CENTRAL BANK OF THE REPUBLIC OF TURKEY

WORKING PAPER NO: 11/01

Augmented Neoclassical Growth Model: A Replication over the 1960-2000 Period

January 2011

Bülent ULAŞAN

Address:
Central Bank of the Republic of Turkey
Head Office
Research and Monetary Policy Department
İstiklal Caddesi No: 10
Ulus, 06100 Ankara, Turkey

Phone:
+90 312 507 54 02

Facsimile:
+90 312 507 57 33

Augmented Neoclassical Growth Model: A Replication over the 1960-2000 Period

Bülent Ulașan†

Abstract

This paper empirically revisits the augmented neoclassical growth model suggested by Mankiw, Romer and Weil (1992, MRW) to answer whether this model is still an appropriate benchmark specification for investigating the relationship between long run economic growth and any particular growth theory for the sample period 1960-2000. For this I replicate MRW using updated and revised data for several samples, and compare my findings with those obtained by MRW. My findings are consistent with the theory and support the results of MRW. I obtain more reasonable coefficient estimates on capital shares. I also check for geographical differences and for outlying countries. I conclude that the inclusion of continental dummy variables does not change the basic results and any outlier effects are small.

J.E.L. Classification : O40, O47.

Key words: Economic Growth, Cross-Country Growth Regression, Conditional Convergence, Outliers.

†Corresponding author: Economist at the Central Bank of the Republic of Turkey. e-mail: Bulent.Ulasan@tcmb.gov.tr. The views expressed in this paper and all remaining errors are solely the author's. The author is thankful to Wyn Morgan, Peter Wright and Mehmet N. Eris for their valuable comments.

1 Introduction

This paper empirically revisits the augmented neoclassical growth model suggested by Mankiw, Romer and Weil (1992, MRW henceforth). I attempt to answer whether this model is still an appropriate benchmark specification in order to investigate the relationship between long run economic growth and any particular growth theory for the 1960-2000 period. For this I replicate MRW using updated and revised data for several samples, and compare my findings with those obtained by MRW.

The framework provided by MRW is the workhorse in the empirical cross-country growth literature and most of the studies in this literature are based on MRW.[1] However, my replication is different from these studies in some aspects: First, most of the studies in the literature estimate MRW over the 1960-1990 or the 1970-1990 period. In this paper, I estimate MRW over the period of 40 years, which should be sufficiently long to reflect long run growth dynamics; second, many previous studies employ a proxy variable for the initial level of human capital stock rather than the saving rate for human capital. In this paper, I strictly follow MRW by employing the secondary school enrolment rate.

My findings are consistent with the theory and support the results of MRW. Both my cross-country regression results and those obtained by MRW show that the investment rates of both physical and human capital significantly contribute to the growth rates of countries, while the rate of population growth has a negative effect on growth in the long run. In addition, I obtain more reasonable coefficient estimates on capital shares. I also check for geographical differences and for outlying countries. I conclude that the inclusion of continental dummy variables does not change the basic results and any outlier effects are small.

This paper is organised as follows: The next section, Section 2, firstly describes the basic framework of MRW and data and then reports and discusses basic findings. Section 3 seeks geographical differences. Section 4 deals with the outlier problem. Section 5 concludes.

[1] Islam (1995) and others further adapt the framework of Mankiw, Romer and Weil (1992) for panel estimation.

2 Replication of Augmented Neoclassical Growth Model by MRW

In this section, I estimate the augmented neoclassical growth model developed by MRW by using updated data over the period 1960-2000 and compare my findings with those obtained by MRW. In other words, I set up this model for the period 1960-2000 since I specifically aim to answer whether augmented neoclassical growth model is an appropriate benchmark model in order to investigate the relationship between long run economic growth and any particular growth theory. This means that the explanatory variable in my cross-country growth regressions is growth rate of output per worker between 1960 and 2000. My data set typically covers 107 countries.[2] The sample of countries is listed in the Appendix.

2.1 Description of Benchmark Model

Mankiw, Romer and Weil (1992) augments the Solow-Swan version of neoclassical growth model by adding the accumulation of human capital. Considering a three factor Cobb-Douglas production function and assuming that labour stock and the level of technology grow exogenously at rates n and g, respectively, these authors produce the following equation for cross-country growth regression[3]

$$logy_i(t) - logy_i(0) = \gamma_0 + \gamma_1 logy_i(0) + \gamma_2 log(n_i + g + \delta) + \gamma_3 logs_{i,K} + \gamma_4 logs_{i,H} + \varepsilon_i \tag{2.1}$$

[2]In this point I follow the standard approach in the literature. More clearly I randomly select countries according to the criterion of data availability. At the first sight this approach seems reasonable. However, missing data, especially if the data of some particular countries are systematically missing (such as very poor countries or countries in transition) is a serious problem as noted by Durlauf et al. (2005).

[3]According to three factor Cobb-Douglas production function, production at time t in country i is given by

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

where the notation here is 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 labour-augmenting technology. MRW also assumes that $\alpha + \beta < 1$, which means that there are decreasing returns to both kinds of capital. Since the formal derivation of the cross-country growth regression depicted in equation (2.1) is well-established in the cross-country growth literature, I do not need to elaborate the derivation in this paper. The reader can apply MRW, Barro and Sala-i-Martin (2004) and Durlauf et al. (2005), amongst others.

In equation (2.1), y_i and $(n_i+g+\delta)$ denote the level of GDP per worker and the sum of rates of population growth, technological progress and depreciation in country i, respectively. Similarly, the terms $s_{i,K}$ and $s_{i,H}$ represent the rates of accumulation of both physical and human capital for country i. As seen in equation (2.1), the augmented neoclassical growth model basically involves regressing growth rates on the log of initial income and a set of long-run equilibrium or steady-state level of income determinants. Put differently, growth rates of output per worker can vary across countries either because of differences in the variables determining their steady-state levels namely saving rates for physical and human capital, and rate of population growth or because of differences in the initial level of output per worker, $logy_i(0)$.

Following MRW, I assume that the sum of rates of depreciation and technological progress is constant and equal to 0.05 across countries and estimate equation (2.1) over the period 1960-2000. For this purpose, I measured $s_{i,K}$ by the ratio of real investment to real GDP and $s_{i,H}$ by the secondary school gross enrolment rate. Using the school enrolment rate as a proxy for the saving rate of human capital is problematic and leads researchers to employ average years of schooling as more reliable variables for human capital. I, however, employ the secondary school enrolment rate in order to follow the theoretical framework more strictly as years of schooling are a stock rather than a flow variable for human capital. In addition, school enrolment rates are available for a larger sample of countries. Data are compiled from standard sources: GDP per capita and investment rates are taken from the Penn World Tables Version 6.1 (Heston, Summers and Aten, 2002); population, labour force and gross secondary school enrolment rates come from the World Bank World Development Indicators (2002; 2006). Using labour force as the total population between ages 15 and 64, per capita GDP is converted to per worker GDP. All of these variables are averaged over the period 1960-2000 except the initial level of income. The variables and their sources are detailed in the Appendix.

In summary, my baseline cross-country growth specification for each country i as follows

$$\begin{aligned} logy_{i,2000} - logy_{i,1960} = {} & \gamma_0 + \gamma_1 logy_{i,1960} + \gamma_2 log(n_i + g + \delta) \\ & + \gamma_3 log(Investment\ rate_i) \\ & + \gamma_4 log(School\ enrolment_i) + \varepsilon_i \end{aligned} \tag{2.2}$$

2.2 Results

Before evaluating the regression results, I want to emphasise two points about the regressions. First, in each regression I check the normality assumption applying median and inter quartile range comparison suggested by Hamilton (1992) which is originally based on Hoaglin, Iglewicz and Tukey (1986) on regression residuals and conclude that residuals are normally distributed. Therefore, we may assume that actual errors are normally distributed (at least approximately).

Second, in each regression I also check the constant error variance assumption by employing the Breusch-Pagan test for heteroscedasticity. The common practice in cross-country growth literature for dealing heteroscedasticity is reporting regression results with the heteroscedasticity consistent (White-robust) standard errors since they work well regardless of heteroscedasticity in the actual errors. However, these standard errors are consistent but not unbiased. More clearly they are justified only asymptotically. In small samples, heteroscedasticity consistent standard errors may have distributions that are not close to those of usual standard errors which means that they may be larger or smaller than the usual ones. As pointed out by Wooldridge (2003) heteroscedasticity consistent standard errors are generally found to be larger than the usual standard errors. This can affect the subsequent statistical inference such that one can conclude that a variable is statistically insignificant according to t-test based on the heteroscedasticity consistent standard errors even if that variable is significant (at least marginally) in the case of usual t-test. Therefore, there is no reason to use heteroscedasticity consistent standard errors as long as the homoscedastic error variance assumption holds and the errors are normally distributed. Hence, I carry out regression analysis employing t-statistics based on the usual standard errors unless I reject the homoscedasticity assumption. I report t-statistics based on the heteroscedasticity consistent standard errors only for the regressions in which the assumption of homoscedastic error variance is rejected.

