LONG–RUN GROWTH RATE IN INCOMES OF THE POOR : A CROSS-COUNTRY ANALYSIS

Dissertation submitted to Jawaharlal Nehru University
in partial fulfilment of the requirements
for the award of the degree of

MASTER OF PHILOSOPHY

PIYALI DAS

Centre for Economic Studies and Planning
School of Social Studies
JAWAHARLAL NEHRU UNIVERSITY
New Delhi 110067
2006

CENTRE FOR ECONOMIC STUDIES & PLANNING
SCHOOL OF SOCIAL SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI-110067, INDIA

CERTIFICATE

Certified that the dissertation entitled "LONG–RUN GROWTH RATE IN INCOMES OF THE POOR: A CROSS-COUNTRY ANALYSIS" submitted by Piyali Das in partial fulfillment for the award of the degree of Master of Philosophy (M.Phil.) of this University, is her original work and may be placed before the examiners for evaluation. This dissertation has not been submitted for the award of any other degree of this University or of any other University.

Pradipta Chaudhury
Prof. Pradipta Chaudhury
(Chairperson)

Subrata Guha
Dr. Subrata Guha
(Supervisor)

Phone : 91-11-26717676, 26717557 Ext. 4421/26704421 Cable : JAYENU Fax : 91-11-26717603, 26717585, 26717586

To,

Baba, Ma, Dida

and Dadubhai

Contents:

Acknowledgment

Words have always been a constraint to express my gratitude and this occasion is no exception. All that I can bring myself to state is that I am immensely grateful to my guide Dr. Subrata Guha for his constant guidance, advice and patience in guiding me throughout this one year period. I am all the more thankful to him for undertaking the painful corrections of every single mistake committed in this work and always being there as a guide in spite of hard times. Sir, once again, my heartfelt thanks to you.

For this contribution I am also and shall ever remain grateful to all my teachers through school and college who taught me and have influenced my life in ways more than one. I feel specially indebted to Ms. Pramanik at St. Agnes' Convent, Howrah, who brought "English" to my life without which I could not have written this dissertation.

The inspiration behind my research has been my beloved Dadubhai who I can only behold in my most cherished and loving memories and whose absence in my life will be regretted forever.

To Ma and Baba who devoted their lives to bring me up to this stage and have been patient and understanding in letting me undertake my research. And my wonderful family members whose constant support has been invalueable for this work.

I am thankful to Sachin, for his friendly disposition and timely help. My roommate Somdatta for enlivening my sagging spirit and helping me keep the tensions at bay. Esha for her caring ways. And Sisir for all the support and help without which I could not have completed this on time.

My thanks to all the office bearers at JNU and others at various places like the xerox centres who I could trust with my work and to many others who have been kind and helpful to me.

This acknowledgement would remain incomplete if it is not mentioned how the presence of certain inspiring ideas and the examples set by some people that have helped me survive through bad times.

Piyali Das

CHAPTER 1

Introduction

Poverty is a social concept. A construct of human society. A consequence of the human enterprise which is purely societal in nature. One of the ills that still plague the nations. The damages it can inflict are multifarious. The fight to eradicate poverty is a solemn one that binds nations, institutions and groups but in spite of such efforts poverty still persists. Therefore the weapons used to fight it are at best "ineffective". And this ineffectiveness stems from a lack of proper knowledge of the cause of poverty.

The study here, like many before it, is devoted to unearth the factors which cause this misery. It is not sufficient to demarcate factors which cause this misery. Remedies which are short-lived are like an analgesic which though reduce the pain does not do away with the disease altogether. Therefore poverty being a chronic disease demands that the remedial measures be of an enduring nature which can break the vicious cycle of poverty and better the lot of the poor. This calls for policy prescriptions which can lead to a sustained improvement in the condition of the poor.

1. I Focus of the study

Conceptually, poverty can be studied under two heads: Absolute and Relative. Absolute poverty characterizes the group of individuals or households living below a pre-defined threshold of income or consumption called "poverty line". Relative poverty refers to a situation where the income or consumption level of a group of individuals or households lies below that of another group. For example, the section of the population which has an income or consumption level below the average level of income or consumption for the overall economy is often referred to as relatively poor compared to the section having more than the average. Whereas the absolute poverty addresses the question of "how many" fall below the poverty line, the concept of relative poverty addresses the question of "by how much" the income or consumption falls below that of a reference group in the population.

Absolute poverty analysis involves the definition of a poverty line which can be based on things like income, consumption or nutritional levels. The way to go about in such an analysis is to first define a "poverty line" and then use it to check the number of individuals residing below it (in some cases also measure the deficit between the their

income or expenditure level and the poverty line). Xavier Sala-i-Martin (2002) makes use of such an analysis. In such analyses the crucial things are the concepts of poverty line and the particular index in use.

Such analyses become debatable in the light of the fact that there are more than one ways to define the poverty line and the estimates with regard to the number of poor varies and are subjective in nature. For instance the World Bank defines the poverty line on the basis of $1 a day or $2 a day, whereas in India it was originally defined on the basis of expenditure required to maintain a calorie intake of 2400 calories per capita per day for the rural poor and 2100 calories for that of the urban poor. But as has been pointed out by Prof. Utsa Patnaik (2005) where, in practice official estimates of the poverty lines have corresponded to the expenditure level necessary to purchase a progressively decreasing amount of calories.

As the concept of relative poverty calls for no such consideration as poverty line and poverty index, the subjective differences do not much arise. We know with respect to whom an individual may be considered poor but we may disagree about whether she is "poor". Poverty therefore can be looked at both from an absolute and from a relative aspect. One can analyse factors which reduce absolute poverty pulling them out from below the poverty line. Alternatively one can analyse factors which improve the welfare of the poorest sections in any society and make the development process more inclusive. The latter is the broad question that this study seeks to address.

In a related paper Nancy Birdsall and Juan Luis Londono (1997) look at the determinants of the growth rate of per capita income of the poorest quintiles for countries. In discussing the evolution of the policy stance of the World Bank with regard to poverty alleviation, they find how the World Bank approach towards poverty reduction in Latin America has proved to be insufficient. The three pillars of such a strategy were: acceleration of economic growth, provision of basic social services to raise human capital accumulation and creation of social safety nets. Their findings show that other than the growth rate of per capita income of the economy other factors human capital accumulation (mainly in the form of education) and asset position of the poor do affect the growth rate of per capita income the poorest.

That the rise in per capita income of the economy can explain to a great extent the rise in per capita income of the poorest quintile is well documented in the paper by Dollar and Kraay (2002), where they go on to provide empirical evidence of a one-to-one relationship between the two. In their paper they move a step further and ask whether the factors that explain the cross-country differences in changes in average incomes also have similar effects on the poorest fifth of the population.

The focus of this paper is similar in approach to the above two but it differs in methodology and focus. The methodological difference lies in the span of the period, which is taken to be at least twenty years for each country. In the paper by Birdsall and Londono (1997) they consider a set of 92 countries from the Deininger and Squire (1996) "high quality" dataset. In that they calculate the growth rate of per capita incomes for a five-year span. The Dollar-Kraay study also considers a five-year span for each country to study the effect of various factors. Whereas the two studies consider the effect of various factors on the welfare of poor without discriminating between short and long periods the interest here is to look at factors which can bring about a *sustained* improvement in the welfare of the poor, with the welfare being measured by the per capita growth rate in the income of the poor.

In relative poverty analysis the relative poor is at the helm of the analysis. For a particular nation the individuals are ranked as per income or consumption and then they are grouped into quintiles, deciles or percentiles. And then depending on which group is considered to be the poorest-the lowest ten percent, twenty percent, forty percent -as the poor the analysis is then aimed at that particular subdivision of the population. The focus is on the poorest quintile of a nation. In this paper the poor are those who belong to the lowest quintiles of the nations. (Though of course the standard of living of the poor of Norway may be higher than that of Sri Lanka, this is so because the per capita incomes of the nations vary). The purpose is to find out the factors which can lead to the improvement of the welfare of the poorest quintiles of nations, with welfare being measured by the per capita income of the poorest quintile.

1.II Why study only the growth rate of the poorest quintile ?

A pertinent question might now arise as to why not deal with the factors which affect the per capita growth rate of the economy as a whole since presumably per capita income of all the sections of the population rise with the level of development of the economy as a whole. There can be more than one answer to this question. In the first place, though growth rate of the economy is a major determinant, it cannot explain cross-country differences completely. A closer scrutiny of the per capita growth rate of the poorest quintile reveals that the growth rate depends on the per capita growth rate of the economy and the share of income accruing to the poorest quintile. Hence attention must be paid to factors that affect the growth rate of the economy and those that affect inequality of the economy, which is partly reflected in the share of the income accruing to the poorest quintile[1]. Clearly therefore restricting attention to the overall growth rate of the economy might not reveal the whole picture. This is borne out by the following set of ordinary least square regressions of the growth rate of per capita income in each quintile on the growth rate of per capita income in the remaining four quintiles. [See Table1] In this table *rpcyi* (i = 1,2,...,5,) denotes the rate of growth of per capita income of the ith quintile and *rpcyni* (i = 1,2,...,5,) denotes the rate of growth of per capita income of the remaining quintiles except the ith quintile.

The variable *rpcyi* and *rpcyni* are calculated as follows. Firstly, the per capita income of the ith quintile (pcyi) and the same for the rest of the 80 percent of the population (pcyni) is calculated as

$$pcyi = (qi)\ Y/\ 0.2\ P = 5.qi.(Y/P) = 5.qi.pcy$$

$$pcyni = (qni)\ Y/\ 0.8\ P = 1.25.qni.pcy$$

where, pcyi is the per capita income of the ith quintile, qi is the percentage share of total income accruing to the ith quintile, Y is the overall income or GDP of the economy, P is the total population of the economy and pcy is the per capita income of the economy as a whole. The variable qni is the percentage share of total income accruing to the whole population except the ith quintile. Note that the first quintile (q1) is referred as the poorest quintile. Similarly, pcyni is the per capita income of the whole population except the ith quintile.

[1] Analysis of the two concepts cannot be carried out separately as changes in inequality affect growth and vice versa

After calculating pcyi and pcyni, the growth rate of per capita income of the ith quintile (rpcyi) and the same for the rest of the 80 percent of the population (rpcyni) is calculated as follows

$$rpcyi = [\{(pcyi \text{ at } t_1) / (pcyi \text{ at } t_0)\} \wedge \{1/n\}] - 1$$

$$rpcyni = [\{(pcyni \text{ at } t_1) / (pcyni \text{ at } t_0)\} \wedge \{1/n\}] - 1,$$

where t_1 and t_0 are the final and initial years of the specified time period and 'n' is the duration of the time period under consideration, i.e., $n = t_1 - t_0$. rpcyni (the growth rate of per capita income of the whole economy except the ith quintile) is similarly calculated for the remaining eighty percent of the population.

Ideally the regressor in each equation should have been the per capita growth rate of the economy (*rpcy*) as a whole. However, *rpcyi* (i = 1, 2,..., 5,) is a component of *rpcy,* this could lead to problems of endogeneity. Also the growth rate of per capita income of the remaining 80 percent is a good indicator of the per capita growth rate of the economy as a whole.

The inequality dataset used to calculate the above growth rates has been taken from the Deininger and Squire and the WIID (World Income Inequality Dataset) from WIDER. As a measure of inequality the share of income accruing to the quintiles is picked up from the dataset. The countries and their respective twenty year time periods have been selected depending on certain criteria regarding unit of analysis, income definition, area and population coverage of the survey. In so doing a trade-off between the number of countries and the time period was encountered. This led to a dataset of 36 countries having the desirable properties.

