976, published by the World Bank.

- Rich: This variable is defined as a dummy variable which is equal to 1 for any countries that are classified as high-income countries by the World Bank in 1972 (and 0 otherwise) based on the income range of these countries. There are 13 high-income countries in our sample. The data is taken from the World Tables 1976, published by the World Bank.

- L_s/L_u: This variable is defined as the log of the ratio of the amount of skilled labor to unskilled labor during the period 1970-2000. The amount of skilled labor is defined as the percentage of population who have attained certain level of education multiplied by the quantity of labor. The data on the percentage of population with certain education level is taken from Barro and Lee (2001) while the data on labor force is taken from PWT6.1.

Since there are three levels of education (primary, secondary, and tertiary), we could construct three different measures of skilled labor. Nonetheless, we follow Duffy, Papageorgiou, and Perez-Sebastian (2004) and Caselli and Coleman (2006) in considering six alternative measures of skilled labor: a) workers who have attained complete tertiary education (L_{s0}), b) workers who have attained at least some tertiary education (L_{s1}), c) workers who have attained at least complete secondary education (L_{s2}), d) workers who have attained at least some secondary education (L_{s3}), e) workers who have attained at least complete primary education (L_{s4}), and f) workers who have attained at least some primary education (L_{s5}). Given these six measures, the corresponding measures of unskilled labor can be calculated residually. For example, if skilled labor is defined as in (a), then unskilled labor is defined as any workers who have not completed tertiary education. Similarly, if skilled labor is defined as in (b), then unskilled labor is defined as any workers who have not attained any tertiary education. Of all these alternative measures of skilled labor plus workers who have not received any education at all (L_u), workers who have attained at least some and complete primary education (L_{s5} and L_{s4})account for a large bulk of all workers in our 46-country sample over the period 1970-2000 (see Table 3.1).

Table 3.1: Relative Size of Alternative Measures of Skilled Labor

Year	L_{s0}	L_{s1}	L_{s2}	L_{s3}	L_{s4}	L_{s5}	L_u
1970	50.81 (1.74)	83.67 (2.86)	200.59 (6.85)	345.59 (11.80)	659.13 (22.51)	964.33 (32.93)	624.19 (21.32)
1980	102.92 (2.49)	171.51 (4.14)	415.33 (10.03)	644.32 (15.56)	870.53 (21.03)	1,228.43 (29.67)	707.19 (17.08)
1990	179.53 (3.33)	292.73 (5.43)	555.56 (10.30)	863.10 (16.01)	1,165.35 (21.61)	1,582.95 (29.36)	752.57 (13.96)
2000	255.07 (3.75)	417.24 (6.14)	743.11 (10.93)	1,127.74 (16.59)	1,507.42 (22.17)	2,043.99 (30.06)	704.67 (10.36)
Avg	147.08 (3.05)	241.29 (5.01)	478.65 (9.94)	745.19 (15.48)	1,050.61 (21.82)	1,454.92 (30.22)	697.16 (14.48)

Notes: Entries in the cells and parentheses are the number of workers (in thousands) and their percentages (in percentage points), respectively.

3.4 Empirical Analysis

Using cross-country data for 46 countries during the period 1970-2000, we conduct an empirical analysis of the Kremer-Chen model based on Eqs.(3.8) and (3.9). In particular, Eq.(3.8) is estimated by the instrumental variable (IV) method, where (L_s/L_u) is instrumented by Gini, Poor, and Rich. In other words, Eq.(3.9) serves as the first-stage regression while Eq.(3.8) the second-stage regression.[13]

3.4.1 Basic Analysis

Since there are six alternative measures of skilled and unskilled labor, we estimate our model specification using all of them (one at a time) and report the results in Tables 3.2–3.7. Table 3.2 shows the estimation results when skilled labor is defined as workers who have attained complete tertiary education. In Column (1b), which corresponds to the first-stage regression, we see that the coefficients of Gini and Poor enter with the expected signs and significant at the 1% level. The coefficient of Rich is positive and significant at the 5% level. Since these coefficients are also

[13] Since *Invest* and $(n + g + \delta)$ are assumed to be exogenous, their coefficients will enter the first-stage regression as well to ensure that *Educ* is estimated with the optimal set of instruments; see Chapter 5 of Wooldridge (2002). However, these exogenous variables have little meaning in the first-stage regression. Hence, their coefficients will be suppressed from the first-stage regression results.

jointly significant at the 5% level, we proceed with the second-stage regression. In Column (1a), which corresponds to the second-stage regression, we observe that the coefficients of (L_{s0}/L_{u0}), Invest, and $(n+g+\delta)$ enter with the anticipated signs and significant at the 1% level. These results clearly lend support to the Kremer-Chen model.

Table 3.2: Baseline Estimation with L_{s0}/L_{u0}

Dep.Variable	*Inc2000* (1a)	L_{s0}/L_{u0} (1b)	*Inc2000* (2a)	L_{s0}/L_{u0} (2b)
Constant	4.542^{**} (2.13)	7.473^{**} (2.09)	4.530^{*} (1.94)	9.090^{**} (2.26)
Gini	–	-1.471^{***} (-3.09)	–	-1.808^{***} (-3.32)
Poor	–	-1.592^{***} (-4.69)	–	-1.420^{***} (-3.86)
Rich	–	0.631^{**} (2.57)	–	0.670^{**} (2.43)
L_{s0}/L_{u0}	0.858^{***} (5.62)	–	0.806^{***} (5.20)	–
Invest	0.392^{**} (2.18)	–	0.386^{**} (2.07)	–
$(n+g+\delta)$	-2.223^{***} (3.26)	–	-2.193^{***} (-2.74)	–
Latin	–	–	-0.227 (-0.95)	–
Asia	–	–	0.046^{**} (0.18)	–
Adj. R^2	0.63	0.54	0.65	0.53
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

One may argue that our evidence is tempered by the failure to control for institutional and cultural factors that might differ across regions. To accommodate this objection, we add two regional dummy variables for Asian and Latin American countries and reestimate the model. As shown in Column (2b), the inclusion of these regional dummies does not affect the signs and significance levels of Gini, Poor, and Rich. However, the magnitude of Gini does rise substantially. Since these coefficients are also jointly significant, we proceed with the second-stage regression. In Column (2a), we see that the magnitudes, signs, and significance levels of (L_{s0}/L_{u0}), Invest, and $(n+g+\delta)$ are not affected by the inclusion of these regional dummies. In addition, the coefficient of Latin enters with the negative sign but is insignificant, and the coefficient of Asia enters with the positive sign and significant at the 5% level. We take these results as evidence in favor of the Kremer-Chen model.

Table 3.3 reports the estimation results when skilled labor is defined as workers who have attained at least some tertiary education. In Column (1b), we observe that the coefficients of Gini, Poor, and Rich enter with the correct signs and significant at the 1% level. Since these coefficients are also jointly significant at the 5% level, we proceed with the second-stage regression. In Column (1a), we see that the coefficients of (L_{s1}/L_{u1}), Invest, and $(n + g + \delta)$ enter with the correct signs and significant at least at the 5% level. As before, we add regional dummies to the specification; see Columns (2a) and (2b). Once again, we see that the inclusion of these regional dummies does not affect the magnitudes, signs, and significance levels of the coefficients of most variables. An exception to these results is the coefficient of Gini: its size has risen appreciably. In this case, however, the coefficients of regional dummies are insignificant. Overall, the findings appear to lend support to the Kremer-Chen model.

Table 3.4 shows the results when skilled labor is defined as workers who have attained at least complete secondary education. In Column (1b), we see that the coefficients of Gini, Poor, and Rich enter with the correct signs and significant at the 1% level. Since these coefficients are also jointly significant at the 5% level, we proceed with the second-stage regression. In Column (1a), we see that the coefficients of (L_{s2}/L_{u2}), Invest, and $(n + g + \delta)$ enter with the correct signs and significant at least at the 5% level. Now we add regional dummies to the specification; the results are reported in Columns (2a) and (2b). As before, the inclusion of these regional dummies does not affect the magnitudes, signs, and significance levels of the coefficients of most variables. However, the magnitude of $(n + g + \delta)$ falls substantially and it becomes insignificant. We interpret these findings as evidence that lends support to the Kremer-Chen model.

Table 3.3: Baseline Estimation with L_{s1}/L_{u1}

Dep.Variable	*Inc2000* (1a)	L_{s1}/L_{u1} (1b)	*Inc2000* (2a)	L_{s1}/L_{u1} (2b)
Constant	3.819^{**} (2.07)	9.465^{**} (2.57)	3.551^{*} (1.72)	11.312^{***} (2.75)
Gini	–	-1.610^{***} (−3.28)	–	-1.987^{***} (−3.56)
Poor	–	-1.723^{***} (−4.93)	–	-1.529^{***} (−4.05)
Rich	–	1.022^{***} (4.04)	–	1.061^{***} (3.75)
L_{s1}/L_{u1}	0.684^{***} (6.26)	–	0.651^{***} (5.66)	–
Invest	0.433^{***} (2.74)	−0.117 (−0.58)	0.419^{**} (2.50)	−0.037 (−0.17)
$(n+g+\delta)$	-2.121^{**} (−3.45)	2.100^{**} (2.43)	-2.207^{***} (−3.03)	2.378^{**} (2.44)
Latin	–	–	−0.143 (−0.65)	0.312 (1.05)
Asia	–	–	0.086 (0.36)	−0.112 (−0.37)
Adj. R^2	0.71	0.64	0.71	0.64
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3.4: Baseline Estimation with L_{s2}/L_{u2}

Dep.Variable	*Inc2000* (1a)	L_{s2}/L_{u2} (1b)	*Inc2000* (2a)	L_{s2}/L_{u2} (2b)
Constant	4.945^{**} (2.59)	7.140^{*} (1.90)	5.707^{***} (2.69)	5.388 (1.28)
Gini	–	-1.489^{***} (−2.97)	–	-1.447^{**} (−2.54)
Poor	–	-1.379^{***} (−3.86)	–	-1.431^{***} (−3.71)
Rich	–	0.933^{***} (3.61)	–	1.122^{***} (3.88)
L_{s2}/L_{u2}	0.770^{***} (6.34)	–	0.715^{***} (5.96)	–
Invest	0.416^{**} (2.68)	–	0.468^{***} (2.95)	–
$(n+g+\delta)$	-1.504^{**} (−2.34)	–	−1.187 (−1.61)	–
Latin	–	–	−0.191 (−0.91)	–
Asia	–	–	−0.209 (−0.95)	–
Adj. R^2	0.72	0.60	0.73	0.60
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3.5 documents the results when skilled labor is defined as workers who have attained at least some secondary education. Column (1b) shows that the coefficients of Gini, Poor, and Rich enter with the correct signs and, except for Gini, are significant at the 1% level; note that the coefficient of Gini is significant at the 5% level. Column (1a) shows that the coefficients of (L_{s3}/L_{u3}), Invest, and $(n+g+\delta)$ enter with the correct signs and, except for $(n+g+\delta)$, are significant at the 1% level; note that the coefficient of $(n+g+\delta)$ is insignificant. Columns (2a) and (2b) show the results when regional dummies are added to the specification. Once again, the presence of these dummies does not affect the sizes, signs, and significance levels of the coefficients of most variables. There are two exceptions to these results: the magnitude of Rich rises whereas the magnitude of $(n+g+\delta)$ declines. Unlike the previous exercises, the coefficients of regional dummies are significant here (albeit at the 10% level only). We take these results as evidence supporting the Kremer-Chen model.

Table 3.5: Baseline Estimation with L_{s3}/L_{u3}

Dep.Variable	*Inc2000* (1a)	L_{s3}/L_{u3} (1b)	*Inc2000* (2a)	L_{s3}/L_{u3} (2b)
Constant	6.139^{***} (3.06)	4.339 (1.07)	7.115^{***} (3.40)	1.531 (0.36)
Gini	–	-1.291^{**} (−2.40)	–	-1.270^{**} (−2.18)
Poor	–	-1.162^{***} (−3.03)	–	-1.227^{***} (−3.11)
Rich	–	1.083^{***} (3.90)	–	1.419^{***} (4.80)
L_{s3}/L_{u3}	0.753^{***} (6.34)	–	0.644^{***} (6.17)	–
Invest	0.585^{***} (4.02)	–	0.656^{***} (4.69)	–
$(n+g+\delta)$	−0.653 (−0.92)	–	−0.273 (−0.37)	–
Latin	–	–	-0.325^{*} (−1.70)	–
Asia	–	–	-0.361^{*} (−1.80)	–
Adj. R^2	0.73	0.60	0.77	0.64
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3.6 reports the results when skilled labor is defined as workers who have attained at least complete primary education. Columns (1a) and (1b) show that the coefficients of Gini, Poor, Rich, (L_{s4}/L_{u4}), and Invest enter with the correct signs and significant at the 1% level. These results continue to hold when regional dummies are included; see Columns (2a) and (2b). In addition, the magnitudes of these coefficients are fairly robust to the addition of regional dummies. In both cases, with and without regional dummies, the coefficient of $(n + g + \delta)$ enters with the wrong sign and insignificant. Finally, the coefficients of regional dummies are insignificant. We conclude that these results lend support to the Kremer-Chen model.

Table 3.6: Baseline Estimation with L_{s4}/L_{u4}

Dep.Variable	*Inc2000* (1a)	L_{s4}/L_{u4} (1b)	*Inc2000* (2a)	L_{s4}/L_{u4} (2b)
Constant	8.526^{***} (3.80)	3.183 (0.80)	9.075^{***} (3.78)	1.074 (0.24)
Gini	–	-1.629^{***} (−3.05)	–	-1.417^{**} (−2.30)
Poor	–	-1.500^{***} (−3.95)	–	-1.630^{***} (−3.92)
Rich	–	0.841^{***} (3.06)	–	0.953^{***} (3.06)
L_{s4}/L_{u4}	0.769^{***} (6.52)	–	0.753^{***} (6.12)	–
Invest	0.440^{***} (2.92)	–	0.474^{***} (3.00)	–
$(n + g + \delta)$	0.311 (0.39)	–	0.520 (0.60)	–
Latin	–	–	−0.068 (−0.31)	–
Asia	–	–	−0.160 (−0.72)	–
Adj. R^2	0.73	0.69	0.73	0.68
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3.7 shows the results when skilled labor is defined as workers who have attained at least some primary education. In this case, however, only 43 observations are available because the (L_{s5}/L_{u5}) data are undefined for three countries (Finland, Japan, and New Zealand) due to division by 0. Columns (1a) and (1b) show that the coefficients of Poor, Rich, (L_{s5}/L_{u5}), and

Invest enter with the correct signs and significant at the 1% level. However, the coefficients of Gini and $(n+g+\delta)$ are insignificant (although Gini enters with the correct sign). When regional dummies are added to the specification (see Columns (2a) and (2b)), we notice quite a few drastic changes in terms of the significance of the coefficients (Gini and $(n+g+\delta)$ become significant) and the magnitudes of the coefficients (L_{s5}/L_{u5}, Poor, and Rich). Since these results are sensitive to the addition of regional dummies, we take them as evidence against the Kremer-Chen model.

Table 3.7: Baseline Estimation with L_{s5}/L_{u5}

Dep.Variable	*Inc2000* (1a)	L_{s5}/L_{u5} (1b)	*Inc2000* (2a)	L_{s5}/L_{u5} (2b)
Constant	12.067^{***} (3.16)	-9.249 (-1.64)	12.743^{***} (3.66)	-8.429 (-1.43)
Gini	–	-0.625 (-0.83)	–	-1.417^{*} (-1.80)
Poor	–	-2.047^{***} (-3.85)	–	-1.662^{***} (-3.27)
Rich	–	1.254^{***} (3.10)	–	1.773^{***} (4.20)
L_{s5}/L_{u5}	0.639^{***} (4.58)	–	0.575^{***} (4.89)	–
Invest	0.524^{**} (2.47)	–	0.521^{***} (2.72)	–
$(n+g+\delta)$	2.077 (1.44)	–	2.187^{*} (1.69)	–
Latin	–	–	-0.696^{**} (-2.67)	–
Asia	–	–	-0.109 (-0.37)	–
Adj. R^2	0.47	0.72	0.60	0.78
Obs.	43	43	43	43

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

One may justifiably argue that these poor results are due to the reduction in sample size. In response to this objection, we repeat the analysis in Tables 3.2–3.6 with 43 observations. It turns out that the reduced sample size does not affect the results in those tables (see Tables D1–D5 in Appendix D). Therefore, we conclude that the threshold level of skilled labor cannot be represented by workers who have attained at least some primary education; they need to be more educated than this in order to be considered skilled labor.

The preceding discussion suggests that all of the above measures of skilled labor (except the last one) are empirically plausible. However, the maximum impact of Gini on Inc2000 (which is calculated from the coefficients of Gini in the first-stage regression and (L_s/L_u) in the second-stage regression) is achieved when skilled labor is defined as those who have attained at least some tertiary education. This definition of skilled labor makes much sense because the ability to think and learn complex concepts (such as learning a new computer language) is probably highly correlated with the ability to pursue college education. Accordingly, we take (L_{s1}/L_{u1}) as the best measure of skilled-unskilled labor ratio to be employed in the subsequent robustness analysis.

3.4.2 Robustness Analysis

One may argue that our baseline results might be due to an exceedingly small sample size, 46. This problem, in turn, arises because the data on Gini is not available for many countries in early years. Therefore, one way to increase the sample size would be to curtail the sample period to 1980-2000. However, doing so will increase the sample size only marginally; the sample size becomes 61 instead of 46. Another way to increase the sample size would be to work with panel data (as opposed to cross-sectional data). So we construct a panel data of countries with a five-year interval during 1970-2000, where Gini, Poor, and Rich are measured at 1970, 1975, $\ldots$, 1995, Invest and $(n + g + \delta)$ are measured as averages of 1971-1975, 1976-1980, $\ldots$, 1996-2000, and Income and (L_{s1}/L_{u1}) are measured at 1975, 1980, $\ldots$, 2000. Including only those data for which there are at least two consecutive observations, we end up with an unbalanced panel of 53 countries and 226 observations.

Given this substantially expanded number of observations, we reestimate our model by the pooled IV method and document the results in Table 3.8(a). Columns (1a) and (1b) show that the coefficients of Gini, Poor, Rich, (L_{s1}/L_{u1}), Invest, and $(n + g + \delta)$ enter with the anticipated signs and significant at the 1% level. Compared to the corresponding coefficients in Table 3.3, however, the coefficients of Gini and Poor do not appear to be robust.[14] When we add regional dummies to the specification, we see that most results remain intact (see Columns (2a) and (2b)). However, the magnitudes of Gini and $(n + g + \delta)$ have changed considerably. It should also be noted that the coefficients of Latin and Asia are significant. Compared to the corresponding coefficients in Table 3.3, we see that there is a considerable change in the magnitudes of Gini, $(n + g + \delta)$, and Latin. Despite all of these changes, the coefficients of key variables continue to deliver the same message. Therefore, we conclude that these findings lend further support to the Kremer-Chen model.