Table 1 presents the OLS estimates of equation (2.2). In column 1, the model is estimated for a sample of 107 countries whose data are available over the 1960-2000 period. All variables have the expected signs and are found to be strongly significant. In the literature, some studies exclude oil producing countries from cross-country samples since a substantial part of GDP in these countries depends on the usage of their oil resources rather than value added. Column 2 displays regression results after five oil producing countries are excluded from the full sample.[4] The estimation results of the non-oil sample are only slightly different from those of the full sample and hence I can conclude that five oil producing countries in the full sample are not changing the basic results. Therefore, I prefer to employ the full sample for my cross-country empirical investigation.

[4]These countries are, Algeria, Indonesia, Nigeria, Oman and Venezuela.

Table 1: Augmented Neoclassical Growth Model: OLS Estimates

Dependent Variable: Log Difference real GDP per worker over the 1960-2000 period

	My Estimation Results: 1960-2000 period							Estimates of MRW† 1960-1985 Period		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sample	Full	NonOil	OECD[a]	NonOECD	High Income	Low Income	p-value[b]	NonOil	Intermediate	OECD
log GDP per worker 1960	-0.442	-0.423	-0.568	-0.448	-0.465	-0.393	0.68	-0.288	-0.366	-0.398
	(6.54)	(5.95)	(3.63)	(6.30)	(4.05)	(2.56)		(4.68)	(6.24)	(5.67)
$\log(n_i + g + \delta)$	-1.049	-0.906	-0.262	-1.023	-1.302	-0.735	0.73	-0.506	-0.545	-0.863
	(2.83)	(2.36)	(0.39)	(1.82)	(3.07)	(0.96)		(1.75)	(2.36)	(2.56)
log of Investment rate	0.411	0.410	0.588	0.390	0.581	0.389	0.55	0.524	0.538	0.332
	(4.47)	(4.43)	(1.63)	(3.07)	(3.05)	(3.24)		(6.03)	(3.98)	(1.91)
log of School enrolment	0.450	0.449	0.408	0.445	0.333	0.476	0.74	0.231	0.270	0.228
	(5.18)	(5.08)	(0.80)	(5.79)	(1.98)	(4.08)		(3.89)	(3.12)	(1.57)
Constant	2.782	3.014	6.492	2.834	2.522	3.225	0.70	1.874	2.498	4.155
	(2.56)	(2.72)	(2.31)	(1.76)	(1.63)	(1.32)		(2.22)	(3.15)	(4.16)
p-value for heteroscedasticity[c]	0.66	0.60	0.07	0.13	0.75	0.31		0.24	0.01	0.20
Number of observations	107	102	26	81	52	55		98	75	22
Adjusted R^2	0.61	0.61	0.60	0.56	0.61	0.58		0.46	0.43	0.65
Implied λ^*	0.0146	0.0137	0.0210	0.0149	0.0156	0.0125		0.0136	0.0182	0.0203

Note: t-statistics are in parenthesis. In the regressions where the heteroscedasticity test is failed to pass at 15 % level t-statistics based on heteroscedastic-consistent (White-robust) standard errors are reported. All variables, except initial level of income are averaged over the 1960-2000 period. The variable, $(n_i + g + \delta)$ refers to sum of rates of population growth, technical progress and depreciation.

† Estimates of augmented neoclassical growth model in table V, page 426 of Mankiw, Romer and Weil (1992). Dependent variables for these regressions are log difference GDP per worker over the period 1960-1985.

[a] Czech Republic, Germany, Poland and Slovak Republic are excluded due to missing data.

[b] The p-value refers to the hypothesis that individual coefficients are the same for the low and high income samples.

[c] Breusch-Pagan test for heteroscedasticity in which the null refers to the homoscedastic errors.

* The rate of convergence in the neighborhood of steady-sate

In columns 3 and 4, the full sample is divided according to the membership of the OECD. Estimation results for the 26 OECD countries and for remaining countries are given in columns 3 and 4, respectively. As can be seen, the results for the OECD sample are not very precise since, except for the initial level of income, all variables are found to be statistically insignificant. An important reason is that the sample size for this regression is small. Therefore, the regression result is very sensitive to including or excluding observations.[5] Another, and more important, reason is that the relatively high coefficient of initial income and high level of R^2 imply greater absolute convergence for the OECD countries.[6]

The results for the remaining non-OECD countries indicate that all variables are strongly significant with anticipated signs. In columns 5 and 6, I divide the full sample into low-income and high-income countries according to initial income level. For this, I calculated the median of the GDP per worker in 1960 and classified the countries with initial income above the sample median as high income, while those with initial income below the sample median as low income countries. As can be seen, almost half the sample consists of high income countries since the median is calculated for all countries. Another distinguishing feature of this separation is that a majority of countries located in East Asia and Pacific belong to the low income country sample since these countries start the sample period with a relatively low GDP per worker.[7] Estimation results for both groups are displayed in columns 5 and 6.

[5]While estimating the augmented neoclassical growth model for the OECD sample, I consider all OECD members except Germany, Poland, Czech Republic and Slovakia. These countries are omitted due to the missing data. However, one can prefer to select the OECD sample which consists of only larger countries as in the case of MRW, or to employ an OECD sample including only members since the foundation of OECD or use the OECD sample based on only high-income members which means that low-income members such as Turkey and Mexico excluded as in the case of Barro and Sala-i-Martin (2004). I also check these possibilities and conclude that all variables are insignificant except initial level of income in each case. However, in some cases (for instance when we estimate the model for 22 high-income OECD countries) I find that the coefficient of school enrolment rate is negative. On the other hand in each case the remaining non-OECD samples yield the regression results which are very close to those reported in Table 1

[6]Testing of absolute convergence hypothesis for the OECD sample yields the following cross-country growth regression (heteroscedastic-consistent t-statistics are in parentheses)

$$logy_{i,2000} - logy_{i,1960} = \underset{(3.94)}{5.247} - \underset{(3.21)}{0.451} logy_{i,1960} \quad \bar{R}^2 = 0.40$$

[7]According to my criteria I classified the following countries located in East Asia and Pacific as low income countries: China, Indonesia, Malaysia, Papua New Guinea, Philippines, Singapore, South Korea, and Thailand. Similarly, a majority of the countries located in Sub-Saharan Africa are defined as low income countries whereas the opposite situation

All variables, except the intercept terms and population growth for the low income sample, have the expected sign and are strongly significant. However, the coefficients of the initial income, population growth and investment rate are higher in the high income sample than those in low income group while the opposite situation is true for the coefficient of secondary school enrolment rate. In order to check parameter stability for these two income groups, a joint test for null hypothesis of equality of all coefficients across two samples concludes that the null cannot be rejected with a high probability level.[8] In addition, the same test is carried out across pairs of coefficients. As can be seen from p values in column 7, test results show parameter stability for each individual coefficients between two income groups. Therefore, it is possible to conclude that parameters are stable across high income and low income countries.

In general, the regression results presented in Table 1 are consistent with the theory and support the results of MRW. In order to facilitate comparison of my results with those of MRW, I also present their regression results in the last three columns in Table 1.[9] Both my cross-country regression results and those obtained by MRW show that investment rates of both physical and human capital significantly contribute to the growth rates of countries while the rate of population growth has a negative effect on growth in the long run. The outstanding differences between my estimates and those by MRW are that my regression results show that the effects of investment in human capital are greater than those found by MRW, while the contribution of the investment rate in physical capital is found to be lower than MRW. In addition, I conclude a higher effect of population growth on long run growth. In the cross-country growth literature, researchers generally find a weak negative relation between population growth and long run growth. We expect a negative association between these two due to the simple reason that it is very difficult to keep a high level of capital per worker for a given saving rate while the number of workers is growing at a higher rate. However, it should be remembered that the negative effect of population growth on GDP growth may be larger due to the other factors, such as environmental factors and access to safe water as noted by Temple (1999). In addition, especially in Western Europe the aging population strains the social security system, neg-

is true for the many countries in Latin America and Caribbean. See the list of countries in the full sample depicted in the Appendix for further information low and high income countries.

[8]The F-statistics and p value are for this joint test is F(5, 97)=0.40 and 0.85, respectively.

[9]Indeed I estimated the augmented neoclassical growth model by employing the original data and samples of MRW over the 1960-1985 period and reported.

atively affecting the public budget and labour force participation because of higher health care costs of elderly people and the growing number of retirees in population. In particular, the substantially larger and strongly significant coefficient of population growth in the high income sample support this claim since a considerable part of this sample consists of Western European countries. Therefore, I believe that my results related to population growth seem more plausible.

