Economics
The Open-Access, Open-Assessment E-Journal

Discussion Paper
No. 2011-37 | September 9, 2011 | http://www.economics-ejournal.org/economics/discussionpapers/2011-37

Cross-Country Growth Empirics and Model Uncertainty: An Overview

Bülent Ulaşan

Central Bank of the Republic of Turkey

Please cite the corresponding journal article:
http://dx.doi.org/10.5018/economics-ejournal.ja.2012-16

Abstract The aim of this paper is to provide an overview of empirical cross-country growth literature. The paper begins with describing the basic framework used in recent empirical cross-country growth research. Even though this literature was mainly inspired by endogenous growth theories, the neoclassical growth model is still the workhorse for cross-country growth empirics. The second part of the paper emphasises model uncertainty, which is indeed immense but generally neglected in the empirical cross-country growth literature. The most outstanding feature of the literature is that a large number of factors have been suggested as fundamental growth determinants. Together with the small sample property, this leads to an important problem: model uncertainty. The questions which factors are more fundamental in explaining growth dynamics and hence growth differences are still the subject of academic research. Recent attempts based on general-to-specific modeling or model averaging are promising but have their own limits. Finally, the paper highlights the implications of model uncertainty for policy evaluation.

Submitted as Survey and Overview

JEL O40, O47

Keywords Economic growth; convergence; cross-country growth; regression; model uncertainty; policy evaluation

Correspondence Bülent Ulaşan, Research and Monetary Policy Department, Central Bank of the Republic of Turkey, Istiklal Cad. No:10, Ulus, 06100, Ankara, Turkey; e-mail: Bulent.Ulasan@tcmb.gov.tr

The views expressed in this paper are solely the author's and do not reflect those of the Central Bank of the Republic of Turkey. The author is thankful to Wyn Morgan, Peter Wright and Mehmet N. Eris for their valuable comments. Responsibility for any remaining errors lies solely with the author.

When I finished that 1956 paper, I had no idea that it would still be alive and well 50 years later, more or less part of the folklore.

— Robert M. Solow (2007, p.4)

1 Introduction

Why do growth rates vary across countries? Why are some countries are growing rapidly and some are not growing at all? Do countries convergence or diverge in terms of per capita income? Which factors are effective in promoting economic growth? These questions are the central motivation of the recent empirical cross-country growth literature. Following the seminal studies by Kormendi and Meguire (1985), Barro (1991) and Mankiw, Romer and Weil (1992), a large number of empirical works have emerged over the last few decades. Undoubtedly, this renewed interest arose from recent developments in the theory of endogenous growth and the increasing availability of multi-country growth data sets, e.g. the Penn World Tables, (Summers and Heston (1988, 1991)).

The recent empirical cross-country growth studies are, however, mainly based on the extended versions of the neoclassical growth model in spite of the contribution of recent endogenous growth models. The most important reason is that endogenous growth theories have less explanatory power for understanding cross-country growth differences even though growth economists have reached a general consensus that whilst basic technological change is the most important determinant for understanding why the world economy as a whole has been growing indefinitely in per capita terms. For instance Barro (1997, p.8) argues that "[I]t is surely an irony that one of the lasting contributions of endogenous growth theory is that it stimulated empirical work that demonstrated the explanatory power of the neoclassical growth model." Therefore, the neoclassical growth model firstly developed by Solow (1956) and Swan (1956) is a starting point of most cross-country growth studies.

In this respect, the pioneering work by Mankiw, Romer and Weil (1992) augments the neoclassical growth model with the inclusion of human capital. The prominent aspect of this study is that it provides a coherent theoretical framework for empirical cross-country growth studies and a large body of empirical cross-country growth literature is based on Mankiw, Romer and Weil (1992). Islam (1995) and others further adapt the framework of Mankiw, Romer and Weil (1992) for panel estimation.

Conceptually, there are two main empirical approaches in the literature, namely growth accounting and growth regression, quantifying the following relation:

$$Output = F(Production\ Factors,\ Technology)$$

It is obvious that the accumulation of production factors (namely physical and human capital and population) and technological progress (whether exogenous or endogenous) are *proximate* determinants of economic growth. In other words, even though these factors explain a considerable part of cross-county growth differences and in spite of the common consensus concerning these factors as potential growth determinants, these facts bear a pertinent question: *Why are countries different in terms of proximate growth determinants?* That is why beyond the proximate determinants, explaining the *fundamental* sources of growth differences across economies is the main objective of the empirical cross-country growth literature. In doing so, the cross-country growth studies apply a wide range of new growth theories. A typical study, firstly presents a new growth theory, then suggests a proxy variable for that theory and finally concludes a cross-country growth regression including this new theory as well as the proximate determinants.[1]

The most outstanding characteristic of these new growth theories is that they are open-ended, such that the inclusion of one growth theory does not preclude that the causal role of others, as pointed out by Brock and Durlauf (2001). This means that unlike the proximate determinants of growth, there is no common consensus among the new growth theories. Whilst almost all studies include the same proximate determinants, the new growth theories change from study to study. In other words, there is no clear answer as to which of these new growth theories is more important. Wacziarg (2002, p.907) nicely summarises this phenomenon:

> All-encompassing hypotheses concerning the sources of economic growth periodically, and with the support of adequately chosen cross-country correlations, enjoy their fifteen minutes of fame. Over the last few decades, the list of proposed panaceas for growth in per-capita income has included high rates of physical-capital investment, rapid human-capital accumulation, low income inequality, low fertility, being located far from the equator, a low incidence of tropical diseases, access to the sea, favorable weather

[1] Typically such cross-country growth regressions include the initial income level, the rate of population growth, the investment ratio and a measure of human capital such as primary and secondary school enrolment rate, as well as some proxy variables for the new theory. Regression of this kind is also known as "Barro type regression" due to the pioneering work by Barro (1991).

> patterns, hands-off governments, trade-policy openness, capital-market developments, political freedom, economic freedom, ethnic homogeneity, British colonial origins, a common-law legal system, the protection of property rights and the rule of law, good governance, political stability, infrastructure, market-determined prices (including exchange rates), foreign direct investment, and suitably conditioned foreign aid. This is a growing and non-exhaustive list.

As a consequence of open-ended nature of new growth theories, a vast number of explanatory variables appears in the empirical cross-country growth literature.[2] This implies that identification of explanatory variables in a growth regression is an important task in order to highlight the exact contribution of new theories to understand economic growth. In other words, the results of empirical cross-country growth studies are very sensitive to model selection and hence presenting results of a single model is often misleading about the sources of economic growth. It is however, practically infeasible to run a cross-country growth regression encompassing all variables suggested by new growth theories due to the large number of growth variables and the small number of observations, that is the number of countries in the world is limited. Furthermore, when running cross-country growth regressions, we are missing observations of many countries due to data availability.[3] Under these circumstances, the problem of model uncertainty is an important econometric defect in cross-country growth studies. The lack of robustness has also serious implications for providing strong policy recommendations, the ultimate goal of this literature.

The objectives of this overview are threefold: First, it describes the general theoretical framework which constitutes the basis for the most empirical cross-country growth works. Second, it addresses the model uncertainty problem which is indeed immense but generally ignored in the empirical cross-country growth literature. Third, the paper highlights the importance of model uncertainty for policy evaluation.

The rest of the paper is organised as follows. Section 2 describes the basic framework for the recent cross-country growth literature. Section 3 deals with the model uncertainty problem and discusses its possible solutions.

[2] Another reason for this proliferation is the difficulties arising from construction of proxy variables for new growth theories. For instance, a theory pointing out that openness to international trade is important for economic growth does not provide a clear answer as to how we measure openness.

[3] As noted by Sala-i-Martin (2001), empirical cross-country growth works are subject to small sample econometrics. Therefore, the econometric problems discussed in cross-country growth empirics are common to other applied studies with small samples.

Section 4 briefly evaluates the policy implications that can be drawn from cross-country growth studies in the presence of model uncertainty. Section 5 concludes.

2 Empirical Framework

In this section, we provide a theoretical framework for cross-country growth regression. Special emphasis is focused on Mankiw, Romer and Weil (1992) since this study suggests a benchmark equation for much of the subsequent cross-country growth literature.

Following the standard notation, we denote the level of output by $Y(t)$, labour stock by $L(t)$ and level of labour-augmenting technology by $A(t)$ at time t. Assuming that production function exhibits constant returns to scale and labour and technology grow exogenously at rates n and g such that $L(t) = L(0)e^{nt}$ and $A(t) = A(0)e^{gt}$, output per unit of labour; $y(t) = Y(t)/L(t)$ and output per unit of effective labour; $\tilde{y}(t) = Y(t)/A(t)L(t)$ are defined. As indicated by many authors both Solow-Swan or Ramsey-Cass-Koopmans versions of the neoclassical growth model for closed economies conclude that the growth rate of per capita output is inversely related to initial level of per capita output.[4] This implies that

$$\lambda = -\frac{\partial(\dot{\tilde{k}}(t)/\tilde{k}(t))}{\partial log\tilde{k}(t)} \tag{1}$$

where a dot over the variable indicates the derivative of that variable with respect to time, $\tilde{k}$ denotes the physical capital stock per unit of effective labour and λ measures the speed of convergence (defined as how much the growth rate decreases when the capital stock increases proportionally). Notice that in equation (1), the speed of convergence is defined with a negative sign since the derivative is negative due to the marginal diminishing return to capital. Therefore, λ must be positive, and its size depends on the parameters of the model. The other important point is that λ is not constant. This means that λ decreases monotonically while capital stock converges to its steady-state value. Put differently λ is implicitly a function of $\tilde{k}(t)$ and becomes zero when the capital stock reaches its steady-state level. Therefore, we denote speed of convergence in the neighborhood of steady-sate by λ^*. Since the production function is assumed to have constant returns to scale, equation

[4] See for instance, Barro and Sala-i-Martin (1992), Mankiw et al. (1992), Mankiw (1995), Islam (1995), Durlauf et al. (2005).