Table 3.8(a): Panel Data Estimation

Dep.Variable	*Income* (1a)	L_{s1}/L_{u1} (1b)	*Income* (2a)	L_{s1}/L_{u1} (2b)
Constant	6.692^{***} (6.05)	0.741 (0.32)	6.798^{***} (6.39)	1.744 (0.74)
Gini	–	-0.720^{**} (-2.31)	–	-1.161^{***} (-3.37)
Poor	–	-1.314^{***} (-7.04)	–	-1.267^{***} (-5.92)
Rich	–	0.956^{***} (7.12)	–	1.112^{***} (7.42)
L_{s1}/L_{u1}	0.787^{***} (12.63)	–	0.633^{***} (9.78)	–
Invest	0.426^{***} (3.11)	–	0.621^{***} (4.89)	–
$(n + g + \delta)$	-1.054^{***} (-3.05)	–	-0.730^{**} (-2.01)	–
Latin	–	–	-0.230^{*} (-1.90)	–
Asia	–	–	-0.335^{***} (-3.12)	–
Adj. R^2	0.59	0.52	0.70	0.53
Obs.	226	226	226	226

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Kristin Forbes (2000) argues that the negative impact of income inequality on growth could be attributed to omitted variable bias. To overcome this problem, she estimates her model based on

[14] A coefficient is said to be robust if its sign, magnitude, and significance level is similar to the corresponding coefficient in the baseline estimation. In terms of magnitude, their similarity is determined by whether their difference is within one standard deviation. In the baseline estimation, one standard deviation of *Gini*, *Poor*, *Rich*, (L_{s1}/L_{u1}), *Invest*, $(n + g + \delta)$, *Latin*, and *Asia* is 0.25, 0.31, 0.46, 0.98, 0.52, 0.14, 0.46, and 0.47, respectively.

the fixed effect (FE) method of panel data. Following her argument, we conduct the first-stage regression of the FE model. Since the FE model controls for any time-constant country-specific effects, we drop any dummy variables that we have employed so far (they will be captured by country-specific effects). Therefore, we end up with four explanatory variables for the first-stage regression: Gini, Rich, Invest, and $(n + g + \delta)$.[15] As shown in Column (1b) of Table 3.8(b), however, the coefficient of Gini is not significant. Furthermore, Gini and Rich are jointly insignificant, and this prevents us from conducting the second-stage regression. We take these findings as evidence against the Kremer-Chen model.[16]

Table 3.8(b): Panel Data Estimation

Dep.Variable	L_{s1}/L_{u1} (1b)	*Income* (2a)	L_{s1}/L_{u1} (2b)
Constant	-12.914^{***} (-6.72)	5.691^{***} (3.29)	12.142^{**} (2.24)
Gini	-0.059 (-0.15)	–	-2.482^{***} (-3.22)
Poor	–	–	-1.407^{***} (-3.45)
Rich	0.362^{**} (2.06)	–	1.144^{***} (3.78)
L_{s1}/L_{u1}	–	0.466^{***} (3.83)	–
Invest	–	0.817^{***} (3.32)	–
$(n + g + \delta)$	–	-0.801 (-1.31)	–
Latin	–	-0.349^{*} (-1.80)	–
Asia	–	-0.470^{**} (-2.20)	–
Adj. R^2	0.26	0.76	0.66
Obs.	226	226	226

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

15 Note that *Rich* is not time-constant; it varies over time as some countries shift their position between middle- and high-income countries.

16 In practice, we could have proceeded by estimating the random effect (RE) model. In principle, however, the use of a RE method is unwarranted for two reasons. First, the use of a panel IV estimation based on a RE method is inappropriate if (L_{s1}/L_{u1}) is correlated with individual country effects. Second, the unwarranted use of a panel IV estimation based on a FE method precludes us from evaluating the appropriateness of a RE method through a Hausman test.

Hauk and Wacziark (2004) conduct a Monte Carlo study of the augmented Solow growth model using a few alternative estimation techniques. They find that the between effect (BE) estimator outperforms other estimators in terms of the degree of bias of the estimated coefficients. Based on this consideration, we reestimate our model using a BE method and report the results in Columns (2a) and (2b) of Table 3.8(b). There we see that the coefficients of Gini, Poor, Rich, (L_{s1}/L_{u1}), and Invest enter with the correct signs and significant at the 1% level. Compared to the corresponding coefficients in Table 3.3, we see that there is an appreciable change in the magnitude and significance of $(n + g + \delta)$, significance of regional dummies, and magnitude of Gini. Overall, however, these results lend support to the Kremer-Chen model.

The fact that our dependent variable, Inc2000, is measured in 2000, while some of our explanatory variables (Invest and n)[17] are measured as averages over the period 1970-2000 may make us susceptible to simultaneity bias (i.e., the direction of causality may run from these variables to Inc2000 instead). In the growth empirics, this endogeneity issue is usually taken care of by instrumenting the relevant regressors with their lagged values (see, for example, Barro and Sala-i-Martin (2004)). In our context, if the baseline results are sensitive to these changes, then this is taken as evidence that the suspected regressors (Invest and n in our context) are endogenous.

Before we do that, however, it is imperative that we test the endogeneity of Invest and n using a Hausman test. First, we estimate the second-stage regression with and without instrumenting Invest and n with their respective lagged values, which are measured as averages over the period 1965-1995. Second, we test whether the difference between estimates obtained from the regression with and without instrumentating Invest and n is statistically significant. (Note that (L_{s1}/L_{u1}) is always instrumented by Gini, Poor, and Rich by theoretical implication.)

17 Although (L_{s1}/L_{u1}) data is averaged over the period 1970-2000, this should not pose any simultaneity problem because it is instrumented by *Gini*, *Poor*, and *Rich* that are measured at the beginning of the period of analysis.

Unfortunately, the Hausman test fails to deliver any results in the basic specification because the variance-covariance matrix is not positive definite. In the augmented specification, the test fails to deliver useful results because the test statistic is negative.

As an alternative to the Hausman test, we adopt an auxiliary regression approach.[18] This method can be summarized in the following steps. First, we run the first-stage regression for each (L_{s1}/L_{u1}), Invest, and $(n + g + \delta)$. (Note that the first-stage regression needs to conducted for (L_{s1}/L_{u1}) as well.) Second, we extract residuals obtained from each first-stage regression. Third, we run the second-stage regression with the inclusion of these residuals using the method of ordinary least squares (OLS). Finally, we test whether the estimated coefficients from the residuals are jointly significant. Performing all of these steps, we find that the estimates are jointly significant. We take this result as evidence that (L_{s1}/L_{u1}), Invest, and $(n+g+\delta)$ are endogenous.

Given the above results, we repeat our baseline estimation with Invest and n instrumented by their respective lagged values and report the results in Table 3.9. In Columns (1a) and (1b), we find that the coefficients of all variables enter with the expected signs and significant. These results continue to hold when we add regional dummies to the basic specification (see Columns (2a) and (2b)). Compared to the corresponding coefficients in Table 3.3, we see that the results are quite robust. We take these results as evidence that our baseline results are not affected much by the endogeneity of Invest and n.

Previous work that employs the (L_s/L_u) data points out that the way skilled and unskilled labor is defined suffers from an aggregation problem. For example, the (L_{s1}/L_{u1}) data that we use treat workers with different levels of education equally. If labor is paid according to its marginal revenue product, then workers with a higher level of education should be given a greater weight than workers with a lower level of education. To overcome this aggregation problem, we follow Duffy, Papageorgiou and Perez-Sebastian (2004) and Caselli and Coleman (2006) in weighting

[18] I am grateful to Dr. Carter Hill for suggesting this method; see Baum, Schaffer, and Stillman (2003) for reference.

the (L_{s1}/L_{u1}) data according to the marginal revenue product of labor (see Appendix E for details). Unfortunately, the weighting procedure requires some additional data on the return to education and on the duration of education at various levels. It turns out that data on the return to schooling is not available for many countries, and this results in the reduction of our sample size to 30. Therefore, we opt to work with the panel data of countries. Utilizing the same panel data set as before, but interacting it with data on the return to education and the duration of education, yields an unbalanced panel of 32 countries and 145 observations.

Table 3.9: Estimation with Instrumented *Invest* and n

Dep.Variable	*Inc2000* (1a)	L_{s1}/L_{u1} (1b)	*Inc2000* (2a)	L_{s1}/L_{u1} (2b)
Constant	2.138 (1.11)	7.945** (2.07)	2.237 (1.05)	10.953** (2.68)
Gini	–	−1.435*** (−2.88)	–	−1.943*** (−3.61)
Poor	–	−1.402*** (−3.64)	–	−0.981** (−2.32)
Rich	–	1.050*** (4.15)	–	0.961*** (3.44)
L_{s1}/L_{u1}	0.583*** (4.84)	–	0.507*** (3.44)	–
Invest	0.687*** (3.11)	–	0.771*** (2.96)	–
$(n+g+\delta)$	−2.375*** (−3.89)	–	−2.209*** (−2.86)	–
Latin	–	–	−0.194 (−0.85)	–
Asia	–	–	−0.065 (−0.24)	–
Adj. R^2	0.71	0.64	0.70	0.67
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

With this weighted (L_{s1}/L_{u1}) data, we reestimate our model by the pooled IV method, the results of which are documented in Table 3.10. Columns (1a) and (1b) show that the coefficients of all variables enter with the correct sign and, except for Gini and $(n+g+\delta)$, are significant at the 1% level. The coefficient of $(n+g+\delta)$ is significant at the 10% level only whereas the coefficient of Gini is insignificant. Compared to the corresponding coefficients in Table 3.2, we see that there is a huge change in the magnitude and significance of Gini, and the magnitude of Poor, Invest, and $(n+g+\delta)$. When we add regional dummies to the specification, we see that the magnitude of Poor declines considerably, and the coefficient of $(n+g+\delta)$ loses its significance; see Columns (2a) and (2b). The coefficients of other variables remain stable in terms of magnitudes, signs, and significance levels. Compared to the corresponding coefficients in Table 3.2, we see that there is a significant change in the magnitude and significance of Gini and $(n+g+\delta)$, and the magnitude of Poor and Invest. Since the coefficient of Gini is insignificant with and without regional dummies, these results are taken as evidence against the Kremer-Chen model.[19]

Table 3.10: Panel Data Estimation with Weighted L_{s1}/L_{u1}

Dep.Variable	*Income* (1a)	L_{s1}/L_{u1} (1b)	*Income* (2a)	L_{s1}/L_{u1} (2b)	L_{s1}/L_{u1} (3b)	L_{s1}/L_{u1} (4b)
Constant	6.374^{***} (4.69)	-0.461 (-0.16)	6.365^{***} (4.05)	0.803 (0.27)	-14.219^{***} (-7.02)	11.849 (1.50)
Gini	–	-0.395 (-1.02)	–	-0.410 (-0.86)	-0.142 (-0.31)	-1.380 (-1.26)
Poor	–	-0.980^{***} (-4.17)	–	-0.657^{**} (-2.34)	–	-0.844 (-1.39)
Rich	–	0.996^{***} (6.27)	–	0.816^{***} (3.97)	0.128 (0.73)	0.656 (1.42)
L_{s1}/L_{u1}	0.935^{***} (4.69)	–	0.864^{***} (4.75)	–	–	–
Invest	0.948^{***} (5.59)	–	0.989^{***} (5.80)	–	–	–
$(n+g+\delta)$	-0.780^{*} (-1.82)	–	-0.691 (-1.21)	–	–	–
Latin	–	–	-0.031 (-0.13)	–	–	–
Asia	–	–	-0.143 (-0.44)	–	–	–
Adj. R^2	0.47	0.40	0.54	0.41	0.37	0.52
Obs.	145	145	145	145	145	145

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

One way to interpret these unfavorable results is that the Kremer-Chen model is rejected when it is confronted with a better (weighted) human capital data. But this view cannot be tested. A

[19] We obtain similarly poor results when we employ FE and BE methods; see Columns (3b) and (4b) in Table 3.10.

testable interpretation of these poor results is that they are driven by the reduction in sample size from 226 to 145. To test this hypothesis, we reestimate our model using the original unweighted (L_{s1}/L_{u1}) data with 145 observations. The results, which are reported in Table D6 of Appendix D, show that the coefficient of Gini is insignificant. Therefore, we conclude that our results are sensitive to sample size.

According to the Kremer-Chen model, the impact of income inequality on long-run income per capita is unlikely to be applicable to poor and rich countries. For poor countries, this is the case because the income effect dominates; for rich countries, this is the case because the substitution effect becomes weaker (see Section 3.2). If so, then it follows there exists a threshold level of initial income below and above which inequality does not matter. This can be done by relaxing the definition of poor and rich countries to include some middle-income countries.

Renaming our previous measures of Poor and Rich as Poor1 and Rich1, we include two additional categories of poor and rich countries: Poor2 is defined as the sum of Poor1 and lower middle-income countries, and Rich2 is the sum of Rich1 and the top one-third of upper middle-income countries. (Poor2 contains 10 countries and Rich2 19 countries.) A priori, we expect that the magnitude and significance of Gini would decline with these new measures, Poor2 and Rich2, because this implies the strengthening of the income effect and the weakening of substitution effect. Columns (1)–(4) of Table 3.11 report how the estimates of Gini change when we vary the threshold levels of poor and rich countries. We observe that the magnitude and significance of Gini alternate between ups and downs. We take these results as evidence against the Kremer-Chen model.

Another way to test the above hypothesis is by including the interactive terms, Gini$\times$Poor and Gini$\times$Rich, in the first-stage regression. If the interactive terms are significant, then this implies that the impact of Gini on (L_{s1}/L_{u1}) is different between poor, rich, and middle-income countries.

As shown in Column (5) of Table 3.11, however, the interactive terms are insignificant. We take this result as evidence against the Kremer-Chen model. However, there is a caveat: this poor result might be due to the high correlation between Poor and Gini×Poor (their correlation coefficient is 0.99), and Rich and Gini×Rich (their correlation coefficient is 0.99).

Table 3.11: Estimation with Alternative Measures of
Poor and Rich, Interactive Terms, and Fertility Differential

Specification	*Poor1, Rich1* (1)	*Poor2, Rich1* (2)	*Poor1, Rich2* (3)	*Poor2, Rich2* (4)	*Gini×Poor, Gini×Rich* (5)	*Fertd* (6)
Gini	-1.987^{***} (-3.56)	-1.688^{**} (-2.56)	-2.555^{***} (-4.59)	-2.34^{***} (-3.55)	-2.136^{***} (-2.88)	-0.604 (-1.27)
Gini×Poor	–	–	–	–	-1.023 (-0.64)	–
Gini×Rich	–	–	–	–	1.012 (0.78)	–
Fertd	–	–	–	–	–	-0.415^{*} (-1.69)
(L_{s1}/L_{u1})	0.651^{***} (5.66)	0.583^{***} (4.43)	0.748^{***} (5.97)	0.708^{***} (4.85)	0.657^{***} (5.77)	-0.049 (-0.18)
Obs.	46	46	46	46	46	43

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS. The coefficients of (.) and Fertd come from the second-stage regression, where Inc2000 is the dependent variable. The coefficients of other variables come from the first-stage regression, where (.) is the dependent variable. Regional dummies are always included in the second-stage regression. t-statistics are given in parentheses. t-statistic in Column (6) is based on White's heteroskedasticity-consistent standard error. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

According to the Kremer-Chen model, the impact of income inequality on long-run income per capita works through the fertility channel. That is, income inequality results in a fertility differential, and the fertility differential in turn results in human capital differential, (L_s/L_u). If so, the impact of inequality on (L_s/L_u) will be weakened if a measure of fertility differential is added to the specification. To test this hypothesis, a measure of fertility differential, Fertd, is added to the second-stage regression.[20] A priori, we expect that the presence of Fertd would yield one of the following: either the magnitude and significance level of Gini will decline or the coefficient of Fert is insignificant. As shown in Column (6) of Table 3.11, the coefficient

[20] Following Kremer and Chen (2002), *Fertd* is the fitted value of the overall fertility rate, *Fert*, when it is regressed against educational attainment, *Educ*. Data on *Fert* is taken from Barro and Lee (1994) while data on *Educ* is taken from Barro and Lee (2001).

of Fertd is significant, the coefficient of Gini is insignificant, and the coefficient of (L_{s1}/L_{u1}) enters with the incorrect sign and insignificant. We conclude that these results lend support to the Kremer-Chen model.

As a further robustness check, we test whether the coefficients of Gini and (L_{s1}/L_{u1}) are sensitive to the inclusion of additional control variables. In practice, this is hard to implement because different papers employ different control variables in their robustness tests. To make this exercise more manageable, we employ three criteria in choosing these additional control variables. First, the variables must not be implied by other theoretical models in this area. Second, the variables should make an intuitive sense. Third, the variables should be frequently employed.

Applying these criteria, we obtain five control variables: government consumption expenditures (Govcons), government consumption expenditures net of education spending (Govconse), government expenditures on defense (Govdef), the price of investment (PPI), and life expectancy (Life). Each of these variables is added one at a time in the second-stage regression.[21] As shown in Table 3.12, in terms of their signs and significance, the coefficients of Gini and (L_{s1}/L_{u1}) are robust to the inclusion of these variables. In terms of their magnitudes, however, the coefficient of (L_{s1}/L_{u1}) is sensitive to the inclusion of all control variables (the coefficient of Gini is robust). However, since the main results continue to hold, we take them as further support for the Kremer-Chen model.

Previous work that analyzes the inequality-growth relationship expresses Gini in levels (as opposed to logs). To test whether our results are robust to this variable definition, we reestimate our baseline estimation in Table 3.3 with Gini expressed in levels, Ginilev. As shown in Table D7 in Appendix D, the coefficients of all variables retain their signs and significance. We take these results as further evidence in favor of the Kremer-Chen model.

---

[21] Each of these variables is measured as the average for 1970-1984 or 1970-1985, depending on the availability of data. All of these data are taken from Barro and Lee (1994).

Table 3.12: Estimation with Alternative Control Variables

Specification	*Govcons* (1)	*Govconse* (2)	*Govdef* (3)	*PPI* (4)	*Life* (5)
Gini	-1.942^{***} (-2.72)	-1.925^{**} (-2.70)	-1.950^{***} (-2.73)	-1.902^{***} (-3.02)	-1.645^{***} (-2.73)
(L_{s1}/L_{u1})	0.457^{***} (3.42)	0.433^{***} (3.28)	0.425^{***} (3.13)	0.373^{***} (2.92)	0.285^{**} (2.64)
Control	-1.656 (-0.92)	-2.467 (-1.07)	-0.929 (-0.28)	0.484^{**} (2.24)	0.046^{**} (2.53)
Obs.	45	45	45	46	46

Notes: Except for dummies and additional control variables, all variables are expressed in logs. Estimation is done by 2SLS. The coefficient of (.) and additional control variables come from the second-stage regression, where Inc2000 is the dependent variable. The coefficients of other variables come from the first-stage regression, where (.) is the dependent variable. Regional dummies are always included. t-statistics are given in parentheses. t-statistic in Column (5) is based on White's heteroskedasticity-consistent standard error. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Given our small sample size and the asymptotic inference of IV estimation, it is tempting to complement our IV analysis with the inference based on bootstrap standard errors. However, this temptation should be resisted in light of the evidence that bootstrap standard errors perform poorly in models with endogenous regressors; see Mackinnon (2002).

In addition to estimating a structural model, we also estimate a reduced-form version of the Kremer-Chen model where Eq.(3.9) is substituted into Eq.(3.8). As reported in Columns (1) and (2) of Table D8, however, the results of this reduced-form equation are not so clear-cut. That is, although the coefficient of Gini enters with the expected sign in both columns, it is significant in Column (1) only.

3.5 Discussion

As implied by Eqs.(3.8) and (3.9), income inequality affects income per capita through human capital differential, (L_s/L_u). Since each of these variables is measured in logs, the linkage between any two variables is in the form of an elasticity. Consider Columns (1a) and (1b) in Table 3.3 again. The coefficients of Gini and (L_{s1}/L_{u1}) are –1.61 and 0.68, respectively. This implies that a 10% increase in inequality would decrease the human capital differential by 16.1% and a 10% increase in the human capital differential would raise income per capita by 6.8%. Taken together, both coefficients imply that a 10% increase in inequality would reduce income per capita by 10.95% (i.e., $6.8\% \times (-16.1\%/10\%) = -10.95\%$.)

The above estimate is based on the impact of inequality on income through the human capital differential channel. To measure the net effect of inequality on income, we look at the reduced-form estimate described in Table D8. The counterpart of Columns (1a) and (1b) in Table 3.3 is Column (1) in Table D8. There, we see that the coefficient of Gini is –1.003. This implies that a 10% increase in inequality is expected to decrease income per capita by 10.03%. Hence, the magnitudes are quite similar.

Now we compare our finding on the impact of income inequality on income with the findings in previous studies. In order to make our quantitative effect of inequality on income comparable to that in previous work, we need to make a few adjustments. First, we replace our dependent variable, Inc2000, with Growth. Second, we express Gini in levels (instead of logs). Finally, we assess how much a one-standard-deviation change in Gini affects Growth. With these modifications, we reestimate our reduced-form model in Column (1) of Table D8. We find that the coefficient of Gini is –0.03 and significant (see Column (3) in Table D8). This result implies that an increase in income inequality by one standard deviation (i.e., 10 in our sample) is expected to decrease economic growth by 30%.