On the other hand, strong evidence is found in each regression for the hypothesis of conditional convergence in the manner that an economy with a lower initial value of output per worker tends to generate higher growth rate of output per worker when other determinants are controlled. For instance, according to the full sample, the logarithm of real GDP per worker in 1960 has a cross-country mean of 8.295 and a standard deviation of 0.854. Therefore, the regression result based on the full sample indicates that a one-unit standard deviation decrease in the logarithm of initial income would increase the subsequent growth rate by 0.377 points over the 1960-2000 period (-0.442*-0.854=0.377). This is equivalent to a rise in annual growth rate of 0.9 percentage point over the same period (0.377/40=0.009). Comparing to MRW, the implied convergence rates in all samples are found to be very close to those estimated by MRW.

However, some authors such as Barro (1991), Easterly and Levine (1997), and others point out that this convergence result is generally quadratic rather than linear. If this argument is true, the subsequent growth rates firstly rise and then decrease with the initial level of income. Put differently, the subsequent growth rate is a concave function of initial income. This implies that the conditional convergence effect will be weaker for very poor countries while stronger for middle-income countries. In order to check for a possible non-linear relationship between initial income and growth rate, I also include the square of logarithm of initial income in the cross-country growth regressions in Table 1 and reestimate. Table 2 presents regression results. As shown in Table 2, I could not find any statistically significant quadratic relationship between the initial level of income and subsequent growth. In each sample, except for the NonOECD and high income samples, the coefficients of both initial income and initial income squared have the wrong signs if the argument above is true. Therefore, I conclude that the conditional convergence hypothesis is linear. The graphical visualization of partial association between growth rate and initial income shows a clear linearity (Figure 1).

For a further evaluation of the augmented neoclassical model, following MRW I imposed a restriction on equation (2.2) such that $\gamma_2+\gamma_3+\gamma_4=0$. As can be seen in equation (2.1), the sum of coefficients of $log(n_i+g+\delta)$, $logs_{i,K}$,

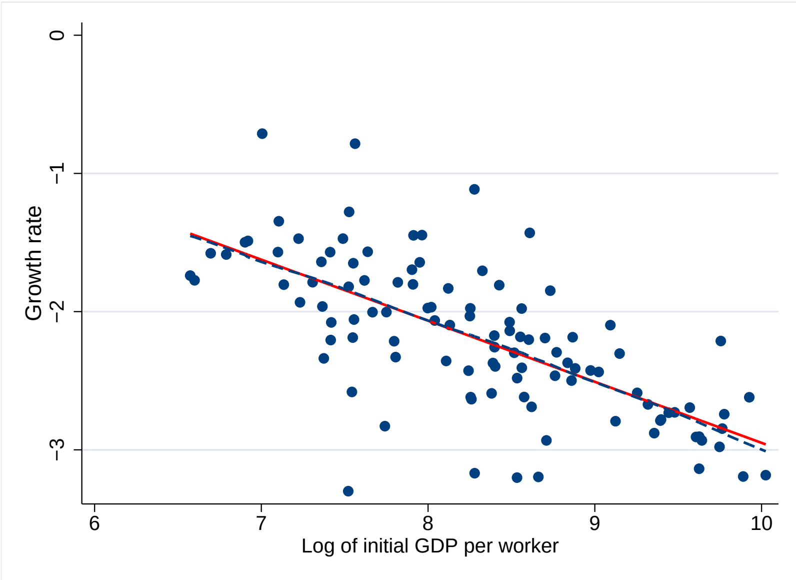


Figure 1: Growth Rate versus Initial Income: Partial Relation

and $logs_{i,H}$ should be equal to zero.[10] Therefore, this restriction implies that equation (2.2) can be expressed as

$$\begin{aligned} logy_{i,2000} - logy_{i,1960} = \gamma_0 + \gamma_1 logy_{i,1960} \\ + \gamma_3[log(Investment\ rate_i) - log(n_i + g + \delta)] \\ + \gamma_4[log(School\ enrolmet_i) - log(n_i + g + \delta)] + \varepsilon_i \end{aligned} \tag{2.3}$$

The restricted regression results are presented in Table 3. Before estimating the restricted model, this restriction is tested for each sample and p-values for test of restriction are given in Table 3. As can be seen, this restriction is not rejected in all samples. The implied estimates of physical capital share (α), human capital share (β) and convergence rate (λ^*) are given in the last three rows of Table 3. Again I present restricted regression results of MRW through column 7 and column 9 of Table 3. The results show that estimation of the restricted model slightly improves the coefficients of investment rates for both physical and human capital. All variables are found to be highly significant with expected signs in each sample except the OECD sample.

[10] Proof is available in MRW

Table 2: Quadratic Augmented Neo-classical Growth Model: OLS Estimates

Dependent Variable: Log Difference real GDP per worker over the 1960-2000 period

	My Estimation Results: 1960-2000 period						Estimates of MRW† 1960-1985 Period		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sample	Full	NonOil	OECD[a]	NonOECD	High Income	Low Income	NonOil	Intermediate	OECD
log GDP per worker 1960	-0.086	-0.152	-2.424	0.494	1.747	-0.432	-0.340	-0.358	0.598
	(0.10)	(0.17)	(0.78)	(0.41)	(0.38)	(0.09)	(0.45)	(0.52)	(0.21)
Square of log GDP per worker 1960	-0.021	-0.016	0.102	-0.059	-0.121	0.003	0.003	-0.000	-0.058
	(0.41)	(0.31)	(0.58)	(0.78)	(0.48)	(0.01)	(0.07)	(0.01)	(0.35)
$\log(n_i + g + \delta)$	-1.086	-0.936	-0.427	-1.051	-1.268	-0.735	-0.500	-0.546	-0.795
	(2.84)	(2.35)	(0.53)	(1.84)	(2.93)	(0.95)	(1.65)	(2.41)	(2.00)
log of Investment rate	0.420	0.417	0.578	0.408	0.587	0.388	0.524	0.538	0.325
	(4.42)	(4.36)	(1.54)	(2.90)	(3.05)	(2.80)	(5.99)	(3.92)	(1.82)
log of School enrolment	0.441	0.442	0.369	0.432	0.319	0.476	0.233	0.270	0.209
	(4.88)	(4.84)	(0.67)	(5.43)	(1.86)	(3.80)	(3.61)	(2.54)	(1.32)
Constant	1.222	1.826	14.388	-0.958	-7.488	3.371	2.061	2.469	-0.139
	(0.31)	(0.45)	(1.15)	(0.18)	(0.36)	(0.19)	(0.73)	(0.86)	(0.01)
Number of observations	107	102	26	81	52	55	98	75	22
p- value for heteroscedasticity[b]	0.67	0.62	0.07	0.12	0.65	0.30	0.24	0.01	0.25
Adjusted R^2	0.61	0.61	0.59	0.55	0.60	0.57	0.46	0.43	0.63

Note: t-statistics are in parenthesis. In the regressions where the heteroscedasticity test is failed to pass at 15 % level t-statistics based on heteroscedastic-consistent (White-robust) standard errors are reported. All variables, except initial level of income variables are averaged over the 1960-2000 period. The variable, $(n_i + g + \delta)$ refers to sum of rates of population growth, technical progress and depreciation.

† Estimation results based on the original data and samples of MRW.

[a] Czech Republic, Germany, Poland and Slovak Republic are excluded due to missing data.

[b] Breusch-Pagan test for heteroscedasticity in which the null refers to the homoscedastic errors.