(1) can be applied for the output per unit of effective labour, i.e. speed of convergence can be alternatively defined for $\tilde{y}(t)$:

$$\lambda = -\frac{\partial(\dot{\tilde{y}}(t)/\tilde{y}(t))}{\partial log\tilde{y}(t)} \tag{2}$$

Equation (1) implies that the first-order Taylor approximation of $log\tilde{k}(t)$ around the steady state yields

$$\dot{\tilde{k}}(t)/\tilde{k}(t) \cong -\lambda^*[log(\tilde{k}(t)/\tilde{k}(t)^*)] \tag{3}$$

Similarly, equations (2) and (3) imply that

$$\dot{\tilde{y}}(t)/\tilde{y}(t) \cong -\lambda^*[log(\tilde{y}(t)/\tilde{y}(t)^*)] \tag{4}$$

As can be seen, equations (3) and (4) are first-order differential equations. Equation (4) can be written more explicitly as follows

$$\frac{dlog\tilde{y}(t)}{dt} = \lambda^* log\tilde{y}(t)^* - \lambda^* log\tilde{y}(t) \tag{5}$$

Solving (5) gives

$$log\tilde{y}(t) = (1-e^{-\lambda^* t})log\tilde{y}(t)^* + e^{-\lambda^* t}log\tilde{y}(0) \tag{6}$$

Equation (6) can be expressed for output per labour instead of output per unit of effective labour as follows

$$logy(t) - logA(t) = (1-e^{-\lambda^* t})log\tilde{y}(t)^* + e^{-\lambda^* t}logy(0) - e^{-\lambda^* t}logA(0) \tag{7}$$

and so

$$logy(t) = gt + (1-e^{-\lambda^* t})log\tilde{y}(t)^* + (1-e^{-\lambda^* t})logA(0)) + e^{-\lambda^* t}logy(0) \tag{8}$$

Subtracting the logarithm of the initial level of output per capita from both sides of equation (8) and dividing by time t yields the following growth equation

$$t^{-1}(logy(t) - logy(0)) = g + \eta[log\tilde{y}(t)^* + logA(0) - logy(0)] \tag{9}$$

where $\eta = t^{-1}(1-e^{-\lambda^* t})$. The left-hand side of equation (9) shows the growth rate of output per labour between 0 and t.[5] As seen in equation

[5] Notice that the growth rate in equation (9) is defined per unit of time. If the unit of time is a year, the left-hand side of equation (9) measures the average growth rate of output per labour *annually*. On the other hand one can construct the growth rate as the log difference between initial and end of period values such that $logy(t) - logy(0)$ since equation (9) is based on the log-linear approximation of output per unit of effective labour in the vicinity of steady state. As long as it is explicitly expressed, both approaches are in essence the same, and choosing between these two depends on the researchers' preferences.

(9), the growth rate of per capita output may be decomposed into two main factors. The first one is the growth rate of technological progress, g. The second one is the distance between initial level of output per unit of effective labour and its steady state value, $log\tilde{y}(t)^* - log\tilde{y}(0)$. In order to show the second factor more explicitly, equation (9) can be written as

$$t^{-1}(logy(t) - logy(0)) = g + \eta[log\tilde{y}(t)^* - log\tilde{y}(0)] \tag{10}$$

As shown in equation (10), the growth rate of per capita output is inversely related to the initial level of output per unit of effective labour while it is positively related to the steady state level of output per unit of effective labour and hence its determinants. As time approaches infinity, i.e. as an economy converges to its steady state, the effect of the second factor vanishes, and at the steady state, is equal to zero. This means that in the long run, the growth rate of per capita output is determined by the rate of technological progress, g.

If we assume that rate of technological progress, g and the determinants of the steady state level of output per unit of effective labour are constant across countries, then each economy approaches the same steady state in the long run. That is why, countries with a lower initial level of output per unit of effective labour grow faster than those with a higher initial level of output per unit of effective labour during the transition period due to the diminishing returns to capital. This result is known as the *absolute convergence hypothesis* and predicts that poor countries tend to catch up with rich ones. However, if the countries have different values of g and determinants of steady state value of output per unit of effective labour, then steady states will be different across countries. Therefore, each economy will converge to its own steady state rather than a common steady state, and the speed of this convergence will be inversely related to the distance of the initial level from the steady state. This property is again a result of the assumption of diminishing returns to capital, so that economies which have less capital per head relative to its steady state level tend to have higher rates of return and so faster growth. In this situation, the neoclassical growth model implies *conditional convergence* instead of absolute convergence in the sense that an economy with a lower initial value of per capita output tends to generate higher growth rate of per capita output if g and determinants of the steady state value of output per unit of effective labour are the same across countries or their effects are controlled.

If convergence hypothesis defined above is true, then we expect a negative association between the level of initial income and subsequent growth rate across countries. In order to test the convergence hypothesis, researchers run the growth-initial income regressions. Therefore, in the literature both absolute and conditional convergence are sometimes referred as β-convergence due to the coefficient of the initial income level (namely β) in the cross-country growth regression (see, for instance, Sala-i-Martin (1996)). Notice that both absolute β-convergence and conditional β-convergence imply that the initial conditions of countries do not matter for their steady state levels of income. The only difference between these two convergence concepts is that the later also allows the structural heterogeneity across countries, that is structurally similar countries converge to the similar income level in the long run. Hence, absolute convergence and conditional convergence coincide if all countries are structurally the same. On the other hand, Quah (1996) criticises the concept of β-convergence and suggest to the concept of σ-convergence measuring the relative dispersion of per capita income level across countries. Whilst the sentiment of Quah (1996) rings true, β-convergence is still an important concept since it is a necessary condition for σ-convergence.[6]

Equations (9) and (10) are the basis for the estimation of cross-country growth regressions in the empirical growth literature. Adding an error term μ, which is independent from all right-hand side variables, yields the following cross-country growth regression

$$t^{-1}(logy_i(t) - logy_i(0)) = g_i + \eta log\tilde{y}_i(t)^* - \eta logy_i(0) + \eta logA_i(0) + \mu_i \quad (11)$$

where subscript i denotes the country i. This last equation is the basic cross-country growth regression in discrete time which is derived from continuous time neoclassical Solow-Swan growth model.

In this context, the seminal study by Mankiw, Romer and Weil (1992, MRW hereafter) augments the Solow-Swan version of neoclassical growth model by adding the accumulation of human capital. They assume a Cobb-Douglas production function such that production at time t in country i is given by

$$Y_i(t) = K_i(t)^\alpha H_i(t)^\beta (A_i(t)L_i(t))^{1-\alpha-\beta} \quad (12)$$

where the notation here is again standard such that Y is output, K is physical capital, H is the stock of human capital, L is labour, and A is level of technology. MRW (1992) assumes that $\alpha + \beta < 1$, which means that there are decreasing returns to both kinds of capital. Labour stock and the level

[6] Closely related to β-convergence, another concept of convergence is *club convergence* suggested by Durlauf and Johnson (1995) and Galor (1996). This concept says that initially and structurally similar countries converge to similar steady states. See Islam (2003) and Durlauf and Johnson (2008) for nice surveys on convergence debate.

of technology are assumed to grow exogenously at rates n and g, respectively as before.

The production function in equation (12) can be written in its intensive form. More clearly, it can be expressed in terms of per unit of effective labour as it shows constant returns to scale property.

$$\tilde{y}(t) = \tilde{k}(t)^{\alpha}\tilde{h}(t)^{\beta} \tag{13}$$

where $\tilde{h}$ is the stock of human capital per unit of effective labour and the remaining variables are as before. The model assumes that a constant fraction of output is invested in both physical and human capital such that s_K is the fraction of income invested in physical capital and s_H is the fraction of income invested in human capital. Defining δ as the depreciation rate of both physical and human capital, yields

$$\dot{\tilde{k}}(t) = s_K\tilde{y}(t) - (n + g + \delta)\tilde{k}(t) \tag{14}$$

$$\dot{\tilde{h}}(t) = s_H\tilde{y}(t) - (n + g + \delta)\tilde{h}(t) \tag{15}$$

Equations (14) and (15) imply that the economy converges to a steady state defined as follows

$$\tilde{k}(t)^* = \left(\frac{s_K^{1-\beta} s_H^{\beta}}{n + g + \delta}\right)^{1/1-\alpha-\beta} \tag{16}$$

$$\tilde{h}(t)^* = \left(\frac{s_K^{\alpha} s_H^{1-\alpha}}{n + g + \delta}\right)^{1/1-\alpha-\beta} \tag{17}$$

Substituting equations (16) and (17) into the production function gives the steady state level of output per unit of effective labour:

$$\tilde{y}(t)^* = \left[\frac{s_K^{\alpha} s_H^{\beta}}{(n + g + \delta)^{\alpha+\beta}}\right]^{1/1-\alpha-\beta} \tag{18}$$

Using the definition of speed of convergence expressed in equation (2) with the equations from (13) to (18), the convergence coefficient in the vicinity of the steady state can be defined by[7]

$$\lambda^* = (1 - \alpha - \beta)(n + g + \delta) \tag{19}$$

[7] Since the derivation of convergence coefficient in the augmented neoclassical growth model is available elsewhere, we do not need to elaborate it here. The reader can apply, *inter alia*, Mankiw (1995) or Barro and Sala-i-Martin (2004) for details in derivation.