Note (refer Table 1), first of all, that the *rpcyni* is a significant factor explaining growth rate in each quintile. With reference to Table 1, in all the regressions 1 to 5, *rpcyni* (i= 1, 2, ..., 5) is significant at 99 percent. Secondly, it is observed that there is an almost one-for-one relationship between *rpcyni* and *rpcyi*. In all the regression the coefficient is close to 1. Finally, it is observed that the degree to which the growth rate of per capita income of the economy explains that of the various quintiles, varies across the quintiles. For instance the values of R- square is 0.52, 0.82, 0.9, 0.94 and 0.77 for the first, second, third, fourth and fifth quintiles respectively. Therefore, the power of *rpcyn1*

to explain *rpcy1* is significantly less than that for the remaining quintiles. The graphs of *rpcyni* and *rpcyi* (see figure 1 to 5) clearly reveals that, except for the first quintile, the growth rate of per capita income of the remaining eighty percent has an almost one-to–one relation with that of the other quintiles.

The above result is graphically depicted in the figures 1 to 5, where *rpcyni* has been taken along the X-axis and *rpcyi* is taken along the Y-axis. The scatter plot in addition with a trendline fitted to the scatter reveal that there exist an almost linear relation between *rpcyni* and *rpcyi* for all the quintiles except the first one.

These preliminary findings provides for the justification for research on the contributory factors of the per capita growth rate of the poor. The objective in this paper is to explore the factors other than the per capita income of the economy, which has explanatory power for the growth rate of per capita income of the poorest quintile.

1.III Plan of the dissertation

The lay out of the dissertation is as follows. Chapter 2 discusses the relevant literature, methodology and the dataset used in the analysis. It consists of three subsections: Section 2.I deals in relevant literatures, section 2.II provides a briefing on related papers, section 2.III explains methodology and finally section 2.IV outlines the dataset. The empirical results and the interpretations are provided in chapter 3 and finally chapter 4 concludes.

Table 1

Linear regression results of *rpcyi* on *rpcyni*, i = 1, 2, ..., 5.

Dependent Variables →	1 rpcy1	2 rpcy2	3 rpcy3	4 rpcy4	5 rpcy5
rpcyn1	0.9824# . [6.84]	.	.	.	
rpcyn2	.	1.0356# . [13.77]	.	.	
rpcyn3	.	.	1.0307# . [23.91]	.	
rpcyn4	.	.	.	1.0307# . [23.2]	
rpcyn5	.	.	.	.	0.8217# [12.34]
Const	-0.0027 [-0.21]	-0.0027 [-0.42]	-0.0015 [-0.39]	-0.0019 [-0.51]	0.0133+ [2.26]
R2	0.5246	0.8163	0.9018	0.9395	0.7758
Obs	35	35	35	35	35

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity. #, + and * implies significant at 99%, 95% and 90% respectively.
Source: Authors calculations.

Figure 1

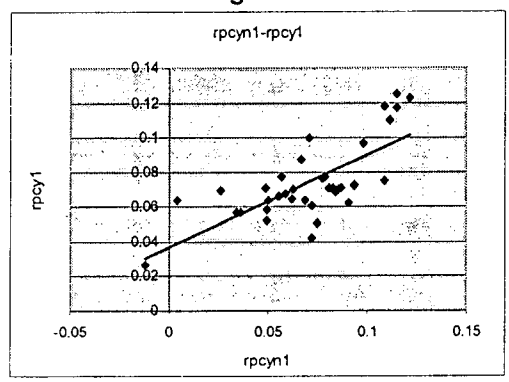


Figure 2

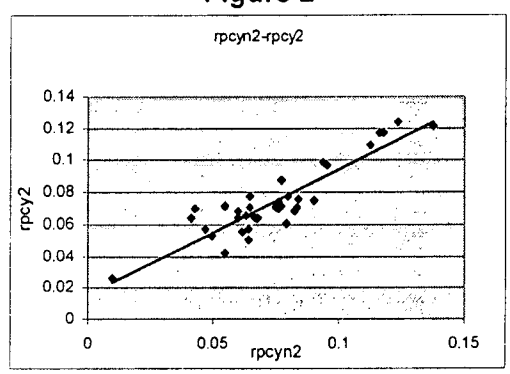


Figure 3

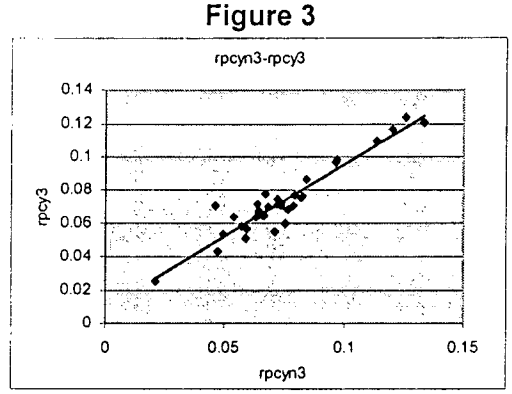


Figure 4

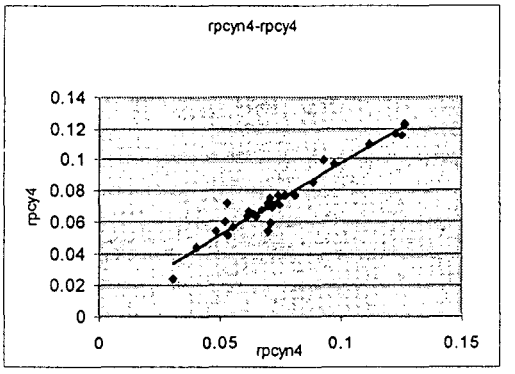
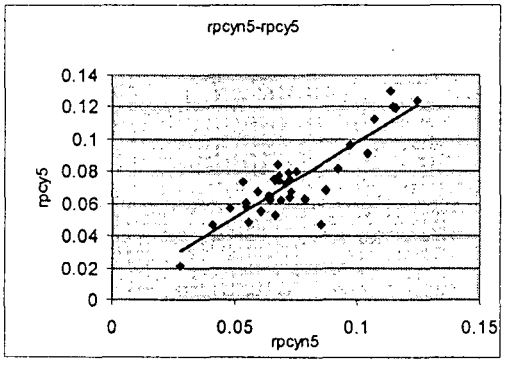


Figure 5



CHAPTER 2

2.I A brief summary of the relevant literature

In chapter 1 it was seen that the overall growth rate of the economy is a significant factor affecting the growth rate of per capita income of the poorest sections of the population. It was due to this fact that the World Development Report (1990) on poverty, recommended a new strategy emphasizing on the "acceleration of economic growth". However the fact is that even though the economic growth is a significant determinant of the growth rate of poor, it tends to capture a relatively small part of the growth rate of the poor. It is observed from the preliminary exercise in chapter 1 that the economy's growth rate explains a relatively greater part of the growth rate of the non-poor. Now the question that comes to mind next is what other factors can explain the growth rate of the poor, other than the growth rate of the economy?

The importance of human capital accumulation in steering the economic growth process has been well documented in the economic growth literature. And health and education are an integral part of human capital. Much as it is important for the non-poor, the importance of these two are monumental when it comes to the issue of the welfare of the poor. "Health" for a poor is the fuel to labour and "education" provides the skill to direct his labour power towards more rewarding ends.

That Health is an important factor in generating income is evident if we consider the growth experience of the east Asian economies of Singapore, Thailand, South Korea and Taiwan. As Sen and Dreze (1995) has pointed out, "These countries (high growth east Asian countries) have typically had much better levels of health conditions even before their period of rapid economic growth. Greater provision of medical facilities in the east-Asian countries, particularly in terms of preventive health care, is an important factor in explaining this contrast. In the case of China, the expansion of rural health care has been one of the most remarkable achievements of the pre-reform period. It has proved to be an asset of great value in the economic reforms. This is an important issue not only for the quality of life, but also for economic performance, since morbidity and undernourishment can be serious barriers to productive work and economic performance."

Education is inevitable for growth of the poor. As education can take many forms it is important to figure out what sort education are we referring to. When the concern is the poor, the real debate then is, between primary and secondary education. The importance of basic education is well-founded if we look at the Chinese experience where primary education is in a way providing a square meal to the poor. Coming now to secondary education, it is common knowledge that an added qualification can add to existing income. Therefore secondary education can help raise income. To raise the income of the poor basic education may be necessary but not sufficient. The sufficiency condition could be met by secondary education.

To gauge the influence that education can have on the growth process, a case in point is the East Asian miracle. The contribution of primary education and state action has been considerable in these countries where modern industries have flourished. An effective public policy goal had been to ensure that the majority of the young population had the capability to read, write, communicate and interact in a way that is essential for modern industrial production, Sen and Dreze (1995, pp.38-9). To quote the authors, "The economies of South Korea and China which have succeeded in flooding the world market with goods the making of which requires no great university training, but is helped widespread basic education that enables people to follow precise guidelines and maintain standards of quality. It may be much less glamorous to make simple pocketknives and reliable alarm clocks than to design state-of-the-art computer programmes, but the former gives the Chinese poor a source of income that the latter does not provide. It is in the making of these unglamorous products that, the market for which is very large across the world, that a high level of basic education is a major asset for China- and for many other high growth economies of East and Southeast Asia. The social opportunities offered by market-based economic growth are severely limited in a society in which very large numbers cannot read or write or count, cannot follow printed or hand written instructions, cannot operate comfortably in a modern industry, and so on. Inequality in basic education thus translates into inefficiency as well as further inequality in the use of new economic opportunities". Thus even though opening up of the economy may throw open a slew of market opportunities the poor may remain oblivious to them due to lack of necessary education.

Asset position of the poor can be an important factor holding back the growth rate of income. Credit plays an important role in income generation. And asset position of the borrower plays a significant role in the accessibility of credit. Generally land as an asset is used as a collateral for availing credit. Therefore, the growth rate of income of the poor would greatly depend on the initial asset position, more specifically on the land holding. This brings us to another related issue of "land reforms". As has been pointed out by Sen and Dreze (1995) extensive land reforms has been at the root of the growth process of some East-Asian countries. Therefore it can be said that institutional changes like "land reforms" if brought about would favour the poor and give them an asset, which in turn could help in raising their income level.

The importance of the asset position of the poor has also been discussed by Deininger and Squire (1997). To quote the authors, "While the results confirm a negative link between initial income inequality and subsequent growth", they suggest that this relationship is not very strong. "By contrast, initial inequality of assets, as measured by the distribution of land, exerts a significant negative effect on subsequent growth." The mechanism through which this effect of asset inequality can wreak havoc on the income generation possibility of the poor is mediated through the financial channel. In the absence of a decent asset position an individual is denied access to credit. This would keep him from making the necessary investments and thereby lead to retarded growth. In this article they look at the effects that income inequality may have on economic growth of the economy. As the contention paper is the growth rate of the poorest quintile of a nation, apparently, an objection is not ill-founded.

Sen and Dreze (1995, p.41), also harp on the same issue of asset position in discussing the importance of land reforms programme in the growth process of the East Asian countries. Reference is made of the extensive land reform carried out in Japan, South Korea, Taiwan and China. To quote the authors, "The advantages of the abolition of landlordism from the point of view of equity are obvious enough, but it has also much to contribute to the general incentive to expand production and to making it easier for agricultural producers to respond to the opportunities offered by a freer market". They add, "Significantly enough, one of the least successful growth performers among the East-Asian economies, namely, the Philippines, is also an example of an extensive failure to carry out adequate land reform".