Table 3: Augmented Neo-classical Growth Model: Restricted OLS Estimates

Dependent Variable: Log Difference real GDP per worker over the 1960-2000 period

	My Estimation Results: 1960-2000 period						Estimates of MRW[†] 1960-1985 Period		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sample	Full	NonOil	OECD[a]	NonOECD	High Income	Low Income	NonOil	Intermediate	OECD
log GDP per worker 1960	-0.438	-0.422	-0.539	-0.450	-0.467	-0.391	-0.298	-0.372	-0.402
	(6.55)	(6.03)	(5.53)	(5.00)	(4.09)	(2.58)	(4.93)	(6.38)	(5.81)
log of Investment rate - $\log(n_i+g+\delta)$	0.417	0.412	0.548	0.392	0.606	0.387	0.501	0.506	0.395
	(4.60)	(4.52)	(1.91)	(3.78)	(3.25)	(3.27)	(6.09)	(4.48)	(2.61)
log of School enrolment - $\log(n_i+g+\delta)$	0.465	0.452	0.218	0.457	0.391	0.469	0.235	0.266	0.241
	(5.75)	(5.47)	(0.92)	(4.96)	(2.65)	(4.34)	(3.98)	(3.08)	(1.69)
Constant	3.227	3.124	4.707	3.326	3.449	2.879	2.457	3.090	3.554
	(6.61)	(6.14)	(5.72)	(4.97)	(3.72)	(2.58)	(5.19)	(6.48)	(5.61)
Number of observations	107	102	26	81	52	55	98	75	22
p-value for heteroscedasticity[b]	0.74	0.62	0.48	0.15	0.62	0.29	0.32	0.02	0.23
Adjusted R^2	0.62	0.61	0.60	0.56	0.61	0.59	0.47	0.44	0.66
p-value for test of restriction	0.65	0.91	0.46	0.75	0.36	0.90	0.36	0.36	0.36
Implied α	0.32	0.32	0.42	0.30	0.41	0.31	0.48	0.44	0.38
Implied β	0.35	0.35	0.17	0.35	0.27	0.38	0.23	0.23	0.23
Implied λ^*	0.0144	0.0137	0.0194	0.0149	0.0157	0.0124	0.0142	0.0186	0.0206

Note: t-statistics are in parenthesis. In the regressions where the heteroscedasticity test is failed to pass at 15 % level t-statistics based on heteroscedastic-consistent (White-robust) standard errors are reported. All variables, except initial level of income variables are averaged over the 1960-2000 period. The variable, $(n_i + g + \delta)$ refers to sum of rates of population growth, technical progress and depreciation.

[†] Estimates of restricted augmented neoclassical growth model in table VI, page 429 of Mankiw, Romer and Weil (1992). Dependent variables for these regressions are log difference GDP per worker over the period 1960-1985.

[a] Czech Republic, Germany, Poland and Slovak Republic are excluded due to missing data.

[b] Breusch-Pagan test for heteroscedasticity in which the null refers to the homoscedastic errors.

[*] The rate of convergence in the neighborhood of steady-sate

However, Table 3 indicates that my results are different from those found by MRW in some respects. First, compared with the MRW results I find a stronger effect of investment in human capital and a weaker effect of investment in physical capital on economic growth. Second, I find that the effect of the accumulation of human capital is stronger than the contribution of investment in physical capital on economic growth. Third, the implied capital shares based on my regressions are substantially different from those estimated by MRW. As a natural result of first two findings I estimate a relatively larger share for human capital. However, conventionally capital shares are one-third (Mankiw (1995)); therefore, it is possible to conclude that my estimates of α and β are more reasonable. However, an exception to these findings are the regressions based on the high income and OECD samples. My findings therefore imply two important conclusions for economic growth. First, accumulation of human capital is very important, especially in poor countries. Second, physical capital accumulation is more important in richer countries.

There can be several reasons for these results. First, the secondary school enrolment rate is a crude proxy for accumulation of human capital and the strong relation between the school enrolment rate and economic growth in regressions in Table 1 and 3 may reflect other macroeconomic policies and factors which are excluded in the analysis. This is more likely in my regressions because, differently from MRW, I examine the effect of human capital on growth by using the secondary school enrolment rate over the schooling age population rather than over the economically active population. I check this possibility and find that the secondary school enrolment rates are highly correlated with the average inflation rate and some institutional quality measures such as rule of law, bureaucratic quality, corruption over the 1960-2000 period. However, even though this explanation is reasonable for the full sample results, it is not likely to explain the higher coefficient estimate of school enrolment rate in low income sample since these correlations are weaker for the low income countries.

Second, as noted by Caselli (2005), a considerable part of world's physical capital is produced in technologically advanced countries. This implies that, whilst the share of investment expenditures in GDP is a reasonable proxy in high income countries, the ratio of imported capital goods in GDP is a more plausible proxy for the physical capital saving rate in low income countries. More importantly, Caselli (2005) also emphasises the importance of technological progress embodied in capital goods. If physical capital in high income countries includes greater technological progress, then the effect of the investment rate on growth will be higher in these countries. Moreover, as pointed out by Mankiw (1995), physical and human capital are generally complementary inputs in the production process and so it is highly likely that this relation is stronger in high income countries due to the vintage physical capital. Therefore, the higher coefficient estimate of the investment rate in high income sample might be partly attributed to the accumulation of human capital.

Third, the strong association between the school enrolment rate and growth may be a result of a number of unrepresentative countries. In Section 4, I investigate the outlying countries and highlight how Tanzania is an unrepresentative country. When I omit this country from the regression analysis, I conclude that the effect of the investment rate becomes stronger relative to the school enrolment rate.

Finally, the higher coefficient estimate on the school enrolment rate in the low income sample is a result of sub-Saharan Africa. As mentioned before, almost all countries in sub-Saharan Africa belong to the low income sample. Similarly, the majority of countries in East Asia and the Pacific are in this sample. Sub-Saharan Africa has a substantially lower school enrolment rate with an average value of 17 percent, than the sample mean which is equal to 48 percent over the 1960-2000 period. By contrast, the average school enrolment rate in East Asia and the Pacific is 68 percent. It is a well-known fact that sub-Saharan Africa experienced very poor growth performance over the 1960-2000 period while the countries of East Asia and the Pacific recorded very much faster growth rates during the same period. Thus, with the poor performing countries of sub-Saharan Africa with low school enrolment and the well performing countries of East Asia and the Pacific associated with greater schooling, the higher coefficient estimate is inevitably obtained.

Another distinguishing difference between my results and MRW estimates is that the R^2 is higher in my regressions, since the variation in the explanatory variables, especially in the school enrolment rates are considerably higher over the 1985-2000 period than between 1960-1985.

It is obvious that the most important reasons for the differences between our results and those obtained by MRW are that I estimate the augmented neoclassical growth model employing the updated and revised data for different samples over a different time period. In addition, in order to make comparison concrete, I have also estimated the augmented neoclassical growth model using my data and the original data of MRW for the same samples of countries. The estimation results are given in the Appendix (Table 7). However, selecting the same samples does not remove all of these differences. This implies that the differences between findings and those by MRW are partly the result of different time period and partly the result of different data. Yet, in spite of these differences it is possible to conclude that my results confirm the work of MRW and are consistent with the existing cross-country growth literature.

Before closing this paper, I consider two further checks on my estimation, namely the effect of geographical differences across countries and investigating possible outlying observations.

3 Geographical Differences

In the cross-country growth literature, some studies (Barro (1991), De Long and Summers (1991), Levine and Renelt (1992), Barro and Lee (1994), Easterly and Levine (1997), Barro and Sala-i-Martin (2004) *inter alia*) include dummy variables for sub-Saharan Africa, Latin America and the Caribbean and East Asia. The reason for employing these dummy variables is that estimated growth models are not adequate to explain different growth performances across these regions. Due to the poor growth performance in Africa and Latin America one can expect a significantly negative coefficient on dummy variables for these two regions whereas the opposite situation is true for East Asia and Pacific.

In order to investigate this claim, in particular for sub-Saharan Africa, I add to my baseline model three regional dummies. First column of Table 4 reports the results of this estimation. It can be seen that all dummy variables have the anticipated signs and jointly significant. This implies that these three regions exhibit different growth performances compared to the rest of the world.

In column 2, I omit the dummy variable for East Asia and Pacific. Notice that for this regression the reference countries now include the countries in East Asia and Pacific as well as others. The regression result shows that dummy variables for both sub-Saharan Africa and Latin America are individually and jointly significant with the negative signs. This finding clearly shows that both regions experienced a slower growth performance compared to other countries once the reference sample includes East Asia and Pacific. In column 3, I drop the Latin American dummy but keep the Asian dummy with African dummy and conclude that only dummy for East Asia and Pacific is significant. Finally, in column 4, I allow only dummy variable for sub-Saharan Africa and find that its coefficient estimate is not significant.

What can be inferred from these findings and should one employ regional dummies in the cross-country growth regressions? Undoubtedly, the cross-country growth regressions in Table 4 show clear evidence for different growth performances in these three regions. However, care must be taken with the reference group of countries. Yet, I think that the importance of sub-Saharan African dummy may have been exaggerated in the previous literature since I find a significant statistical relationship between growth and this dummy in only regression in Column 2.[11]

To investigate this further, I estimate the benchmark model for sub-Saharan Africa and rest of the world separately in columns 5 and 6 and carry out a parameter stability test across two samples. However, I could not reject the null hypothesis of equality of all coefficients between Africa and non-Africa samples (The F-statistics and p value for the parameter stability test are $F(5,97)=0.95$ and $p=0.45$, respectively). Moreover, the same test for individual coefficients concludes parameter stability for each variable across two samples as it can be seen from p values in column 7. This implies that sub-Saharan Africa does not exhibit a different growth performance compared to the rest of the world on average.