λ^* measures how rapidly a country's output per unit of effective labour approaches its steady state value in the neighbourhood of the steady state. For instance, if we assume that the sum of rates of population growth, technological progress and depreciation is seven percent and capital shares are one-third, then λ^* would be equal to 0.023. This means that 2.3 percent of the gap between a country's steady state and its current income level is eliminated each year and halfway to convergence takes approximately 30 years, in the absence of any other shocks.[8]

In order to get a cross-country growth regression, we need to substitute expression (18), i.e. steady state level of output per unit of effective labour into equation (11). This produces

$$\begin{aligned} t^{-1}(logy_i(t) - logy_i(0)) = g &+ \eta\frac{\alpha}{1-\alpha-\beta}logs_{i,K} + \eta\frac{\beta}{1-\alpha-\beta}logs_{i,H} \\ &- \eta\frac{\alpha+\beta}{1-\alpha-\beta}log(n_i+g+\delta) - \eta logy_i(0) \\ &+ \eta logA_i(0) + \mu_i \end{aligned} \tag{20}$$

As can be seen from the last equation, MRW assume that rates of technological progress and of depreciation are constant across countries. On the other hand, logarithm of initial level of technology is assumed to be different across countries and be equal to the sum of a fixed parameter, a and a country specific shock, ε_i such that

$$logA_i(0) = a + \varepsilon_i \tag{21}$$

According to MRW, the level of initial technology represents not only the technology but also the resource endowment, institutions, climate and so on. Therefore, initial differences across countries are reflected by the term ε_i. Substituting equation (21) into equation (20) yields the following cross-country growth regression;

$$\begin{aligned} t^{-1}(logy_i(t) - logy_i(0)) = g + \eta a - \eta logy_i(0) - \eta\frac{\alpha+\beta}{1-\alpha-\beta}log(n_i+g+\delta) + \\ \eta\frac{\alpha}{1-\alpha-\beta}logs_{i,K} + \eta\frac{\beta}{1-\alpha-\beta}logs_{i,H} + \mu_i + \eta\varepsilon_i \end{aligned} \tag{22}$$

[8] According to the equation (6), the half-way convergence to steady state requires the condition $1 = 2e^{-\lambda^* t}$. Therefore, the half-life convergence to steady state is $log(2)/\lambda^*$. Similarly, elimination of a three-quarter gap must satisfy the condition $1 = 4e^{-\lambda^* t}$, and takes $2log(2)/\lambda^*$. For instance in the example above, three-quarters of convergence to steady state takes $2log(2)/0.023 = 60.3$ years.

The most critical assumption of MRW (1992, p.411) is that "[t]he rates of saving and population growth are independent of country-specific factors shifting the production function." This means that $s_{i,K}$, $s_{i,H}$, and n_i are independent from the country specific shocks ε_i and thus, a cross-country growth regression expressed as in equation (22) can be estimated by OLS.

The cross-country growth regression in equation (22) can be written in its reduced form as follows

$$\varrho_i = \pi_0 + \pi_1 logy_i(0) + \pi_2 log(n_i + g + \delta) + \pi_3 logs_{i,K} + \pi_4 logs_{i,H} + \upsilon_i \quad (23)$$

where

$$\begin{aligned}
\varrho_i &= t^{-1}(logy_i(t) - logy_i(0)) \\
\pi_0 &= g + \eta a \\
\pi_1 &= -\eta \\
\pi_2 &= -\eta \frac{\alpha + \beta}{1 - \alpha - \beta} \\
\pi_3 &= \eta \frac{\alpha}{1 - \alpha - \beta} \\
\pi_4 &= \eta \frac{\beta}{1 - \alpha - \beta} \\
\upsilon_i &= \mu_i + \eta \varepsilon_i
\end{aligned}$$

Equation (22) and its reduced form in (23) are the basis of the augmented neoclassical growth model. MRW estimated the augmented neoclassical growth model for 98 countries (oil producing countries are excluded) over the 1960-1985 period. The share of investment in GDP and the fraction of working-age population enrolled in secondary school are used as proxy variables for s_K and s_H, respectively. All right-hand side variables, except the initial level of GDP per worker are entered into the regression as period averages instead of their initial value.[9] Regression results show that the average growth rate of GDP per worker is positively correlated with the investment to GDP ratio and secondary school enrolment rate and negatively with the initial income level and population growth. Moreover, MRW estimate the augmented neoclassical model imposing the restriction that coefficients on $log(n + g + \delta)$, $logs_K$ and $logs_H$ add up to zero. Finding that this restriction is not rejected, MRW conclude the regression estimates of $\lambda^* = 0.0142$,

[9] Note that theory does not provide a clear answer for choosing between period averages and initial values, since these variables are considered as constant over the period and exogenous. Yet, the common practice in the literature is to use average values over the period.

$\alpha = 0.48$ and $\beta = 0.23$, which denote the convergence rate, physical and human capital shares in the vicinity of the steady-state, respectively. According to MRW, their estimation results produce a lower convergence rate than the standard neoclassical growth model excluding human capital and remove some anomalies which are not captured by the standard model. In other words, with the inclusion of human capital, differences in saving, education and population growth produce a consistent explanation for cross-country growth variations.

Even though MRW provide a coherent framework to explain cross-country growth differences, it is subject to a number of criticisms. The most important one is that it is unlikely that variations in the initial level of technology are uncorrelated with the right-hand side variables. As mentioned before, the initial level of technological efficiency is omitted from the cross-country regression since it is not observed. Yet, if initial income, saving rates and population growth are correlated with the initial technological efficiency, then coefficient estimates of regressors will be biased. As suggested by Islam (1995), one solution to this problem is to employ panel data estimation methods. Since the initial level of technology is time invariant, it can be considered as a country fixed effect. Islam (1995), Caselli et al. (1996) Lee et al. (1997) amongst others apply the augmented neoclassical growth model on panel data. An outstanding result of these studies is that they find a higher rate of conditional convergence compared to cross-sectional studies. In addition, they generally conclude that other explanatory variables, especially human capital, either are insignificant or have wrong signs.

The second criticism is that the secondary school enrolment rate is not an appropriate proxy for the investment rate in human capital. An important issue concerning school enrolment rate is that this variable is sometimes used as a proxy for level of human capital sometimes as a measure of change in human capital. It is however more appropriate to use school enrolment rate as a flow variable for human capital as indicated by Barro (1991) and Barro and Lee (1994b). Indeed, in the cross-country growth literature there are many studies (such as Barro (1991), Levine and Renelt (1992), Sala-i-Martin (1997a,b), Sala-i-Martin et al. (2004)) as well as Mankiw et al. (1992) that employ school enrolment rate as a proxy for accumulation of human capital and find that school enrolment rate is positively and significantly associated with economic growth. However, there are also other studies strongly criticizing these findings. For instance Bils and Klenow (2000) argue that strong empirical relation between growth and school enrolment rate is spurious since it is more likely that both variables are correlated with other omitted factors such as openness to international trade or institutions. In addition, according to these authors there is the possibility that this relation reflects reverse causality. Similarly, Pritchett (2001) points out secondary school enrolment rate is an extremely poor proxy for growth in average years of schooling because school enrolment rates, especially those in developing countries, substantially increase over the time period in the cross-country growth analysis.

Due to these criticisms, there is a tendency in the literature about the schooling years per person published by Barro and Lee (1994a, 2000) as a more reliable measure for the level of human capital. However, some studies such as Benhabib and Spiegel (1994) and Pritchett (2001) employ average years of schooling as a measure of human capital stock and conclude that the relationship between change in years of schooling and growth of per capita income is insignificant and mostly negative. One possibility for this adverse relation is outlier effect. Temple (1999b) concludes a positive and significant relation between change in schooling year and growth when a number of extreme observations are omitted. Another possibility is that these studies are based on growth accounting framework rather than standard cross-country growth regression and hence their regression results may have suffered from omitted variable bias. However, in spite of these possibilities, an important conclusion from these studies is that neither school enrolment rates nor average years of schooling are good proxies for human capital. The most important reason is that they do not directly measure cognitive skills of labour force. This leads some researchers to employ alternative variables measuring directly the quality of labour force such as teacher-student ratio or math and science test scores.[10] On the other hand, some authors (for example Temple (1999a), Benhabib and Spiegel (1994), Krueger and Lindahl (2001), Bils and Klenow (2000), Klenow and Rodríguez-Clare (1997), Hall and Jones (1999), Acemoglu (2009)) suggest measures of human capital based on returns to schooling or some measures based on the findings of other micro studies, specifically Mincerian approach to human capital.[11]

There is no doubt that the relation between growth and schooling (hence human capital) is a complex one. We expect a positive relation between these two due to the fact that education directly increases productive skills of labour force. In addition, schooling can stimulate economic growth through other channels such as reducing corruption, better conflict management, in-

[10] For instance, Hanushek and Kimko (2000) employing international math and science test scores from 31 countries conclude a significant and positive correlation between this variable and growth. Similarly, Jones and Schneider (2006) find that national average IQ test score is positively correlated with growth. Their finding is robust such that IQ test score passes a Bayesian model averaging test at 99.8 significance level.

[11] According to this approach, human capital is an exponential function of years of schooling.

creasing health quality and so on. However, an increase in school attainment is necessary but not sufficient condition for accumulation of human capital. As pointed out by Pritchett (2001) institutional environment and demand for human capital are important factors. Yet, it is more likely that schooling significantly contributes to the level of human capital since it teaches how to learn and thus help to adapt and use new technological advances (Phelps (1995)). Therefore, variables measuring school attainment can still be used as a proxy for human capital, especially in the absence of better data.[12] MRW argue that if secondary school enrolment rate is proportional to saving rate for human capital (a reasonable assumption), then it can be used in the cross-country growth regressions. Of course problems such as data quality or measurement error associated with school enrolment rate are important. However, it may be worth reminding that all proxy variables in the cross-country literature are not free of these problems. As claimed by Mankiw (1995), many variables in the literature are crude proxies at best.

Thirdly, the assumption that the rate of technological progress is constant across countries is criticised. According to MRW, technology differs across countries due to the differences in initial level of technology, not differences in technological improvements. Put differently, they consider that technology is a public good freely and equally spreading over the world. Therefore, differences in growth are a result of differences in saving rates and population growth. However, as argued by Temple (1999), there is no logical reason to expect that countries with initially different levels of technology experience the same rate of technological improvement. For instance, Bernard and Jones (1996) and Lee, Pesaran and Smith (1997) indicate that rates of technological progress vary across countries, even among industrial ones. Therefore, it seems difficult to explain growth miracles after the Second World War, such as Japan and South Korea, as purely a result of capital accumulation. On the other hand, one may conclude that this assumption is less unrealistic in the long run. More clearly, even if the level of technology is different across countries, in the long run the rate of technological progress will be the same over the world since countries try to access all technology available everywhere.[13]

[12] Recently, Cohen and Soto (2007) have provided a new data set for average years of schooling over the period 1960-2000 as an alternative to schooling years published by Barro and Lee (2000). These authors replicate the previous studies which concluded negative and insignificant relation between growth and schooling and find that their new series is positively and significantly correlated with growth.