The empirical evidence of the importance of asset inequality has been presented by Li, Squire and Zou (1998). For a set of 49 countries using the Deininger and Squire (1996) dataset on income inequality they show that income inequality (measured by Gini) is significantly affected by asset inequality.

The poor generally are the people who cannot take care of themselves. Therefore they need the help of an agency. No agency can do better than the government. The success of East Asia is to a great extent thanks to the effective role of government in the provision of basic health and education needs. Therefore government expenditure on health and education would definitely play an important role in raising the growth rate of the poor. The role of government cannot remain restricted to health and education when we are talking about the growth rate of income of the poor. The case in point is one of the various government aided employment generation programmes. These programmes help generate employment among the poor and less developed, which effectively generates income for the poor. For example, in India there are various govt. run poverty alleviation programmes like the SGSY, SGRY, SJSRY and the recent National Rural Employment Guarantee Act to be implemented from 2006.

Dreze and Sen (1995) refer to the role of the State in bringing about the desired growth progress while discussing the growth process of the four tiger economies of South-East Asia. They add, "Besides state action in basic education, preventive health care and land reform in understanding and interpreting the 'economic miracle' in East Asia, these roles of public action have to be viewed along with the part played by governments in directly promoting industrial expansion and export orientation and in guiding the pattern of industrialization ".

"A person is as much alive as he can communicate", this quotation by L Ron Hubbard is probably an all time truth. Being able to communicate, voice one's opinions and wants have been very important to acquire bargaining power. A one word which confers the ability to communicate and aquire bargaining power is "voice". Though this factor to be discussed is more political than economic, still it deserves some attention because this would ensure that the poor has access to his needs and hence his survival. This brings to the fore yet another issue of "gender inequality".

Gender inequality is a major hindrance when it comes to the issue of growth It cannot be called "growth of poor" if the pro-poor growth affects only the male poor. The growth process gets stunted in a society where females are treated unequally and prevented from participation in the work force. Such practices aggravate the impoverishments of the poor. Revisiting the growth experience of the East Asian economies shows how they could reduce the gender gap and forge ahead. Without much imagination it can be realized that removal of the social barriers to participation of women in economic activities can also have some positive influence on the growth rate of income of the poorest quintile. As Sen and Dreze (1995 , p.40) put it, "The high performing Asian economies have been able to reduce the gender gap in basic education much more rapidly than had happened anywhere else. By doing this they could reduce the relative disadvantage of women in various social opportunities including economic participation".

Political rights and rule of law play crucial roles for the poor. These help the poor voice their needs and prevent them from exploitation in the hands of the privileged. Civil liberties have been found to be a significant determinant of income inequality, Li, Squire and Zou (1998). This finding is acceptable in the light of the fact that the political clout of the rich helps them to bend policies in their favour which in turn helps tilt the income share in their favour and thereby perpetrates income inequality. Besides civil liberties, (which also includes political rights, Li, Squire and Zou (1997)) is found to significantly affect the income of the poor.

2. II Related papers

Two papers that are closely related to this study are one by Birdsall and Londono (1997), and the other by Dollar and Kraay (2002). Both these papers undertake a cross-country analysis of the factors that affect the growth rate[2] of per capita income of the poorest quintile. Birdsall and Londono undertake a linear regression analysis to arrive at the factors responsible for the per capita growth rate of the poorest. The second paper deals with panel data to analyse the relevant factors.

[2] Dollar and Kraay (2002) looks at the change in per capita incomes of the poorest quintile and not actually growth rate.

From the "high-quality" Deininger and Squire dataset (1996), Birdsall and Londono (1997) select those countries for which Lorenz curves are available for two periods of time separated by at least five years with the income estimates per capita in international purchasing power prices. For a sample of 52 countries they collect information on physical capital investment, the education of the labour force, land distribution and trade indicators.

With the help of cross-country growth regression they arrive at the following factors that stand out to be statistically significant: capital accumulation, initial education, initial income inequality, initial asset position, educational inequality (acting as a measure of the distribution of human capital), change in income gini. Their findings reveal that the income growth of the poor depends heavily on overall capital accumulation, which can be taken as proxy for overall economic growth. Besides the initial inequalities in the distribution of land and of human capital is seen to have a negative effect that is almost twice as great for this group as for the population as a whole. Therefore the important factors that they find can affect the income of the poorest are capital accumulation, initial inequalities in the distribution of land and education. Hence the initial asset position of the poor both in terms of physical assets (land) and human capital assets play important roles.

Turning now to the paper by Dollar and Kraay (2002). In this paper they test for the hypothesis: whether the income *level* of the poor has a strong relation with the overall income *level* of the economy. For the analysis they draw their dataset on income distribution from four different sources. Their primary source is the UN-WIDER World Income Inequality Database which is a substantial extension of the income distribution dataset constructed by Deininger and Squire (1996). A total of 706 of their country-year observations are obtained from this source. In addition they obtain 97 observations from the "high quality" data sample of the Deininger and Squire (1996) dataset that do not appear in the UN-WIDER dataset. Their third dataset is taken from Chen and Ravallion (2000) who construct measures of income distribution and poverty from 265 household surveys in 83 developing countries. They obtain an additional 118 observations only. Finally they augment the inequality dataset with 32 observations primarily from developed countries not appearing in the above three sources, that are reported in

Lundberg and Squire (2000). This results in an overall sample of 953 observations covering 137 countries over the period 1950-1999.

As the dataset so formed was highly unbalanced and had irregularly spaced panel of observations they had to filter the dataset. For each country they pick observations for two years subject to the constraint that at least five years separate the observations. In the course of the econometric analysis the sample gets restricted to 92 countries.

To test their hypothesis they regress the logarithm of per capita income of the poor on the logarithm of average per capita income and a set of additional control variables. Their coefficients of interest are that of the logarithm of average per capita income and those of the control variable which are determinants of the income of the poor other than per capita income of the country.

They finally arrive at the conclusion that there is a one-to-one relationship between growth of income of the poor and the growth of mean income. Such a finding is suggestive of the fact that a range of policies and institutions that are associated with higher growth will also benefit the poor proportionately. However, they do not rule out the fact that growth from different sources can have differential impact on the poor. Therefore they next take up a number of the policies and institutions that have been identified as pro-growth in the empirical growth literature, and examine whether there is any evidence that any of these variables has disproportionate effects on the poorest quintile. The five indicators that they focus on are: inflation, which Fischer (1993) finds to be bad for growth ; government consumption, which Easterly and Rebelo (1993) find to be bad for growth; exports and imports relative to GDP, which Frenkel and Romer (1999) find to be good for growth; a measure of financial development, which Levine Loayza and Beck (2000) have shown to have important causal effects on growth; and a measure of the strength of property rights or rule of law. The importance of property rights for growth has been established by, among others, Knack and Keefer (1995).

They then run a regression of the log of income of the poor on these variables including the log of average income. That way whatever effect these variables may have on the income of the poor, which is mediated through the average growth channel is

represented by the growth rate itself and hence the coefficient would represent the effect of these variables mediated through the distribution channel. The regression results of each of these factors with that of the overall growth rate as the regressors reveal the following results. In the case of trade volumes they find a small, negative and statistically insignificant effect on the income share of the bottom quintile; the same is true for government consumption as a share of GDP and inflation, where higher values of both are associated with lower income share of the poorest quintile, although again insignificantly. The point estimates of the coefficients of the measure of financial development and that of rule of law indicate that both of these variables are associated with higher income shares in the poorest quintile, but again insignificantly. When they include all the five regressors together the coefficients on each are similar to those in the simpler regressions. However, government consumption as a share of GDP now has an significant effect on the income share of the poorest with a negative sign. In addition, inflation continues to have a negative effect, which just falls short of significance at the 10 percent level. Dollar and Kraay's results suggest that the growth from different sources does not appear to have a significant differential impact on growth in incomes of the poorest.

Finally, they consider a number of other factors that may have direct effects on incomes of the poor through their effect on income distribution. They consider four such variables: primary educational attainment, social spending, agricultural productivity and formal democratic institutions. They find that while the years of primary education and relative productivity in agriculture both enter positively, neither is significant at conventional levels. In the regression with social spending they also include the overall government consumption in order to capture both the level and compositional effects of public spending. Overall government spending remains negatively associated with incomes of the poor and the share of this spending devoted to health and education does not enter significantly. This they reason out as not surprising because in many developing countries, these social expenditures often benefit the middle class and the rich primarily and the simple share of public spending on the social sectors is not a good measure of whether government policy and spending is particularly pro-poor. Finally they find the measure of formal democratic institutions to enter positively and significantly (at the 10% level). However this result is not very robust. In their large sample of developed and developing countries measures of formal democratic

institutions tend to be significantly correlated with other aspects of institutional quality, especially the rule of law index considered earlier. When they include the other determinants of growth in the regression, the coefficient on the index of democratic institutions is no longer significant.

In the paper by Birdsall and Londono (1997) primarily look at the determinants of the growth rate of income of the poorest quintile. Their study is motivated by the finding that even though the Latin American and Caribbean countries embraced the World Bank prescriptions for growth and poverty reductions, these countries have failed to reach the desired ends, so that clearly ones needs to probe into the other factors which could fill the lack and help these countries grow and reduce poverty. With that end in mind and with a time period spanning at least five years for each country, they analyse by regressing the growth rate of income of the poorest quintile on various factors. Their findings reveal that capital accumulation, educational inequality, change in income gini and initial asset position significantly affect the growth rate of the poor. As capital accumulation is a robust determinant of economic growth, their result conforms to the empirical finding that the overall growth rate per capita of the economy is a significant factor affecting the growth rate of the poorest quintile. An important result that they arrive at is that of the initial asset position significantly affecting the growth rate of income of the poor.

The paper by Dollar and Kraay (2002) studies the determinants of the per capita income of the poorest quintiles. Using a five year time period for their analysis they arrive at the fact that the change in per capita income of an economy has a one-to-one relationship with the change in per capita income of the poor. Next they test for other factors like inflation, government consumption etc. which have come up in the pro-growth literature, to see whether the sources of growth have a differential effect on the growth of the poor. They arrive at the conclusion that these factors affect the poor much the same way as they affect everyone else in the economy. They next seek to analyse factors like primary educational attainment, social spending etc., which can have direct effect on the growth rate of the poor, not mediated through the overall growth channel. This analysis fails to come up with any significant factors which are mediated through the distribution channel.

This dissertation is similar in spirit to the above although differs in methodology and the question it poses. The main point of difference with the above mentioned papers is the time period of analysis which was at least five years for the two above and is at least twenty years in this paper. The question posed here is to see if there exist factors which can directly and significantly affect the growth of the poor and to analyse the factors affecting growth in incomes of the poorest where the influence is not channelised through overall growth for the economy as a whole.

2.III Methodology

In this study the "poor" are defined as those who belong to the poorest twenty percent of the population after arranging the whole population either in ascending or descending order of income. The convention followed here is that the first quintile of the population refers to the poorest and the fifth to the richest. The objective of this research is to analyse the determinants of the growth rate of per capita income in the poorest quintile of a nation. The analysis is carried out by running a cross-country regression of the growth rate of per capita income in the poorest quintile on various factors, which are commonly cited in the pro-poor growth literature and are known to significantly affect the growth rate of the poor.