Persson and Tabellini (1994) employ income share of the third quintile as a measure of income equality. In Column (1) of their Table 5, they report that the coefficient of this variable is 0.189 and significant. This implies that an increase in income equality by one standard deviation (i.e., 3 in their sample) is expected to increase growth by 0.57%. Perotti (1996) employs income share of the third and fourth quintiles as a measure of income equality. In Column (1) of his Table 4, he reports that the coefficient of this variable is 0.118 and significant. This suggests that a rise in income equality by one standard deviation (i.e., 0.16 in his sample) is expected to increase growth by 0.58%. Alesina and Rodrik (1994) employ the Gini coefficient in 1960 (measured in levels) as a measure of income inequality. In Column (6) of their Table 1, they report that the coefficient of Gini is –5.23. This implies that a rise in Gini by one standard deviation (i.e., 0.16 in their sample) is expected to decrease growth by almost 0.84%. Sylwester (2000) employs the Gini coefficient in 1970 (also measured in levels) as a measure of income inequality. In Column (1) of his Table 1, he reports that the coefficient of Gini is –0.0007. This suggests that a rise in inequality by one standard deviation (i.e., 10 in his sample) is expected to decrease growth by 0.7%. Finally, Forbes (2000) employs lagged values of Gini (also measured in levels) as a measure of income inequality. In Column (4) of her Table 3, she reports that the coefficient of Gini is 0.0013. This implies that a rise in inequality by one standard deviation (i.e., 10 in her sample) is expected to increase growth by 1.3%. Taken together, these findings indicate that the impact of income inequality on economic growth takes a very close range of values, from as low as 0.57% (Persson and Tabellini's findings) to as high as 0.84% (Alesina and Rodrik's finding). Therefore, our estimate is too high.

3.6 Conclusion

In this paper, we conduct an empirical analysis to test the implications of the Kremer-Chen model based on a cross-section of 46 countries during the period 1970–2000. Our baseline analyses in Tables 3.2–3.6 (with alternative proxies for human capital differential) yield the results that income inequality exerts a negative impact on long-run per capita income, and it does so through the human capital differential channel (which has a positive impact on income). These findings continue to hold when we a) employ the pooled IV method of panel data, b) employ the BE method of panel data, c) add the fertility differential variable to the specification, d) employ alternative control variables, and d) express the Gini variable in levels. However, we obtain unfavorable results when we a) employ the FE method of panel data, and b) vary the threshold levels of poor and rich countries.

At first glance, it seems that poor results are also obtained when we a) employ the weighted measure of human capital differential, and b) include the interactive terms, Gini$\times$Poor and Gini$\times$Rich, into the specification. However, there is evidence that poor results in (a) are due to a reduced sample size whereas poor results in (b) are due to the multicollinearity problem. Given these qualifications, the preponderance of evidence on the robustness check is skewed toward our baseline results. Therefore, we conclude that our analysis yields results that lend support to the Kremer-Chen model.

Chapter 4 A Reexamination of the Convergence Hypothesis

4.1 Introduction

According to Solow (1956), if economies are structurally similar,[22] then poor economies tend to grow faster than rich ones; if this tendency exists, then economies tend to converge to each other in terms of the level of per capita income over time because economic growth tends to slow down the richer the economies become. If convergence occurs among economies, then all economies will eventually enjoy similar standards of living regardless of their initial positions. Given this staggering welfare implication, the issue is whether this convergence hypothesis is borne out by the experience of modern economies.

One of the earliest empirical studies on this issue is Baumol (1986). He finds that a small sample of countries exhibits convergence in the level of per capita income over the period 1870-1979. His analysis, however, has been criticized by De Long (1988) on the ground of sample selection bias. That is, Baumol happens to pick a sample of rich countries in 1979. When De Long repeats the Baumol's analysis using a sample of rich countries in 1870, he finds that convergence does not occur.

Next, Mankiw, Romer, and Weil (1992) (MRW hereafter) discover that these earlier studies do not include any control variables in their growth regressions as implied by the Solow model. When these control variables are added to the growth regression, MRW (1992) find that convergence does take place among a large sample of countries. Their convergence findings, known as conditional convergence, have been confirmed by Islam (1995), Caselli, Esquivel and Lefort (1996), and Lee, Pesaran and Smith (1997), among others.[23]

---

22 Economies are said to have similar structures if they have similar steady-state determinants such as the rate of technological progress, the saving rate, the population growth rate, and the capital depreciation rate.

23 Of course, these researchers improve on MRW's analysis by using more sophisticated estimation methods such as fixed effects estimation and generalized method of moments. Furthermore, unlike MRW, they find that countries appear to converge at a much faster rate.

Durlauf and Johnson (1995) question the use of a linear growth regression model on a sample of apparently heterogeneous countries. One way to capture this country heterogeneity is to rely on multiple-regime models such as the model of Azariadis and Drazen (1990). According to this model, countries converge to multiple steady states depending on their levels of human capital. The idea is that there exist some threshold levels of human capital below which productivity is stagnant and above which productivity accelerates. As a result, economic growth is sluggish (accelerated) when human capital is below (above) the threshold. It follows then that countries whose level of human capital is below (above) the threshold level will converge to a low (high) steady-state level of income. When Durlauf and Johnson partition countries into a few regimes based on this threshold variable using a regression tree method, they find evidence in support of multiple regimes. Their convergence findings, known as club convergence, have been confirmed by others using a variety of statistical methods; see Durlauf, Johnson, and Temple (2005) (hereafter DJT). Since then, club convergence has become conventional wisdom in the growth literature.

Recently, however, researchers have even expressed reservations about the use of any growth regression models to study convergence among countries (again, see DJT (2005) for a lengthy survey of econometric issues in growth economics). They claim that regression-based convergence studies are inadequate to establish convergence because they do not analyze the entire growth distribution. To illustrate, consider a standard growth regression model

$$\ln\left(y_{it}/y_{i0}\right) = \alpha + \beta \ln\left(y_{i0}\right) + x_{it}^{'}\delta + \varepsilon_{it},$$

where y_{it} is the level of per capita income for country i at time t, y_{i0} is the level of per capita income for country i at time 0 or the level of initial per capita income (hence the left-hand-side variable is the growth rate of per capita income), x_{it} is a vector of control variables, ε_{it} is an error term, α is the intercept, β is the coefficient of interest (i.e., the convergence coefficient), and δ is a vector of parameters of control variables.

Since convergence implies that economic growth tends to slow down the richer a country becomes, a negative and statistically significant coefficient of the initial income ($\beta < 0$) in a growth regression model is usually taken as evidence of convergence. However, this coefficient merely captures the conditional mean of the growth rates; thus, it is inadequate to establish evidence of convergence.

Apparently, an analysis of the entire growth distribution requires a departure from a regression-based model. This can be done by resorting to methods such as the stochastic dominance method.[24] In brief, this method works as follows: given the growth distribution for two groups of countries, X and Y, then convergence is said to occur if the growth distribution of X fails to stochastically dominate that of Y and vice versa. If the growth distribution of X dominates that of Y or vice versa, then they diverge from each other.

To begin with, the growth distribution of X and Y can be represented by their probability density functions (PDFs).[25] X is said to stochastically dominate Y if the PDF of X lies to the right of the PDF of Y (see Figure 4.1). If their PDFs have the same means, then Asia still stochastically dominates Africa as long as the area under the PDF of X is smaller than the area under the PDF of Y (see Figure 4.2). This implies that divergence has occurred between X and Y. However, a regression-based analysis would conclude that convergence has occurred between X and Y since it is based on the mean of the distribution.

---

24 Another method which analyzes the entire distribution is called the transition matrix method due to Quah (1996).

25 To define a PDF, let X denote a random variable X and x the value that X may take. Then, the PDF of X, denoted as $f(x)$, is the probability that X takes the value x, $f(x) = P[X = x]$. If X is a continuous variable, then $f(x) = 0$. Hence, the PDF of X for a continuous variable is defined over a range of values that X may take. A related concept is known as cumulative distribution function (CDF). The CDF of X, denoted by $F(x)$, is defined as the probability that X takes a value less than or equal to x, $F(x) = P[X \leq x]$. They are related to each other in the sense that a CDF is the integral of a PDF (or a PDF is the derivative of a CDF). Since the integral of a function always exists but the derivative of a function may or may not exist, the concept of CDF is more general. Therefore, stochastic dominance techniques will be discussed in terms of a CDF in the following sections.

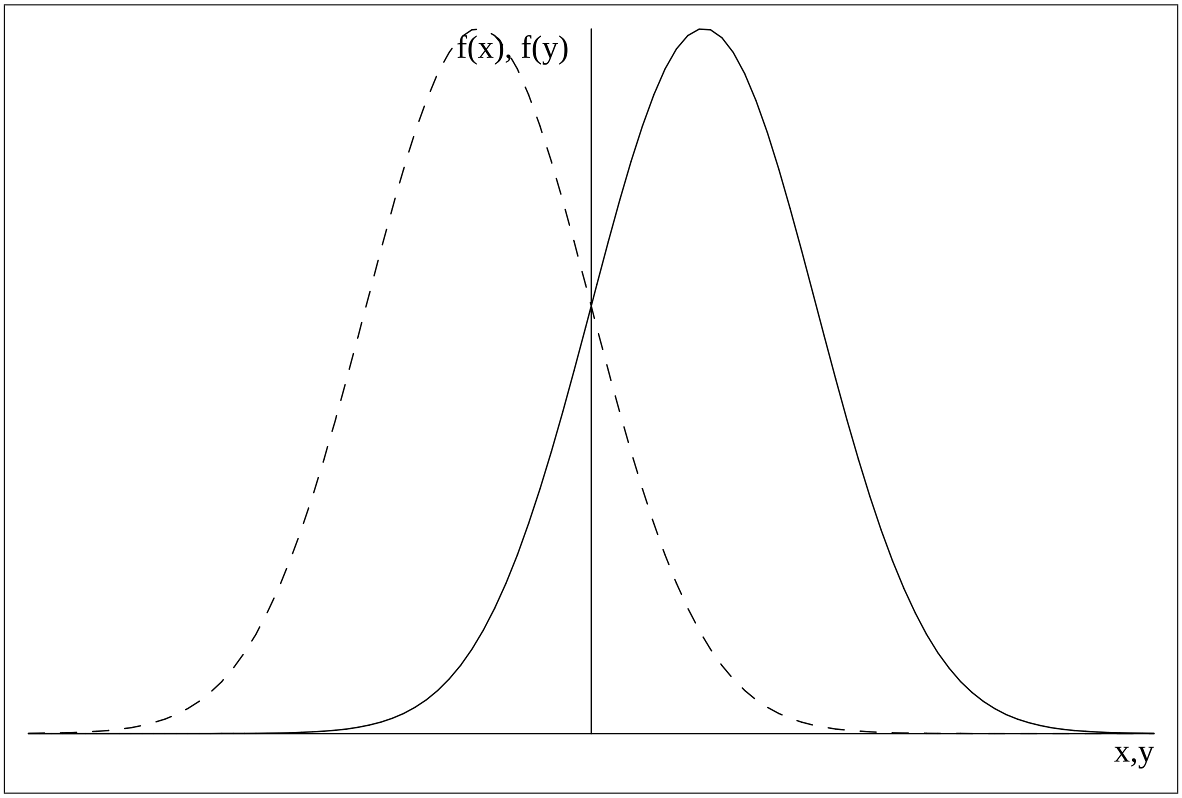

Figure 4.1: f(x) (solid line) has a larger mean than f(y) (dashed line)

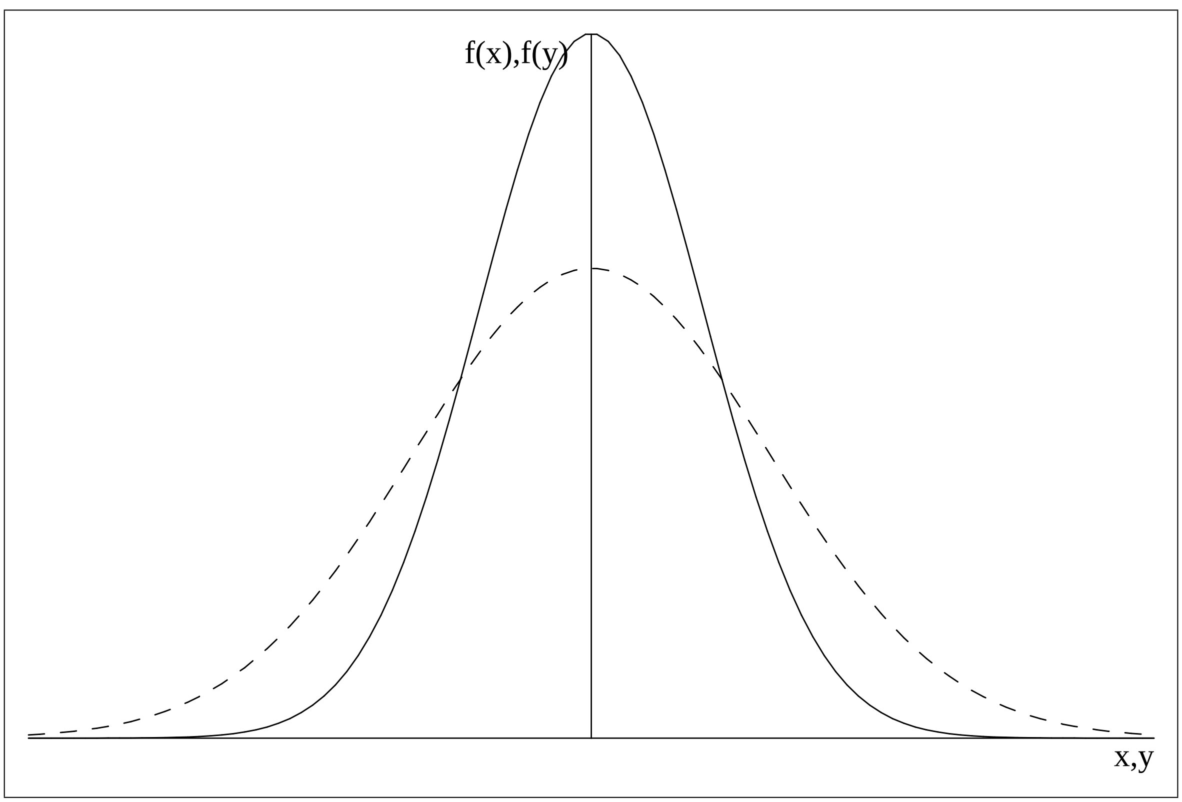

Figure 4.2: f(x) (solid line) has a smaller variance than f(y) (dashed line)

To date, the application of stochastic dominance techniques to the convergence analysis is very limited.[26] This limited application includes Anderson (2004a) and Maasoumi, Racine, and Stengos (2006) (hereafter MRS). Anderson (2004a) studies how the growth distribution of a large sample of countries evolves over multiple time periods. In most cases, he finds evidence in favor of convergence. His analysis is fairly brief in two senses. First, he considers the growth distribution of the world only; it remains to see how different regions of the world evolve over time. Second, the fact that he merely considers the world distribution precludes him from conducting interregional analyses. MRS (2006) analyze how the growth distribution of OECD vs. Non-OECD countries evolves over multiple time periods. In terms of absolute convergence, they find that Non-OECD countries fail to converge to OECD countries in many different time periods. In terms of conditional convergence, they find that Non-OECD countries appear to be converging to OECD countries. While their analysis is more elaborate, it remains to see how other country groups (Africa, Latin America, and Asia) fare against OECD countries and against one another.

This paper tries to fill the gap accordingly. In particular, we employ a particular stochastic dominance test due to Davidson and Duclos (2000) to study convergence among many different groups of countries. Our main findings can be decomposed into interregional and intertemporal analyses. For interregional analysis, we obtain evidence in support of conditional convergence in most cases considered. (We obtain inconclusive evidence in two out of 10 cases considered.) For intertemporal analysis, we obtain the following. First, African countries are converging to each other. Second, Asian countries are diverging from each other. Third, OECD countries are diverging from each other. Finally, all countries in the world are diverging from each other. (The

[26] Of course, here we refer to growth convergence studies among countries. For if we are talking about income convergence studies, there are quite a few such studies that use stochastic dominance methods. These include Bishop, Formby, and Thistle (1992), who study the distribution of income between South and Non-South U.S. states over time; Anderson (1996), who analyzes the distribution of household income in Canada over time; Maasoumi and Heshmati (2000), who investigate the Swedish income distribution over time; and Anderson (2004b), who analyzes the world income distribution over time. Stochastic dominance methods have also been applied to poverty convergence studies such as Bishop, Formby, and Smith (1993), who analyze the extent of poverty across 10 Western countries, and Madden and Smith (2000), who study the extent of poverty in Ireland.

evidence is mixed for Latin American countries.) These intertemporal findings suggest that, except for Africa, there is an intraregional divergence among countries. We take these results as evidence in favor of the club convergence hypothesis. For Africa, however, it seems that the continent as a whole is converging to a low steady-state equilibrium. Coupled with interregional findings, these intertemporal findings suggest that, once structural features are held constant, there is an intraregional (but not interregional) mobility among countries. We take these results as evidence in support of the club convergence hypothesis, which in turn, is evidence in support of the phenomenon of multiple regimes among countries.

This paper is organized as follows. Section 4.2 introduces the most frequently used measures of stochastic dominance. Section 4.3 discusses the specific stochastic dominance test employed in this paper. Section 4.4 presents the results of our convergence studies. Section 4.5 synthesizes our findings with the convergence literature. Section 4.6 concludes.

4.2 Stochastic Dominance

It is convenient to discuss the method of stochastic dominance in the context of convergence among countries. Suppose we want to analyze whether there is a tendency for two groups of countries, say the OECD and Non-OECD countries, to converge to each other in terms of growth rates over time. To apply the stochastic dominance method, we need to consider the distribution of the growth rate of per capita income for both the OECD and Non-OECD countries over a period of time, say, 1960–2000. Convergence is said to occur if a) the growth distribution of the OECD countries fails to stochastically dominate that of the Non-OECD countries, and b) the growth distribution of the Non-OECD countries fails to stochastically dominate that of the OECD countries.

How do we measure the stochastic dominance of one growth distribution over another growth distribution? There are three frequently used measures of stochastic dominance: first-order stochastic dominance (FSD), second-order stochastic dominance (SSD), and third-order stochastic dominance (TSD). As their names partially imply, all of these measures can be expressed in terms of various orders of the integral of a PDF.

FSD can be expressed in terms of the first-order integral of a PDF:

$$D^1(z) = \int_0^z f(x)\,dx,$$

where x is the growth rate of per capita income, z is the upper limit of x under consideration, $f(x)$ is the PDF of x, and $D^1(z)$ is the CDF of x up to z. (A CDF is usually denoted by $F(x)$; here it is denoted by $D^1(z)$ to signify that it is the first-order integral of a PDF.) Graphically, $D^1(z)$ can be shown as a curve whose value ranges between zero and one (see Figure 4.3).