[13] Of course, the fact that the level of technology grows at the same rate across countries in the long run does not necessarily mean that one can assume a common rate of

Finally, some authors (for example Hall and Jones (1999), Frankel and Romer (1999), and Acemoglu et al. (2001)) suggest that the theoretical framework provided by MRW can be used for income level regression instead of growth regressions.[14] Even though this suggestion seems reasonable since the primary objective of growth studies is to explain growth and ultimately income differences across countries, the disadvantages of income level regressions are twofold. First, the possible endogenity problem between the dependent variable and regressors is more obvious and finding good instruments in order to solve this problem is almost impossible. Second, this approach explicitly requires the assumption that countries are in their steady states.

Despite these problems, the large body of empirical cross-country growth literature consists of extended versions of the baseline specification in equation (23). A recent extension of this specification occurs through adding proxy variables suggested by the new growth theories as

$$\varrho_i = \pi_0 + \pi_1 log y_i(0) + \pi_2 log(n_i + g + \delta) + \pi_3 log s_{i,K} + \pi_4 log s_{i,H} + \psi Z_i + \upsilon_i \quad (24)$$

where Z is a vector of additional explanatory variables offered by new growth theories and ψ is a coefficient vector of additional covarities. The extended versions of the augmented neoclassical growth model in (24) can be rewritten in its generic form which is sometimes useful as follows

$$\varrho_i = \gamma + \pi X_i + \psi Z_i + \upsilon_i \quad (25)$$

where γ is constant term and X_i is a vector of explanatory variables suggested by the augmented neoclassical growth model, i.e. proximate determinants of growth and π is a vector of coefficient parameters of X variables. However, there are some unclear points in the extended versions of MRW, as argued by Temple (1999a).

Firstly, whether recent extensions attempt to explain differences in initial level of technology or to allow differences in the rate of technological progresses across countries is unclear. Put differently, whether Z_i determines the steady state level of income or long run growth rate is not defined. As can

technological progress for any given sample. See Temple (1999) and Aghion and Howitt (1999) for further discussion.

[14] Taking the logarithm of the steady state value of output per unit of effective labour expressed in equation (18) and rearranging it, produce the following level regression:

$$log \frac{Y_i(t)^*}{L_i(t)^*} = a + gt - \frac{\alpha+\beta}{1-\alpha-\beta} log(n_i + g + \delta) + \frac{\alpha}{1-\alpha-\beta} log s_{i,K} + \frac{\beta}{1-\alpha-\beta} log s_{i,H} + \varepsilon_i$$

where $log(Y_i(t)^*/L_i(t)^*)$ is the steady-state level of per worker income, and the other variables are as before.

be easily seen, the only difference between the cross-country growth specification by MRW and its recent extensions is that the term $g+\eta a+\eta\varepsilon_i$ in equation (22) is replaced by the term $g+\eta a+\psi Z_i+\eta\varepsilon_i$ in equation (24). This would lead one to interpret the introduction of ψZ_i to relate to initial differences in the level of technology, since extended versions of MRW ignore the fact that the term $log(n_i+g+\delta)$ should be replaced with the term $log(n_i+g_i+\delta)$ if new growth theories affect the rate of technological progress. Of course, allowing new growth theories to affect the rate of technological progress is not easy since this makes the cross-country growth regression nonlinear via the term $log(n_i+g_i(Z_i)+\delta)$.[15]

Secondly, if Z_i is correlated with the initial level of technology, then Z_i enters the cross-country growth regression with the expected sign even if Z_i does not have any effect on the long run growth rate. For instance, if initially more efficient countries are more open to international trade, then an openness variable will have a positive sign even though openness does not affect growth in the long run. Yet, in spite of these facts, one can claim that Z_i has an effect on the long run growth rate. The reason is that many cross-country growth works cover 20 or 30 years and it is not reasonable to assume that countries experience the same rate of technological progress over the time period. As suggested by Durlauf, Johnson and Temple (2005), whether Z_i affects the income level or growth rate in the long run depends on the researcher's prior beliefs.

However, even if it is plausible to assume that the interpretation of Z_i depends on the researcher's beliefs, another important problem related to Z_i remains. As mentioned earlier, while the X_i variables are generally constant in empirical cross country studies, there is no consensus about the Z_i variables in the literature. Therefore selecting Z_i variables is problematic and the selection differs from one study to another and thus raise the model uncertainty problem in cross-country growth regressions. We turn to this next.

3 Model Uncertainty and Cross-Country Growth Regressions

It is obvious that one of the most fundamental and controversial problem with cross-country growth regressions is model uncertainty and this issue has been acknowledged by many authors since the important work by

[15] Rodrìguez (2007) recently attempts to fill this gap and empirically analyses the nonlinearities in the growth process.

Levine and Renelt (1992).[16] Indeed, model uncertainty is a crucial problem for any kind of empirical work in economics. However, the degree and solution to this problem become more severe and difficult in the context of cross-country growth regression since, as pointed out by Brock and Durlauf (2001, p.234), growth theories are fundamentally open-ended in the sense that " [t]he idea that the validity of one causal theory of growth does not imply the falsity of another." Thus new growth theories suggest a wide range of different explanations for cross-country growth differences such as quality of institutions, political stability, resource curse, population heterogeneity, the role of geography and so on. For instance, a recent survey by Durlauf, Johnson and Temple (2005) concludes that 145 different proxies have been found to be statistically significant in at least one study. This implies that identification of explanatory variables is a very important task and thus the problem of omitted variable bias in a particular cross-country regression is immense.

However, it is impossible to simply run a cross-country growth regression including all variables suggested by new growth theories due to large number of growth variables and the small sample at hand, that is the limited number of countries in the world. Furthermore, the number of countries in a particular cross-country growth regression is considerably less than the actual number of countries because of data availability. Of course, in the empirical cross-country growth literature, there is no study attempting to employ all possible variables. Rather, many studies chose a subset of explanatory variables and then report a selected model with the results of diagnostic test to provide robust evidence for one or more of the variables of interest. However, during the last decade this approach has been criticised since the results of these studies are very sensitive to included and/or excluded variables. The main difficulty in these studies is that several different models may all provide reasonable representations of the data, but lead to very different conclusions about what causes economic growth. Under these conditions, presenting results of a single preferred model can often be misleading.

Brock, Durlauf and West (2003, p.268) characterise the model uncertainty in a more general context. These authors suggest that "[i]t is useful in specifying a model space to consider several distinct levels of model uncertainty and build up the space sequentially." They then highlight three basic aspects of model uncertainty: First and most importantly, "theory uncertainty" stems from disagreements over alternative theories used to explain the phenomenon. Of course, this disagreement is closely related to the absence of strong empir-

[16] See for instance Mankiw (1995), Sala-i-Martin (1997a,b), Temple (1999a, 2000), Brock and Durlauf (2001).

ical evidence that would be conclusive for ranking alternative theories; The second is “specification uncertainty”. Many empirical proxies for a particular variable give rise to this kind of uncertainty. Therefore, specification uncertainty is sometimes referred to as “proxy uncertainty”. However, specification uncertainty encompasses the possible nonlinearities and lag length of variables as well as proxy uncertainty; The third is “heterogeneity uncertainty” stemming from the heterogeneity among different observations. For example, in the growth context, the effect of a particular theory and/or variable on Kenya will undoubtedly be different from that on the United Kingdom. This is why one needs to clarify whether there is heterogeneity in the growth process among the countries or regions being considered. Different specifications of heterogeneity among countries and regions produce different models and raise model uncertainty.

In short, theory, specification and heterogeneity uncertainties related to the model selection process produce different models.[17] Therefore, specifying the model space is the first step in handling the model uncertainty problem. However, the specification of the model space is generally based on the researcher’s judgment. For example, whilst one researcher may interpret model uncertainty as proxy uncertainty, another may emphasise only heterogeneity uncertainty in the context of cross-country growth study.

Levine and Renelt (1992) is the first study to take into account model uncertainty in the empirical cross-country growth literature. Employing a variant of Leamer’s extreme bounds analysis (Leamer (1983), Leamer and Leonard(1983)), these authors test the robustness of coefficient estimates for a large number of policy indicators as other explanatory variables alter. To illustrate the basic mechanism of a modified version of extreme bounds analysis (EBA hereafter) employed by Levine and Renelt (1992), consider the generic representation of cross-country growth regression expressed in equation (25) in the following form

$$\varrho_i = \gamma + \pi X_i + \delta p_i + \psi Z_i + \upsilon_i \tag{26}$$

where X is the vector of variables always included in the regressions and consists of the initial level of real GDP per capita in 1960, the investment share of GDP, the initial secondary school enrolment rate, and the average annual rate of population growth; Z is a subset of variables chosen from a pool of over 50 variables suggested by previous growth studies and p is the variable of interest.

[17] As mentioned before, cross-country growth regression is a very good case for all levels of model uncertainty. However, it is worth recalling that other applied works in economics are not free from model uncertainty as defined here.

In order to carry out an EBA test, Levine and Renelt (1992) firstly run the benchmark regression including only X variables and the variable of interest, p. In the second step, the authors compute the regression results for all possible linear combinations of one to three variables from the pool of variables and determine the highest and lowest values for the coefficient estimate of variable of interest, $\hat{\delta}$, and its the corresponding standard error, $\hat{\sigma_{\delta}}$. The rationale for choosing up to three Z variables from the pool is to avoid the possible multicollinearity problem which inflates the standard errors of coefficient estimates. In addition, Levine and Renelt (1992) restrict the number of pool variables to seven which are used as fiscal, trade, monetary, macroeconomic uncertainty, and political stability indicators in the literature.[18] Otherwise, it is highly likely that the variable of interest loses its significance. Finally, for every variable of interest, they further restrict the pool of variables such that some variables are dropped from the pool if they measure the same phenomenon with respect to the variable of interest. Therefore, in this study, EBA is restricted such that total number of explanatory variables included in any regression to be eight or less. In this regard, Levine and Renelt (1992) identifies the upper extreme bound as the highest value of $\hat{\delta}$ plus two times its standard error and define the lower extreme bound as the lowest value of $\hat{\delta}$ minus two times its standard error over all possible models for the variable of interest and then conclude the EBA test such that the variable of interest, p is robust if its coefficient is significant and has the same sign at the extreme bounds ($\hat{\delta} \pm 2\hat{\sigma_{\delta}}$). If the coefficient of variable of interest does not remain significant and/or changes its sign, then EBA test indicates that this variable is fragile.