To proceed, the first step is to calculate the growth rate of per capita income of the poor. Using the income distribution data the per capita income of the poorest quintile is calculated as follows,

$$pcy1 = (q1.Y)/\ 0.2\ P = 5.q1.(Y/P) = 5.q1.pcy \qquad (3)$$

where, pcy1 is the per capita income of the poorest quintile, q1 is the percentage share of total income accruing to the poorest quintile, Y is the overall income or GDP of the economy, P is the total population of the economy and pcy is the per capita income of the economy as a whole.

The next step is to find out the relevant average annual growth rates of per capita income. Due to paucity of income distribution data instead of taking the trend growth rate, the average annual compound growth rate is taken. Therefore to calculate

the average annual per capita growth rate of the poor the per capita income of the poor at the beginning and at the end of a certain pre-specified time period is considered. The equation below shows the formula applied to find out the growth rate. To maintain compatibility the growth rate of the economy, which is to be represented by the growth rate of per capita income of the remaining eighty percent of the economy, is calculated in the same manner.

$$rpcy1 = [\{(pcy1 \text{ at } t_1) / (pcy1 \text{ at } t_0)\} \wedge \{1/n\}] - 1 \quad (4)$$

$$rpcyn1 = [\{(pcyn1 \text{at } t_1) / (pcyn1 \text{at } t_0)\}^{\{1/n\}}] - 1 \quad (5)$$

where, rpcy1 is the average annual growth rate of per capita income of first quintile for a period spanning 'n' years, rpcyn1 is the average annual growth rate of per capita income of the remaining eighty percent of the population in the economy for the same period, t_1 and t_0 are the final and initial years of the specified time period and 'n' is the duration of the time period under consideration.

Unlike previous studies by Dollar and Kraay (2002), Birdsall and Londono (1997), where the time period is taken to be five years, a longer period is considered. In this paper the interest has been to probe into the factors that can contribute towards a sustained improvement in the welfare of the poor. Therefore the attention is on the long run growth of per capita income of the poor. Hence a time period of at least twenty years is considered.

Variables used in the analysis

The next step in the analysis is to select the variables to represent factors directly affecting the growth rate of per capita income of the poor.

(i) Health

Life expectancy and *infant mortality rate* are the two health indicators considered in this study. Besides these two indicators there are other indicators too which have been used in empirical studies involving health. For instance Easterly (1999) used indicators like hospital beds per capita, physicians per capita, nurses per capita , access to sanitation, access to safe water etc. to represent health. These indicators are

input variables whereas *infant mortality rate* and *life expectancy* are outcomes or output variables. Due to the shortage of income distribution data, the number of regressors which could be used at a time in this study is fairly small, therefore the output variables are chosen over the input variables as they are more comprehensive than the input variables. Besides, keeping in mind the fact that the focus of this study are the poor and that the provision of health services in a country are not always well-targeted towards the poorer sections, it seems more appropriate to work with the outcome variables.

(ii) Education

As has been pointed out earlier in this chapter, education, both primary and secondary are essential factors to alleviate poverty. A choice problem was encountered when deciding on the education variables. The choice is between the *enrolment rate* and that of the *average years of schooling in the population.* The *enrolment rate* did not seem to be a good factor reflecting education due to the fact that mere enrolment in a school does not necessarily mean attainment of education. Hence the variable representing education is taken to be the *average years of primary education in the population* and *average years of secondary schooling in the population.* Both these variables have been taken to check which one of the two turns out to be more significant in the determination of the rate of growth of per capita income of the poor.

The effect of education on the growth rate of the poor was checked by considering the effect of initial levels of primary education represented by *primaryyrs* and that of secondary education represented by *secondaryyrs.* As the regressand in the analysis is growth rate, the effect of *change* in the variables instead of the effect of initial *level* of the variables is observed. Therefore the effect of changes in the years of education, for both primary and secondary over the relevant time period, on the growth rate of the poor are considered. Accordingly the variables *cpy* and *csy* are created to represent the *change* in primary years of education and *change* in the secondary years of education, respectively. *cpy* is calculated by taking the difference between the average years of primary education at the final year of the time period and that at the initial year of the period and *csy* is calculated similarly. A point of caveat about use of *changes* is that the causality, may not run from *csy* or *cpy* to the growth rate. For instance it could very well be the case that due to rise in growth rate of the poor, their education level is rising, so that then there is a case of reverse causality. To tackle such issues *changes* in primary

and secondary years of education, over the initial five and ten years of the relevant time period is considered. The idea being that if such *changes* over the five and ten years is found to be significant then it can be said that the *change* factor bears significance over the entire period. As a matter of interest and also because the dataset contains both developed and developing countries, it could be the case that in some countries secondary education is more important than primary education. Therefore to incorporate this possibility in the analysis a variable named *psr, defined as* the ratio of primary years of education to that of the secondary years of education, at the initial year of the time period is used in the analysis.

(iii) Asset Inequality

As cited in the relevant literature, Deininger and Squire (1996), Li, Squire and Zou (1998), Birdsall and Londono (1997), the asset position of the poor is represented by land as an asset. And to represent the asset inequality the common variable in use is the Gini coefficient of land distribution. Due to unavailability of the land GINI data[3] a variable called *resgini* was created to substitute for the lack of the data. The idea behind *resgini* is as follows. The total income of an individual consists of his wage income and his asset income and the income gini coefficient is a fall out of this additive income. So it might be the case that if the effect of wage income inequality is taken out from the income gini then what remains could be an approximation of the asset inequality[4].

The University of Texas Inequality Project (UTIP) has produced an alternative global inequality data set, based on the Industrial Statistics database published annually by the United Nations Industrial Development Organization (UNIDO). The UNIDO Industrial Statistics Database provides data for 29 industries in the manufacturing sector. It provides information on the number of establishments, employment, wages and salaries, output, value added, gross fixed capital formation, number of female employees and production indexes.

Using the UNIDO Industrial Statistics Galbraith and Kum (2004) constructs an index of manufacturing pay inequality within each country which they call UTIP-UNIDO .

[3] The dataset used by Deininger and Squire (1998) has been constructed by them from various data sources is not publicly available. Also Li, Squire and Zou (1998) mention the LandGINI data used by them to be a personal dataset of Deininger and Squire.

[4] Share of labour income in the whole income will also matter, but due to lack of appropriate data the analysis could not be undertaken.

Thus UTIP-UNIDO is a set of measures of the dispersion of pay across industrial categories in the manufacturing sector.

Using the above information the income gini from the Deininger and Squire dataset is regressed on the UTIP-UNIDO index of pay inequality and the residuals are collected from the regression. These residuals are named "resgini" which forms the asset inequality indicator representing the asset inequality variable in this analysis.

(iv) Government expenditure

To access the importance of the government expenditure on the social sector, data on the aggregate percentage of government spending on health, education and social security and welfare was collected for each of the 36 countries. For each country's time period the average of the middle most three years and in some cases two years (due to the unavailability of the data the) is calculated which finally forms the *govex1* variable. Thus *govex1* in this study represents the percentage of government expenditure targeted towards the social sector.

From the government expenditure data, the percentage of government expenditure on the *social security and welfare* is considered. The variable *sw* is calculated by taking the average of the percentage of government expenditure directed towards this particular sector. This variable is relevant to this analysis because it captures the effect of the government expenditure targeted specially towards the poor.

(v) Gender Inequality

In the case of gender inequality the analysis focuses on the education based gender inequality. This is so because of the fact that education is inevitable in empowering women and men alike, therefore if the access to it is denied then it would inevitably lead to a gender based inequality among the masses. For this the data on the ratio of female to male attainment of schooling in the population is considered. Specifically, the ratio for *no-schooling, primary schooling and secondary schooling*– represented by *gu, gp, gs* respectively, is taken to analyse the effect of gender inequality on the growth rate of income of the poor. Thus a higher *gu* and lower *gp* and *gs* would stand for higher gender inequality. As has been discussed in the case of education, here

too the effect of *changes* of the above mentioned ratios are considered and then to check if the *change* variables are of any significance the *change* over the initial five and ten years is considered.

(vi) Income Inequality

As has been cited in the literature income inequality can have a significant effect on the growth rate of income of the poor. Perotti (1996) and Barro (1999) find evidence of a negative effect of income inequality on growth whereas Forbes (2000) and Li and Zou (1998) both find positive effects of income inequality on growth. Birdsall and Londono (1997) have also considered the effect of this factor on the incomes of the poor and they have also found it to be significant. Ideally a variable representing *change* in the income inequality would be a good regressor but then the problem would be that there can arise a problem of multicollinearity between the variable *rpcyn1* and the variable representing *change* in income inequality, when both these variables are used as reggressors. Therefore to avoid this problem the initial *level* of income inequality is considered. The variables used to represent initial income inequality are *q1* and *rpnp*. The variable *q1* represents the percentage share of income accruing to the poorest quintile and *rpnp* is the ratio of the income of the poor to that of the non-poor.

(iii) Human Rights

This factor consists of two: *voice* and *rule of law*. *Voice* is an index of formal democratic institutions and greater values of it indicate more democracy. The *rulelaw*, is an index with higher values indicating stronger rule of law. It is to be noted that *rulelaw* would better be described as part of property law. As rule of law captures the extent of justice delivered the citizens it distantly reflects how well the human rights of the people of a nation are protected by the legal system.

A detailed description of all the variables used in the analysis have been put in an appendix to Chapter 2 given at page no. 41.

The basic regression equation for the analysis is as follows,

$$(rpcy1)_i=\alpha(rpcyn1)_i+\beta X_i+u \qquad (6)$$

where, (rpcy1)i is the growth rate of per capita income of the poorest quintile in the *i*th country in the sample; $(rpcyn1)_i$ is the growth rate per capita income of the remaining eighty percentage of the population in the *i*th country in the sample; X_i represents the vector of control variables which affects rpcy1 other than rpcyn1, u_i represents the stochastic disturbance term , α and β are the elements of the vector of parameters to be estimated. The coefficients α and β, the elements of the parameter vector, are estimated by the method of ordinary least squares (OLS). As the idea is to look at the factors that directly affect the growth rate of the poor and not through ones which are mediated via the overall growth rate of the economy, rpcyn1 enters as a control variable in all the regression equations.

The term rpcyn1, is a proxy for the growth rate of the economy as a whole. This is done in order to minimize a possible endogeneity problem, between rpcy1 and rpcy, the rate of growth of per capita income.

The OLS method gives the *best* linear unbiased estimate of the coefficients, based on the assumption that Var (u) = σ^2, where σ represents some arbitrary constant, that is, the variance of the error term is constant (homoscedastic). If the error term does not have a constant variance then it is said that there is heteroscedasticity. The presence of heteroscedasticity affects the accuracy of the estimation results. Therefore after each regression analysis the test for heteroscedasticity has been done.

The statistical package STATA has been used to estimate the various regression equations. And each regression analysis has been followed by the test for heteroscedasticity. STATA by default uses the Breusch-Pagan/ Cook-Weisberg test for heteroscedasticity. A brief outline of the test is as follows. This test is designed to detect any linear form of heteroscedasticity. Consider any k variable linear regression model

$$Y_i = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + u_i ,$$

We assume the variance of the stochastic error term u_i, that is, σ_i^2 is a function of the regressors, given by

$$\sigma_i^2 = f\left(\alpha_1 + \alpha_2 Z_{2i} + \ldots + \alpha_m Z_{mi}\right).$$

Some or all of the X's can serve as Z_i's. For the sake of simplicity we assume

$$\sigma_i^2 = \alpha_1 + \alpha_2 Z_{2i} + \ldots + \alpha_m Z_{mi} .$$

Now if $\alpha_{2.} = ... = \alpha_{m.} = 0$; $\sigma_i^2 = \alpha_1$, which is a constant. Therefore to test homoscedasticity one can test the hypothesis that $\alpha_{2.} = ... = \alpha_{m.} = 0$. This is the basic idea behind the Breusch-Pagan/ Cook-Weisberg test. The actual procedure is to estimate the multiple linear regression by OLS and obtain the residuals $\hat{u}_1, \hat{u}_2, ..., \hat{u}_n$. This done compute the maximum likelihood estimator of σ^2, which is given by $\sum \frac{\hat{u}_i^2}{n} = \tilde{\sigma}^2$.