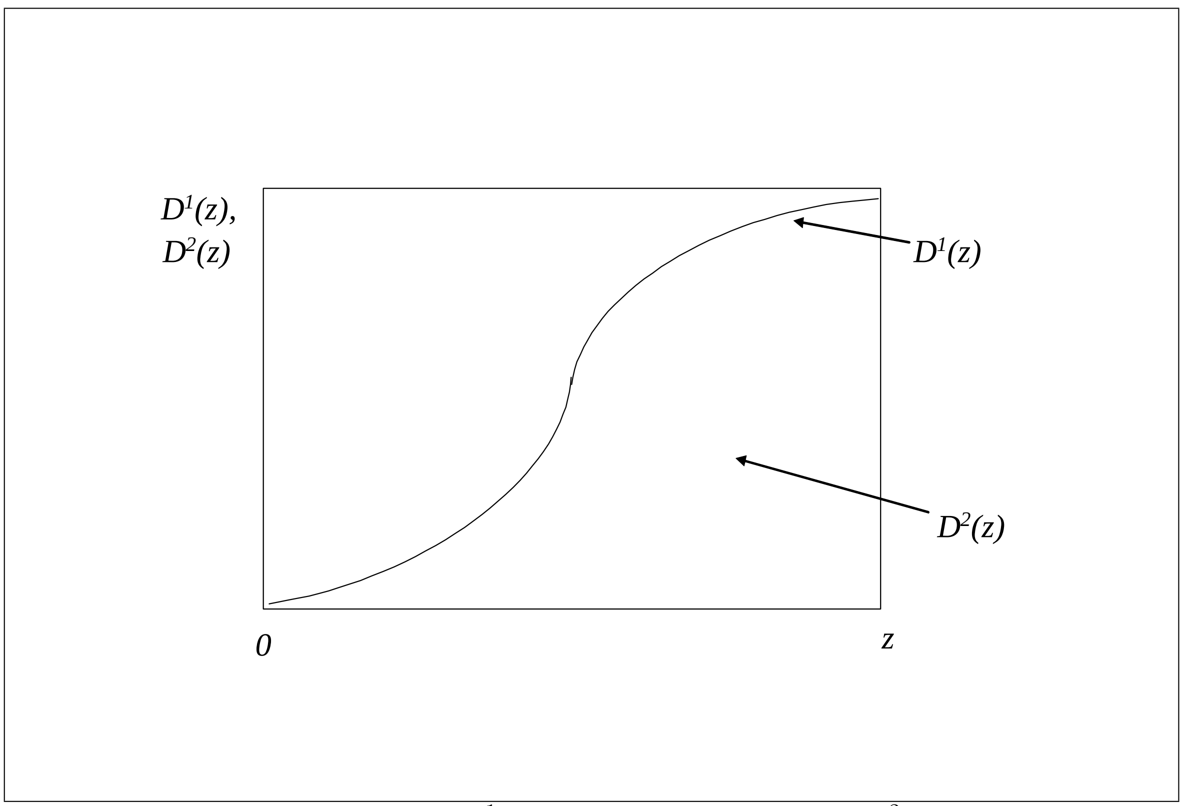

Figure 4.3: Curve $D^1(z)$ and Area Underneath It, $D^2(z)$

SSD can be expressed in terms of the second-order integral of a PDF, or the first-order integral of $D^1(z)$:

$$D^2(z) = \int_0^z D^1(x)\,dx,$$

where $D^2(z)$ is the second-order integral of a PDF. Graphically, $D^2(z)$ can be shown as the area under the curve $D^1(z)$ in Figure 4.3.

TSD can be expressed in terms of the third-order integral of a PDF, which is equivalent to the first-order integral of $D^2(z)$:

$$D^3(z) = \int_0^z D^2(x)\,dx,$$

where $D^3(z)$ is the third-order integral of a PDF. Graphically, $D^3(z)$ can be shown as the volume under the surface $D^1(z_1, z_2)$ (see Figure 4.4).

To measure the stochastic dominance of one growth distribution over another growth distribution, we need two growth distributions. For concreteness, let $D_X^1(z)$ represent the CDF of the OECD countries and $D_Y^1(z)$ the CDF of the Non-OECD countries. Then, the stochastic dominance of the OECD countries over the Non-OECD countries can be established as follows.

First, FSD of the OECD countries over the Non-OECD countries is achieved if $D_X^1(z) \leq D_Y^1(z)$ for all nonnegative z. Graphically, this can be shown by the curve $D_X^1(z)$ being no higher than the curve $D_Y^1(z)$ for all values of z (see Figure 4.5). We take this result as evidence of divergence between the OECD and Non-OECD countries. If the two distributions cross each other for some values of z, then FSD cannot be established. We take this finding as lack of evidence of divergence between the OECD and Non-OECD countries. However, we can still find evidence of divergence by resorting to SSD.

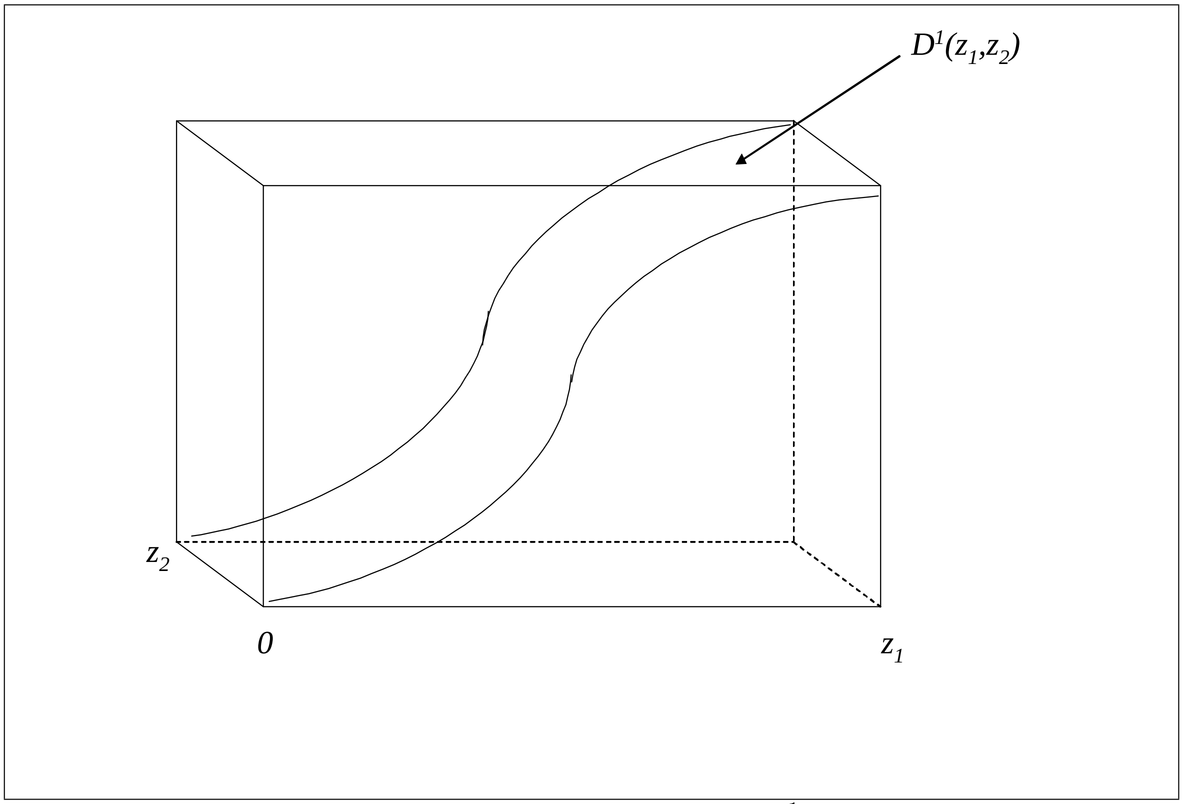

Figure 4.4: Volume under the Surface $D^1(z_1, z_2)$

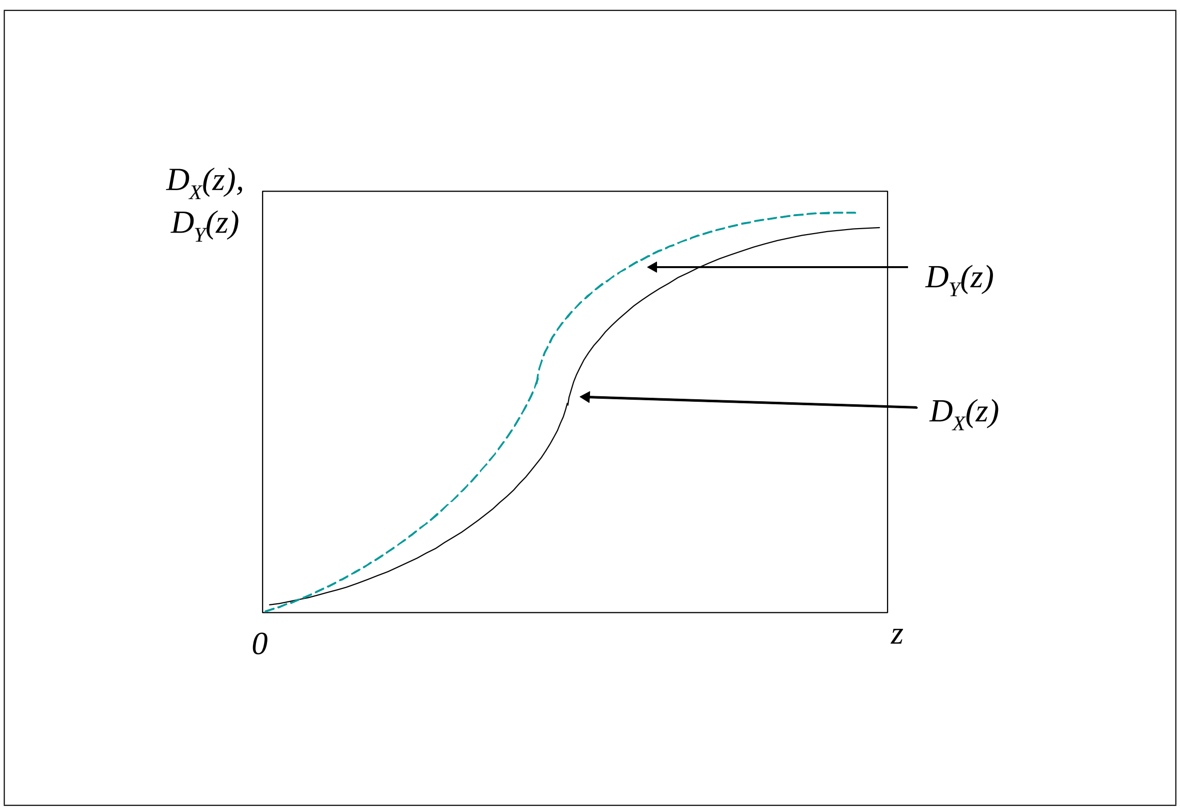

Figure 4.5: X FSD Y since $D_X(z) \leq D_Y(z)$

Second, SSD of the OECD countries over the Non-OECD countries is achieved if $D_X^2(z) \leq D_Y^2(z)$ for all z. Graphically, this can be shown by the area under the curve $D_X^1(z)$ being no greater than the area under the curve $D_Y^1(z)$ for all values of z (see Figure 4.6). As before, this result can be taken as evidence of divergence between the OECD and Non-OECD countries. If the area under $D_X^1(z)$ is greater than the area under $D_Y^1(z)$ for some z, then SSD cannot be established; we conclude that there is lack of evidence of divergence. In the latter case, we may still find evidence of divergence by resorting to TSD.

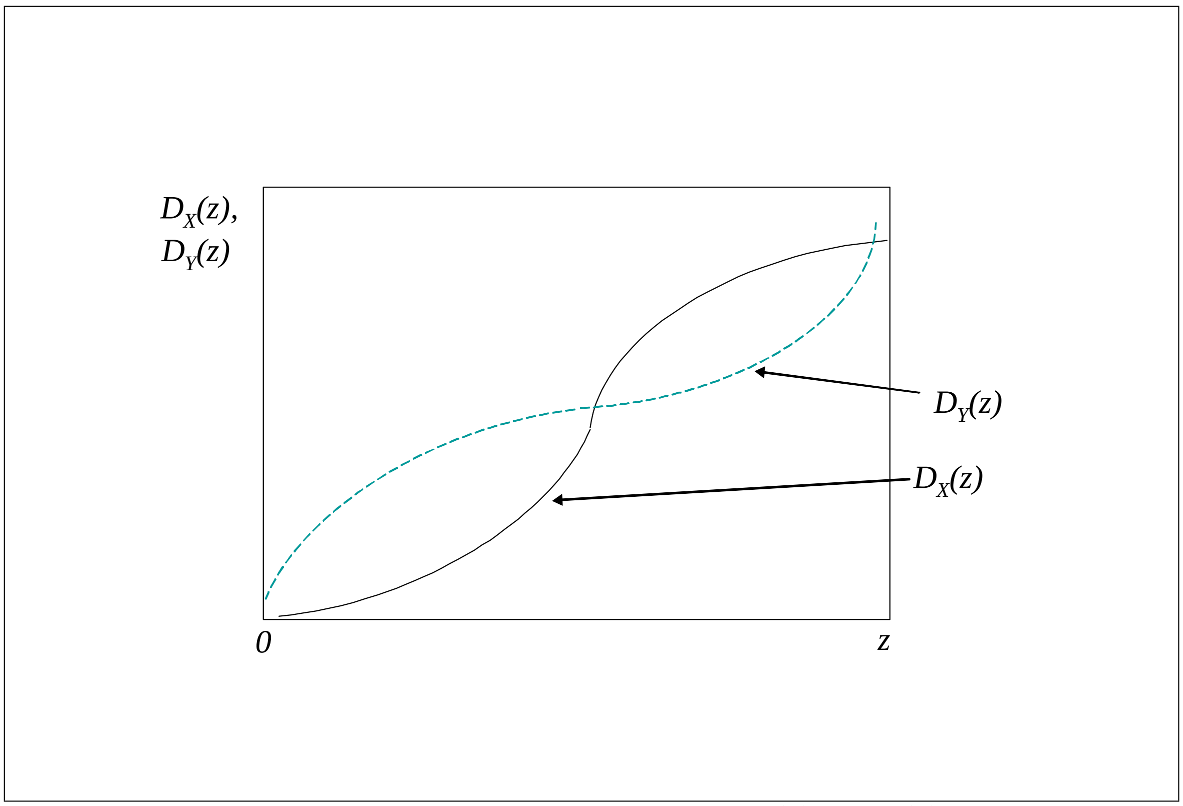

Figure 4.6: X SSD Y since Area under $D_X(z) \leq D_Y(z)$

Third, TSD of the OECD countries over the Non-OECD countries is achieved if $D_X^3(z) \leq D_Y^3(z)$ for all z. Graphically, this can be shown by the volume under the surface $D_X^1(z_1, z_2)$ being no larger than the volume under the surface $D_Y^1(z_1, z_2)$ for all values of z (see Figure 4.7). We conclude that divergence has taken place between the OECD and Non-OECD countries. If the volume under the surface $D_X^1(z_1, z_2)$ is larger than the volume under the surface

$D_Y^1(z_1, z_2)$ for some z, then TSD cannot be established. We conclude that there is lack of evidence of divergence. In this case, we may resort to the fourth-order stochastic dominance (and so on).

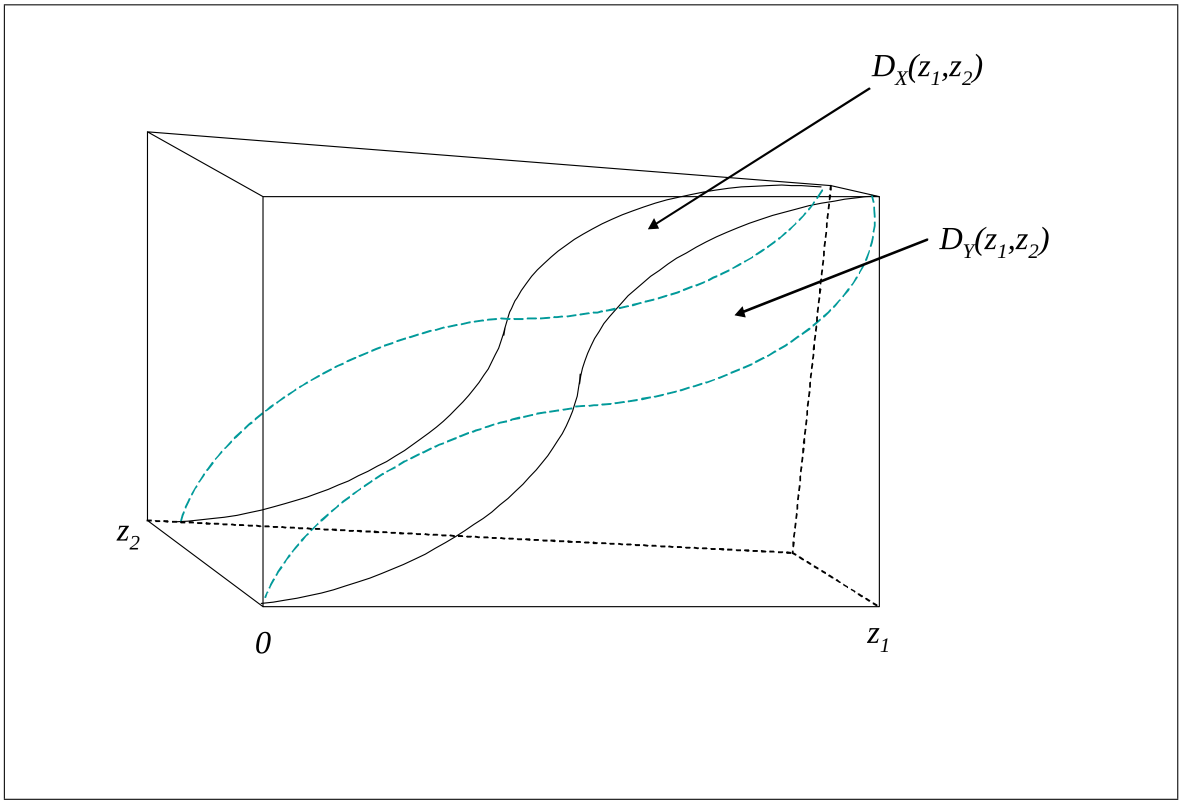

Figure 4.7: X TSD Y since Volume under $D_X(z_1, z_2) \leq D_Y(z_1, z_2)$

The stochastic dominance test was introduced by Beach and Davidson (1983) for FSD. Subsequent work by McFadden (1989) and Bishop, Formby and Thistle (1992) extended the test to SSD. More recently, the test has been extended to TSD by Anderson (1996), Davidson and Duclos (2000), and Barrett and Donald (2003) (hereafter, the last two will be respectively referred to as DD and BD). Of the first two, a recent Monte Carlo study by Tse and Zhang (2004) favors the test developed by DD (2000). Of the last two, there appears to be a trade-off between (a) the ability to take into account the covariance structure between two distributions, and (b) the ability to conduct tests at all points. DD (2000) dominate BD (2003) on the former but are dominated by BD (2003) on the latter. That is, DD (2000) take into account the covariance structure of distributions but fail to conduct tests at all points while BD (2003) can conduct tests at all points but cannot consider the covariance structure of distributions. A Monte Carlo experiment conducted by BD (2003) does not provide clear-cut guidance as to which test is better. However, Anderson (2004a) notes that the problem associated with (b) is trivial. For this reason, we are going to employ the DD (2000) test.

4.3 The Davidson-Duclos Test

In Section 4.2, we discussed how stochastic dominance can be expressed up to the third order. To generalize, let s be a positive integer up 3, $s = \{1, 2, 3\}$.[27] Then, s-order dominance can be expressed in terms of (s-1)-order dominance:

$$D^s(z) = \int_0^z D^{s-1}(x)\, dx.$$

By the inductive method, the above expression is equivalent to

$$D^s(z) = \frac{1}{(s-1)!}\int_0^z (z-x)^{s-1}\, dF(x)$$

If we have N observations of x_i, then the estimate of $D^s(z)$ can be written as

$$\begin{aligned} \widehat{D}^s(z) &= \frac{1}{(s-1)!}\int_0^z (z-x)^{s-1}\, d\widehat{F}(x) \\ &= \frac{1}{N(s-1)!}\sum_{i=1}^{N} (z-x)^{s-1}\, I(x_i \leq z) \qquad (4.1) \\ &= \frac{1}{N(s-1)!}\sum_{i=1}^{N} (z-x_i)_+^{s-1} \end{aligned}$$

where $I(.)$ in the second equality is an indicator function whose value is equal to one when its argument is true and zero otherwise, and the subscript "$_+$" in the third equality indicates that the calculation will only be carried out if $z > x_i$.