In short, Levine and Renelt (1992) investigate the robustness of the relationship between growth and a variable of interest according to the stability of the sign and statistical significance of the estimated coefficient over all possible models. Using more than 50 variables over the 1960-1989 period, Levine and Renelt (1992) find that only the initial level of income and the share of investment in GDP are robustly correlated with growth. In other words, except for these two, they conclude that all variables are fragile.[19]

Sala-i-Martin (1997a,b) criticises Levine and Renelt (1992) and argues that the EBA test is too extreme as they conclude a variable is fragile if

[18] The pool of variables are the average rate of government consumption expenditures to GDP, the ratio of exports to GDP, the average inflation rate, the average growth rate of domestic credit, the standard deviation of inflation, the standard deviation of domestic credit growth and an index of the number of revolutions and coups.

[19] In addition, Levine and Renelt (1992) carry out the same analysis for the investment rate and conclude that only trade ratio is robustly and positively associated with investment.

the coefficient estimate loses its statistical significance and/or changes its sign even in one regression. Sala-i-Martin (1997a,b) suggests that one should consider the whole distribution of $\hat{\delta}$, and assign a level of confidence for the robustness test instead of labelling a variable as robust or fragile according to extreme bounds. In order to compute the cumulative distribution function of $\hat{\delta}$, he calculates the weighted averages of all estimates of δ and its corresponding standard error for each model as follows

$$\hat{\delta} = \sum_{i=1}^{M} \omega_i \hat{\delta}_i \tag{27}$$

$$\hat{\sigma_\delta} = \sum_{i=1}^{M} \omega_i \hat{\sigma}_{\delta,i} \tag{28}$$

where $\hat{\delta}$ and $\hat{\sigma_\delta}$ are the weighted averages of the coefficient of variable of interest and of its standard error over all possible models, respectively. The weights, ω_i are the critical point of the analysis and calculated as a proportion of the integrated likelihoods of each model as follows

$$\omega_i = \frac{\ell_{\delta i}}{\sum_m^M \ell_m} \tag{29}$$

where ℓ_m is the likelihood of each of the M models. Notice that $\sum_m^M \omega_i = 1$.

As can be seen, the weighting scheme gives higher weights to the regressions or models which are more likely to be the true model. In order to measure the robustness of each variable, Sala-i-Martin (1997a) calculates the cumulative distribution function as follows: First, he assumes that $\hat{\delta}$ has a normal distribution over models, so he uses the normal distribution tables; Second, he relaxes the assumption that $\hat{\delta}$ has a normal distribution and calculates the cumulative distribution function as the weighted sum of a normal cumulative distribution function. The weights are again proportional to likelihoods. For the sake of comparability with Levine and Renelt (1992), Sala-i-Martin (1997a,b) allows the model to include three fixed variables namely income level in 1960, life expectancy in 1960 and primary school enrolment rate in 1960. Combining these fixed variables with a variable of interest and a group of three variables from the pool consisting of 59 variables, Sala-i-Martin (1997a) estimated nearly 2 million regressions. Differently from Levine and Renelt (1992), he tries every three combinations of doubtful variables in order to reduce computational burden. Therefore, his regressions always contain seven explanatory variables. Sala-i-Martin (1997a,b) argues that if 95 percent of cumulative distribution function of $\hat{\delta}$ lies on each side of zero, then that variable can be considered robust. Put differently, a variable is robust if its statistical significance and sign hold over 95 percent of all possible models. Unlike Levine and Renelt (1992), Sala-i-Martin (1997a) concludes that 21 of 59 variables are robustly correlated with growth. In his subsequent work, Sala-i-Martin (1997b) introduced the average investment rate between 1960 and 1990 as an additional fixed regressor. The reason for including average investment rate in the later study is to highlight the channels through which the variable of interest affects growth, namely via effects on the level of efficiency. Therefore, Sala-i-Martin (1997b) estimates two million more regressions combining four fixed variables with the variable of interest and again trios from remaining 59 variables and concludes that 17 of 59 variables are robustly correlated with growth.[20]

Even though the Levine and Renelt (1992) and Sala-i-Martin (1997a,b) versions of EBA provide useful information concerning the model uncertainty problem in the cross-country growth literature, these studies are subject to important drawbacks. Firstly, EBA is heavily criticised by McAleer et al. (1985) and Hendry and Mizon (1990). One criticism is that extreme bounds depend on the selection of doubtful variables. In other words, different selections yield different extreme bounds. Generally, most of EBA applications classify the variables as fixed and doubtful variables as in the case of Levine and Renelt (1992). This classification is sometimes arbitrary, even though it is reasonable and defendable in the study by Levine and Renelt(1992). Secondly, extreme bound levels can come from models which are unreasonable in some ways or even clearly poor. For instance McAleer (1994) criticises Levine and Renelt (1992) since they present summary statistics of extreme bounds without diagnostic tests and also ignoring functional form misspecification. Therefore, Granger and Uhlig (1990) propose *reasonable EBA* such that extreme bounds may come from models having R^2 values very close to maximum achievable value of R^2 over the model space. If this is done, then models with relatively low goodness-of-fit will be eliminated. Similarly, Temple (2000) suggests reporting a table listing models with the results of diagnostic tests instead of presenting only upper and lower extreme bounds. Thirdly and perhaps most importantly, if one of the doubtful variables is important in explaining the dependent variable, then fragile results are inevitably obtained. More clearly, while testing for the sensitivity of a particular variable of interest over all possible models, that key variable will be

[20] Sala-i-Martin (1997a,b) also carried out his approach for fixed variables in order to gain same confidence about their robustness and found that initial income level, initial primary school enrolment, initial life expectancy and average investment rate are all robust determinants of growth.

sometimes omitted. Models excluding key varible(s) certainly affect the sign and statistical significance of $\hat{\delta}$. Therefore, it is possible to conclude that EBA is useful but not efficient and so overstates model uncertainty. On the other hand one may argue that Sala-i-Martin (1997a,b) version of EBA is more reasonable than Levine and Renelt (1992), but statistical properties of this approach, especially the weighting scheme of models, are unclear since they are not based on a formal statistical theory (as Barro and Sala-i-Martin (2004) point out).

In summary, both versions of EBA fail to provide satisfactory solutions to the problem of identifying the true determinants of growth. Two approaches recently appeared in the literature. The general-to-specific modelling (GETS henceforth) approach[21] is based on the idea that the true model can be characterised by a sufficiently rich regression. This means that a regression including all possible regressors has all the information about the dependent variable. However, the information presented by the *general regression* can be represented by a parsimonious regression called the *specific regression*. Of course, this specific regression must have some desirable properties such that it must be well defined, it should encompass every other parsimonious regression and so on. In short, the GETS approach starts with the general model and then searches for a specific model comparing all possible models in the model space according to some statistical criterion. Bleaney and Nishiyama (2002), Hoover and Perez (2004), and Hendry and Krolzig (2004) apply this approach to cross-country growth regressions.

The paper by Bleaney and Nishiyama (2002) is, in essence, based on the encompassing test among three non-nested models for cross-country growth regressions suggested by Sachs and Warner (1997), Barro (1997) and Easterly and Levine (1997). Even though these three models have some common explanatory variables, Bleaney and Nishiyama (2002) conclude that none of them dominates each other according to non-nested hypothesis testing procedures. This means that a model encompassing these three models fits the data better. Therefore, they combine the explanatory variables of the three models and eliminate them according to the GETS approach to derive a specific model which passes a battery of statistical tests successfully. According to Bleaney and Nishiyama (2002), this model cannot be improved by adding or omitting any variable, and can be used as a benchmark model in order to test new growth theories.

Hoover and Perez (2004) and Hendry and Krolzig (2004) apply the GETS methodology directly to the data set employed by Sala-i-Martin (1997a,b) af-

[21] It is sometimes referred to as London School of Economics (LSE) methodology.

ter some adjustment.[22] In both studies, a linear model including the number of revaluations and coups, the ratio of equipment investment, fraction of confucians, fraction of open years according to the Sachs and Warner (1995) criteria and fraction of protestants as explanatory variables of growth is estimated.[23] An interesting point concerning the results of these two studies is that the R^2 values of the regressions are found to be 0.42 and 0.44, respectively. This implies that the selected models explain less than 50 percent of the cross-country growth differentials. In addition, theoretically important variables, such as initial income level and variables relating to human capital, are not included in the final model.

One important criticism of the GETS methodology is that there can be several simplification paths from the general model and there is no guarantee that a particular simplification path leads to the true model. That is why, the GETS approach is sometimes referred as "sophisticated data mining", as Hendry (1995) points out. However, Hoover and Perez (2004) and Hendry and Krolzig (2004) argue that the GETS approach employed in their papers is based on multiple-path searching program in order to handle this objection. In other words, both studies implement the GETS approach by employing the automated search algorithm first suggested by Hoover and Perez (1999) and improved by Krolzig and Hendry (2001), in order to take into account competing models derived from different search paths and to select one on the basis of encompassing tests. In particular, the *PcGets* algorithm developed by Hendry and Krolzig (2005) is effective in reducing searching costs when the initial model is more general than needed.

The selection process of the specific model is based on six stages: First, assuming the true model is nested in a sufficiently rich model, a general unrestricted model (GUM) is formulated. In the second and third stages, a set of mis-specification tests and selection criteria are applied for final selection between mutually encompassing congruent models and then the GUM is estimated to check the congruence of the specification. Therefore, after the second and third stages, the GUM is reformulated as a baseline general model for the next steps. Fourth, a pre-search reduction process is carried out. In other words, the highly insignificant variables are eliminated using a less stringent significance level in order to simplify large dimensional problems. Thus, this stage is optional, not necessary. The fifth and main stage consists

[22] The original data set used by Sala-i-Martin (1997a,b) contains 64 variables (including the dependent variable) for 138 countries. After a number of variables and countries are dropped from the data set in order to provide a complete data matrix, the resulting data set includes 126 countries and 61 variables and the dependent variable.