More importantly, previous studies concluded strong and significant relation between regional dummy variables, especially dummies for Latin America and Africa, and growth may be attributable to the importance of model selection problem. Put differently, I believe that employing new growth theories in order to explain different growth performances of different geographical regions is more informative and useful than employing simple dummy variables.

Finally, in column 8, I include a landlocked country dummy that takes the value of 1 for countries that do not have access to international waters in order to check the growth performance of these countries compared with others. Landlocked countries may face higher costs for any kind of international activities, especially international trade. Therefore, this dummy variable has been extensively used in cross-country growth work. However, the regression in column 8 indicates that there is no significant difference in the growth performance of landlocked countries. In the last column of Table 4.4, I omit the landlocked countries in Europe such as Austria, Luxembourg, Switzerland and Hungary, from the dummy variable, since becoming landlocked for these countries may not create an important disadvantage. When I use this dummy variable, the regression result shows an improvement on both coefficient and t-statistics of landlocked dummy, but it is still insignificant.

[11] Studies such as by Barro (1991), Barro and Lee (1994), Easterly and Levine (1997) consistently conclude that the African dummy is significant and negative.

Table 4: Augmented Neo-classical Growth Model and Geographical Dummy Variables

Dependent Variable: Log Difference real GDP per worker over the 1960-2000 period

	(1) Full	(2) Full	(3) Full	(4) Full	(5) Africa	(6) NonAfrica	(7) p-value[a]	(8) Full	(9) Full
log GDP per worker 1960	-0.414	-0.421	-0.431	-0.449	-0.406	-0.474	0.36	-0.441	-0.443
	(6.23)	(6.31)	(6.44)	(6.61)	(2.41)	(5.56)		(6.49)	(6.53)
$\log(n_i + g + \delta)$	-0.970	-0.863	-1.172	-1.076	-1.421	-1.104	0.32	-1.072	-1.059
	(2.61)	(2.35)	(3.20)	(2.90)	(1.02)	(2.52)		(2.84)	(2.85)
log of Investment rate	0.340	0.359	0.364	0.404	0.296	0.643	0.64	0.414	0.413
	(3.73)	(3.95)	(3.96)	(4.39)	(2.12)	(3.65)		(4.47)	(4.47)
log of School enrolment	0.382	0.375	0.402	0.398	0.455	0.265	0.31	0.440	0.436
	(3.90)	(3.81)	(4.06)	(3.94)	(3.10)	(1.91)		(4.75)	(4.75)
Sub-Saharan Africa	-0.228	-0.299	-0.079	-0.137	-	-	-	-	-
	(1.52)	(2.08)	(0.59)	(1.01)					
Latin America & Caribbean	-0.237	-0.294	-	-	-	-	-	-	-
	(2.11)	(2.75)							
East Asia & Pacific	0.198	-	0.290	-	-	-	-	-	-
	(1.51)		(2.31)						
Landlocked country dummy	-	-	-	-	-	-	-	-0.039	-
								(0.36)	
Landlocked country dummy [b]	-	-	-	-	-	-	-	-	-0.062
									(0.50)
Constant	2.648	3.079	2.211	2.747	1.229	3.231	0.31	2.715	2.765
	(2.43)	(2.90)	(2.03)	(2.53)	(0.35)	(2.15)		(2.45)	(2.53)
p-value for heteroscedasticity[c]	0.86	0.91	0.83	0.66	0.34	0.10	-	0.72	0.74
F-value[d]	3.68	4.32	3.19	-	-	-	-	-	-
Number of observations	107	107	107	107	34	73	-	107	107
Adjusted R^2	0.64	0.64	0.63	0.61	0.42	0.53	-	0.61	0.61

Note: t-statistics are in parenthesis. In the regressions where the heteroscedasticity test is failed to pass at 15 % level t-statistics based on heteroscedastic-consistent (White-robust) standard errors are reported. All variables, except initial level of income variables are averaged over the 1960-2000 period. The variable, $(n_i + g + \delta)$ refers to sum of rates of population growth, technical progress and depreciation.

[a] The p-value refers to the hypothesis that individual coefficients are the same for Africa and non-Africa samples.

[b] The landlocked countries in Europe are dropped.

[c] Breusch-Pagan test for heteroscedasticity in which the null refers to the homoscedastic errors.

[d] Test for joint significance of the regional dummies.

4 Outliers

An important concern related to the cross-country growth works is that results may be partly driven by outlying countries. Undoubtedly this concern is very important since cross-country growth studies are based on small samples. As Temple (1999, 2000) points out we should make sure that our results reflect the tendencies of a majority of data not those of a minority of observations if we want to reach useful generalizations about growth.

Before proceeding for outlier checking, I want to make clear the terminology since the definition of outlier is sometimes unclear and confused. An outlier is simply an observation which is considerably different from the remaining observations in the sample (Hawkins (1980), Barnett and Lewis (1994)). This difference between outlying observation and others may occur either in the dependent variable or in the explanatory variable(s) or in the both. An outlier in the dependent variable yields a large residual and is some times referred to as outlier in the response variable (Chatterjee and Hadi (1988, 2006)) or outlier in y-direction or vertical outlier (Rousseeuw and Leroy (1987)). Hence, one can easily detect a single outlier in the response variable by simply checking residuals. Yet, an outlier may also arise in the explanatory variable(s) and take a place far from the bulk of data of observed explanatory variables in the sample. This kind of outlier is also known as the outlier in the predictors or design outlier. Since outlier in the predictors are far away from the bulk of data, they have high leverage values and are some times referred to as high leverage data points in order to distinguish them from the outliers in the response variable. Therefore, detecting an outlying observation by checking residuals is often misleading. The reason is that a high leverage data point pulls the OLS regression line towards itself and yields small residuals. In addition, this data point causes the larger residuals for other observations.

On the other hand, an influential observation is the data point that has individually or collectively substantial influence on the regression results with respect to other observations in the sample (Belsley et al. (1980)). Thus, removing an influential observation from the sample changes the fitted regression equation considerably. Two points in this definition deserve special emphasis. First, an observation can be influential individually or together with a group of other observations. This implies that while removing a single observation from the sample does not change the regression result, dropping that observation with other observation(s) can substantially affect fitted regression equation. Second, the term of influence in the definition is partly subjective such that an influential observation can affect the coefficient estimates of variable(s) or their standard deviations, and hence t-statistics,

R^2 and so on. This means that all influential observations do not have equal influence on the regression results as argued by Chatterjee and Hadi (1988). One can effect the magnitude and/or sign of coefficient estimates while the other can be influential on the statistical significance or goodness of fit. Therefore, one should particularly keep in mind how outliers affect the regression results in terms of the objective of empirical research.

Even though outliers are not necessarily influential observations, they are generally have influence on the regression results. Therefore, an important primary task of regression analysis is to detect influential outlying observations in the sample in order to conduct useful and reliable generalizations. However, identification of outliers in the multiple regression is not easy especially if the sample includes more than one outlier. As a starting point, checking residuals (especially studentized ones) and leverages after the regression is always suggested.

Many outlier identification methods have been suggested in statistics.[12] It is possible to classify these methods under the two main categories. The first and most common one is the regression diagnostics such as Cook's distance, DFITS statistics and Welsh distance. These diagnostics basically take into account the changes in the fitted regression equation after a single observation is removed and hence they directly measure the influence of each individual observation. Therefore, these diagnostics are some times referred as the direct outlier detecting methods (Rousseeuw and Leroy (1987)). Among these statistics, DFITS suggested by Belsley et al. (1980) is the most widely used in the regression analysis and shows the effect of each observation in the sample on the overall fitted regression analysis. In addition, Belsley et al. (1980) proposed a similar measure which is known as DFBETA statistics. Differently from DFITS, DFBETA reveals the influence of each observation on a particular explanatory variable and thus it is very useful when the primary interest of researcher is focused on a specific variable in the regression equation.

As a rule of thumb, observations having large values of these diagnostics are considered as influential outliers in the response variable and/or in the predictors. Moreover, several cutoff points for them are suggested in the literature. The choice of cutoff points depends on the sample size and number of explanatory variables. Generally, for small samples a high cutoff point is plausible (Bollen and Jackman (1990)). However, instead of using a particular cutoff level, it is better to examine these diagnostics graphically and identify observations with unusual patterns, as pointed out by Chatterjee

[12]Outlier detection methods from the Bayesian perspective are also suggested in the statistical literature. See, Guttman et al. (1978), Pettit (1992), Hoeting (1994)

and Hadi (1988).