[27] As a matter of fact, s can be any positive integer. However, the stochastic dominance test employed in this paper merely considers up to TSD.

4.3.1 Economic Interpretations

The first three orders of stochastic dominance can be given distinct economic interpretations. In doing so, we will adopt the Foster-Greer-Thorbecke (FGT) classification of poverty. According to FGT, poverty can be classified into three indices: poverty incidence, poverty intensity, and poverty severity. It turns out that these indices correspond to the estimate of $D^s(z)$ in Eq.(4.1) above for $s = \{1, 2, 3\}$.

To illustrate, note that if $s = 1$, then Eq.(4.1) reduces to

$$\widehat{D}^1(z) = \frac{1}{N}\sum_{i=1}^{N}(z - x_i)_+^0 = \frac{k}{N},$$

where k is the i^{th} observation when $x_i = z$ ($i = 1, 2, ..., k, ..., N$). In the poverty dominance analysis, x_i is the level of income for household i, z is the poverty line, k is the number of households such that $x_i \leq z$ (or the number of poor households), and N is the number of households. Therefore, k/N captures the fraction of households who are poor. For this reason, this measure is called the poverty incidence index. In our analysis, x_i is the growth rate of per capita GDP for country i, z is the upper limit of x set by the researcher, k is the number of countries such that $x_i \leq z$, and N is the number of countries. Hence, k/N captures the fraction of countries whose $x_i \leq z$. By analogy, this measure is called the growth incidence index. Accordingly, FSD can be interpreted as the divergence incidence index.

If $s = 2$, then Eq.(4.1) reduces to

$$\widehat{D}^2(z) = \frac{1}{N}\sum_{i=1}^{N}(z - x_i)_+ = \frac{1}{N}\left[(z - x_1) + (z - x_2) + ... + (z - z)\right].$$

In the poverty dominance analysis, $\widehat{D}^2(z)$ captures the average of the sum of the gap of the poor's income from the poverty line. The advantage of this measure over the previous one is that, instead of simply reporting how many people are poor, it measures how poor the poor are (i.e., how far they fall below the poverty line). For this reason, this measure is called the poverty intensity index. By analogy, this measure is called the growth intensity index in our analysis. Accordingly, SSD can be interpreted as the divergence intensity index.

If $s = 3$, then Eq.(4.1) reduces to

$$\widehat{D}^3(z) = \frac{1}{N}\sum_{i=1}^{N}(z - x_i)_+^2 = \frac{1}{N}\left[(z - x_1)^2 + (z - x_2)^2 + ... + (z - z)^2\right].$$

In the poverty dominance analysis, $\widehat{D}^3(z)$ captures the average of the sum of the squared gap of the poor's income from the poverty line. The advantage of this measure over the preceding one can be inferred from the fact that the gap between the poverty line and a particular household income is weighted more the smaller the income is. Hence, instead of simply reporting how poor the poor are, it measures how severe their plight is. For this reason, this measure is called the poverty severity index. By analogy, this measure is called the growth severity index in our analysis. Accordingly, TSD can be interpreted as the divergence severity index.

4.3.2 Estimation and Inference

Consider two country groups, X and Y, where the number of countries in X and Y are denoted by, respectively, N and M. Then, the estimates of $D_X^1(z)$ and $D_Y^1(z)$ can be written as

$$\widehat{D}_X^s(z) = \frac{1}{N(s-1)!}\sum_{i=1}^{N}(z - x_i)_+^{s-1} \; and \; \widehat{D}_Y^s(z) = \frac{1}{M(s-1)!}\sum_{i=1}^{M}(z - y_i)_+^{s-1} \tag{4.2}$$

where x_i is the growth rate for country i in X and y_i is the growth rate for country i in Y.

To test for stochastic dominance, DD (2000) propose two types of tests; the first is based on the Wald statistic and the second is based on the t-statistic. In both tests, we test the following hypotheses:

$$H_0 : \widehat{D}_X^s(z) - \widehat{D}_Y^s(z) \leq 0 \; vs. \; \widehat{D}_X^s(z) - \widehat{D}_Y^s(z) > 0 \tag{4.3}$$

The null hypothesis states that the growth distribution of X and Y weakly diverges from each other in the sense that X may potentially dominate Y while the alternative hypothesis states the growth distribution of X and Y strictly diverges from each other in the sense that Y strictly dominates X (see the operators "$\leq$" and "$>$"). If the null hypothesis can be rejected, then we conclude that Y dominates X. If the null hypothesis cannot be rejected, then there is a possibility that X dominates Y. In this ambiguous case, we may reverse the position of X and Y. If the null hypothesis can be rejected, then we conclude that X dominates Y. If the null hypothesis cannot be rejected once again, then we conclude that X and Y converge to each other.

The null hypothesis can be rejected if the s-order dominance can be established. To determine whether the s-order dominance can be established, we utilize a modified Wald statistic proposed by Wolak (1989):

$$W^s(z) = \frac{\left[\widehat{D}_X^s(z) - \widehat{D}_Y^s(z) - u\right]^2}{Var\left[\widehat{D}_X^s(z) - \widehat{D}_Y^s(z)\right]}, \tag{4.4}$$

where $\widehat{D}_X^s(z)$ and $\widehat{D}_Y^s(z)$ are given by Eq.(4.2), $Var\left[\widehat{D}_X^s(z) - \widehat{D}_Y^s(z)\right]$ is given by

$$Var\left(\widehat{D}_X^s(z) - \widehat{D}_Y^s(z)\right) = Var\left[\widehat{D}_X^s(z)\right] + Var\left[\widehat{D}_Y^s(z)\right] - 2\,Cov\left[\widehat{D}_X^s(z), \widehat{D}_Y^s(z)\right], \tag{4.5}$$

$$\text{where } Var\left(\widehat{D}_X^s(z)\right) = \frac{1}{N}\left[\frac{1}{((s-1)!)^2}\frac{1}{N}\sum_{i=1}^{N}\left[(z-x_i)_+^{s-1}\right]^2 - \left[\widehat{D}_X^s(z)\right]^2\right], \tag{4.6}$$

$$Var\left(\widehat{D}_Y^s(z)\right) = \frac{1}{M}\left[\frac{1}{((s-1)!)^2}\frac{1}{M}\sum_{i=1}^{M}\left[(z-y_i)_+^{s-1}\right]^2 - \left[\widehat{D}_Y^s(z)\right]^2\right], \tag{4.7}$$

$$Cov\left(\widehat{D}_X^s\left(z\right), \widehat{D}_Y^s\left(z\right)\right) = \frac{1}{N}\left[\frac{1}{\left(\left(s-1\right)!\right)^2}\frac{1}{N}\sum_{i=1}^{N}\left(z-x_i\right)_+^{s-1}\left(z-y_i\right)_+^{s-1} - \widehat{D}_X^s\left(z\right)\widehat{D}_Y^s\left(z\right)\right],^{28} \tag{4.8}$$

and u is some weight variable. Wolak (1989) shows that, under the null hypothesis, $W^s(z)$ is asymptotically distributed as a mixed chi-squared random variable (without u, $W^s(z)$ is asymptotically distributed as a pure chi-squared random variable).

BD (2003) note that Eq.(4.4) implies that we need to calculate the solutions to a large number of quadratic programming problems in order to estimate the weight variable, u. This means that we need to use a simulated Wald statistic. The procedure is as follows: pick some values of z, estimate the weight variable, calculate the Wald statistic, and then calculate its associated p-value. The decision rule can be stated as follows: if p-value $<$ 0.05, then the null hypothesis can be rejected at the 5% level and we conclude that X and Y diverge from each other.

An alternative to using the Wald statistic would be to use the t-statistic:

$$t^s\left(z\right) = \frac{\left[\widehat{D}_X^s\left(z\right) - \widehat{D}_Y^s\left(z\right)\right]}{\sqrt{Var\left[\widehat{D}_X^s\left(z\right) - \widehat{D}_Y^s\left(z\right)\right]}}, \tag{4.9}$$

and test the hypotheses of the form

$$H_0 : \widehat{D}_X^s\left(z\right) - \widehat{D}_Y^s\left(z\right) = 0 \; vs. \; \widehat{D}_X^s\left(z\right) - \widehat{D}_Y^s\left(z\right) \neq 0. \tag{4.10}$$

DD (2000) show that, under the null hypothesis, $t^s(z)$ is asymptotically distributed as a standard normal variable. This implies that we can use a regular t test. As noted by Tse and Zhang (2004), however, this means that we need to test the hypothesis for all values of z, which

[28] If we assume that X and Y are independent, then $Cov(.) = 0$. If we assume that X and Y are dependent, then $Cov(.) \neq 0$. When we incorporate the covariance term, however, the matrix of $Var(.)$ in Eq.(4.5) turns out to be not positive definite in some cases. Hence, we follow the standard practice of assuming $Cov(.) = 0$.

is impossible. To get around this problem, DD (2000) suggest that we pick some values of z, calculate their corresponding t-values, but use a conservative critical value from a studentized maximum modulus (SMM) distribution, which was tabulated in Stoline and Ury (1979).

Instead of testing the null hypothesis in (4.10), we may test the null hypothesis in (4.3). In this case, instead of testing the null hypothesis based on individual t-values, we pick the maximal t-value (MT) out of those individual t-values, compare it against a critical value from the SMM, and then calculate its associated p-value. As before, if p-value < 0.05, then the null hypothesis in (4.4) can be rejected at the 5% level and we conclude that X and Y diverge. In this paper, we follow BD (2003) by using the MT statistic approach.

4.4 Empirical Analysis

Our analysis is based on the growth rates of GDP per capita for a sample of 100 countries during the period 1960–2000. Data are obtained from the Penn World Table, version 6.1 (PWT6.1). To conform to stochastic dominance theory, these data have been adjusted to be nonnegative. This is done by shifting the PDF of a growth distribution to the right in such a way that the minimum value of growth rates is set to zero.

A few points should be mentioned about the data. First, data for Singapore and Taiwan are available up to, respectively, 1996 and 1998 only; we extrapolate the missing observations by assuming that each country grew at the same rate as Hong Kong. This assumption is based on the fact that these countries are at a similar stage of economic development. Second, data for Germany are available during the period 1970-2000 only; we extrapolate the missing observations by assuming that Germany grew at the same rate as Japan. This assumption is based on the fact that both of these war-devastating countries grew at similar rates during the period 1970-1980.

4.4.1 Basic Analysis

Now we apply this stochastic dominance method to the study of growth convergence among countries. Our study is divided into interregional and intertemporal analyses. To conduct the interregional analysis, our 100-country sample is partitioned into four regions: Africa (36 countries), Latin America (21 countries), Asia (20 countries), and OECD countries (24 countries).

4.4.1.1 Interregional Analysis

We begin by testing whether the null hypothesis can be rejected between any two different regions. If these two regions are Africa and the rest of the world (ROW), then the test is implemented as

$$H_0 : \widehat{D}^s_{\text{Africa}}(z) - \widehat{D}^s_{\text{ROW}}(z) \leq 0 \; vs. \; \widehat{D}^s_{\text{Africa}}(z) - \widehat{D}^s_{\text{ROW}}(z) > 0.$$

The null hypothesis says that Africa and ROW weakly diverge from each other in the sense that Africa may potentially dominate ROW, while the alternative hypothesis says that Africa and ROW strictly diverge from each other in the sense that ROW surely dominates or outperforms Africa. Therefore, if the null can be rejected, then we conclude that ROW dominates Africa. If the null cannot be rejected, we may reverse the position of Africa and ROW. Now if the null can be rejected, then we conclude that Africa dominates ROW. If the null cannot be rejected, then we conclude that Africa and ROW converge to each other. Instead of stating "X and Y diverge in the sense that X dominates Y," we will simply state "X dominates Y" in the following discussion. We employ three types of tests: the conservative MT statistic (CMT), the simulated MT statistic (SMT), and the Wald statistic (Wald).

We begin by testing whether the null hypothesis can be rejected between Africa and other country groups; the results of our analysis are presented in Table 4.1. For Africa vs. ROW, we find that the null can be rejected at any order at 1% for all tests. Hence, we conclude that ROW dominates Africa. For Africa vs. OECD countries, we find that the null can be rejected at any order at 1% for all tests. Hence, we conclude that OECD countries dominate Africa. For Africa vs. Asia, again the null can be rejected at any order at 1% for all tests. Hence, we conclude that

Asia dominates Africa. For Africa vs. Latin America, the null can be rejected at FSD at 1% but cannot be rejected at SSD and TSD even at 5% based on CMT. Using SMT and Wald statistics, however, the null can be rejected at any order at least at 5%. Hence, we have mixed evidence that Latin America dominates Africa.

Table 4.1: Africa vs. Non-Africa

		Original Position		
Region	Test	FSD	SSD	TSD
ROW	CMT	0.000**	0.000**	0.000**
	SMT	0.000**	0.000**	0.000**
	Wald	0.000**	0.000**	0.000**
OECD	CMT	0.000**	0.000**	0.000**
	SMT	0.000**	0.000**	0.000**
	Wald	0.000**	0.000**	0.000**
Asia	CMT	0.000**	0.000**	0.000**
	SMT	0.000**	0.000**	0.000**
	Wald	0.000**	0.000**	0.000**
Latin	CMT	0.010**	0.068	0.108
	SMT	0.007**	0.021*	0.032*
	Wald	0.012*	0.026*	0.031*

Notes: Africa vs. Non-Africa refers to the null hypothesis that Africa weakly dominates Non-Africa. If the null can be rejected, then we conclude that Non-Africa strictly dominates Africa. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

We proceed with the test of whether the null hypothesis can be rejected between Latin America and other country groups (the results are reported in Table 4.2). For Latin America vs. ROW, the null can be rejected at FSD at 5% but cannot be rejected at SSD and TSD even at 5% for all tests. When we reverse their position, we find that the null cannot be rejected at any order even at 5% for all tests. Taken together, both results imply that Latin America and ROW converge to each other. For Latin America vs. OECD countries, the null can be rejected at any order at 1% for all tests. Hence, we conclude that OECD countries dominate Latin America. We obtain similar results for Latin America vs. Asia. That is, the null can be rejected at any order at 1% for all tests. Hence, we conclude that Asia dominates Latin America.

Table 4.2: Latin America vs. Non-Latin America

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	0.014*	0.296	0.670	0.990	0.988	0.997
	SMT	0.016*	0.086	0.195	0.879	0.583	0.605
	Wald	0.011*	0.084	0.187	0.857	0.550	0.542
OECD	CMT	0.000**	0.000**	0.000**	–	–	–
	SMT	0.000**	0.000**	0.000**	–	–	–
	Wald	0.000**	0.000**	0.000**	–	–	–
Asia	CMT	0.000**	0.000**	0.000**	–	–	–
	SMT	0.000**	0.000**	0.000**	–	–	–
	Wald	0.000**	0.000**	0.000**	–	–	–

Notes: Latin America vs. Non-Latin America refers to the null hypothesis that Latin America weakly dominates Non-Latin America. If the null can be rejected, then we conclude that Non-Latin America strictly dominates Latin America. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

Next, we test whether the null hypothesis can be rejected between Asia and other country groups (see Table 4.3). For Asia vs. ROW, the null cannot be rejected at any order even at 5% for all tests. When we reverse their position, we find that the null can be rejected at any order at 1% for all tests. Collectively, both results indicate that Asia dominates ROW. For Asia vs. OECD countries, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, we find that the null can be rejected at FSD at 5% for all tests but cannot be rejected at SSD and TSD at 5% for all tests. Both of these results show that Asia and OECD countries converge to each other.

Finally, for OECD countries vs. ROW (see Table 4.4), we find that the null cannot be rejected at any order even at 5% for all tests. When we reverse their position, we find that the null can be rejected at any order at 1% for all tests. Hence, we conclude that OECD countries dominate ROW.

To recap, our interregional analysis produces the following findings. First, convergence occurs between a) Latin America and ROW, and b) Asia and OECD countries. Second, divergence occurs between seven pairs of regions: a) Africa and ROW, b) Africa and OECD countries, c) Africa and Asia, d) Latin America and OECD countries, e) Latin America and Asia, f) Asia and ROW, and g)

OECD countries and ROW. Third, there is inconclusive evidence (of convergence or divergence) for Africa and Latin America.

Table 4.3: Asia vs. Non-Asia

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	1.000	1.000	1.000	0.000**	0.000**	0.000**
	SMT	1.000	1.000	0.997	0.000**	0.000**	0.000**
	Wald	0.904	0.788	0.698	0.000**	0.000**	0.000**
OECD	CMT	0.652	0.765	0.776	0.036*	0.502	0.908
	SMT	0.394	0.340	0.292	0.035*	0.206	0.418
	Wald	0.389	0.312	0.276	0.037*	0.191	0.380

Notes: Asia vs. Non-Asia refers to the null hypothesis that Asia weakly dominates Non-Asia. If the null can be rejected, then we conclude that Non-Asia strictly dominates Asia. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

Table 4.4: OECD vs. ROW

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	1.000	1.000	1.000	0.000**	0.000**	0.000**
	SMT	0.982	1.000	0.996	0.000**	0.000**	0.000**
	Wald	0.909	0.742	0.679	0.000**	0.000**	0.000**

Notes: OECD vs. ROW refers to the null hypothesis that OECD weakly dominates ROW. If the null can be rejected, then we conclude that ROW strictly dominates OECD. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

4.4.1.2 Intertemporal Analysis

Instead of making interregional comparisons, we could make intertemporal comparisons (i.e., we compare the same country group over two time periods). We test whether the growth distribution of a given country group during the periods 1961-1980 and 1981-2000 diverge. The test is implemented as follows:

$$H_0 : \widehat{D}^s_{61-80}(z) - \widehat{D}^s_{81-20}(z) \leq 0 \; vs. \; \widehat{D}^s_{61-80}(z) - \widehat{D}^s_{81-20}(z) > 0.$$

The null says that 1961-1980 and 1981-2000 weakly diverge in the sense that 1961-1980 may potentially dominate 1981-2000 and the alternative says that 1961-1980 and 1981-2000 strictly diverge in the sense that the second period dominates the first period. Therefore, if the null can be rejected, then we conclude that 1981-2000 dominates 1961-1980. If the null cannot be rejected, then we may reverse their position. Now if the null can be rejected, then we conclude that 1961-1980 dominates 1981-2000. If the null cannot be rejected, then we conclude that the region under study converges over time.

For Africa, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, we find that the null can be rejected at FSD at 5% for all tests but cannot be rejected at SSD and TSD at 5% for all tests (see Table 4.5, which reports all results on intertemporal analysis). Taken together, both results indicate that African countries have converged to each other over time. For Latin America, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, we find that the null can be rejected at any order at least at 5% for all tests. Hence, the first period dominates the second period. We take this result as evidence that Latin American countries have diverged from each other over time.

For Asia, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, we obtain similar results (i.e., the null cannot be rejected at any order even at 5% for all tests). We take these results as evidence that Asian countries have converged to each other over time. For OECD countries, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, we find that the null can be rejected at FSD at least at 5% for all tests but cannot be rejected at SSD and TSD even at 5% for all tests. Taken together, both results imply that OECD countries have converged to each other over time.

Finally, we test whether the null hypothesis can be rejected for a large sample of 100 countries. We find that the null cannot be rejected at any order even at 5%. When we reverse their position, we find similar results (i.e., the null cannot be rejected at any order even at 5% for all tests). Hence, the "world" has converged over time.