[23] Hendry and Krolzig (2004) also apply the GETS methodology on the data set used by Fernández et al. (2001).

of multiple-path reduction searches. In this stage, many possible reduction paths are undertaken from each feasible initial deletion and each reduction is diagnostically evaluated for the congruence of the final model. That is, after a particular reduction path, if all diagnostic tests are successfully passed and all remaining variables are statistically significant, then that model is considered as a terminal specification. Next another reduction path is searched and hence another terminal model is selected and so on. After all possible paths are investigated and all terminal models are determined, encompassing tests are carried out for each union of terminal models to find an undominated encompassing contender. The union of surviving terminal models which is referred to as the smaller GUM is employed for a new multiple-path reduction search. The search process continues until a unique model, called the specific model, emerges. In the sixth and final stage, the significance of every variable in the final model is evaluated in two overlapping sub-samples for reliability of the specific model.[24]

The second approach is Bayesian model averaging (BMA hereafter) which was developed by, *inter alia*, Madigan and Raftery (1994), Hoeting (1994), Chatfield (1995), Draper (1995), Raftery et al. (1997).[25] The basic idea of BMA is to incorporate the model uncertainty into statistical inference such that the true model is considered as an unobservable random variable. In this regard, BMA takes into account model uncertainty over a variable of interest making inferences based on the weighted averages of all possible models. Therefore, differently from the GETS approach, the main aim of BMA is to provide a better parameter estimate of the variable of interest rather than to find the best model. Fernández et al. (2001), Brock and Durlauf (2001), Sala-i-Martin et al. (2004), Masanjala and Papageorgiou (2005, 2007), Durlauf et al. (2006, 2007), Ulașan (2008), are examples of the application of BMA in the cross-country growth context.[26]

Obviously, both approaches are valuable statistical techniques for tackling model uncertainty and have their own advantages and disadvantages.[27] There

[24] While applying the *PcGets* algorithm, one can set any selection criteria for the significance levels, from strong to weak. The program also provides two basic strategies for these, namely liberal and conservative strategies. Both strategies are based on the critical values depending on sample size and for large samples on the number of possible explanatory variables. If there are many potentially irrelevant variables and few relevant variables, the conservative strategy is suggested. Conversely, for few irrelevant and many relevant variables, liberal strategy is better (Granger and Hendry (2005)).

[25] The basic paradigm for BMA was presented by Leamer (1978). See Hoeting et al. (1999) for the historical development of BMA.

[26] The approach in Sala-i-Martin (1997a,b) is close in spirit to that of BMA.

[27] There is a vast statistical literature debating classical versus Bayesian approaches on model uncertainty and model selection problem. See, for instance Chatfield (1995),

is no doubt that the GETS approach is particularly useful if one needs a specific model for some purpose, e.g. forecasting.[28] On the other hand, one advantage of BMA is that it provides a better framework for policy evaluation as discussed in the next section.

4 Model Uncertainty and Policy Evaluation in Cross-Country Growth Regression

Undoubtedly, the most important aim of cross-country growth studies is to explain growth differences across countries and to suggest policy implications which may be effective in promoting growth. Brock and Durlauf (2001, p.230) argue that "[I]n emprical macroeconomics, efforts to explain cross-country differences in growth behavior since World War II become a predominant area of research. The implications of this work for policymakers are immense... [I]n turn, the academic community has used this new empirical work as the basis for strong policy recommendations." However, as indicated by Brock and Durlauf (2001), Brock et al. (2003), Easterly (2005) and Rodrik(2005) this literature largely fails with respect to the perspective of policy evaluation. While Rodrik (2005) points out the endogeneity problem between the policy variable and economic growth, Easterly (2005) argues that the strong effects of policies obtained from cross-country growth regressions are mainly a result of extreme observations. Brock and Durlauf (2001) and Brock et al. (2003) emphasise the difficulty of macroeconomic policy evaluation in the presence of model uncertainty.

According to Brock and Durlauf (2001) and Brock et al. (2003), policy analysis can be carried out on the basis of two factors, namely the policy maker's preferences and a conditional distribution of the outcome of interest given the policy and available information. The authors argue that standard practice in the cross-country growth literature is uninfor-

Hoover and Perez (1999), Pötscher (1991), Granger and Hendry (2005), Hansen (2005). It is obvious that the solution of the matter is beyond the scope of this paper. Yet, we just remind that classical econometric model selection methods such as the GETS approach suffer from four conceptual errors namely parametric vision, the assumption of true data generating process, evaluation based on fit, ignoring the effect of model uncertainty on subsequent statistical inference as noted by Chatfield (1995) and Hansen (2005). Although BMA directly takes into account the impact of model uncertainty on inference, specifying appropriate priors over different models is challenging.

[28] Another advantage of the GETS approach is that it is labour saving as noted by Hendry and Krolzig (2005). For instance, according to Hendry and Krolzig (2004), implementation of GETS approach to Sala-i-Martin (1997a,b)'s data set by *PcGets* takes approximately two hours, including stacking the data.

mative from the perspective of policy evaluation since it fails to appropriately define the policy maker's preferences and ignores model uncertainty. Hence, Brock and Durlauf (2001) and Brock et al. (2003) propose that cross-country growth work for policy recommendations requires an explicit decision-theoretic formulation. Using the findings of modern statistical decision theory[29], these authors integrate model uncertainty into policy analysis.[30] In this section, we briefly summarise the implications of model uncertainty for policy evaluation in the context of cross-country growth regressions.

Recall the generic representation of cross-country growth regression expressed in equation (26). The key question in the context of policy evaluation is how a policy maker can use the cross-country growth regressions in order to formulate policy recommendations for enhancing the growth in country i.[31] Suppose that variable p in equation (26) represents a policy variable of interest which can be controlled by the policy maker. The standard answer to this question in the growth literature is to make policy suggestions according to the hypothesis tests for the coefficient corresponding with the policy variable of interest. More precisely, a policy maker recommends a change in the magnitude of the policy variable p for stimulating growth in country i according to the statistical significance of δ, typically assessed at 5 percent level, using a single model and a given data set. Obviously, this policy evaluation is conditional on the model employed by policy maker as well as data set.

The first problem with this kind of policy analysis in the context of cross-country growth regressions is that it neglects theory, specification and heterogeneity uncertainties. Secondly, even if model uncertainty can be eliminated, policy analysis based on statistical significance is problematic from the perspective of policy maker's preferences. In order to explain these problems more clearly, following the notation of Brock and Durlauf (2001) we define

[29] Wald (1950), Brainard (1967), Chamberlain (2000), Sims (1980), Berger (1985), Manski (2000), Heckman (2001), Sims (2002) are few examples.

[30] Although Brock and Durlauf (2001) and Brock et al. (2003) focus mainly on the cross-country empirical growth work, the framework developed by these authors are explicitly subject to other macroeconomic empirical analysis in formulating policy recommendations in the presence of model uncertainty. For a more general context concerning the issue see Brock and Durlauf (2006) and Brock et al. (2007). In terms of policy analysis, a related direction of the literature is carried by Hansen and Sargent (2001) that emphasise the robust control theory to analyse macroeconomic policy under the model uncertainty.

[31] As noted by Eris (2005), the term "policy maker" is used in a broader sense in the manner that he or she may be an economist suggesting a government to implement a particular policy, say openness to international trade, using some cross-country growth regression.

the policy maker's preferences in terms of utility (or objective) function as

$$V(\varrho_i, O_i) \tag{30}$$

where ϱ_i is growth rate of per capita GDP in country i, as previously defined, and O_i indicates the set of characteristics in country i affecting policy maker's utility. In the context of policy maker's utility function, implementing or suggesting a policy change which is effective for enhancing growth depends on comparisons of policy maker's utilities in alternative settings. More clearly, if the policy maker believes that a particular policy variable has some effect in increasing growth, then he faces two options: either implementing or not implementing a policy change. Therefore, denoting the level of policy variable by p, policy maker's decision set will be $A = \{0, dp_i\}$, where dp_i represents the policy change and for simplicity it is assumed to be positive. The objective of empirical work is to develop a decision rule which is conditional on observable data D. Since the cross-country growth regression in equation (26) is linear, the effect of a marginal change in p is δ. Therefore, the growth rate in country i will be $\varrho_i + \delta dp_i$ in the case of a policy change while it is ϱ_i in the absence of policy implementation. Policy evaluation requires comparison of expected utilities of policy maker with and without policy change

$$E(V(\varrho_i + \delta dp_i, O_i)|D) - E(V(\varrho_i, O_i)|D) \tag{31}$$

where E represents the expected value operator. The conventional wisdom in the empirical cross-country growth literature is to compute this comparison selecting one model as if it is true model and applying a statistical significance test. A statistically insignificant coefficient is taken to mean that a particular policy is not important for economic growth while the statistical significance is used as strong evidence that the policy is important for economic growth. This kind of decision rule is implicitly assumed that the policy maker's utility function is defined by

$$E(V(\varrho_i + \delta dp_i, O_i)|D) - E(V(\varrho_i, O_i)|D) = [\widehat{\delta}(dp_i) - 2\hat{\sigma}_\delta(dp_i)] \geq 0 \tag{32}$$

where $\widehat{\delta}$ and $\hat{\sigma}_\delta$ denote the OLS coefficient estimate of policy variable p and its corresponding standard error, respectively. Obviously both are conditional on a particular model. Then one would suggest policy implementation in the form of dp_i if the t-statistic in OLS regression is equal or greater than 2 (2 is selected according to typical assessment of statistical significance level at 5 percent).

Policy analysis based on significance level is, however, troublesome in many ways even if the model used in OLS regression is true as argued by

Brock and Durlauf (2001), Durlauf (2002) and Brock, Durlauf and West (2003). We emphasise two important problems: First, the policy maker evaluates a particular policy only using the mean and variance of the policy variable. However, the whole probability distribution of δ might be important for policy analysis. For instance, a policy maker may be more sensitive to negative growth rates than positive ones or the effect of growth on poverty can be asymmetric and a typical policy maker tries to act in socially acceptable way. Second, even if the policy maker takes into account only the mean and variance of policy variable of interest, policy analysis based on statistical significance considers the effect of policy change on the component of growth rather than the effect of the policy change on growth per se. In other words, a statistically significant coefficient of estimate shows the marginal effect of the policy variable on growth and does not provide a clear answer whether policy change should be implemented.