Even though these diagnostic measures are useful, their efficiency substantially decreases if the sample includes more than one outlier. The reason is that all diagnostic measures mentioned above are based on the removing of single observation and they are no longer powerful in the case of multiple outliers due to the masking and swamping effects. When the sample includes more than one outlier, some outliers may be hidden by the others and this effect is known as masking effect. On the other hand, swamping effect arises due to the fact that outliers, especially those with high leverage values, make other observations lie far from the fitted regression equation by pulling the regression equation towards themselves. Therefore, the best solution of these problems is to calculate diagnostic measures based on the deletion of all subsets of observations. However, this is almost practically infeasible not only because deciding the number of subsets is difficult but also computation is immense due to the larger number of subsets.[13]

The second and in the case of multiple outliers, more efficient class of outlier detection methods is the robust regressions. These methods in essence suggest employing robust regression techniques which are resistant to outliers. On the contrary to the common view, robust regressions do not simply ignore outliers. Rather, one can identify outliers by comparing residuals obtained from a robust regression with those derived from the OLS. Therefore, regression diagnostics and robust regressions basically serve the same purpose only from the opposite side as argued by Rousseeuw and Leroy (1987). In statistics many robust regression techniques such as Least Median of Squares, Least Trimmed Squares are suggested (See, Rousseeuw and Leroy (1987), Rousseeuw and van Zomeren (1990), Atkinson (1994)). However, robust regression gives us an idea such that we can apply weighted least squares analysis based on the identification of the outliers. If this can be done, the results of weighted least squares will be less sensitive to the outlying observations and more plausible with respect to those obtained from usual least squares.

Hadi (1992b) suggests another measure in order to identify influential outliers in the data. This diagnostic which measures overall potential influence of the ith observation is defined as

$$H_i = \frac{k}{1-h_i}\frac{d_i^2}{1-d_i^2} + \frac{h_i}{1-h_i}, i = 1, ..., n \tag{4.1}$$

[13]For instance, if the consideration of all subsets includes only 2 observations and the sample size is 107 as in the case of our full sample, there are 5,671 possibilities. When we consider 3 out of 107 observations, the number of possibilities is 198,485.

where k is the number of explanatory variables (including constant), $d_i = e_i/\sqrt{SSE}$ and h_i is the ith normalized residual and leverage, respectively. As can be seen, this diagnostic measure is the sum of two components. The first term on the right-hand side of the equation (4.1) is a function of the ith normalized residual weighted by the ith leverage value and measures outlyingness in the response variable. The second component is also known as the potential function and measure the outlyingness in the predictors.

The diagnostic measure proposed by Hadi (1992b) has several desirable properties compared to the traditional diagnostics. First, it measures overall potential influence of an observation. As can be seen, outlying observations in either response variable and/or predictors will have large values of Hadi's measures. Second, Hadi's measure identifies potentially influential observations on several regressions rather than the focusing on a single regression. Third, this measure is an additive function of both residual and leverage values.[14] Fourth, Hadi's measure is monotonically increasing function of both residuals and leverages.

Even though Hadi's measure is superior than the traditional measures, it is still designed for a single observations. In order to highlight multiple outliers in the data set, Hadi proposed two ways. Firstly, he suggests a simple graph which is based on the formula in equation (4.1). This graph which is also known as "potential-residual plot" is a more powerful tool in order to detect both single and multiple influential cases. Secondly and more importantly, he proposed a practical method to search multiple outliers (Hadi (1992a, 1994), Hadi and Simonoff (1993)).

In the light of these explanations, I apply the method suggested by Hadi (1992a) for identification of outlying observations in my baseline cross-country growth regression. However, before applying Hadi methodology, I investigate outlying observations by employing diagnostic plots.

The diagnostic plot suggested by Gray (1986) is commonly used in statistical literature for a quick way of checking influential observations and shows leverage versus the residual squared. Figure 2(a) plots leverages against the normalised residual squared of baseline cross-country growth regression based on the full sample. In this figure two reference lines parallel to horizontal and vertical axes show the mean values of leverage and normalised residual squared, respectively. Observations that are located far away from reference lines are of great concern for us. As can be easily seen Tanzania

[14] Traditional diagnostics such as Cook's distance and DFITS are the multiplicative function of residual and leverage values. Since observations with large values yields either large residuals or large leverages or both, multiplicative diagnostics would be small if one of these quantities (residuals and leverages) is small and thus could fail to detect influential outliers.

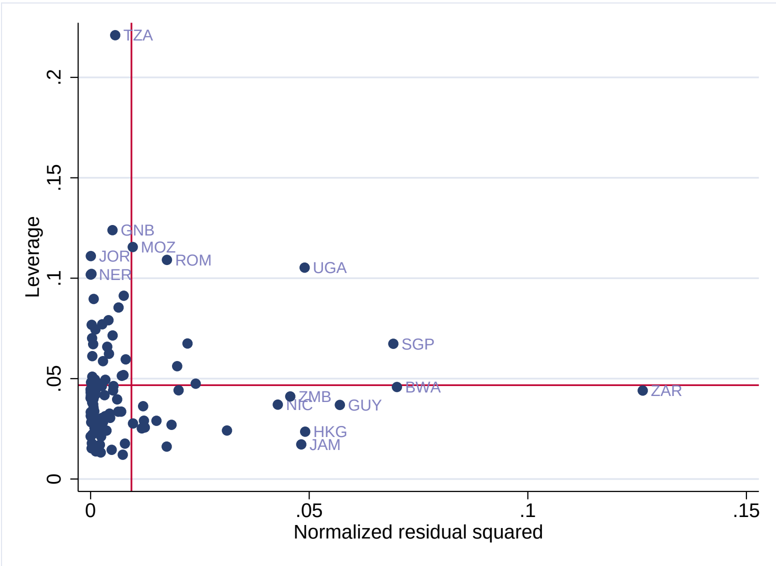
is the most influential country with the highest leverage in my sample. The feature of this country is that it has the smallest school enrolment rate in the sample and hence it is an outlier in the saving rate for human capital. In addition, the countries Uganda, Romania, Mozambique, Guinea-Bissau, Jordan and Niger have moderately high level of leverages. Figure 2(a) also shows that Congo Democratic Republic (former Zaire) is the country with the highest residual. Congo Democratic Republic has the lowest growth performance in my sample and hence this country is clearly an outlier in the growth rate. Yet, the leverage value of this country is very close to the sample mean. Therefore, it is possible to conclude that Congo Democratic Republic is not influential in spite of its very high residual. In addition to this country, Figure 2(a) indicates the countries Guyana, Botswana, Jamaica, Zambia, Nicaragua, Uganda, Hong Kong and Singapore whose residuals are relatively higher. However, neither of them has a high leverage value. Therefore, graphical inspection indicates that only Tanzania may be a candidate as a potential influential outlier. Even though the leverage of Congo Democratic Republic is low, this country is of concern due to its high level of the residual. In addition to leverage versus plot I also present Hadi's potential-residual plot in figure 2(b). As can be seen, potential-residual plot exactly similar to leverage-residual plot and hence supports our findings from figure 2(a).

When I apply Hadi method on my data set, I conclude only Tanzania as a potential outlying country. As can be seen, Hadi measure confirms our findings from graphical inspection.

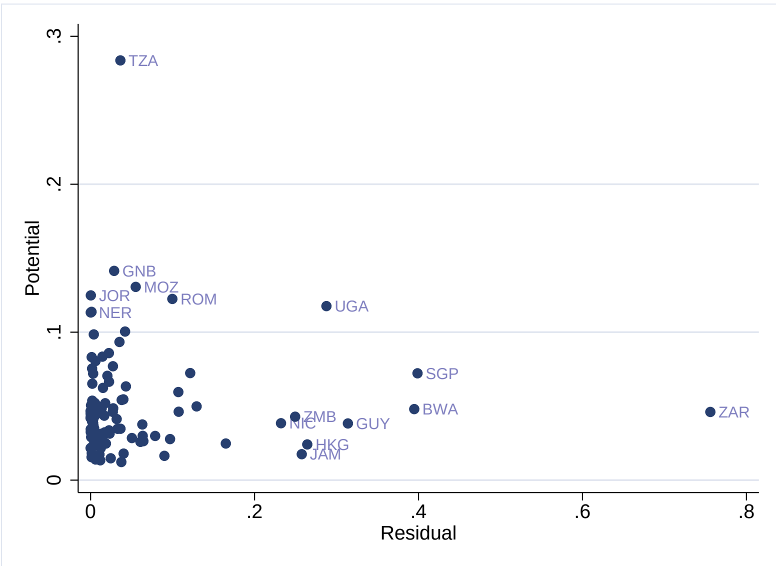
The most common approach for solving outlier problems is dropping these observations from the sample. Omitting Tanzania yields the following cross-country growth regression (t-statistics are in parentheses)

$$\begin{aligned} logy_{i,2000} - logy_{i,1960} = {} & \underset{(2.55)}{2.779} - \underset{(6.55)}{0.443} logy_{i,1960} \\ & - \underset{(2.88)}{1.070} log(n_i + g + \delta) \\ & + \underset{(4.44)}{0.445} log(Investment\ rate_i) \\ & + \underset{(4.48)}{0.421} log(School\ enrolment_i), \quad \overline{R}^2 = 0.62 \end{aligned} \tag{4.2}$$

Compared to my previous regression presented the first column of Table 4.1, dropping Tanzania from the sample slightly changes the coefficient estimates of initial income and population growth. Therefore, there is not an important change for the implied convergence rate. Yet now, the effect of saving rate



(a) Leverage versus Residual plot



(b) Potential-Residual Plot

Figure 2: Benchmark Growth Model: Diagnostic Plots for Outliers

for physical capital is stronger than the saving rate for human capital. In addition, R^2 of the model increases by one percent.