Construct the variable $p_i = \frac{\hat{u}_i^2}{\tilde{\sigma}^2}$ and regress p_i on the Z's. Obtain the ESS (error sum of square) and define $\Theta = \frac{1}{2} ESS$. Assuming u_i s are normally distributed, one can show that if there is homoscedasticity and if the sample size n increases indefinitely then Θ follows χ^2 distribution with (m-1) degrees of freedom. Therefore, if the computed Θ exceeds the critical χ^2 value at the chosen level of significance one can reject the hypothesis of homoscecdasticity. To mitigate the problem of heteroscedasticity, I have estimated the regressions with robust standard errors wherever heteroscedasticity was detected.

Another problem which can adversely affect the estimation results is multicollinearity. This problem exists when the regressors in a regression equation are linearly related. A consequence of multicollinearity is that the estimators develop large variances and covariances due to which the confidence interval of the estimator tends to become much larger, leading to an acceptance of the zero null hypothesis. As multicollinearity results in higher variance of the estimator the value of the t-statistic goes up ($t = \frac{\hat{\beta}}{se(\hat{\beta})}$, where $\hat{\beta}$ is the estimated β and $se(\hat{\beta})$ is the standard error of $\hat{\beta}$) whereby the value of t goes down thereby rendering some of the coefficients insignificant. Hence it is important that a multicollinearity test is undertaken in order to avoid such undesirable consequences.

The multicollinearity test in STATA is done with the help of Variance Inflation Factor (v.i.f.). The variance of the OLS estimator for a typical regression coefficient (say, β_i) is given by,

$$\text{Var}(\beta_i) = \sigma^2 / s_{ii}(1 - R_i^2)$$

Where $s_{ii} = \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2$ and R_i^2 is the unadjusted R^2 when we regress *Xi* against all the other explanatory variables in the model, that is, against a constant, *X*2, *X*3, ..., *Xi*-1,*Xi*+1,..., *Xk*. Suppose there is no linear relation between *Xi* and the other explanatory variables in the model. Then, R_i^2 will be zero and the variance of $\hat{\beta}_i$ will be σ^2 / s_{ii}. Dividing this into the above expression for Var (β_i), we obtain the Variance Inflation Factor and Tolerance as

$$VIF(\hat{\beta}_i) = \frac{1}{1 - R_i^2} \text{ and Tolerance } (\hat{\beta}_i) = \frac{1}{VIF} = 1 - R_i^2.$$

The higher the variance of $\hat{\beta}_i$ and the greater the chance of finding β_i insignificant, which means severe multicollinearity effects are present in the model. Thus, these measures can be useful in identifying multicollinearity. In a multiple linear regression the method to detect multicollinearity is to choose each explanatory variable as the dependent variable and regress it against a constant and the remaining explanatory variables. This process would yield *k*–1 values for VIF. If any of them is high, we conclude multicollinearity is present. Unfortunately, there is no theoretical way to say what the threshold value should be to judge that VIF is high. The common practice is to take 10 as the threshold value of V.I.F. Assuming that, mulcollinearity was not found in a single regression reported in this study.

Owing to the limitations of the dataset, which is a widely known fact that income inequality data are hard to come by, and the imposed desirability conditions for the research work the dataset for this analysis got limited to a total of 36 countries. Due to this a single regression with all the variables could not be run because then the degrees of freedom would get reduced considerably which in turn could affect the accuracy of the results. With a dataset having a number of observations less than 30 the standard deviation of the t-statistic, which is the conventionally used test statistic, goes up. So the regression analysis had to be restricted to a maximum of four regressors besides the variable *rpcyn1*.

2.IV Data

2.IV. A The Income distribution dataset

Unlike the national accounts data the income distribution database is not easily available. This is because there is no set standard for the various necessary conceptual definitions. Besides the survey method varies, both internally and internationally within countries. Therefore owing to the data coming from diverse sources there is lack of uniformity in the in the basis of data collection. To make this point clear let us take an example from the dataset. The qualitative variable "unit of analysis" has two attributes - person, household. Due to the difference in the data sources for each country or across country, the unit of analysis could be either person or household. Therefore the original dataset had to cleaned up and an attempt has been made to maintain some amount of uniformity among the countries selected for analysis. This has been done under the guidance of a pre-determined preferred set of the attributes; the preferred set being taken from the dataset itself. In the following lines it is explained step wise how the set of 31 countries is arrived at.

The dataset on income distribution has been taken from two sources. The majority of the datapoints are taken from the World Income Inequality Database version 2a[5] (WIID2a) compiled by World Institute for Development Economics Research, United Nations University (UNU-WIDER) and the other being the "high quality" income distribution data by Deininger and Squire (1996). The WIID is a substantial extension of the income distribution dataset constructed by Deininger and Squire (1996). The data from WIDER is augmented by selecting datapoints from the "high quality" dataset of Deininger and Squire (1996). The income inequality dataset used in this analysis contains 72 datapoints covering 36 countries, of which the entries for 31 countries have been taken from WIDER and the rest from the Deininger and Squire (1996) dataset.

[5] The version 1.0 of WIID (World Income Inequality Dataset) was initially compiled over 1997-1999 for the UNU/WIDER-UNDP project "Rising Income Inequality and Poverty Reduction: Are They Compatible?" directed by Giovanni Andrea Cornia, the former Director of UNU/WIDER which was published in September 2000. The current update, WIID2a, is part of the UNU/WIDER project "Global Trends in Inequality and Poverty" directed by Tony Shorrocks, the Director of UNU/WIDER.

2. IV. A. I. The WIDER dataset

The World Income Inequality Database (WIID version 2a) provides income distribution data for 154 countries. The income distribution data is given in the form of gini, quintiles, deciles and percentiles data. Besides there is data on the survey methods adopted. This is provided under the headings of area coverage , population coverage , age coverage and the income definition .

The list of variables in the WIDER dataset is as follows:

- **Year**
- **Gini coefficient** in percentage points as calculated by WIDER. If deciles or quintiles were not available this will be equal to the reported Gini.
- **Reported Gini** = the Gini as reported by the source (if no Gini were reported by the source, this will include the Gini as calculated by WIDER or Deininger & Squire for the old databases using POVCAL, a program estimating the Gini coefficient using parametric extrapolation).
- **Q1-Q5, D1-D10, P5, P95** = quintile, decile, percentile group shares.
- **AreaCovr** = area coverage. The land area, which was included in the original sample surveys etc.
- **PopCovr** = population coverage. The population covered in the sample surveys in the land area (all, rural, urban etc) which was included.
- **AgeCovr** = age coverage. Age limits imposed on the sample population. This is not explicitly given e.g. for the wage earning population, which by definition excludes children and most elderly people, unless special restrictions are used in the sample.
- **IncSharU** = income sharing unit/statistical unit.
- **UofAnala** = unit of analysis, indicates whether the data has been weighted with a person or a household weight.
- **Equivsc** = equivalence scale used.
- **IncDefn** = income/expenditure definition.
- **Source1** = the source from which the observation value was obtained.
- **Survey/Source2** = if the survey underlying the estimates is known this variable includes the name of the survey, otherwise it includes the source that Source1 cites as the (primary) source.
- **Quality** = quality classification.

1 is for observations where the underlying concepts are known.

2 is for observations where the quality of *either* the income concept *or* the survey is problematic or unknown or they have not been able to verify the estimates (the sources were not available to them)

3 is for observations where both the income concept and the survey are problematic or unknown.

4 is for observations classified as memorandum items; some of the observations origin from the older compilations of inequality data have been given this rating since the data lying behind the observations often are unreliable.

The method of selection of countries

Step 1

As per the dataset if the unit of analysis is "household" it means that the size of the households and the needs of different sized households have not been taken into account. If the unit is "person" it means that the needs of different sized households have been taken into account. In the income distribution analysis it is desirable to take "person" as the unit of analysis. This is because if unit of analysis is "household" and we have two households having the same income then obviously the data would place both of these in the same group. This might be flawed because the per capita income for the households might vary considerably. Therefore a subset is taken out of WIID2a having "person" as the unit of analysis.

Step 2

The next important thing was to keep only those countries which had data for all the deciles or all the quintiles. This way there is room for undertaking comparative analysis across the quintiles. Finally when the countries had been selected on the basis of the desired criteria the deciles data was converted into quintiles.

Step 3

In the dataset after sorting for unit of analysis each country had more than two observations to choose from. Next the countries were sorted according to some criteria based on certain qualitative variables like area coverage, age coverage, population

coverage and income definition, used in the survey. Among the various types of area coverage like urban, rural, cities, metropolitan etc. the countries with only "all" as the area coverage were selected. A somewhat similar criterion was used for and age coverage population coverage where once again the criterion of "all" as the age coverage and population coverage was taken. With regard to income definition care has been taken that the data points had "income" as the income definition. In the dataset for some countries the income definition for the survey was "consumption" instead of "income". For this analysis only those countries were selected for which the income definition was "income" and *not* "consumption".

Step 4

In the dataset for each country multiple year entries were available . For each country two years were selected such that, the difference between them is at least twenty years. While doing so care is taken that the data definitions-regarding area coverage, population coverage and income definition, match as far as possible for the initial year and the final year. This is done to maintain comparability between the two selected years. Also certain restrictions have been imposed on the choice of initial and terminal years with the initial year entry allowed not earlier than 1955 and that for the terminal years the entries are all post-1980.

Step 5

Where ever possible the "quality" attribute has been tried to be maintained in the range of 1 and 2. The countries where the quality appears to be 3 or 4, it must be the case that it was not possible to get any better match after controlling for the other selection criteria on area coverage, age coverage, population coverage and income definition.

2.IV.A.II The Deininger and Squire dataset

A total of 31 countries is available from the WIDER dataset. As econometric analysis does not respond well to datasets having observations less than 30 an attempt was made to augment this dataset. This was done by taking data points from "high quality" dataset of Deininger and Squire (1996).

The "high quality" dataset sets certain quality standards for the inequality database in order to allow better cross-country comparisons. It is important that the observations should be based on household surveys, on comprehensive coverage of population and on the comprehensive coverage of income sources, [Deininger and Squire (1996)]. With regard to surveys it is important that the data on inequality be drawn from household surveys and not from the information from the national accounts or some assumption regarding a general functional form of the income distribution. The consideration with respect to the coverage of the population is also important . This is so because use of a non-representative subset of the population can easily result in biased estimates. Therefore it is required that the data on inequality, even if drawn from household surveys must be based on a representative sample covering all of the population. Finally as the authors point out, it is important that the measures of inequality income and expenditure be based on a comprehensive coverage of different income sources as well as of population groups.

The list of variables and their description as provided in the guide to the dataset are as follows :

1. Quality

The abbreviations used in the "quality" column are as follows:

accept: Included in our high quality data set.

nn: Based on a survey of less than national coverage.

cs: Estimate that was not included due to availability of an estimate from a consistent source.

ps: Estimate that was not included as there is no clear reference to the primary source.

est: Estimate based on national accounts or surveys of less than full national coverage.

wg: Estimate excluded because it was based on the income earning population only or derived from non-representative tax records.