Table 4.5: 1961-1980 vs. 1981-2000

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
Africa	CMT	0.996	0.773	0.524	0.028*	0.482	0.642
	SMT	0.895	0.322	0.155	0.025*	0.186	0.203
	Wald	0.868	0.296	0.148	0.036*	0.165	0.191
Latin	CMT	1.000	1.000	1.000	0.009**	0.006**	0.020*
	SMT	0.995	0.987	0.933	0.007**	0.004**	0.003**
	Wald	0.879	0.722	0.623	0.003**	0.004**	0.005**
Asia	CMT	0.871	0.932	0.870	0.736	1.000	1.000
	SMT	0.608	0.441	0.348	0.467	0.754	0.840
	Wald	0.568	0.400	0.312	0.480	0.717	0.651
OECD	CMT	0.996	0.917	0.851	0.010**	0.486	0.745
	SMT	0.897	0.434	0.319	0.008**	0.182	0.235
	Wald	0.875	0.408	0.292	0.014*	0.166	0.225
World	CMT	1.000	1.000	1.000	0.084	0.183	0.220
	SMT	0.990	0.880	0.752	0.063	0.057	0.065
	Wald	0.869	0.709	0.635	0.061	0.054	0.056

Notes: 1961-1980 vs. 1981-2000 refers to the null hypothesis that 1961-1980 weakly dominates 1981-2000. If the null can be rejected, then we conclude that 1981-2000 strictly dominates 1961-1980. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

In summary, our intertemporal analysis yields two main results. First, intraregional convergence occurs for Africa, Asia, OECD countries, and the world. Second, intraregional divergence occurs for Latin America only.

4.4.2 Robustness Analysis

Hitherto, our convergence dominance analysis has been based on actual growth rates of per capita GDP. Since no attempt is made to control for structural features of the economies (in the spirit of conditional convergence), our findings can merely be construed in terms of absolute versus club convergence. If this deficiency can be overcome, then our results can be interpreted in terms of conditional versus club convergence. MRS (2006) overcome this deficiency by utilizing residual (instead of actual) growth rates. Their method can be summarized as follows. First, they run a linear parametric regression in the manner of MRW (1992). Second, they employ a specification test developed by Hsiao, Li, and Racine (2006) to determine whether this linear specification is rejected. If the model is not rejected, then they take residuals from this regression and conduct a stochastic dominance analysis. If the model is rejected, then they proceed with a nonparametric regression, take residuals from this regression, and conduct a stochastic dominance analysis.

Following their footsteps, we run a MRW regression for a large sample of countries. Due to data constraints, we end up with 76 countries with the following regional decompositions: Africa (16 countries), Latin America (21 countries), Asia (18 countries), and OECD countries (22 countries). To enable us to conduct interregional and intertemporal stochastic dominance analyses, we divide the sample according to three periods of analysis: 1960-2000, 1961-1980, and 1981-2000. The MRW specification based on the 76-country sample is rejected by the Hsiao-Li-Racine test for the periods 1960-2000 and 1961-1980, but not for the period 1981-2000. Therefore, for the periods 1960-2000 and 1961-1980, we proceed by running a nonparametric regression model and extracting its residual growth rates. For the period 1981-2000, we obtain residual growth rates from the MRW regression.[29] As before, these residuals have been adjusted to be nonnegative.

4.4.2.1 Interregional Analysis

Using these residual growth rates, we begin by testing whether the null hypothesis can be rejected between Africa and other country groups. For Africa vs. ROW, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, we obtain similar results: the null cannot be rejected at any order at 5% (see Table 4.6). Taken together, these results imply that Africa converges to ROW. Note that these results contradict our baseline results (results

[29] I am grateful to Dr. Jeff Racine for providing me with the data on residual growth rates from parametric and nonparametric regressions.

obtained based on actual growth rates), where we conclude that ROW dominates Africa. We take these results as evidence in favor of the conditional convergence hypothesis (and against the club convergence hypothesis).

Table 4.6: Africa vs. Non-Africa

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	0.615	0.941	0.916	0.950	1.000	1.000
	SMT	0.382	0.408	0.301	0.713	0.784	0.815
	Wald	0.368	0.370	0.291	0.702	0.659	0.572
OECD	CMT	0.781	0.698	0.744	0.976	1.000	1.000
	SMT	0.506	0.231	0.173	0.797	0.852	0.899
	Wald	0.478	0.218	0.179	0.778	0.678	0.569
Asia	CMT	0.696	0.890	0.896	0.953	1.000	1.000
	SMT	0.433	0.359	0.293	0.720	0.828	0.851
	Wald	0.327	0.330	0.273	0.719	0.683	0.591
Latin	CMT	0.369	0.809	0.789	0.825	0.976	0.999
	SMT	0.219	0.331	0.238	0.548	0.535	0.636
	Wald	0.253	0.296	0.232	0.559	0.511	0.576

Notes: Africa vs. Non-Africa refers to the null hypothesis that Africa weakly dominates Non-Africa. If the null can be rejected, then we conclude that Non-Africa strictly dominates Africa. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

We repeat the analysis for Africa vs. OECD countries. For all tests, the null cannot be rejected at any order at 5%. Since we obtain similar results when we reverse their position (see Table 4.6), we conclude that Africa converges to OECD countries (note that these results stand in stark contrast to those in our baseline analysis). As before, we take these results as evidence in support of the conditional convergence hypothesis.

When we conduct the analysis for Africa vs. Asia (for original and reversed positions), we obtain results which imply that Africa converges to Asia (see Table 4.6). We interpret these findings as evidence in favor of conditional convergence. Finally, we conduct the analysis for

Africa vs. Latin America (again for original and reversed positions). We obtain results which suggest that Africa converges to Latin America (see Table 4.6). These results are taken as evidence in support of conditional convergence.

We proceed with the test of whether the null hypothesis can be rejected between Latin America and other country groups. For Latin America vs. ROW, the null cannot be rejected at any order at 5% for all tests. Since these results continue to hold when their position is reversed (see Table 4.7), they suggest that Latin America converges to ROW (note that these results concur with those we obtain in the baseline analysis). We construe these results as evidence that lends support to the conditional convergence hypothesis.

Table 4.7: Latin America vs. Non-Latin America

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	0.150	0.482	0.741	0.618	1.000	1.000
	SMT	0.097	0.150	0.203	0.352	0.810	0.816
	Wald	0.106	0.129	0.191	0.366	0.661	0.595
OECD	CMT	0.116	0.093	0.182	0.454	1.000	1.000
	SMT	0.080	0.029*	0.040*	0.279	0.875	0.943
	Wald	0.085	0.025*	0.036*	0.213	0.673	0.592
Asia	CMT	0.425	0.488	0.498	0.822	1.000	1.000
	SMT	0.257	0.192	0.142	0.549	0.918	0.919
	Wald	0.200	0.157	0.130	0.557	0.736	0.646

Notes: Latin America vs. Non-Latin America refers to the null hypothesis that Latin America weakly dominates Non-Latin America. If the null can be rejected, then we conclude that Non-Latin America strictly dominates Latin America. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

For Latin America vs. OECD countries, the null cannot be rejected at any order for CMT test. When we employ SMT and Wald tests, however, the null cannot be rejected at FSD at 5% but can be rejected at SSD and TSD at 5%. These results are anomalous given the fact that lower-order dominance implies higher-order dominance but not the other way around (see Table 4.7). The anomaly of these results precludes us from drawing any definitive conclusion from this exercise. At best, the evidence is inconclusive.

Next, we repeat the analysis for Latin America vs. Asia. For all tests, the null cannot be rejected at any order at 5%. Similar results are obtained when their position is reversed (see Table 4.7). Since these results imply that Latin America converges to Asia (note that these results differ from those obtained in the baseline analysis), they are construed as evidence in support of conditional convergence.

Now we test whether the null hypothesis can be rejected between Asia and other country groups. For Asia vs. ROW, the null cannot be rejected at any order at 5%, and similar results are obtained when their position is reversed (see Table 4.8). Since these results suggest that Asia converges to ROW (and they contradict our baseline results), they are taken as evidence in favor of club convergence.

Table 4.8: Asia vs. Non-Asia

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	0.986	1.000	1.000	0.641	0.881	0.921
	SMT	0.838	0.846	0.851	0.399	0.364	0.357
	Wald	0.826	0.703	0.616	0.348	0.340	0.309
OECD	CMT	0.985	0.740	0.772	0.640	0.972	0.999
	SMT	0.833	0.286	0.229	0.391	0.522	0.674
	Wald	0.813	0.258	0.229	0.391	0.493	0.606

Notes: Asia vs. Non-Asia refers to the null hypothesis that Asia weakly dominates Non-Asia. If the null can be rejected, then we conclude that Non-Asia strictly dominates Asia. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

For Asia vs. OECD countries, we obtain results similar to those obtained between Asia vs. ROW: the null cannot be rejected at any order at 5% in the original position, and similar results are obtained when their position is reversed (see Table 4.8). These results, which agree with our baseline results, imply that Asia converges to OECD countries. We take these results as evidence in favor of conditional convergence.

Finally, we test whether the null hypothesis can be rejected between OECD countries and ROW. As shown in Table 4.9, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, the results are not so clear-cut. For CMT test, the null can be rejected at FSD and SSD at 5% test but cannot be rejected at TSD at 5%. For SMT and Wald tests, the null can be rejected at all orders at 5%. We take these mixed results as inconclusive evidence in favor of conditional convergence.

Table 4.9: OECD vs. ROW

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	0.578	1.000	1.000	0.019*	0.046*	0.091
	SMT	0.390	0.812	0.921	0.017*	0.016*	0.027*
	Wald	0.415	0.708	0.633	0.028*	0.016*	0.024*

Notes: OECD vs. ROW refers to the null hypothesis that OECD weakly dominates ROW. If the null can be rejected, then we conclude that ROW strictly dominates OECD. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

One may argue that the contradicting results between baseline and robustness analyses might be due to the difference in sample size (100-country vs. 76-country sample). To accommodate this objection, we repeat our baseline analysis with this reduced 76-country sample. It turns out that our baseline results are broadly robust to the difference in sample size (see Tables F1–F5 in Appendix F). That is, the baseline results are similar except for a) Africa vs. Latin America (where there is evidence of convergence between them) and b) Latin America vs. ROW (where there is inconclusive evidence of convergence or divergence). Therefore, we conclude that these contradicting results are due to whether or not structural features of the economies are held constant. Once these structural features are accounted for, evidence of interregional divergence disappears in most cases considered. We take all of these findings as overwhelming evidence in favor of conditional convergence.

4.4.2.2 Intertemporal Analysis

As before, we proceed by making intertemporal comparisons of each country group and report their results in Table 4.10. For Africa, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, we obtain similar results for CMT test. For SMT and Wald tests, however, the null can be rejected at FSD at 5% but cannot be rejected at SSD and TSD at 5% (note that these results accord with our baseline results). Therefore, we conclude that African countries are converging to each other over time. We take these results as evidence in favor of conditional convergence.

Table 4.10: 1961-1980 vs. 1981-2000

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
Africa	CMT	0.996	0.936	0.830	0.051	0.400	0.652
	SMT	0.894	0.443	0.308	0.045*	0.131	0.196
	Wald	0.870	0.416	0.276	0.038*	0.123	0.183
Latin	CMT	1.000	1.000	1.000	0.045*	0.079	0.110
	SMT	0.976	0.942	0.921	0.043*	0.025*	0.030*
	Wald	0.877	0.716	0.634	0.009**	0.027*	0.029*
Asia	CMT	1.000	1.000	1.000	0.020*	0.014*	0.043*
	SMT	0.983	0.996	0.975	0.018*	0.007**	0.011*
	Wald	0.881	0.716	0.622	0.013*	0.007**	0.011*
OECD	CMT	1.000	1.000	1.000	0.000**	0.000**	0.000**
	SMT	1.000	0.994	0.935	0.000**	0.000**	0.000**
	Wald	0.909	0.703	0.612	0.000**	0.000**	0.000**
World	CMT	1.000	1.000	1.000	0.000**	0.000**	0.000**
	SMT	0.993	0.988	0.971	0.000**	0.000**	0.000**
	Wald	0.879	0.703	0.622	0.000**	0.000**	0.000**

Notes: 1961-1980 vs. 1981-2000 refers to the null hypothesis that 1961-1980 weakly dominates 1981-2000. If the null can be rejected, then we conclude that 1981-2000 strictly dominates 1961-1980. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

For Latin America, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, we find the following. For SMT and Wald tests, the null can be rejected at all orders at 5%. For CMT test, however, the null can be rejected at FSD at 5% but cannot be rejected at

SSD and TSD at 5%. We take these mixed findings (which contradict our baseline findings) as inconclusive evidence in favor of conditional convergence.

For Asia, the null cannot be rejected at any order at 5% for all tests. When we reverse their position, the null can be rejected at all orders at least at 5% for all tests. These findings, which contradict our baseline findings, suggest that Asian countries are diverging from each other over time. We take these findings as evidence in favor of club convergence.

For OECD countries, the null cannot be rejected at any order for all tests. When we reverse their position, the null can be rejected at all orders at 1% for all tests. As in the case of Asia, these findings imply that OECD countries are diverging from each other, thereby providing support for club convergence.

Finally, we conduct the analysis for the world. We obtain results similar to those found in the OECD countries (i.e., the null cannot be rejected in the original position but can be rejected in the reversed position). As in the OECD case, these results suggest that countries in the world are diverging from each other. In this case, too, there is support for club convergence.

As before, the contradicting results between these baseline and robustness analyses can be attributed to the difference in sample size. To confirm whether or not this is the case, we repeat our baseline analysis with the 76-country sample. It turns out that the baseline results are sensitive to the sample size difference for the case of a) OECD countries and b) the world. That is, there is evidence of divergence in each case (see Table F5 in Appendix F).

The purpose of this exercise is to show that, once we control for the difference in sample size, any remaining differences can be conveniently accounted for by the structural features of the economies. Holding these structural features constant, evidence of intraregional convergence vanishes in the case of Asia, OECD countries, and the world. We take these results as evidence in favor of club convergence. For Africa, however, evidence of intraregional convergence remains intact. We interpret this result as evidence that Africa is converging to a low steady-state equilibrium.

4.5 Discussion

As mentioned earlier, the convergence literature distinguishes between absolute, conditional, and club convergence. Absolute convergence refers to the tendency for countries to converge to a single steady state without conditioning on any control variables. Conditional convergence refers to the tendency for countries to converge to a single steady state conditional on the structural features of the economies. Club convergence refers to the tendency for countries to converge to different steady states even when the structural features of the economies are controlled for.

In the stochastic dominance context, using actual growth rates mimics the notion of absolute convergence whereas using residual growth rates mimics the idea of conditional convergence. Since there exists no method for studying club convergence within the framework of stochastic dominance, we coarsely partition our large sample of countries (100 or 76) into four regional groups: Africa, Latin America, Asia, and OECD countries. In the baseline analysis, we analyze convergence among countries based on actual growth rates. Since this analysis does not control for structural features of the economies, our findings can merely be construed in terms of absolute versus club convergence. In the robustness analysis, we repeat the analysis using residual growth rates. In this latter analysis, our findings can be construed in terms of conditional versus club convergence.

Our analysis in the preceding section has produced a number of intriguing findings. In interregional analysis, we find that most regions converge to each other in a pairwise comparison (after controlling for structural features of the economies). We take these findings as evidence in favor of conditional convergence. In intertemporal analysis, we obtain the following results (again, after controlling for structural features of the economies). First, African countries are converging to each other. Second, Asian countries are diverging from each other. Third, OECD countries are diverging from each other. Finally, countries in the world are diverging from each other. Except for Africa, these intertemporal findings suggest that there is an intraregional divergence among countries. We take these results as evidence in favor of club convergence. For Africa, however, it seems that the continent as a whole is converging to a low steady-state equilibrium.

It seems that there is conflicting evidence from interregional and intertemporal analyes: in the former, countries appear to exhibit conditional convergence; in the latter, countries appear to exhibit club convergence. How do we reconcile these seemingly paradoxical findings? It turns out that there is a simple, plausible explanation for them: taken together, both findings suggest that there is an intraregional (but not interregional) mobility among countries. We take these results as evidence in support of the club convergence hypothesis, which in turn, is evidence in support of the phenomenon of multiple regimes among countries.

Now we compare our findings with the previous findings in the convergence literature. Since the contribution of Baumol (1986), regression-based convergence studies have undergone three stages of refinements. First, De Long (1988) improves on Baumol's analysis by choosing countries that appear to converge as of the initial period of analysis. He finds no evidence of absolute convergence among countries. Second, MRW (1992) refine De Long's study by including control variables as implied by the Solow model. They find evidence of conditional convergence among countries. Third, Durlauf and Johnson (1995) improve on MRW's method by partitioning countries into multiple regimes as implied by the Azariadis-Drazen model. They find evidence of club convergence among countries. Subsequent work proves that evidence of club convergence is broadly robust across a host of different methods and specifications; see DJT (2005). Accordingly, the notion of club convergence is almost universally taken as a norm in the convergence literature.

Despite the overwhelming evidence on club convergence, regression-based convergence studies share one common flaw: the convergence coefficient merely captures the conditional mean of the growth rates. In contrast, a stochastic dominance approach to convergence analysis examines the entire growth distribution. Anderson (2004a), who pioneers the work in this area, conducts a convergence dominance analysis for a large sample of countries over multiple time periods: 1970-1990, 1971-1991,. . . , 1975-1995. In most cases considered, he finds evidence of convergence (see Table 2 in his paper). Since his analysis is confined to the entire world only, and no attempt is made to control for steady-state determinants, it is not possible to interpret his results beyond absolute convergence.

MRS (2006) take the analysis one step forward by breaking the world into OECD countries and ROW and analyzing whether convergence occurs a) between OECD countries and ROW, b) within OECD countries over two time periods, and c) within ROW over two time periods. For all the three cases considered, they find no evidence of absolute convergence but there is evidence in favor of conditional convergence. Taken together, their results can be taken as evidence against club convergence. However, since their analysis is limited to two regions of the world only, the data may not be able to uncover the true growth behavior of countries. In our analysis, where the world is broken up into four country groups, evidence in favor of club convergence is restored.

4.6 Conclusion

In this paper, we revisit the convergence hypothesis by using a stochastic dominance method developed by Davidson and Duclos (2000). An advantage of this method over the conventional regression-based models is that the convergence behavior of countries is examined using the entire growth distribution as opposed to the conditional mean of the growth rates. Therefore, the results are more informative.

Our findings can be decomposed into interregional and intertemporal analyses. For interregional analysis, we find that most regions converge to each other in a pairwise comparison. Since these findings control for structural features of the economies, we take these findings as evidence in favor of conditional convergence. For intertemporal analysis, we obtain the following.

First, African countries are converging to each other. Second, Asian countries are diverging from each other. Third, OECD countries are diverging from each other. Finally, countries in the world are diverging from each other. Except for Africa, these intertemporal findings suggest that there is an intraregional divergence among countries. We take these results as evidence in favor of club convergence. For Africa, however, it seems that the continent as a whole is converging to a low steady-state equilibrium. Coupled with interregional findings, these intertemporal findings suggest that, once structural features are held constant, there is an intraregional (but not interregional) mobility among countries. We take these results as evidence in support of the club convergence hypothesis, which in turn, is evidence in support of the phenomenon of multiple regimes among countries.

While our convergence results are based on the entire growth distribution, they are not without drawbacks. For one thing, our results are based on an exogenous partitioning of countries; it is not clear how we can endogenously divide countries into different groups in the manner of club convergence.

Chapter 5 Conclusion

This dissertation is motivated by the need to ascertain whether the world economy, conditional on the structural features of the economies, is characterized by a single or multiple regimes. If the world is characterized by a single regime, then economic policies need to be directed toward structural features (technologies, saving rates, population growth rates, etc.). If the world is characterized by multiple regimes, then economic policies need to be directed toward initial conditions (initial stock of human capital, initial distribution of income, etc.). There are two ways to address this issue. The first approach is to conduct a direct test on multiple-regime models; this approach is adopted in Chapters 2 and 3. The second approach is to build upon the work of the convergence hypothesis; this approach is employed in Chapter 4.

To begin with, today we have observed a growing number of multiple-regime models. Of these, the model developed by Galor and Zeira (1993) is particularly influential because it shows the linkage between income inequality and economic growth. However, except for Perotti (1996) who tests this model in passing, this model has not been comprehensively tested.

Chapter 2 of this dissertation is devoted to the empirical analysis of this model. Our baseline estimation based on a sample of 46 countries during the period 1970-2000 produces results that lend support to the model. These results are broadly robust to different model specifications, different samples, and alternative control variables. We take these results as evidence in favor of the Galor-Zeira model. Since the model belongs to the family of multiple-regime models, our results can be further interpreted as support for the multiple-regime models.