However, removing outlying countries from the sample may not be a good solution in the cross-country growth regressions. The most important reason is that some countries can behave as outliers due to the fact that a relevant variable has been omitted from the specified model. This last point is closely related to model uncertainty problem and hence removing some observations may be considered another kind of data mining (Chatfield (1995)). Since the estimated cross-country growth regression is proposed as a benchmark model in order to investigate the relationship between any particular growth theory and long-run economic growth, I prefer keeping Tanzania in the full sample. However, detecting outliers in the cross-country growth works is a noteworthy task. Temple (1999, p.127) points out "[T]he identification of possible outliers will not only render generalizations more robust, but will also highlight countries with atypical growth experiences, ones that are particularly likely to reward further study." This is particularly very important when testing new growth theories and/or investigating the relationship between economic growth and a policy variable.[15]

5 Conclusion

In this paper, I replicate the augmented neoclassical growth model developed by MRW using an updated and revised data set over the period 1960-2000. My results support findings of MRW, yet are different in some aspects: First, I find a stronger effect of investment in human capital and a weaker effect of investment in physical capital on economic growth compared to MRW; Second, the effect of accumulation of human capital is stronger than the contribution of investment in physical capital on economic growth. This finding is more obvious in the low income countries while results obtained from high income sample are more similar to MRW; Third, my coefficient estimates of physical and human capital shares are more reasonable than MRW. The regression results based on the largest sample indicates that the shares of physical and human capital are 32 and 35 percent, respectively. Comparing those obtained by MRW, my empirical results are more consistent with three factor Cobb-Douglas production function, $Y = K^{1/3}H^{1/3}L^{1/3}$.

I also check the effect of geography by employing three region dummy variables and conclude that inclusion of these dummies does not alter my main conclusions. In addition, my findings are not mainly driven by outliers.

To conclude, the estimated growth model in this paper appears appropriate for investigating the relation between growth and any particular growth theory.

[15]For instance, Easterly (2005) argues that the strong effects of policies obtained from cross-country growth regressions are mainly a result of extreme observations.

References

A. C. Atkinson. Fast very robust methods for the detection of multiple outliers. *Journal of the American Statistical Association*, 89(428):1329–1339, December 1994.

V. Barnett and T. Lewis. *Outliers in Statistical Data.* Wiley, New York, 3rd edition, 1994.

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

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

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

D. A. Belsley, E. Kuh, and R. E. Welsch. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* Wiley, New York, 1980.

K. A. Bollen and R. W. Jackman. Regression diagnostics: An expository treatment of outliers and influential cases. In J. Fox and J. S. Long, editors, *Modern Methods of Data Analysis*, pages 257–291. Sage Publications, Newburry Park, CA, 1990.

F. Caselli. Accounting for cross-country income differences. In P. Aghion and S. N. Durlauf, editors, *Handbook of Economic Growth*, volume 1A, chapter 9, pages 680–741. North-Holland, Amsterdam, 2005.

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

S. Chatterjee and A. S. Hadi. *Regression Analysis by Example.* Wiley, Hoboken, New Jersey, fourth edition, 2006.

S. Chatterjee and A. S. Hadi. *Sensitivity Analysis in Linear Regression.* Wiley, New York, 1988.

J. B. De Long and L. H. Summers. Equipment investment and economic growth. *Quarterly Journal of Economics*, 106(2):445–502, May 1991.

S. N. Durlauf, P. A. Johnson, and J. R. Temple. Growth econometrics. In P. Aghion and S. N. Durlauf, editors, *Handbook of Economic Growth*, volume 1A, chapter 8, pages 555–677. North-Holland, Amsterdam, 2005.

W. Easterly. National polices and economic growth: A reappraisal. In P. Aghion and S. N. Durlauf, editors, *Handbook of Economic Growth*, volume 1A, chapter 15, pages 1015–1059. North-Holland, Amsterdam, 2005.

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

J. B. Gray. A simple graphic for assessing influence in regression. *Journal of Statistical Computation and Simulation*, 24:121–134, 1986.

I. Guttman, R. Dutter, and P. R. Freeman. Care and handling of univariate outliers in the general linear model to detect spuriousity. *Technometrics*, 20:187–193, 1978.

A. S. Hadi. Identifying multiple outliers in multivariate data. *Journal of the Royal Statistical Society, Series B (Methodological)*, 54(3):761–771, 1992a.

A. S. Hadi. A modification of a method for the detection of outliers in multivariate samples. *Journal of the Royal Statistical Society, Series B (Methodological)*, 56(2):393–396, 1994.

A. S. Hadi. A new measure of overall potential influence in linear regression. *Computational Statistics and Data Analysis*, 14:1–27, 1992b.

A. S. Hadi and J. S. Simonoff. Procedures for the identification of multiple outliers in linear models. *Journal of the American Statistical Association*, 88(424):1264–1272, December 1993.

L. C. Hamilton. *Regressions with Graphics: A Second Course in Applied Statistics*. Duxbury Press, Belmont, CA, 1992.

D. M. Hawkins. *Identications of Outliers*. Chapman and Hall, London, 1980.

A. Heston, R. Summers, and B. Aten. *Penn World Table Version 6.1*. Center for International Comparisons at the University of Pennsylvania (CICUP), October 2002. Available at http://pwt.econ.upenn.edu/.

D. C. Hoaglin, B. Iglewicz, and J. W. Tukey. Performance of some resistant rules for outlier labeling. *Journal of the American Statistical Association*, 81(396):991–999, December 1986.

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

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

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

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

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

L. I. Pettit. Bayes factors for outlier models using the device of imaginary observations. *Journal of the American Statistical Association*, (87):541–545, 1992.

P. J. Rousseeuw and A. M. Leroy. *Robust Regression and Outlier Detection.* Wiley, New York, 1987.

P. J. Rousseeuw and B. C. van Zomeren. Unmasking multivariate outliers and leverage points. *Journal of the American Statistical Association*, 85 (411):633–639, September 1990.

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

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

J. M. Wooldridge. *Introductory Econometrics: A Modern Approach.* South-Western, Mason, Ohio, 2003.

World Bank. *World Development Indicators (CD-ROM data).* Washington, DC, 2002.

World Bank. *World Development Indicators (CD-ROM data).* Washington, DC, 2006.

6 Appendix: Descriptions and Sources of Variables used in Cross-Country Growth Regression Analysis

Real GDP per capita (RGDPCH) : 1996 international prices, chain series. **Source**: Global Development Network Growth Database (2005) which rely on Heston, Summers and Aten (2002)

Population (TP) : Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. **Source**: The World Bank World Development Indicators (2002, 2006).

Labour force (LF) : Labour force or economically active population defined as the total population between ages 15 and 64. **Source**: The World Bank World Development Indicators (2002, 2006).

Share of labour force (SLF) : Share of labour force in total population. The exact calculation is LF/TP.

Real GDP per worker (PWGDP) : 1996 international prices, chain series. The exact calculation is $PWGDP = RGDPCH * (1/SLF)$.

Growth : Average growth rate of real GDP per worker over the 1960-2000 period. The exact calculation is $log(PWGDP2000/PWGDP1960)$, where PWGDP1960 and PWGDP2000 is the real GDP per worker in 1960 and 2000, respectively. Because of missing variables, for the countries Bahamas, Belize, Haiti, Hungary, Malta, Oman, Puerto Rico, Sierra Leone, Sudan and Tunisia, 1961 values are used instead of 1960 values.

Initial income (PWGDP1960) : Real GDP per worker in 1960. Because of missing variables, for the countries Bahamas, Belize, Haiti, Hungary, Malta, Oman, Puerto Rico, Sierra Leone, Sudan and Tunisia, 1961 values are used instead of 1960 values.

Population growth (n) : Average rate of population growth between 1960 and 2000. The exact calculation is $(1/40) * log(TP2000/TP1960)$, where TP1960 and TP2000 are total population in 1960 and 2000, respectively.

(g+δ) : Sum of exogenous rates of technological process and depreciation over the 1960-2000 period and assumed to be equal to 0.05.

(n+g+δ) : Sum of rates of population growth, technical process and depreciation over the 1960-2000 period.