2. Country

3. Code (3-digit country code)

4. Year

5. Gini

6. - 9. Cumulative quintile shares

10. Inc: Whether the Gini coefficient is calculated based on income or expenditure (I = Income, E Expenditure)

11. Pers: Whether the recipient unit is the person or the household

He = Household equivalent (households are weighted by the number of persons);

Pe = Person equivalent (in addition to He, the effective number of members in the household is assumed to be the square root of the actual members).

12. Gross: Whether the income reported is gross or net of taxes (G = Gross; N = Net).

13. and 14. Coverage 1 and 2

IR: Income recipients

EAP = Economically active population

15 and 16. Sources (self-explanatory)

To obtain additional countries from this dataset , the countries which are already present from the WIDER dataset are removed. Next as has been done in the case of WIDER dataset, similar criteria are applied in selecting the countries. In the first place a subset is taken out on the basis of the fact that quality equals "accept". This gives the "high quality" Deininger and Squire dataset. By doing this the data so obtained is one based on the household survey, having a comprehensive coverage of the whole population and is one for which the measurement of income or expenditure is comprehensive. Regarding the income definition the countries are so selected which has the variable "Inc" equal to "Income" instead of "Expenditure".

Following the same procedure as in the case of the countries from the WIDER dataset , the two year entries are chosen such that they are at least twenty years apart and maintaining some parity in definition of the attributes for the two years. This way a conformity was maintained with the WIDER dataset with which the present one would be augmented. The additional countries which have been picked up from the Deininger and Squire dataset are – Taiwan , Philippines, Japan, Australia, Hong Kong, Thailand.

Limitations of the dataset

Eventhough utmost care had been taken to maintain compatibility among the two year entries of each country still there remains certain drawbacks. For instance some compromise had to be made with respect to population coverage mainly due to non-availability of any other satisfactory pair of data. This happened in the case of

Netherlands, Sri Lanka and Zimbabwe where the population coverage for the initial year is "Income Recipients" and it is "All" for the final year. For Panama, Peru and Venezuela the population coverage for the initial year is "Economically Active Population " and it is "All" for the final year. The entry "Economically Active Population" had been taken as the next best entry after "All" and in the cases where that was unavailable the next preferred entry had been "Income Recipients". And further wherever even that was unavailable the country was not selected for the dataset.

Similarly certain other less desirable cases had to be accommodated with regard to income definition, where it was not possible to match the income definition for the initial and the final years. In many cases due to non-availability of a matching income definition for the initial year is taken to be "Income, ..." which implies no information about the income concept and also meant that the income concept might include earnings only, monetary incomes only, or it might be net or gross of taxes has been taken. It is to be noted that for many countries the initial year data are of dubious quality but which improved during the latter years.

Some unwanted adjustments also had to made when augmenting the WIID 2a with the Deininger and Squire dataset. The assumption of unit of analysis to be "person" was relaxed to allow the inclusion of cases with "Household" as the unit of analysis. Hence with a view to maintain a dataset which can allow itself to minimum econometric analyses these undesirable cases had to be included.

Table 2 The Income Inequality Dataset

country	year	q1	q2	q3	q4	q5	areacovr	popcovr	agecovr	incsharu	uofanala	equivsc	incdefn	quality
Bangladesh	1973	0.09	0.14	0.18	0.22	0.38	All	All	All	Household	Person	Household per capita	Income, Gross	2
Bangladesh	1996	0.06	0.11	0.15	0.21	0.48	All	All	All	Household	Person	Household per capita	Income, Gross	2
Brazil	1972	0.02	0.05	0.09	0.17	0.67	All	All	All	Household	Person	No adjustment	Monetary Income	3
Brazil	1992	0.03	0.06	0.11	0.18	0.62	All	All	All	Household	Person	Household per capita	Income, Gross	2
Canada	1971	0.06	0.12	0.16	0.23	0.43	All	All	All	Family	Person	Family per capita	Monetary Income, Disposable	1
Canada	1994	0.08	0.13	0.17	0.23	0.39	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	1
Colombia	1964	0.02	0.05	0.09	0.17	0.68	All	EAP	All		Person		Income	4
Colombia	1995	0.03	0.07	0.11	0.18	0.61	All	All	All	Household	Person	Household per capita	Income, Gross	2
Costa Rica	1971	0.06	0.08	0.10	0.22	0.55	All	All	All	Family	Person	Family per capita	Income, Gross	4
Costa Rica	1996	0.04	0.09	0.13	0.21	0.53	All	All	All	Household	Person	Household per capita	Monetary Income, Gross	3
Ecuador	1968	0.06	0.10	0.16	0.23	0.44	All	All	All		Person		Income, Gross	4
Ecuador	1994	0.02	0.05	0.09	0.16	0.67	All	All	All	Household	Person	Household per capita	Income, Disposable	1
El Salvador	1969	0.04	0.09	0.14	0.23	0.51	All	All	All		Person		Income	4
El Salvador	1995	0.03	0.08	0.13	0.21	0.55	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	2
Finland	1966	0.08	0.13	0.17	0.23	0.39	All	All	All	Household	Person	Household eq, OECDM	Income, Disposable	1
Finland	1995	0.11	0.15	0.18	0.22	0.35	All	All	All	Household	Person	Household per capita	Income, Disposable	1
France	1975	0.07	0.12	0.16	0.23	0.43	All	All	All	Household	Person	Household per capita	Income, Disposable	2
France	1989	0.07	0.13	0.17	0.23	0.40	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	1
Honduras	1968	0.02	0.05	0.10	0.19	0.65	All	All	All	Household	Person	Household per capita	Income	4
Honduras	1996	0.03	0.07	0.11	0.20	0.59	All	All	All	Household	Person	Household per capita	Earnings, Net	3
Hungary	1972	0.10	0.15	0.19	0.23	0.34	All	All	All	Household	Person	Household per capita	Income, Disposable	1
Hungary	1993	0.11	0.15	0.18	0.22	0.33	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	3
Indonesia	1971	0.07	0.11	0.13	0.17	0.52	All	IR	All	Person	Person	No adjustment	Income	3
Indonesia	1993	0.06	0.10	0.14	0.20	0.49	All	All	All	Household	Person	Household per capita	Income, Gross	2
Ireland	1973	0.09	0.13	0.17	0.22	0.40	All	All	All	Household	Person	Household per capita	Income, Disposable	1
Ireland	1987	0.07	0.12	0.16	0.22	0.43	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	1
Korea, Rep	1970	0.06	0.10	0.15	0.22	0.47	All	All	All		Person		Income	4
Korea, Rep	1993	0.06	0.14	0.18	0.24	0.39	All	All	All	Household	Person	Household eq, SR	Monetary Income, Gross	1

country	year	q1	q2	q3	q4	q5	areacovr	popcovr	agecovr	incsharu	uofanala	equivsc	incdefn	quality
Madagascar	1960	0.05	0.08	0.11	0.16	0.60	All	All	All		Person		Income	4
Madagascar	1993	0.02	0.05	0.10	0.19	0.65	All	All	All	Household	Person	Household per capita	Income, Gross	2
Malaysia	1970	0.04	0.08	0.13	0.20	0.55	All	All	All	Household	Person	Household per capita	Income	4
Malaysia	1995	0.04	0.08	0.12	0.20	0.55	All	All	All	Household	Person	Household per capita	Income, Gross	1
Mexico	1963	0.04	0.08	0.12	0.21	0.56	All	All	All	Household	Person	Household per capita	Income	3
Mexico	1994	0.03	0.07	0.11	0.18	0.60	All	All	All	Household	Person	Household per capita	Income, Disposable	1
Netherlands	1967	0.04	0.10	0.15	0.22	0.49	All	IR	All		Person		Income	4
Netherlands	1991	0.08	0.13	0.17	0.22	0.40	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	1
Norway	1970	0.06	0.14	0.19	0.24	0.37	All	All	All	Household	Person	Household eq, OECDA	Income, Disposable	1
Norway	2000	0.10	0.14	0.18	0.22	0.36	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	1
Panama	1969	0.03	0.07	0.11	0.19	0.59	All	EAP	All		Person		Income	4
Panama	1995	0.02	0.06	0.11	0.20	0.60	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	2
Peru	1970	0.02	0.06	0.11	0.20	0.63	All	EAP	All		Person		Income	4
Peru	1994	0.03	0.07	0.12	0.18	0.59	All	All	All	Household	Person	Household per capita	Income, Disposable	2
South Africa	1965	0.04	0.07	0.10	0.14	0.66	All	All	All		Person		Income	4
South Africa	1997	0.02	0.05	0.10	0.21	0.62	All	All	All	Household	Person	Household per capita	Income, Gross	3
Spain	1973	0.07	0.12	0.16	0.22	0.42	All	All	All	Household	Person	Household per capita	Income, Disposable	1
Spain	1996	0.07	0.13	0.17	0.22	0.41	All	All	All	Household	Person	Household eq, OECDM	Income, Disposable	1
Sri Lanka	1963	0.04	0.08	0.13	0.20	0.55	All	IR	All	Person	Person	No adjustment	Income, Gross	3
Sri Lanka	1996	0.05	0.09	0.13	0.19	0.54	All*	All	All	Household	Person	Household per capita	Income, Gross	2
Sweden	1967	0.05	0.14	0.18	0.24	0.40	All	All	All	Family unit	Person	Family unit per capita	Monetary Income, Disposable	1
Sweden	1995	0.09	0.14	0.18	0.23	0.34	All	All	All	Family unit	Person	Family unit per capita	Monetary Income, Disposable	1
U.K.	1961	0.09	0.15	0.18	0.23	0.35	All	All	All	Household	Person	Household eq, HBAI	Income, Disposable	1
U.K.	1991	0.07	0.12	0.17	0.23	0.41	All	All	All	Household	Person	Household eq, HBAI	Income, Disposable	1
United States	1974	0.06	0.12	0.17	0.23	0.42	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	1
United States	1994	0.05	0.11	0.16	0.23	0.44	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	1
Venezuela	1971	0.03	0.06	0.10	0.17	0.65	All	EAP	All		Person		Income	4
Venezuela	1996	0.03	0.08	0.13	0.21	0.54	All	All	All	Household	Person	Household per capita	Monetary Income, Disposable	2
Zimbabwe	1968	0.03	0.05	0.08	0.14	0.69	All	IR	All		Person		Income	4
Zimbabwe	1995	0.01	0.03	0.07	0.14	0.76	All	All	All	Household	Person	Household per capita	Income, Gross	3

country	year	q1	q2	q3	q4	q5	areacovr	popcovr	agecovr	incsharu	uofanala	equivsc	incdefn	quality
Taiwan[1]	1970	0.086	0.1325	0.1706	0.2248	0.3861		All			Person		Income, Net	Accept
Taiwan[1]	1993	0.0713	0.1312	0.1765	0.2344	0.3866		All			Person		Income, Net	Accept
Philippines[1]	1957	0.0654	0.0773	0.136	0.2361	0.4852		All			Household		Income, Gross	Accept
Philippines[1]	1985	0.052	0.091	0.133	0.203	0.521		All			Household		Income, Gross	Accept
Thailand[1]	1962	0.08	0.086	0.1208	0.2154	0.4978		All			Household		Income, Gross	Accept
Thailand[1]	1992	0.037	0.076	0.116	0.186	0.585		All			Household		Income, Gross	Accept
Australia[1]	1969	0.0699	0.1317	0.1763	0.2309	0.3912		All			Household		Income, Gross	Accept
Australia[1]	1990	0.046	0.097	0.155	0.238	0.464		All			Household		Income, Gross	Accept
Hong Kong[1]	1971	0.057	0.101	0.137	0.194	0.511		All			Household		Income, Gross	Accept
Hong Kong	1991	0.0489	0.1018	0.1437	0.2119	0.4937		All			Household		Income, Gross	Accept
Japan[1]	1962	0.0575	0.1128	0.1614	0.2252	0.4431		All			Household		Income, Gross	Accept
Japan[1]	1982	0.059	0.118	0.1717	0.2331	0.4182		All			Household		Income, Gross	Accept

Note: * = excl. northern and eastern provinces.
[1] = from Deininger and Squire.
EAP = Economically Active Population.
IR = Income Recipient.
SR = square root.
OECDM = OECD mod.
OECDA = OECD adaptation.