In recent years, we have observed a growing number of models which attempt to explain the link between income inequality and economic growth. In the Galor-Zeira model, this link is made possible by the existence of imperfect credit markets. In other models, the link is provided by endogenous fertility decisions. One model in this class, due to Kremer and Chen (2002), is particularly interesting because it belongs to the family of multiple-regime models.

Chapter 3 is devoted to the empirical investigation of this model. Our baseline estimation, using a number of alternative proxies for human capital differential variable, delivers results that are consistent with the predictions. These results are also broadly robust to different model specifications, different samples, and alternative control variables. We take these results as evidence in favor of the Kremer-Chen model. As in Chapter 2, since this model belongs to the family of multiple-regime models, our results can be further construed as support for the multiple-regime models.

As stated earlier, another way to address the single-versus-multiple-regime issue is by testing the convergence hypothesis. A conventional approach to study convergence is by using a growth regression model. However, as DJT (2005) put it, the regression-based models on which the analysis is based are questionable. In brief, evidence of convergence is obtained from the coefficient of initial income in a regression model where the growth rate of per capita income is the dependent variable and structural variables (plus initial income) are the explanatory variables. A major drawback of this approach is that evidence of convergence is inferred from the conditional mean of the growth rate. A better way is to infer the evidence of convergence from the entire distribution of the growth rate.

This is the approach taken by the stochastic dominance analysis that we employ in Chapter 4. Our baseline analysis based on a sample of 100 countries yields results that are consistent with the conditional convergence hypothesis. In a more careful robustness analysis (controlling for structural features of the economies), however, we obtain results that are consistent with the club convergence hypothesis. We take these results as evidence in favor of the multiple-regime models.

What do we learn from all of these results? Since all of them provide support for the multiple-regime models, we conclude that initial conditions matter in the growth process. The policy implication that emerges from this conclusion is that there is a need to alter the level of these initial conditions in order for countries to converge to high-income groups. In two of the models studied here (the Galor-Zeira and Kremer-Chen models), the initial condition is essentially the initial distribution of income. If the initial income distribution is relatively equal (unequal), then a particular economy will converge to a high-income (low-income) equilibrium. If we take these models seriously, then an appropriate course of action to be taken is to provide easier access to education for the poor people so that they might be able to escape from the poverty trap.

References

Anderson, Gordon (1996), "Nonparametric Tests of Stochastic Dominance in Income Distributions," *Econometrica*, 64: 1183-1193.

Anderson, Gordon (2004a), "Toward an Empirical Analysis of Polarization," *Journal of Econometrics*, 122: 1-26.

Anderson, Gordon (2004b), "Making Inferences about the Polarization, Welfare, and Poverty of Nations: A Study of 101 Countries form 1970-1995," *Journal of Applied Econometrics*, 19: 537-550.

Alesina, Alberto and Roberto Perotti (1996), "Income Distribution, Political Instability, and Investment," *European Economic Review*, 40: 1203–1228.

Alesina, Alberto and Rani Rodrik (1994), "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 109: 465–490.

Azariadis, C. and A. Drazen (1990), "Threshold Externalities in Economic Development," *Quarterly Journal of Economics*, 105: 501-526.

Barrett, G. and S. Donald (2003), "Consistent Tests for Stochastic Dominance," *Econometrica*, 71: 71-104.

Barro, Robert (2000), "Inequality, Growth, and Investment," *Journal of Economic Growth*, 5: 5–32.

Barro, Robert and Jong-Wha Lee (1994), "Data Set for a Panel of 138 Countries," *Mimeo*: Harvard University, Cambridge, MA.

Barro, Robert and Jong-Wha Lee (2001), "International Data on Educational Attainment: Updates and Implications," *Oxford Economic Papers*, 53: 541–563.

Barro, R. and X. Sala-i-Martin (2004), *Economic Growth*, Second Edition, Cambridge: MIT Press.

Baum, C., M. Schaffer, and S. Stillman (2003), "Instrumental Variables and GMM: Estimation and Testing," *Stata Journal*, 3: 1-31.

Baumol, William (1986), "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show," *American Economic Review*, 76: 1072-1085.

Beach, C. and R. Davidson (1983), "Distribution-Free Statistical Inference with Lorenz Curves and Income Shares," *Review of Economic Studies*, 50: 723-735.

Benabou, Roland (1996), "Inequality and Growth," *NBER Macroeconomics Annual*, 11: 11–74.

Bernanke and Gurkaynak (2001), "Is Growth Exogenous? Taking Mankiw, Romer, and Weil Seriously," *NBER Macroeconomics Annual*, 16.

Bils, Mark and Peter Klenow (2000), "Does Schooling Cause Growth?," *American Economic Review*, 90: 1160–1183.

Bishop, J., J. Formby, and W. Smith (1993), "International Comparison of Welfare and Poverty: Dominance Orderings for 10 Countries," *Canadian Journal of Economics*, 26: 707-726.

Bishop, J., J. Formby, and P. Thistle (1992), "Convergence of the South and Non-South Income distributions: 1969-1979," *American Economic Review*, 82: 262-272.

Caselli, Francesco and Wilbur J. Coleman (2006), " The World Technology Frontier," *American Economic Review*, 96: 499–522.

Caselli, F., G. Esquivel, and F. Lefort (1996), "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics," J*ournal of Economic Growth*, 1: 363-389.

Clarke, George (1995), "More Evidence on Income Distribution and Growth," *Journal of Development Economics*, 47: 403–427.

Dahan, M. and D. Tsiddon (1998), "Demographic Transition, Income Distribution, and Economic Growth," *Journal of Economic Growth*, 3: 29–52.

Davidson, R. and J. Duclos (2000), "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality," *Econometrica*, 68: 1435-1464.

Deininger, Klaus and Lyn Squire (1996), "A New Dataset Measuring Income Inequality," *World Bank Economic Review*, 10: 565–591.

Deininger, Klaus and Lyn Squire (1998), "New Ways of Looking at Old Issues: Inequality and Growth," *Journal of Development Economics*, 57: 259–287.

De La Croix, D. and M. Doepke (2003), "Inequality and Growth: Why Differential Fertility Matters," *American Economic Review*, 93: 1091–1113.

De Long, Bradford (1988), "Productivity Growth, Convergence, and Welfare: Comment," *American Economic Review*, 78: 1138-1154.

Duffy, J., C. Papageorgiou, and F. Perez-Sebastian (2004), "Capital-Skill Complentarity? Evidence from a Panel of Countries," *Review of Economics and Statistics*, 86: 327–344.

Durlauf, S. and P. Johnson (1995), "Multiple Regimes and Cross-Country Growth Behavior," *Journal of Applied Econometrics*, 10: 365-384.

Durlauf, S., P. Johnson, and J. Temple (2005), "Growth Econometrics," In: Aghion, P. and S. Durlauf (eds.), *Handbook of Economic Growth*, 1A: 555–677.

Fernandez, C., E. Ley, and M. Steel (2001), "Model Uncertainty in Cross-Country Growth Regressions," *Journal of Applied Econometrics*, 16: 563-576.

Forbes, Kristin (2000), “A Reassessment of the Relationship between Inequality and Growth,” *American Economic Review*, 90: 869–887.

Galor, Oded (1996), "Convergence: Inferences from Theoretical Models," *Economic Journal*, 106: 1056-1069.

Galor, Oded and Joseph Zeira (1993), “Income Distribution and Macroeconomics,” *Review of Economic Studies*, 60: 35–52.

Galor, Oded and Omer Moav (2004), “From Physical to Human Capital Accumulation: Inequality and the Process of Development,” *Review of Economic Studies*, 71: 1001–1026.

Hansen, Bruce (2000), “Sample Splitting and Threshold Estimation,” *Econometrica*, 68: 575-603.

Hauk and Wacziarg (2004), “A Monte Carlo Study of Growth Regressions,” *NBER Working Paper* No. 296.

Hsiao, C., Li, Q., and J. Racine (2006), "A Consistent Model Specification Test of Mixed Categorical and Continuous Data," forthcoming in *Journal of Econometrics*.

Islam, Nazrul (1995), “Growth Empirics: A Panel Data Approach,” *Quarterly Journal of Economics*, 110: 1127-1170.

Kaldor, Nicholas (1957), “A Model of Economic Growth,” *Economic Journal*, 67: 591—624.

Kremer, M. and D. Chen (1999), "Income Distribution Dynamics with Endogenous Fertility," *American Economic Association P&P*, 89: 155–160.

Kremer, M. and D. Chen (2002), "Income Distribution Dynamics with Endogenous Fertility," *Journal of Economic Growth*, 7: 227–258.

Lee, K., M. Pesaran, and R. Smith (1997), "Growth and Convergence in Multi-Country Empirical Stochastic Solow Model," *Journal of Applied Econometrics*, 12: 357-392.

Levine, Ross (1997), "Financial Development and Economic Growth: Views and Agenda," *Journal of Economic Literature*, 35: 688-726.

Li, H. & H. Zou (1998), “Inequality is not Harmful for Growth: Theory and Evidence,” *Review of Development Economics*, 2: 318–334.

Maasoumi E. and A. Heshmati (2000), “Stochastic Dominance among Swedish Income Distributions,” *Econometric Review*, 19: 287-320.

Maasoumi E., J. Racine, and T. Stengos (2006), “Growth and Convergence: A Profile of Distribution Dynamics and Mobility," forthcoming in *Journal of Econometrics*.

Mackinnon, James (2002), "Bootstrap Inference in Econometrics, *Canadian Journal of Econometrics*, 35: 615–645.

Maden, D. and F. Smith (2000), "Poverty in Ireland, 1987-1994: A Stochastic Dominance Approach," *Economic and Social Review*, 31: 187-214.

Mankiw, N. Gregory, David Romer, and David Weil (1992), "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107: 407–437.

Manuelli, R. and A. Seshadri (2005), "Human Capital and the Wealth of Nations," *Working Paper*, University of Wisconsin-Madison.

McFadden, Daniel (1989), "Testing for Stochastic Dominance," In: Formby, T. and T. Seo, *Studies in the Economics of Uncertainty*, New York: Springer, 113-134.

Moav, Omer (2005), "Cheap Children and Persistence of Poverty," *Economic Journal*, 115: 88–110.

Perotti, Roberto (1996), "Growth, Income Distribution, and Democracy: What the Data Say," *Journal of Economic Growth*, 1: 149–187.

Persson, Torsten and Guido Tabellini (1994), "Is Inequality Harmful for Growth?" *American Economic Review*, 84: 600–621.

Quah, Danny (1996), "Empirics for Economic Growth and Convergence," *European Economic Review*, 40: 1353-1275.

Sala-i-Martin, Xavier (1996), "The Classical Approach to Convergence Analysis," *Economic Journal*, 106: 1019-1036.

Solow, Robert (1956), "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70: 65-94.

Stoline, M. and H. Ury (1979), "Tables of Studentized Maximum Modulus Distribution and an Application to Multiple Comparisons among the Means," *Technometrics*, 21: 87-93.

Sylwester, Kevin (2000), "Income Inequality, Educational Expenditures, and Growth," *Journal of Development Economics*, 63: 379–398.

Temple, Jonathan (1999), "The New Growth Evidence," *Journal of Economic Literature*," 37: 112-156.

Tse, Y. and X. Zhang (2004), "A Monte Carlo Investigation of Some Tests for Stochastic Dominance," *Journal of Statistical Computation & Simulation*, 74: 361-378.

Wolak, Frank (1989), “Testing Inequality Constraints in Linear Econometric Models,” *Journal of Econometrics*, 41: 205-235.

Wooldridge, Jeffrey (2002), *Econometric Analysis of Cross Section and Panel Data*, First Edition, Cambridge, MA: MIT Press.

Appendix A: Algebra of the Galor-Zeira Model

In this appendix, we derive Eq.(2.5). To maximize utility in Eq.(2.1) with respect to b subject to the budget constraint in Eq.(2.2), we begin by substituting (2.2) into (2.1):

$$V = \alpha \log\left[(x + w_u)(1 + r) + w_u - b\right] + (1 - \alpha)\log b. \tag{A1}$$

Differentiating (A1) with respect to b, we obtain the optimal bequest for individuals who choose not to invest in education:

$$b_u(x) = (1 - \alpha)\left[(x + w_u)(1 + r) + w_u\right]. \tag{A2}$$

Substituting (A2) back into (A1) and rearranging, we obtain the optimal utility:

$$V_u(x) = \log\left[(x + w_u)(1 + r) + w_u\right] + \varepsilon, \tag{A3}$$

where $\varepsilon = \alpha \log \alpha + (1 - \alpha)\log(1 - \alpha)$.

To maximize utility (2.1) with respect to b subject to the budget constraint (2.3), we begin by substituting (2.3) into (2.1):

$$V = \alpha \log\left[(x - h)(1 + i) + w_s - b\right] + (1 - \alpha)\log b. \tag{A4}$$

Differentiating (A4) with respect to b, we obtain the optimal bequest for individuals whose $x < h$ but choose to invest in education:

$$b_s(x) = (1 - \alpha)\left[(x - h)(1 + i) + w_s\right]. \tag{A5}$$

Substituting (A5) back into (A4) and rearranging, we obtain the optimal utility:

$$V_s(x) = \log\left[(x-h)(1+i) + w_s\right] + \varepsilon. \tag{A6}$$

To maximize utility (2.1) with respect to b subject to the budget constraint (2.4), we begin by substituting (2.4) into (2.1):

$$V = \alpha \log\left[(x-h)(1+r) + w_s - b\right] + (1-\alpha)\log b. \tag{A7}$$

Differentiating (A7) with respect to b, we obtain the optimal bequest for individuals whose $x \geq h$ and choose to invest in education:

$$b_s(x) = (1-\alpha)\left[(x-h)(1+r) + w_s\right]. \tag{A8}$$

Substituting (A8) back into (A7) and rearranging, we obtain the optimal utility:

$$V_s(x) = \log\left[(x-h)(1+r) + w_s\right] + \varepsilon. \tag{A9}$$

Regardless of the amount of inheritance, the decision to invest in education depends on the relative size of relevant utilities. For individuals whose $x < h$, this decision depends on the relative size of (A3) and (A6). Equating (A3) with (A6), we obtain

$$f = \frac{w_u(2+r) + h(1+i) - w_s}{i-r}, \tag{A10}$$

where f is the amount of inheritance in which individuals are indifferent in their decision of whether or not to invest in education. Thus, individuals whose a) $f \leq x < h$ will invest in education, and b) $x < f$ will not invest in education.

For individuals whose $x \geq h$, the decision to invest in education depends on the relative size of (A3) and (A9). We assume that the wage income of skilled workers net of educational investment is at least as great as the wage income of unskilled workers amounts. This amounts to stipulating that (A9) is at least as great as (A3), which implies

$$w_s - h(1+r) \geq w_u(2+r). \tag{A11}$$

Therefore, these individuals will always invest in education.

Given (A1) through (A11), the dynamics of the model can be expressed by Eq.(2.5) in the text:

$$x_{t+1} = (1-\alpha) \left\{ \begin{array}{c} [(x_t + w_u)(1+r) + w_u] \;\; \text{if} \;\; x_t < f \\ [(x_t - h)(1+i) + w_s] \;\; \text{if} \;\; f \leq x_t < h \\ [(x_t - h)(1+r) + w_s] \;\; \text{if} \;\; x_t > h \end{array} \right\}.$$

Appendix B: Additional Results for the Galor-Zeira Model

Table B1: Estimation of 1980-2000 Sample with 46 Obs.

Dep.Variable	*Inc2000* (1a)	*Educ* (1b)	*Inc2000* (2a)	*Educ* (2b)
Constant	2.990 (1.34)	1.497 (0.71)	4.467^{*} (1.81)	1.782 (0.82)
Gini	–	-0.561^{**} (−2.15)	–	-0.651^{**} (−2.17)
Poor	–	-0.563^{***} (−3.14)	–	-0.404^{*} (−1.76)
Educ	1.680^{***} (3.42)	–	1.545^{**} (2.69)	–
Invest	0.740^{**} (2.29)	–	0.733^{*} (1.84)	–
$(n+g+\delta)$	−0.362 (−0.43)	–	0.014 (0.02)	–
Africa	0.986^{**} (2.26)	–	0.612 (1.21)	–
Latin	–	–	−0.287 (−1.38)	–
Asia	–	–	−0.256 (−1.08)	–
Adj.R^2	0.77	0.65	0.79	0.64
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table B2: Baseline Estimation with *Ginilev*

Dep.Variable	*Inc2000* (1a)	*Educ* (1b)	*Inc2000* (2a)	*Educ* (2b)
Constant	5.706^{**} (2.59)	0.389 (0.30)	7.452^{***} (3.17)	0.716 (0.47)
Ginilev	–	-0.018^{***} (-3.31)	–	-0.021^{***} (-2.94)
Poor	–	-0.821^{***} (-4.88)	–	-0.778^{***} (-4.96)
Educ	2.010^{***} (5.39)	–	1.832^{***} (5.05)	–
Invest	0.359^{*} (2.00)	–	0.350^{**} (2.04)	–
$(n+g+\delta)$	0.370 (0.40)	–	0.804 (0.87)	–
Africa	1.125^{**} (2.08)	–	0.635 (1.10)	–
Latin	–	–	-0.477^{**} (-2.03)	–
Asia	–	–	-0.238 (-0.94)	–
Adj.R^2	0.67	0.69	0.72	0.67
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table B3: Baseline Estimation with Wage Premium

Dep.Variable	*Inc2000* (1a)	*Educ* (1b)	*Inc2000* (2a)	*Educ* (2b)
Constant	3.776 (1.54)	1.950 (0.82)	6.154 (1.70)	2.214 (0.77)
Gini	–	-0.524^{*} (-1.73)	–	-0.498 (-1.39)
Poor	–	-0.689^{***} (-3.58)	–	-0.686^{***} (-3.16)
Prem	–	-0.010^{*} (-1.79)	–	-0.009 (-1.46)
Educ	2.039^{***} (4.23)	–	1.910^{***} (3.66)	–
Invest	0.385^{*} (1.85)	–	0.400^{*} (1.90)	–
$(n+g+\delta)$	-0.278 (-0.25)	–	0.451 (0.33)	–
Africa	–	–	1.019 (1.23)	–
Latin	–	–	-0.402 (-0.92)	–
Asia	–	–	-0.230 (-0.58)	–
Adj.R^2	0.58	0.56	0.60	0.53
Obs.	33	33	33	33

Notes: Except for dummies and Premium, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table B4: Reduced-Form Estimation

Dep. Variable	*Inc2000*		*Growth*
	(1)	(2)	(3)
Constant	10.038^{***} (3.04)	9.168^{**} (2.56)	4.260^{**} (2.26)
Gini	-1.462^{***} (−3.59)	-0.856^{*} (−1.71)	–
Ginilev	–	–	-0.029^{***} (−3.68)
Poor	-1.557^{***} (−5.07)	-1.565^{***} (−4.69)	-0.765^{**} (−2.68)
Invest	0.252 (1.41)	0.186 (0.96)	0.283^{*} (2.02)
$(n+g+\delta)$	-1.457^{*} (−1.96)	−1.120 (−1.30)	−0.149 (−0.25)
Inc70	–	–	-0.411^{***} (−3.68)
Africa	–	−0.690 (−1.43)	−0.093 (−0.29)
Latin	–	-0.591^{**} (−2.14)	–
Asia	–	−0.246 (−0.94)	–
Adj.R^2	0.69	0.70	0.43
Obs.	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by OLS. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix C: Algebra of the Kremer-Chen Model

a) The Derivation of the Inverse Relationship between w_s/w_u and R

Let the aggregate production function for the economy be given by $Y = AL_s^{1/2}L_u^{1/2}$, where Y is the level of output, A is the level of technology, L_s is the stock of skilled labor, and L_u is the stock of unskilled labor. (The use of $1/2$ in the exponents of L_s and L_u is to simplify algebra.) Assuming that labor markets are competitive, then

$$w_s = (1/2)\,A\,(L_u/L_s)^{1/2} = (1/2)\,AR^{-1/2}, \text{ and}$$

$$w_u = (1/2)\,A\,(L_s/L_u)^{1/2} = (1/2)\,AR^{1/2}.$$

Dividing w_s by w_u yields the inverse relationship between wage premium and the skilled-unskilled labor ratio:

$$\frac{w_s}{w_u} = \frac{(1/2)\,AR^{-1/2}}{(1/2)\,AR^{1/2}} = \frac{1}{R}.$$

b) The Derivation of the Inverse Relationship between n and w

Let each household faces the following utility function: $V = \ln(n) + x$, where V is utility, n is the number of children, and x is the amount of consumption. Let the total time endowment for each household be 1 and the total time needed to raise each child be a fraction φ of the total time endowment. Then, the household's budget constraint is $x = (1 - \varphi n)\,w$. Then, the first-order condition for optimal fertility implies that higher wages lead people to have fewer children:

$$n = \frac{1}{\varphi w}.$$

c) The Derivation of Eq.(3.7)

For convenience, let us reproduce Eq.(3.5):

$$\begin{gathered} \gamma_t \leq \theta \text{ if } R_{t+1} = 1 - c_L \\ \gamma_t = \theta \text{ if } R_{t+1} \in (1 - c_H, 1 - c_L) \\ \gamma_t \geq \theta \text{ if } R_{t+1} = 1 - c_H \end{gathered}$$

Based on Eq.(3.5), we state the following propositions.