Investment rate (INV) : Average of investment share in GDP at constant prices over the 1960-2000 period. The data are averages for Tunisia and Sierra-Leone over 1961-2000 period, for Hungary and Malta 1965-2000 period, for Namibia, Cyprus, Botswana, Mauritania, Haiti, Central African Republic, Guyana and Fiji, over 1960-1999 period instead of 1960-2000 period. **Source**: Heston, Summers and Aten (2002). In order to increase number of observations, data of seven countries are filled up by using gross capital formation data from the World Bank World Development Indicators (2002, 2006). These countries are Puerto Rico for 1986-1991 period, Hungary for 1965-69 period, Malta for 1965-1993 period and years 1999, 2000, Sierra Leon for years 1997, 1999 and 2000, Cyprus for 1997-99 period, Angola for 1997-2000 period, Congo Democratic Republic for 1998-2000 period.

School enrolment rate (SCH) : Average gross rate of secondary school enrolment over the 1960-2000 period. Gross secondary school enrollment ratio is defined as the ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to the level of secondary education. For countries Chad, Ethiopia, Portugal, Niger and Mauritania, the variable is calculated over the 1965-2000 period. **Source**: The World Bank World Development Indicators (2002, 2006).

Sub-Saharan Africa dummy (REG_SSA) : A dummy variable takes the value of 1 for the countries in sub-Saharan Africa. **Source**: Global Development Network Growth Database (2005).

Latin American dummy (REG_LAC) : A dummy variable takes the value of 1 for the countries in Latin America and Caribbean. **Source**: Global Development Network Growth Database (2005).

East Asian dummy (REG_EAP) : A dummy variable takes the value of 1 for the countries in East Asia and Pacific. **Source**: Global Development Network Growth Database (2005).

Landlocked Country (LANDLOCK) : A dummy variable for landlocked countries. **Source:** Gallup et al. (1999)

Landlocked Country without Europe (LANDLOCK_WE) : A dummy variable for landlocked countries, except those in Europe (Andorra, Austria, Belarus, Czech Republic, Hungary, Luxembourg, Liechtenstein, Moldova and Switzerland). **Source:** Gallup et al. (1999) and author's calculation.

Table 5: List of Countries in the Full Sample

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

Note: Countries with asterisk are the high income countries according to the median of real GDP per worker in 1960.

Table 6: Summary Statistics

Variable	♯ of Obs.	Mean	Std. Dev.	Min	Max
Growth	118	0.6728	0.6639	-1.3525	2.3247
log PWGDP1960	118	8.3153	0.8390	6.5737	10.0252
$(n+g+\delta)$	191	0.0696	0.0120	0.0465	0.1396
INV	116	0.1568	0.0777	0.0207	0.4120
SCH	125	0.4841	0.3120	0.0444	1.1460
REG_SSA	207	0.2367	0.4261	0	1
REG_LAC	207	0.1884	0.3920	0	1
REG_EAP	207	0.1691	0.3757	0	1
LANDLOCK	208	0.1923	0.3951	0	1
LANDLOCK_WE	208	0.1394	0.3472	0	1

Table 7: Augmented Neo-classical Growth Model: OLS Estimates

Dependent Variable: Log Difference real GDP per worker over the 1960-2000 period

Sample	Full		NonOil		Intermediate		High Income		Low Income	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Our	MRW†	Our	MRW†	Our	MRW†	Our	MRW†	Our	MRW†
log GDP per worker 1960	-0.503	-0.312	-0.483	-0.289	-0.612	-0.384	-0.541	-0.347	-0.382	-0.344
	(6.41)	(5.09)	(5.70)	(4.33)	(6.52)	(5.76)	(4.09)	(3.87)	(2.36)	(2.05)
$\log(n_i + g + \delta)$	-1.214	-0.414	-1.056	-0.345	-0.928	-0.584	-1.228	-0.522	-1.817	0.073
	(2.91)	(1.70)	(2.41)	(1.31)	(2.13)	(2.38)	(2.51)	(1.62)	(1.59)	(0.14)
log of Investment rate	0.429	0.572	0.425	0.560	0.445	0.529	0.471	0.661	0.441	0.534
	(4.46)	(5.11)	(4.38)	(4.94)	(3.90)	(3.85)	(2.22)	(4.30)	(3.57)	(3.60)
log of School enrolment	0.469	0.214	0.467	0.201	0.644	0.288	0.463	0.101	0.455	0.218
	(5.03)	(2.99)	(4.91)	(2.74)	(4.84)	(2.98)	(2.05)	(0.67)	(3.84)	(2.34)
Constant	2.914	1.794	3.169	1.529	4.775	2.710	3.268	2.257	0.449	1.112
	(2.58)	(2.02)	(2.75)	(1.59)	(3.51)	(3.03)	(1.93)	(1.94)	(0.14)	(0.76)
p-value for heteroscedasticity[a]	0.79	0.08	0.90	0.07	0.31	0.01	0.91	0.33	0.98	0.07
Number of observations	94	94	89	89	72	72	45	45	49	49
Adjusted R^2	0.62	0.48	0.61	0.46	0.59	0.43	0.59	0.50	0.59	0.43
Implied λ^*	0.0175	0.0150	0.0165	0.0136	0.0237	0.0194	0.0195	0.0170	0.0120	0.0169

Note: t-statistics are in parenthesis. In the regressions where the heteroscedasticity test is failed to pass at 15 % level t-statistics based on heteroscedastic-consistent (White-robust) standard errors are reported. All variables, except initial level of income are averaged over the 1960-2000 period. The variable, $(n_i + g + \delta)$ refers to sum of rates of population growth, technical progress and depreciation.

† Estimates of augmented neoclassical growth model using the original data of Mankiw, Romer and Weil (1992) over the 1960-1985 period. Dependent variables for these regressions are log difference GDP per worker over the period 1960-1985.

[a] Breusch-Pagan test for heteroscedasticity in which the null refers to the homoscedastic errors.

* The rate of convergence in the neighborhood of steady-sate

Central Bank of the Republic of Turkey
Recent Working Papers

The complete list of Working Paper series can be found at Bank's website (http://www.tcmb.gov.tr).

A New Core Inflation Indicator for Turkey
(Necat Tekatlı Working Paper No. 10/19, October 2010)

A Bayesian Generalized Factor Model with Comparative Analysis
(Necat Tekatlı Working Paper No. 10/18, October 2010)

Measuring the Impact of Monetary Policy on Asset Prices in Turkey
(Murat Duran, Gülserim Özcan, Pınar Özlü, Deren Ünalmış Working Paper No. 10/17, September 2010)

The Trade Credit Channel of Monetary Policy Transmission: Evidence from Non-financial Firms in Turkey
(Pınar Özlü, Cihan Yalçın Working Paper No. 10/16, September 2010)

Economic Uncertainty and Money Demand Stability in Turkey
(K. Azim Özdemir, Mesut Saygılı Working Paper No. 10/15, August 2010)

Effects of Monetary Unions on Inequalities
(Timur Hülagü, Devrim Ikizler Working Paper No. 10/14, August 2010)

Understanding Sectoral Growth Cycles and the Impact of Monetary Policy in the Turkish Manufacturing Industry
(Saygın Şahinöz, Evren Erdoğan Coşar Working Paper No. 10/13, July 2010)

Türkiye İçin Yeni Reel Efektif Döviz Kuru Endeksleri
(Hülya Saygılı, Mesut Saygılı, Gökhan Yılmaz Çalışma Tebliğ No. 10/12, Temmuz 2010)

Türkiye'de Piyasa Göstergelerinden Para Politikası Beklentilerinin Ölçülmesi
(Harun Alp, Hakan Kara, Gürsu Keleş, Refet Gürkaynak Musa Orak Çalışma Tebliğ No. 10/11, Haziran 2010)

Organization of Innovation and Capital Markets
(Cüneyt Orman Working Paper No. 10/10, May 2010)

Welfare Gains from Disinflation in an Economy With Currency Substitution
(H. Murat Özbilgin Working Paper No. 10/09, May 2010)

Interest Rates and Real Business Cycles in Emerging Markets
(S. Tolga Tiryaki Working Paper No. 10/08, May 2010)

Employment and Output Dynamics in a Model with Social Interactions in Labor Supply
(Yusuf Soner Başkaya, Mustafa Kılınç Working Paper No. 10/07, May 2010)

Regulation and the Market for Checks
(Semih Tümen Working Paper No. 10/06, March 2010)

On the Sources of Oil Price Fluctuations
(Deren Ünalmış, İbrahim Ünalmış, D. Filiz Ünsal Working Paper No. 10/05, March 2010)

Optimal Monetary Policy under Sectoral Heterogeneity in Inflation Persistence
(Sevim Kösem Alp Working Paper No. 10/04, March 2010)