Source: (1) World Income Inequality Distribution, Version 2.a, WIDER.
(2) Deininger and Squire, 1996. Available at http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK:20699070~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html

To check the representative power of the 36 countries selected for analysis the countries are classified as per geographical location as provided in the UN Statistical Yearbook. This is shown in the following table. The classification has been done as per the classification in the Statistical Yearbook of the United Nations.

Table 3 Regional Classification of countries.

Africa South Africa, Zimbabwe, Tunisia, Madagascar
North America Canada, Costa Rica, El Salvador, Honduras, Mexico, Panama, United States.
South America Brazil, Colombia, Ecuador, Peru, Venezuela.
Asia Bangladesh, Indonesia, Japan, Korea Republic, Malaysia, Sri Lanka, Thailand, Taiwan, Hong Kong, Philippines.
Europe Finland, France, Ireland, Hungary, Netherlands, Norway, Spain, Sweden, United Kingdom, Poland.
Oceania Australia,

Source : Statistical Yearbook, United Nations.

2. IV. B. The GDP dataset

To calculate the per capita growth rate of the poor and of the non-poor the data on real gross domestic product per capita is taken from the Penn World Tables version prepared by Aten, Heston and Summers. The necessary data on GDP per capita has been taken from the current Penn World Table version 6.1[1] produced by The Center for International Comparisons at the University of Pennsylvania. The dataset provides purchasing power parity converted GDP data for 168 countries for some or all of the years 1950-2000.

[1] This is available at http://pwt.econ.upenn.edu/

The Real GDP per capita denoted as CGDP in the dataset is the PPP converted GDP (per capita) data for all the 168 countries. In the dataset the "Purchasing power parity" is the number of currency units required to buy goods equivalent to what can be bought with one unit of the base country. That is, the PPP is the national currency value of GDP divided by the real value of GDP in international dollars. International dollar has the same purchasing power over total U.S. GDP as the U.S. dollar in a given base year (1996 in PWT 6.1). The "Real Gross Domestic Product per capita (CGDP)" are obtained from an aggregation using price parities and domestic currency expenditures for consumption, investment and government .

The older version of Penn World Table 5.6 is not used and instead the current version 6.1 is used. The reason why this is done is because of the fact that the income inequality dataset had some year entries for some countries for the late 90s. Therefore the data for such terminal years would not be available hence the use of version 6.1 instead of the older version.

2. IV. C The data sources for the other factors used in the analysis:

i. **Health** - The two health variables "infant mortality rate" and the "life expectancy" data have been taken from the international database of the United States census bureau. This dataset is available online from the following url http://www.census.gov/ipc/www/idbnew.htm. For some countries the initial year data was unavailable from this source. These missing data were then picked up from the Barro-Lee dataset (1994). This dataset provides the "infant mortality rate" and the "life expectancy" data for 138 countries during the period 1960-1985.

ii. **Education-** The education variables are taken to be *average years of primary education* and *average years of secondary education* both for the whole population over the age of 25 years. This data has been taken from the Barro-Lee dataset (1998).

iii. **Gender Inequality-** In this paper the focus is on education based gender equality so this data has been taken from the Barro-Lee dataset (1998).

iv. **Asset Inequality-** This data was compiled from a regression of the gini coefficient from the Deininger and Squire (1996) dataset on the pay inequality index prepared by the University of Texas Inequality Project. The data is obtained by taking the residuals of the above regression. The reason why this can be done is due to the fact that pay inequality and income inequality are not unrelated, Galbraith and Kum (2004). They argue that in most countries manufacturing pay is a significant component of all pay and that this pay is everywhere the largest single element in income. Therefore if the income inequality is regressed on the pay inequality then we get an estimate of how much of the income inequality is accounted for by the pay inequality. Hence the residual from such a regression would indicate the amount of income inequality not explained by the pay inequality and would serve to be a measure of asset inequality.

v. **Initial Income Inequality-** The data on this variable has been taken from the Income Inequality dataset compiled for this study , which in turn has been prepared from the WIID2a of UN/WIDER and the Deininger and Squire (1996) dataset.

vi. **Government expenditure-** The variables representing this factor are the percentage of government expenditure on health, education and social security and welfare. The data on these variables have been taken from the various issues Government Finance Statistics Yearbook, IMF.

vii. **Political variables-** For this factor the variables have taken from the dataset used by Dollar and Kraay (2002), the primary source being Kaufmann, Kraay and Zoido-Lobaton (1999).

Appendix to Chapter 2
Variable description

Variables	Description
qi	Percentage share of income of the *i*th quintile in total income of a country where i=1,2,...,5, and quintile 1 stands for the poorest 20 percent.
rpcyi	Rate of growth of per capita income of the *i*th quintile where i=1,2,...,5, and quintile 1 stands for the poorest 20 percent.
rpcyni	Rate of growth of per capita income of the economy except the *i*th quintile where i=1, 2,..., 5, and quintile 1 stands for the poorest 20 percent.
imr	Infant mortality rate, per 1000 women.
le	Life expectancy at birth in years.
primaryyrs	Average years of primary education in the total population over the age of 25 years.
secondaryyrs	Average years of secondary education in the total population over the age of 25 years.
csy	Change in the average years of secondary education in the whole population
csy5	Change in the average years of secondary education in the whole population during the initial five years.
csy10	Change in the average years of secondary education in the whole population during the initial ten years.
gu	The ratio of no schooling in the females to that of the males.
gp	The ratio of average years of primary schooling attained in the females to that of the males.

Variables	Description
gs	The ratio of average years of secondary schooling attained in the females to that of the males.
gpc	The ratio of average years of primary schooling completed in the females to that of the males.
gsc	The ratio of average years of secondary schooling completed in the females to that of the males.
cgu	Change in *gu*.
cgp	Change in *gp*.
cgs	Change in *gs*.
cgpc	Change in *gpc*.
cgsc	Change in *gsc*.
gp5	Change in *gp* during the initial five years.
gp10	Change in *gp* during the initial ten years.
govex1	Percentage of government expenditure directed towards health, education and social security and welfare.
sw	Percentage of government expenditure on social security and welfare.
resgini	Asset inequality indicator.

Variables	Description
voice	Index of formal democratic institutions, greater values indicate more democracy.
rulelaw	Rule of Law Index, higher values indicate stronger rule of law.
gini	Gini coefficient of income inequality.
rpnp	Ratio of poor to non-poor, measured as ratio of percentage share of income of the poorest quintile in total income to that of the rest of the economy.

Source: UNU-WIDER; The Deininger and Squire dataset (1996); Summers and Heston Penn Word Tables, World Bank Data, version 6.1; The international Database, United States Census Bureau; Barro- Lee dataset (1994); Barro- Lee dataset (1996); University of Texus Inequality Project; Industrial Statistics Database, UNIDO; Government Finance Statistics Yearbook, IMF; Dollar and Kraay dataset (2002) compiled from Kaufmann, Kraay and Zoido-Lobaton (1999).

CHAPTER 3

3. Regression Results

The objective of this study is to find out the factors that can affect the growth rate of per capita income of the poor. To analyse this, the average annual compound growth rate of per capita income of the poorest quintile is regressed on some factors which have come up in the literature. The factors are growth rate of the economy, health, education, asset inequality, gender inequality, political variables, government expenditure and initial income inequality. The analysis is carried out by running a simple regression of the growth rate of the poor on the chosen factors as per the following equation (refer to equation 6, chapter 2),

$$(rpcy1)_i = \alpha\,(rpcyn1)_i + \beta X_i + u_i .$$

where $(rpcy1)_i$ represents the growth rate of per capita income of the poorest quintile in the ith country, the X_i are the vector of controls that can affect the growth rate of the poor, u_i is the stochastic error term of the regression equation and $(rpcyn1)_i$ is the growth rate of per capita income of the remaining eighty percent of the population of the ith country. The term $(rpcyn1)_i$ is a proxy for the growth rate of the economy as a whole.

According to both Birdsall and Londono (1997) and Dollar and Kraay (2000) overall growth in per capita incomes in the economy is a significant determinant of the growth in per capita incomes of the poor. This study corroborates their findings. But the explanatory power of the rate of growth of per capita income of the overall economy is found to be relatively small for the rate of growth of per capita income of the poor when compared to that for the higher quintiles (see Table 1, Chapter 1). Clearly then, there must be factors other than the economic growth rate which affects the growth rate of per capita incomes of the poor (represented by the variable *rpcy1*). Therefore, in this study an attempt has been made to find out the effect of the other factors on *rpcy1*. Specifically, the idea has been to find the factors which have a direct effect on *rpcy1* and not via the channel of the overall growth rate of the economy.

This regression is conducted on a dataset consisting of 36 countries with a time period of at least 20 years under consideration for each country. So for each country the factors are also considered over a period of at least 20 years. In this way the study aims to

capture the effect of various factors on the long-run growth rate of incomes of the poor. This way the study is targeted towards the analysis of factors that stand to be responsible for a sustained improvement in the welfare of the poor.

This chapter presents the regression results. As the idea is to look at the factors not mediated through the economic growth rate channel, all the regressions are run, with *rpcyn1* entering as a common control variable. An important point to note is that the dataset has only 36 observations, due to which the maximum number of regressors which has been considered in a regression is at most five (including the constant term). This is done so as not to allow the degrees of freedom in a regression to not fall below 30.

The regression analysis is presented as follows. Section 3.1 presents the results of the regression of rpcy1 on *rpcyn1* and each of the variables taken individually. The variables are grouped as per the factors they represent and the results of these individual regressions are reported in the Tables 4 to 10. Section 3.II presents the results where *rpcy1* is regressed on *rpcyn1* and two variables which are found to be individually significant from section 3.I. Finally in section 3.III the multivariate regression results of *rpcy1* regressed on *rpcyn1* and three more regressors are presented.

3. I Factor wise regressions

To report the results of the individual regressions, the variables used in the analysis are grouped under the factors that the variables represent. This exercise is undertaken to arrive at the variables which most significantly affect the growth rate of the poor. As the dependent variable is a *growth rate*, in some cases the regression analysis has been carried out with the *change* in the level of the variable over the relevant time period. And if the variable in the form of *change* was found to be significant, then instead of the initial *level* of the variable the variable in the form of *change* was taken to be the regressor representing a particular factor. The problem attendant with such an exercise is that one cannot assertively determine the direction of causation. Therefore in such cases the changes over five years and ten years, from the initial year of the time period, is checked. If the regression result with such *changes* (i.e., *change* variables) over five years or ten years was found to be significant, it was only then that the variable in the form of a *change* over the entire time period was considered as a regressor. The line of reasoning

being that, if the *change* in a variable over the initial five or ten years is seen to affect the regressand significantly then it can be said that its *change over the entire time period* can also be taken to significantly affect the regressand. For instance in the case of education the variable *secondaryyrs* (representing initial years of secondary education) was not found to be significant whereas *csy* (representing change in the years of secondary education over the entire time period) was found to be significant. To check if this can be used as a regressor representing education the regression was run with *csyr10* (representing change in the years of secondary education over the first ten years of the time period) . As this was found to be significant (refer to Table 5) *csy* was taken to be a regressor representing education.