Proposition 1 *If $R_\theta \leq 1 - c_H$, then $R_{t+1} = 1 - c_H$.*

Proof. *Suppose that $R_{t+1} > 1 - c_H$. Then, Eq.(3.5) implies that $\gamma_t = \theta$ and $R_\theta > 1 - c_H$. But this is a contradiction. Now suppose that $R_{t+1} < 1 - c_H$. Then, every individual will invest in education.* ■

Proposition 2 *If $R_\theta \geq 1 - c_L$, then $R_{t+1} = 1 - c_L$.*

Proof. *Suppose that $R_{t+1} < 1 - c_L$. Then, Eq.(3.5) implies that $\gamma_t = \theta$ and $R_\theta < 1 - c_L$. But this is a contradiction. Now suppose that $R_{t+1} > 1 - c_L$. Then, every individual will not invest in education.* ■

Proposition 3 *If $R_\theta \in (1 - c_H, 1 - c_L)$, then $R_{t+1} = (R_t^2 + \theta) / (1 - \theta)$.*

Proof. *Suppose that $R_{t+1} = 1 - c_H$. Then, Eq.(3.5) implies that $\gamma_t \geq \theta$ and $R_{t+1} \geq R_\theta > 1 - c_H$. But this is a contradiction. Now suppose that $R_{t+1} = 1 - c_L$. Then, Eq.(3.5) implies that $\gamma_t \leq \theta$ and $R_{t+1} \leq R_\theta < 1 - c_L$. But this is a contradiction too.* ■

Combining all of these propositions yields the evolution of R_t as in Eq.(3.7):

$$R_{t+1} = \left\{ \begin{array}{c} 1 - c_L \text{ if } R_\theta \geq 1 - c_L \\ R_\theta \text{ if } R_\theta \in (1 - c_H, 1 - c_L) \\ 1 - c_H \text{ if } R_\theta \leq 1 - c_H \end{array} \right\}$$

Appendix D: Additional Results for the Kremer-Chen Model

Table D1: Baseline Estimation with L_{s0}/L_{u0}

Dep. Variable	*Inc2000* (1a)	L_{s0}/L_{u0} (1b)	*Inc2000* (2a)	L_{s0}/L_{u0} (2b)
Constant	4.488^{*} (1.99)	7.665^{**} (2.05)	4.307^{*} (1.73)	10.194^{**} (2.32)
Gini	–	-1.516^{***} (-3.02)	–	-1.951^{***} (-3.32)
Poor	–	-1.607^{***} (-4.55)	–	-1.406^{***} (-3.70)
Rich	–	0.611^{**} (2.28)	–	0.604^{*} (1.92)
L_{s0}/L_{u0}	0.893^{***} (5.41)	–	0.838^{***} (4.99)	–
Invest	0.403^{**} (2.13)	–	0.395^{**} (2.03)	–
$(n+g+\delta)$	-2.277^{***} (-3.15)	–	-2.306^{**} (-2.66)	–
Latin	–	–	-0.231 (-0.91)	–
Asia	–	–	0.079 (0.27)	–
Adj. R^2	0.60	0.50	0.62	0.50
Obs.	43	43	43	43

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table D2: Baseline Estimation with L_{s1}/L_{u1}

Dep. Variable	*Inc2000* (1a)	L_{s1}/L_{u1} (1b)	*Inc2000* (2a)	L_{s1}/L_{u1} (2b)
Constant	3.916^{**} (2.02)	9.222^{**} (2.42)	3.585 (1.63)	11.485^{**} (2.57)
Gini	–	-1.614^{***} (−3.16)	–	-2.047^{***} (−3.42)
Poor	–	-1.713^{***} (−4.76)	–	-1.511^{***} (−3.91)
Rich	–	1.001^{***} (3.66)	–	1.030^{***} (3.21)
L_{s1}/L_{u1}	0.733^{***} (6.15)	–	0.699^{***} (5.60)	–
Invest	0.437^{**} (2.64)	–	0.422^{**} (2.44)	–
$(n+g+\delta)$	-2.137^{***} (−3.32)	–	-2.247^{***} (−2.89)	–
Latin	–	–	−0.158 (−0.68)	–
Asia	–	–	0.105 (0.39)	–
Adj. R^2	0.69	0.60	0.70	0.60
Obs.	43	43	43	43

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table D3: Baseline Estimation with L_{s2}/L_{u2}

Dep. Variable	*Inc2000* (1a)	L_{s2}/L_{u2} (1b)	*Inc2000* (2a)	L_{s2}/L_{u2} (2b)
Constant	4.988^{**} (2.47)	7.084^{*} (1.79)	6.000^{**} (2.66)	4.099 (0.89)
Gini	–	-1.487^{***} (−2.81)	–	-1.330^{**} (−2.17)
Poor	–	-1.367^{***} (−3.67)	–	-1.429^{***} (−3.61)
Rich	–	0.961^{***} (3.39)	–	1.255^{***} (3.82)
L_{s2}/L_{u2}	0.815^{***} (6.16)	–	0.743^{***} (5.90)	–
Invest	0.417^{**} (2.54)	–	0.477^{***} (2.92)	–
$(n+g+\delta)$	-1.519^{**} (−2.24)	–	−1.103 (−1.41)	–
Latin	–	–	−0.235 (−1.07)	–
Asia	–	–	−0.270 (−1.12)	–
Adj. R^2	0.70	0.56	0.72	0.57
Obs.	43	43	43	43

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table D4: Baseline Estimation with L_{s3}/L_{u3}

Dep. Variable	*Inc2000* (1a)	L_{s3}/L_{u3} (1b)	*Inc2000* (2a)	L_{s3}/L_{u3} (2b)
Constant	6.405^{***} (3.00)	4.324 (1.03)	7.522^{***} (3.42)	0.067 (0.01)
Gini	–	-1.327^{**} (-2.35)	–	-1.171^{*} (-1.89)
Poor	–	-1.156^{***} (-2.91)	–	-1.212^{***} (-3.03)
Rich	–	1.092^{***} (3.61)	–	1.566^{***} (4.73)
L_{s3}/L_{u3}	0.806^{***} (6.18)	–	0.671^{***} (6.16)	–
Invest	0.601^{***} (3.89)	–	0.672^{***} (4.67)	–
$(n+g+\delta)$	-0.560 (-0.75)	–	-0.133 (-0.17)	–
Latin	–	–	-0.371^{*} (-1.86)	–
Asia	–	–	-0.417^{*} (-1.91)	–
Adj. R^2	0.71	0.57	0.76	0.62
Obs.	43	43	43	43

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table D5: Baseline Estimation with L_{s4}/L_{u4}

Dep. Variable	*Inc2000* (1a)	L_{s4}/L_{u4} (1b)	*Inc2000* (2a)	L_{s4}/L_{u4} (2b)
Constant	8.872^{***} (3.72)	3.094 (0.74)	9.585^{***} (3.74)	0.080 (0.02)
Gini	–	-1.655^{***} (−2.96)	–	-1.361^{**} (−2.06)
Poor	–	-1.486^{***} (−3.78)	–	-1.615^{***} (−3.78)
Rich	–	0.863^{***} (2.89)	–	1.052^{***} (2.97)
L_{s4}/L_{u4}	0.808^{***} (6.33)	–	0.779^{***} (6.00)	–
Invest	0.439^{***} (2.76)	–	0.480^{***} (2.93)	–
$(n+g+\delta)$	0.433 (0.51)	–	0.699 (0.76)	–
Latin	–	–	−0.105 (−0.46)	–
Asia	–	–	−0.212 (−0.87)	–
Adj. R^2	0.71	0.66	0.71	0.66
Obs.	43	43	43	43

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table D6: Panel Data Estimation with Unweighted L_{s1}/L_{u1}

Dependent Variable	*Income* (1a)	L_{s1}/L_{u1} (1b)	*Income* (2a)	L_{s1}/L_{u1} (2b)	L_{s1}/L_{u1} (3b)	*Income* (4a)	L_{s1}/L_{u1} (4b)
Constant	5.814^{***} (4.83)	-0.805 (-0.26)	6.887^{***} (5.40)	0.071 (0.02)	-15.100^{***} (-6.66)	5.507^{*} (1.71)	11.048 (1.40)
Gini	–	-0.292 (-0.73)	–	-0.419 (-0.84)	-0.359 (-0.70)	–	-1.382 (-1.26)
Poor	–	-1.203^{***} (-4.93)	–	-0.986^{***} (-3.36)	–	–	-1.202^{*} (-1.98)
Rich	–	1.088^{***} (6.61)	–	1.011^{***} (4.70)	0.199 (1.02)	–	0.903^{*} (1.95)
L_{s1}/L_{u1}	0.830^{***} (9.90)	–	0.671^{***} (5.95)	–	–	0.564^{**} (2.49)	–
Invest	0.848^{***} (5.40)	–	0.925^{***} (6.53)	–	–	1.159^{***} (3.73)	–
$(n+g+\delta)$	-0.915^{**} (-2.37)	–	-0.379 (-0.85)	–	–	-0.553 (-0.44)	–
Latin	–	–	-0.255 (-1.47)	–	–	-0.339 (-0.81)	–
Asia	–	–	-0.420^{*} (-1.89)	–	–	-0.507 (-0.95)	–
Adj. R^2	0.56	0.44	0.70	0.44	0.39	0.77	0.56
Obs.	145	145	145	145	145	145	145

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table D7: Baseline Estimation with *Ginilev*

Dep. Variable	*Inc2000* (1a)	L_{s1}/L_{u1} (1b)	*Inc2000* (2a)	L_{s1}/L_{u1} (2b)
Constant	3.969^{**} (2.14)	5.341^{*} (1.98)	3.618^{*} (1.74)	6.728^{**} (2.23)
Ginilev	–	-0.041^{***} (−3.44)	–	-0.053^{***} (−3.86)
Poor	–	-1.770^{***} (−5.07)	–	-1.546^{***} (−4.18)
Rich	–	1.001^{***} (3.98)	–	1.010^{***} (3.61)
L_{s1}/L_{u1}	0.700^{***} (6.40)	–	0.664^{***} (5.83)	–
Invest	0.426^{**} (2.67)	–	0.412^{**} (2.45)	–
$(n+g+\delta)$	-2.088^{***} (−3.37)	–	-2.201^{***} (−3.00)	–
Latin	–	–	−0.137 (−0.62)	–
Asia	–	–	0.097 (0.40)	–
Adj. R^2	0.70	0.64	0.71	0.65
Obs.	46	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by 2SLS; columns (a) and (b) report results from the second- and first-stage regressions, respectively. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table D8: Reduced-Form Estimation

Dep. Variable	*Inc2000*		*Growth*
	(1)	(2)	(3)
Constant	9.695*** (3.21)	8.041** (2.36)	4.280** (2.12)
Gini	−1.003** (−2.49)	−0.727 (−1.57)	–
Ginilev	–	–	−0.030*** (−3.64)
Poor	−1.384*** (−4.82)	−1.532*** (−4.91)	−0.775** (−2.66)
Rich	0.622*** (2.99)	0.631*** (2.70)	−0.001 (−0.00)
Invest	0.305* (1.85)	0.234 (1.31)	0.292** (2.13)
$(n+g+\delta)$	−0.834 (−1.17)	−1.140 (−1.42)	−0.135 (−0.23)
Inc70	–	–	−0.410*** (−2.82)
Latin	–	−0.156 (−0.64)	–
Asia	–	0.160 (0.64)	–
Adj. R^2	0.74	0.74	0.43
Obs.	46	46	46

Notes: Except for dummies, all variables are expressed in logs. Estimation is done by OLS. t-statistics are given in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix E: Calculation of the Weighted L_{s1}/L_{u1} Data

For a given skill category indexed by i ($i = 0, 1, ..., 5$), skilled and unskilled workers, L_{si} and L_{ui}, are weighted by a function of the return to education, ϕ, times the duration of education, τ. Data on the return to education by country i, ϕ_i, is taken from Bils and Klenow (2000), who estimate the Mincerian wage regression. Data on the duration of education at various levels j by country i, τ_{ij}, is taken directly from Jong-Wha Lee; this data set is unpublished. The levels of education are no education (no), some primary (sp), complete primary (cp), some secondary (ss), complete secondary (cs), some teritary (st), and complete tertiary education (ct). Lee furnishes the data on the duration of education for cp, ss, and cs only. We assume that the duration of ct is 4 years, st is one-half of ct, and sp is one-half of cp.

Given the data and these plausible assumptions, the weighted L_{s1} and L_{u1} data for country i are constructed as follows,

$$\begin{aligned} L_{i,s1} &= L_{i,st} + \exp\left(\phi_i \tau_{i,ct}\right) L_{i,ct}, \text{ and} \\ L_{i,u1} &= L_{i,no} + \exp\left(\phi_i \tau_{i,sp}\right) L_{i,sp} + \exp\left(\phi_i \tau_{i,cp}\right) L_{i,cp} \\ &\quad + \exp\left(\phi_i \tau_{i,ss}\right) L_{i,ss} + \exp\left(\phi_i \tau_{i,cs}\right) L_{i,cs}. \end{aligned}$$

Note that the exponential terms attached to workers with higher education levels ensure that more educated workers are weighted more heavily than less educated workers. In this manner, the aggregation problem is alleviated. Note also that, if we assume that all exponential terms are equal to unity, we will recover the unweighted data.

Appendix F: Additional Results for the Convergence Hypothesis

Table F1: Africa vs. Non-Africa

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	0.000**	0.000**	0.000**	–	–	–
	SMT	0.000**	0.000**	0.000**	–	–	–
	Wald	0.000**	0.000**	0.000**	–	–	–
OECD	CMT	0.000**	0.000**	0.000**	–	–	–
	SMT	0.000**	0.000**	0.000**	–	–	–
	Wald	0.000**	0.000**	0.000**	–	–	–
Asia	CMT	0.000**	0.000**	0.000**	–	–	–
	SMT	0.000**	0.000**	0.000**	–	–	–
	Wald	0.000**	0.000**	0.000**	–	–	–
Latin	CMT	0.369	0.301	0.367	1.000	1.000	1.000
	SMT	0.220	0.098	0.106	0.982	0.912	0.916
	Wald	0.135	0.097	0.091	0.878	0.728	0.642

Notes: Africa vs. Non-Africa refers to the null hypothesis that Africa weakly dominates Non-Africa. If the null can be rejected, then we conclude that Non-Africa strictly dominates Africa. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

Table F2: Latin America vs. Non-Latin America

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	0.009**	0.033*	0.140	0.997	1.000	1.000
	SMT	0.008**	0.012*	0.036*	–	–	–
	Wald	0.003**	0.012*	0.033*	–	–	–
OECD	CMT	0.000**	0.000**	0.000**	–	–	–
	SMT	0.000**	0.000**	0.000**	–	–	–
	Wald	0.000**	0.000**	0.000**	–	–	–
Asia	CMT	0.003**	0.000**	0.000**	–	–	–
	SMT	0.002**	0.000**	0.000**	–	–	–
	Wald	0.000**	0.000**	0.000**	–	–	–

Notes: Latin America vs. Non-Latin America refers to the null hypothesis that Latin America weakly dominates Non-Latin America. If the null can be rejected, then we conclude that Non-Latin America strictly dominates Latin America. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

Table F3: Asia vs. Non-Asia

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	1.000	1.000	1.000	0.000**	0.000**	0.000**
	SMT	1.000	1.000	0.997	0.000**	0.000**	0.000**
	Wald	0.895	0.774	0.681	0.000**	0.000**	0.000**
OECD	CMT	0.640	0.744	0.757	0.099	0.449	0.871
	SMT	0.385	0.324	0.281	0.068	0.181	0.390
	Wald	0.376	0.298	0.259	0.063	0.166	0.341

Notes: Asia vs. Non-Asia refers to the null hypothesis that Asia weakly dominates Non-Asia. If the null can be rejected, then we conclude that Non-Asia strictly dominates Asia. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

Table F4: OECD vs. ROW

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
ROW	CMT	0.979	1.000	1.000	0.000**	0.000**	0.000**
	SMT	0.853	1.000	0.993	0.000**	0.000**	0.000**
	Wald	0.841	0.734	0.657	0.000**	0.000**	0.000**

Notes: OECD vs. ROW refers to the null hypothesis that OECD weakly dominates ROW. If the null can be rejected, then we conclude that ROW strictly dominates OECD. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

Table F5: 1961-1980 vs. 1981-2000

		Original Position			Reversed Position		
Region	Test	FSD	SSD	TSD	FSD	SSD	TSD
Africa	CMT	0.996	1.000	1.000	0.020	0.329	0.515
	SMT	0.895	0.825	0.845	0.016	0.119	0.161
	Wald	0.872	0.744	0.689	0.024	0.111	0.151
Latin	CMT	1.000	1.000	1.000	0.000**	0.000**	0.000**
	SMT	1.000	0.986	0.936	0.000**	0.000**	0.000**
	Wald	0.905	0.732	0.450	0.000**	0.000**	0.000**
Asia	CMT	0.968	1.000	1.000	0.359	0.677	0.875
	SMT	0.759	0.772	0.779	0.221	0.273	0.353
	Wald	0.753	0.728	0.645	0.208	0.241	0.314
OECD	CMT	1.000	1.000	1.000	0.000**	0.000**	0.000**
	SMT	0.996	0.819	0.822	0.000**	0.000**	0.000**
	Wald	0.898	0.751	0.657	0.000**	0.000**	0.000**
World	CMT	1.000	1.000	1.000	0.000**	0.000**	0.000**
	SMT	1.000	0.968	0.943	0.000**	0.000**	0.000**
	Wald	0.897	0.732	0.644	0.000**	0.000**	0.000**

Notes: 1961-1980 vs. 1981-2000 refers to the null hypothesis that 1961-1980 weakly dominates 1981-2000. If the null can be rejected, then we conclude that 1981-2000 strictly dominates 1961-1980. CMT = conservative maximal t statistic, SMT = simulated maximal t statistic, and Wald = Wald statistic. Figures in the cells are p-values. If p-value < 0.01, then the null is rejected at the 1% level; this is indicated by **. If 0.01 < p-value < 0.05, then the null is rejected at the 5% level; this is indicated by *.

Vita

Nor Azam Abdul Razak was born to Abdul Razak Bakar (passed away) and Hasnah Othman. He has a brother (Mohd Nazri) and two sisters (Rafidah and Roslina). He is married to Nafishah Osman, and they are endowed with four wonderful children—Nor Syuhada, Kamarul Izzat, Nor Amira, and Nor Aleeya. Since October 1997, he has been a faculty member of *Universiti Utara Malaysia* (or the Northern University of Malaysia). In August 2001, he was granted a study leave to pursue a doctoral degree in economics at the Louisiana State University.