The message of these criticisms is that one should define more appropriate utility functions and assess a policy change under alternative policy scenarios.[32] Obviously, this policy evaluation will be based on a particular model only if policy maker is certain that the model at hand is true. Yet, since he is not certain about the true model, this adds another uncertainty to the uncertainty over parameter δ. In the case of model uncertainty, the policy maker will not want to evaluate a policy change according to a particular model. Instead, he or she will want to make expected utility comparison expressed in equation (31), conditioning on data. This means that comparison of expected utilities for a given policy should be based on the assumption that the true model is not known. Since calculation of expected utility information expressed in equation (31) contains all information for policy evaluation, in the absence of information about the true model, this expression explicitly requires accounting for model uncertainty since expected utilities are condi-

[32] Brock and Durlauf (2001) and Brock et al. (2003) explore policy implications of cross-country growth analysis employing some alternative utility functions such as risk neutrality, ambiguity aversion and so on. According to these authors, the utility functions that they examine are not particularly compelling, but they are useful to illustrate in order to interpret growth regressions for policy analysis in the presence of model uncertainty. For instance, these authors indicate that EBA employed by Levine and Renelt (1992) corresponds to an extreme risk aversion utility for policy maker. More compactly, according to EBA a policy change is implemented only if

$$E(V(\varrho_i + \delta dp_i, O_i)|D) - E(V(\varrho_i, O_i)|D) > 0$$

for every model in the model space. See Eris (2005) for a nice treatise showing what kind of decision rules arise under the considerations of different assumptions for the policy maker's utility functions and policy robustness preference parameters accounting model uncertainty.

tional on only data not on possible models. Therefore, this requires us to modify the expected utility comparison equation as

$$E(V(\varrho_i + \delta dp_i, O_i)|D) - E(V(\varrho_i, O_i)|D) = \sum_k P(M_k|D)E(V(\varrho_i + \delta dp_i, O_i)|D, M_k) - \sum_k P(M_k|D)E(V(\varrho_i, O_i)|D, M_k) \tag{33}$$

where $P(M_k|D)$ is the probability that model M_k is the true causal relationship between the growth rate and explanatory variables for given data, D. Therefore, the last equation explicitly accounts model uncertainty and as mentioned before, the aim of any policy relevant empirical work is to compute these expected utilities.

As can be seen, equation (33) illustrates that the expected utility comparison depends on the weighted averages of the coefficient of the policy variable, and expected utility calculations are independent of a particular model. Rather, the true model as an unobservable random variable is integrated to this calculation. Hence, the second important message is that identifying a particular model(s) according to some model selection criteria does not have any intrinsic value from the perspective of policy evaluation in the presence of model uncertainty. In contrast, the standard practice in the literature evaluates a policy change according to a particular model and sometimes compares the coefficient estimates with those obtained from modified specifications of that model in order to provide robustness of data analysis. This kind of policy analysis not only does ignore model uncertainty but also does not provide a clear information for policy evaluation. For instance, if the estimated coefficient of a policy variable is large in one regression while small in another, drawing a conclusion concerning the policy variable of interest is unclear. However, the calculation in equation (33) clearly removes this kind of concerns since each possible model is integrated into the calculation. This methodology, known as “model averaging” in the statistics literature, is a coherent way not only in order to handle model uncertainty but also for policy evaluation.

5 Conclusion

In this paper, we reviewed the recent cross-country growth literature aiming to explain growth differences across countries using regression analysis and other statistical methods. Even though this literature was mainly inspired by endogenous growth theories, the neoclassical growth model, especially its augmented version by Mankiw et al. (1992) is still the workhorse for cross-country growth empirics. For instance Mankiw (1995) argues that "[I]f the goal is to explain why standard of living is higher today than a century ago, then neoclassical model is not very illuminating. . . [A] more challenging goal is to explain the *variation* in economic growth that we observe in different countries in different times (p.280). . . [E]ndogenous growth models provide a plausible description of worldwide advances in knowledge. The neoclassical growth model takes world wide technological advances as given and provides a plausible description of international differences (p.308)."

The most outstanding feature of the recent empirical cross-country growth literature is that a large number of factors have been suggested as fundamental growth determinants. Together with the small sample property, this leads to an important problem, model uncertainty: Which factors are more fundamental in explaining growth dynamics and hence growth differences are still the subject of academic research. Recent attempts based on general-to-specific modeling or model averaging are promising but have their own limits.

Closely related to model uncertainty, and indeed the ultimate goal of the literature is policy evaluation. In spite of the fact that model uncertainty has been recognised since the important work by Levine and Renelt (1992), it is very surprising that cross-country growth studies have been used for policy analysis without paying attention to model uncertainty. It is obvious that any policy recommendation derived from a particular cross-country growth regression is troublesome since in the presence of model uncertainty it is conditional on the selected model.

Although we emphasise model uncertainty in this overview, other econometric problems, especially parameter heterogeneity and outliers are equally important in this literature. Due to these problems, cross-country growth empirics can be considered as a mix of economic theory and statistics and it might be more reasonable to refer to it as "growth econometrics" as Durlauf et al. (2005) point out.

In conclusion, given the challenging econometric problems, the results of cross-country growth studies have been controversial in terms of robustness. The implications of this are threefold: First, it is more plausible to accept cross-country growth studies as a wider picture of growth process. This means that combining findings of this literature with detailed case studies is a worthwhile task. Second, it may be more useful to shift research agenda towards more practical or pragmatic issues rather than the international growth differences as suggested by Pritchett (2000). Third, introducing new statistical tools and better proxy variables will make cross-country growth studies more informative.

References

Acemoglu, D. (2009). *Introduction to Modern Economic Growth*, Princeton University Press, Princeton and Oxford.

Acemoglu, D., Johnson, S. and Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation, *American Economic Review* **91**(5): 1369–1401.

Aghion, P. and Howitt, P. (1999). *Endogenous Growth Theory*, MIT Press, Cambridge, USA and London, UK.

Barro, R. J. (1991). Economic growth in a cross section of countries, *Quarterly Journal of Economics* **106**(2): 407–443.

Barro, R. J. (1997). *Determinants of Economic Growth: A Cross-Country Empirical Study*, MIT Press, Cambridge.

Barro, R. J. and Lee, J.-W. (1994a). Data set for a panel of 138 countries, *Technical report*. Downloadable at http://www.nber.org/pub/barro.lee/.

Barro, R. J. and Lee, J.-W. (1994b). Sources of Economic Growth, *Carnegie-Rochester Conference Series on Public Policy* **40**(3): 1–46.

Barro, R. J. and Lee, J.-W. (2000). International data on educational attainment: Updates and implications, *Working Paper 42*, Center for International Development at Harvard University. Downloadable at http://www.cid.harvard.edu/ciddata/ciddata.html.

Barro, R. J. and Sala-i-Martin, X. (1992). Convergence, *Journal of Political Economy* **100**(2): 223–251.

Barro, R. J. and Sala-i-Martin, X. (2004). *Economic Growth*, MIT Press, Cambridge, USA and London, UK.

Benhabib, J. and Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data, *Journal of Monetary Economics* **34**(2): 143–173.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian analysis*, 2nd edn, Springer-Verlag, New York.

Bernard, A. B. and Jones, C. I. (1996). Technology and convergence, *Economic Journal* **106**(437): 1037–1044.

Bils, M. and Klenow, P. J. (2000). Does schooling cause growth?, *American Economic Review* **90**(5): 1160–1183.

Bleaney, M. and Nishiyama, A. (2002). Explaining growth: A contest between models, *Journal of Economic Growth* **7**(1): 43–56.

Brainard, W. C. (1967). Uncertainty and the effectiveness of policy, *American Economic Review* (2): 411–425.

Brock, W. A. and Durlauf, S. N. (2001). Growth empirics and reality, *World Bank Economic Review* **15**(2): 229–272.

Brock, W. A. and Durlauf, S. N. (2006). Macroeconomics and model uncertainty, *in* D. Colander (ed.), *Post-Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium*, Cambridge University Press, New York, pp. 116–134.

Brock, W. A., Durlauf, S. N. and West, K. D. (2003). Policy evaluation in uncertain economic environments (with discussion), *Brookings Papers on Economic Activity* (1): 235–321.

Brock, W. A., Durlauf, S. N. and West, K. D. (2007). Model uncertainty and policy evaluation: Some theory and empirics, *Journal of Econometrics* **136**(2): 629–664.

Caselli, F., Esquivel, G. and Lefort, F. (1996). Reopening the convergence debate: A new look at cross-country growth empirics, *Journal of Economic Growth* **1**(3): 363–389.

Chamberlain, G. (2000). Econometrics and decision theory, *Journal of Econometrics* **95**(2): 255–283.

Chatfield, C. (1995). Model uncertainty, data mining and statistical inference, *Journal of the Royal Statistical Society. Series A (Statistics in Society)* **158**(3): 419–466.

Cohen, D. and Soto, M. (2007). Growth and human capital: Good data, good results, *Journal of Economic Growth* **12**(1): 51–76.

Draper, D. (1995). Assessment and propagation of model uncertainty, *Journal of the Royal Statistical Society. Series B (Methodological)* **57**(1): 45–97.

Durlauf, S. N. (2002). Policy evaluation and empirical growth resarch, *in* N. Loayza and R. Soto (eds), *Economic Growth: Sources, Trends, and Cycles*, Central Bank of Chile, Santiago, Chile, pp. 163–190.

Durlauf, S. N. and Johnson, P. A. (1995). Multiple regimes and cross-country growth behaviour, *Journal of Applied Econometrics* **10**(4): 365–384.

Durlauf, S. N. and Johnson, P. A. (2008). Convergence, *in* S. N. Durlauf and L. E. Blume (eds), *The New Palgrave Dictionary of Economics*, Palgrave Macmillan, Basingstoke.

Durlauf, S. N., Johnson, P. A. and Temple, J. R. (2005). Growth econometrics, *in* P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, Vol. 1A, North-Holland, Amsterdam, chapter 8, pp. 555–677.

Durlauf, S. N., Kourtellos, A. and Tan, C. M. (2006). Is god in the details? a reexamination of the role of religion in economic growth, *Working paper*, Department of Economics, University of Wisconsin.

Durlauf, S. N., Kourtellos, A. and Tan, C. M. (2007). Are any growth theories robust?, *Working paper*, Department of Economics, Tufts University.

Easterly, W. (2005). National polices and economic growth: A reappraisal, *in* P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, Vol. 1A, North-Holland, Amsterdam, chapter 15, pp. 1015–1059.

Easterly, W. and Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions, *Quarterly Journal of Economics* **112**(4): 1203–1250.

Eris, M. N. (2005). *Essays on Economic Growth and Cross-Country Growth Empirics*, PhD thesis, University of Wisconsin at Madison.

Fernández, C., Ley, E. and Steel, M. (2001). Model Uncertainty in Cross-Country Growth Regressions, *Journal of Applied Econometrics* **16**(5): 563–576.

Frankel, J. A. and Romer, D. (1999). Does trade cause growth?, *American Economic Review* **89**(3): 379–399.

Galor, O. (1996). Convergence? inferences from theoretical models, *Economic Journal* **437**(106): 1056–1069.

Granger, C. W. and Hendry, D. F. (2005). A dialogue concerning a new instrument for econometric modelling, *Econometric Theory* **21**(1): 278–297.

Granger, C. W. J. and Uhlig, H. F. (1990). Reasonable extreme-bounds analysis, *Journal of Econometrics* **44**(1-2): 159–170.

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others?, *Quarterly Journal of Economics* **114**(1): 83–116.

Hansen, B. E. (2005). Challenges for econometric model selection, *Econometric Theory* **21**(1): 60–68.

Hansen, L. P. and Sargent, T. J. (2001). Robust control and model uncertainty. Available at http://homepages.nyu.edu/ ts43/.

Hanushek, E. A. and Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations, *American Economic Review* **90**(5): 1184–1208.

Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture, *Journal of Political Economy* **109**(4): 673–748.

Hendry, D. F. (1995). *Dynamic Econometrics*, Oxford University Press, Oxford.

Hendry, D. F. and Krolzig, H.-M. (2004). We ran one regression, *Oxford Bulletin of Economics and Statistics* **66**(5): 799–810.

Hendry, D. F. and Krolzig, H.-M. (2005). The properties of automatic gets modelling, *Economic Journal* **115**(502): C32–C61.

Hendry, D. F. and Mizon, G. E. (1990). Procrustean econometrics: Or stretching and squeezing data, *in* C. W. J. Granger (ed.), *Modelling Economic Series*, Clarendon Press, Oxford, pp. 121–136.

Hoeting, J. A. (1994). *Accounting for Model Uncertainty in Linear Regression*, PhD thesis, University of Washington.

Hoeting, J., Madigan, D., Raftery, A. and Volinsky, C. (1999). Bayesian model averaging: A tutorial, *Statistical Science* **14**(4): 382–417.

Hoover, K. D. and Perez, S. J. (1999). Data mining reconsidered: Encompassing and the general-to-specific approach to specification search, *Econometrics Journal* **2**(2): 167–191.

Hoover, K. D. and Perez, S. J. (2004). Truth and robustness in cross-country growth regressions, *Oxford Bulletin of Economics and Statistics* **66**(5): 765–798.

Islam, N. (1995). Growth empirics: A panel data approach, *Quarterly Journal of Economics* **110**(4): 1127–1170.

Islam, N. (2003). What have we learnt from the convergence debate?, *Journal of Economic Surveys* **17**(3): 309–362.

Jones, G. and Schneider, W. J. (2006). Intelligence, Human Capital, and Economic Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach, *Journal of Economic Growth* **11**(1): 71–93.

Klenow, P. J. and Rodríguez-Clare, A. (1997). The neoclassical revival in growth economics: Has it gone too far?, *in* B. S. Bernanke and J. Rotemberg (eds), *NBER Macroeconomics Annual 1997*, MIT Press, pp. 73–103.

Kormendi, R. C. and Meguire, P. G. (1985). Macroeconomic determinants of growth: Cross-country evidence, *Journal of Monetary Economics* **16**: 141–163.

Krolzig, H.-M. and Hendry, D. F. (2001). Computer automation of general-to-specific model selection procedures, *Journal of Economic Dynamics and Control* **25**(6-7): 831–866.

Krueger, A. B. and Lindahl, M. (2001). Education for growth: Why and for whom?, *Journal of Economic Literature* **39**(4): 1101–1136.

Leamer, E. E. (1978). *Specifaction Searches: Ad Hoc Inference with Nonexperimental Data*, Wiley, New York.

Leamer, E. E. (1983). Let's take the con out of econometrics, *American Economic Review* **73**(1): 31–43.

Leamer, E. E. and Leonard, H. (1983). Reporting the fragility of regression estimates, *Review of Economics and Statistics* **65**(2): 306–317.

Lee, K., Pesaran, M. H. and Smith, R. (1997). Growth and convergence in a multi-country empirical stochastic solow model, *Journal of Applied Econometrics* **12**(4): 357–392.

Levine, R. and Renelt, D. (1992). A sensivity analysis of cross-country growth regressions, *American Economic Review* **82**(4): 942–963.

Madigan, D. and Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window, *Journal of the American Statistical Association* **89**(428): 1535–1546.

Mankiw, G. N. (1995). The growth of nations, *Brookings Papers on Economic Activity* **1995**(1): 275–310.

Mankiw, G. N., Romer, D. and Weil, D. N. (1992). A contribution to the empirics of economic growth, *Quarterly Journal of Economics* **107**(2): 407–437.

Manski, C. F. (2000). Identification problems and decisions under ambiguity: Empirical analysis of treatment response and normative analysis of treatment choice, *Journal of Econometrics* **95**(2): 415–442.

Masanjala, W. and Papageorgiou, C. (2005). Rough and lonely road to prosperity: A reexamination of the sources of growth in Africa using Bayesian Model Averaging, *Workin paper*, Luisiana State University. Luisiana State University.

Masanjala, W. and Papageorgiou, C. (2007). Initial conditions and post-war growth in sub-Saharan Africa, *Technical report*, Luisiana State University.

McAleer, M. (1994). Sherlock holmes and the search for truth: A diagnostic tale, *Journal of Economic Surveys* **8**(4): 317–370.

McAleer, M., Pagan, A. R. and Volker, P. A. (1985). What will take the con out of econometrics?, *American Economic Review* **75**(3): 293–307.

Phelps, E. S. (1995). The growth of nations: Comment, *Brookings Papers on Economic Activity* **1995**(1): 311–313. discussion of Mankiw paper with the same title.

Pritchett, L. (2000). Understanding patterns of economic growth: Searching for hills among plateaus, mountains, and plains, *The World Bank Economic Review* **14**(2): 221–250.

Pritchett, L. (2001). Where has all the education gone?, *The World Bank Economic Review* **15**(3): 367–391.

Pötscher, B. M. (1991). Effects of model selection on inference, *Econometric Theory* **7**(2): 163–185.

Quah, D. T. (1996). Growth and convergence in models of distribution dynamics, *Economic Journal* **106**(437): 1045–1055.

Raftery, A. E., Madigan, D. and Hoeting, J. A. (1997). Bayesian model averaging for linear regression models, *Journal of the American Statistical Association* **92**(437): 179–191.

Rodrìguez, F. (2007). Cleaning up the kitchen sink: Growth empirics when the world is not simple. Available at the author‘s web page.

Rodrik, D. (2005). Why we learn nothing from regressing economic growth on policies. Available at the author’s web page .

Sachs, J. D. and Warner, A. M. (1995). Economic reform and the process of global integration, *Brookings Papers on Economic Activity* **1995**(1): 1–118.

Sachs, J. D. and Warner, A. M. (1997). Sources of slow growth in african economies, *Journal of African Economies* **6**(3): 335–376.

Sala-i-Martin, X., Doppelhofer, G. and Miller, R. I. (2004). Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach, *American Economic Review* **94**(4): 813–835.

Sala-i-Martin, X. X. (1996). The classical approach to convergence analysis, *Economic Journal* **106**(437): 1019–1036.

Sala-i-Martin, X. X. (1997a). I just run two million regressions, *American Economic Review* **87**(2): 178–183.

Sala-i-Martin, X. X. (1997b). I just run four million regressions, *Working paper 6252*, National Bureau of Economic Research.

Sala-i-Martin, X. X. (2001). Growth empirics and reality: Comment, *World Bank Economic Review* **15**(2): 229–272. discussion of Brock and Durlauff paper with the same title.

Sims, C. A. (1980). Macroeconomics and reality, *Econometrica* **48**(1): 1–48.

Sims, C. A. (2002). The role of models and probabilities in the monetary policy process, *Brookings Papers on Economic Activity* **2002**(2): 1–40.

Solow, R. M. (1956). A contribution to the theory of economic growth, *Quarterly Journal of Economics* **70**(1): 65–94.

Solow, R. M. (2007). The last 50 years in growth theory and the next 10, *Oxford Review of Economic Policy* **23**(1): 3–14.

Summers, R. and Heston, A. (1988). A new data set of international comparisons of real product and price levels estimates for 130 countries, 1950-1985, *Review of Income and Wealth* **34**(1): 1–25.

Summers, R. and Heston, A. (1991). The Penn World Table (Mark 5): An expanded set of international comparisons, 1950-1988, *Quarterly Journal of Economics* **106**(2): 327–368.

Swan, T. W. (1956). Economic growth and capital accumulation, *Economic Record* **32**: 334–361.

Temple, J. (1999a). The new growth evidence, *Journal of Economic Literature* **37**(1): 112–156.

Temple, J. (1999b). A positive effect of human capital on growth, *Economics Letters* **65**(1): 131–134.

Temple, J. (2000). Growth regressons and what the textbooks don't tell you, *Bulletin of Economic Research* **52**(3): 181–205.

Ulaşan, B. (2008). *Trade Openness and Economic Growth: A Cross-Country Empirical Investigation*, PhD thesis, University of Nottingham.

Wacziarg, R. (2002). Review of Easterly's "The Elusive Quest for Growth", *Journal of Economic Literature* **40**(3).

Wald, A. (1950). *Statistical Decision Functions*, Wiley, New York.