Health

With the above guidelines the first set of regressions was run on the health variables. The results are reported in Table 4. Both the variables *imr* and *le*, which represents the initial level of infant mortality rate and initial levels of life expectancy, are found to be individually significant at 99 percent and 98 percent, respectively, with the coefficient of *imr* entering with a negative sign and that of *le* with a positive sign. This conforms with reality that the *imr* and *rpcy1* are inversely related, whereas *le* and *rpcy1* are positively related to each other. The variable *rpcyn1* is found to be highly significant in both the regressions. Therefore it turns out to be true that initial level of health plays a significant role in the growth rate of the poorest quintile.

Education

The education variables are: *primaryyrs, secondaryyrs, cpy, csy, csyr5, csyr10.* The first two represent the initial level of average years of primary schooling and secondary schooling, respectively, in the total population. From Table 5 none of the two are found to be significant.

Next the regressions are run on the *changes* in the average years of primary and secondary education over the entire time period, represented by the variables *cpy* and *csy* respectively. Of these only *csy* is found to be significant at precisely 96 percent.

However keeping in mind the fact that by considering *csy* the effect of change over the entire time period is being considered and that this may not help establish an appropriate direction of causation, the same variable with a lessened time span of five and ten years from the initial year is considered. These variables named *csyr5* and *csyr10* are then used as regressors to check the effect of these on *rpcy1. csyr10* turns out to be significant precisely at 97.9 percent.

Therefore as *csyr10* turns out to be significant, it can be said that as the change in average years of secondary education in the whole population over a period of ten years, imposes a positive effect on rpcy1, this "change" can be a cause influencing *rpcy1* . This serves as a basis to arrive at the fact that *csy* is a significant cause affecting *rpcy1*. Hence it is seen that the change in average years of secondary years of schooling in the whole population plays a significant role in raising the income level of the poorest in an economy. It is a well known fact that the income earning potential higher for higher levels of educational qualifications. And because it is the growth rate of income of the poor which is being dealt with it comes as no surprise that changes in the average years if secondary education would have a significant impact on the growth of income of the poorest.

As has been discussed earlier, another variable of interest is the ratio of the primary years of education to that of the secondary years, for all the countries. This is represented as the variable *psr*. But as shown in Table 5, this variable is not found to be significant at the conventional levels of significance.

Gender Inequality

The study here considers the education based gender inequality. The variables used to represent this factor are: *gu, gp, gs, gpc, gsc, cgu, cgp, cgs, cgpc, cgsc, gp5, gp10.* The regression results reported in Table 6 shows *cgp* to be a significant factor among the factors *gu, gp, gs, gpc, gsc, cgu, cgp, cgs, cgpc* and *cgsc. cgp* is found to be significant at precisely 96.3 percent level of significance. As this represents the *change* in the ratio of the average years of primary schooling attained in females to males, over the entire time period similar problems as mentioned in the case of the education variable *csy* might arise. Therefore to check if *cgp* can be said to impose some influence on *rpcy1,* the regressions number 11 and 12 are run with *gp5* and *gp10* , representing the change in *gp*

for the initial five and ten years respectively. *gp5* is found to be significant at precisely 94.8 percent. Therefore as the change in gender inequality, represented by the change in the ratio of female to male primary years of education, over the initial five years matters for *rpcy1* it can be said that the change in the same ratio over the entire period would also be of influence on *rpcy1*.

Hence given the result from regression number 7 in Table 6 supported by regression 11 it can be said that *cgp*, which is the change in the ratio of the female to male average years of primary schooling, is an important cause affecting the growth rate of income per capita of the poor. It is interesting to note that while primary education does not matter, inequality in primary education does so. The coefficient of *cgp* as reported in regression 7 of Table 6 is positive implying that with the rise (fall) in this ratio, that is, a lessening (worsening) of education based gender inequality, would increase (decrease) the growth rate of income of the poor. The interpretation for this could be that in a society the greater the primary education among the female populace the greater would be the awareness regarding education which can have a spill over effect on the next generation and hence more education among the generations to follow and thereby more income accruing to the next generation, which in turn could have implications for growth. Moreover this variable *cgp* may reflect the general attitude towards females in the society which can have its implications for long-run growth rate of the poor.

Government expenditure

Table 7 reports the result of the regression equation of *rpcy1* on *rpcyn1* and the variable *govex1* which represents the government expenditure on health, education and social security and welfare. It is found to be significant at precisely 93.4 percent with a positive coefficient. To find out the importance of the government expenditure targeted towards the poor, *rpcy1* is regressed on *sw*. The regression of *sw* on *rpcy*, shows that *sw* is significant at precisely 97.3 percent with a positive coefficient. Therefore the results conform with the fact that just as the government expenditure directed in the field of education, health and social security and welfare is important and so is the expenditure directed exclusively towards the poor, with the latter factor being more significant for *rpcy1*.

Since govex1 and sw can also be an indicator of the importance of the welfare of poorer sections in public policy then it might reveal that public policy which is more sensitive to the poor, is able to secure greater improvement in the welfare of the poor.

Human rights variables

The variables of this factor are *rulelaw* and *voice.* Initially the regressions of *rpcy1 on rulelaw* and *voice* are run individually to arrive at the variable which stands out to be most significant. Running the regressions individually on each of these factors one by one does not yield a good result as none of the regressors as shown in equation 1 and 2 of Table 8 are found to be significant.

Asset Inequality

The construction of the variable *resgini* had been discussed in the previous section. To reiterate, it is the residual of the regression of the gini coefficient of income, representing the income inequality, on the index for pay inequality. The regression result of the *rpcy1* on *resgini* is shown in Table 9. The coefficient of *resgini* so obtained is negative in sign though not significant. Although initial asset inequality is an important factor affecting the growth rate of per capita income of the poor it is surprising to see it not appearing as a significant factor in the regression shown in Table 9. However it is important to point out here that the data on asset inequality was unavailable and hence it was derived from a regression analysis. The variable *resgini* is a crude measure of asset inequality as it has been derived as a residual from the regression of income inequality on pay inequality , using the UTIP-UNIDO index of pay inequality, on the assumption that income inequality contains asset inequality and that much of the pay inequality is an important component of income inequality.

Initial income inequality

To see the effect of initial income inequality on the growth rate of income of the poor three regressions are run using three measures for income inequality. (Besides this variable is of interest pertaining to the fact that it can remotely serve as a substitute for asset inequality, as no reliable data on asset inequality was available.). The results are reported in Table 10. The first regression shows the result of using the gini coefficient of

income as a definition for income inequality. Though the coefficient of the variable *gini* enters with a negative sign it is not significant. The second regression uses the ratio of the share of income accruing to the poor to that of the non-poor. In this regression the coefficient of the variable *rpnp* is both negative and significant at 94.1 percent. In the case when *q1* defined as the initial share of income accruing to the poor, is used as variable representing the initial income inequality, the regression equation is seen to yield a significant coefficient with a negative sign.

The negative coefficients of all the variables suggest the fact that there is a negative relation between *rpcy1* and the initial income inequality variables. Therefore the greater is the initial income inequality the lesser would be the growth rate of per capita income of the poor. A plausible explanation to this negative relation could be the fact that this initial inequal distribution of income helps distribute income in favour of the non-poor more than that of the poor, whereby the income of the poor rises less than that of the non-poor. Moreover the initial income gap between the poor and the non-poor gets aggravated over time due the non-poor being able to raise their income levels by utilizing various opportunities like higher education which the poor cannot afford.

Table 4

Linear regression results of *rpcy1* on health variables.

Dep. Var.: ***rpcy1***

Variable	1	2
rpcyn1	0.8267# [6.49]	0.9048# [6.92]
imr	-0.0002 # [-2.72]	.
le	.	0.0009 + [2.41]
Const	0.0202 [1.63]	-0.0545* [-1.97]
R2	0.5979	0.5998
Obs	35	35

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity. #, + and * implies significant at 99%, 95% and 90% respectively.

Source: Authors calculations.

Table 5

Linear regression results of *rpcy1* on education variables.

Dep. Var.: ***rpcy1***

Variable	1	2	3	4	5	6	7	8	9
rpcyn1	0.8638# [6.95]	0.8439# [7.07]	0.9834# [6.65]	0.8968# [6.85]	0.8671# [6.75]	0.8662# [6.84]	0.8196# [7.45]	0.8504# [6.59]	0.9053# [6.67]
primaryyrs	0.0022 [0.94]	.	.	.	.	.	.	0.0013 [0.52]	.
secondaryyrs	.	0.0042 [0.95]	.	.	.	.	.	0.0027 [0.58]	.
cpy	.	.	-0.0002 [-0.06]	.	.	.	.	.	-0.0037 [-0.94]
csy	.	.	.	0.0101# [3.15]	.	.	.	.	0.0111# [3.1]
csyr5	.	.	.	.	0.0132 [1.41]	.	.	.	.
csyr10	.	.	.	.	.	0.0160+ [2.27]	.	.	.
psr	.	.	.	.	.	.	-0.0018 [-1.4]	.	.
Const	-0.0004 [-0.02]	0.0046 [0.38]	-0.0026 [-0.2]	-0.0091 [-0.73]	0.0027 [0.2]	-0.0022 [-0.17]	0.0204* [1.67]	0.0011 [0.06]	-0.0075 [-0.6]
R2	0.4660	0.4668	0.5246	0.5908	0.4859	0.5110	0.5051	0.4701	0.5978
Obs	34	34	35	35	34	34	34	34	35

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity. #, + and * implies significant at 99%, 95% and 90% respectively.

Source: Authors calculations.

Table 6

Linear regression results of *rpcy1* on gender inequality variables.

Dep. Var.: ***rpcy1***

Variable	1	2	3	4	5	6	7	8	9	10	11	12
rpcyn1	0.8986# [4.99]	0.8977# [6.55]	0.9003# [6.16]	0.9194# [6.4]	0.8262# [6.76]	0.9543# [6.3]	0.8159# [4.62]	0.9944# [6.6]	0.8780# [4.88]	0.9832# [6.72]	0.8215# [6.41]	0.8553# [6.44]
gu	-0.0017 [-0.36]	.	.	.	.	.	.	.	.	.	.	.
gp	.	0.0215 [1.25]	.	.	.	.	.	.	.	.	.	.
gs	.	.	0.0039 [0.82]	.	.	.	.	.	.	.	.	.
gpc	.	.	.	0.0189 [1.06]	.	.	.	.	.	.	.	.
gsc	.	.	.	.	0.0007 [0.32]	.	.	.	.	.	.	.
cgu	.	.	.	.	.	-0.0039 [-0.81]	.	.	.	.	.	.
cgp	.	.	.	.	.	.	0.0364+ [2.18]	.	.	.	.	.
cgs	.	.	.	.	.	.	.	-0.0028 [-1.1]	.	.	.	.
cgpc	.	.	.	.	.	.	.	.	0.0147 [1.33]	.	.	.
cgsc	.	.	.	.	.	.	.	.	.	-0.0006 [-0.32]	.	.
gp5	.	.	.	.	.	.	.	.	.	.	0.0766+ [2.02]	.
gp10	.	.	.	.	.	.	.	.	.	.	.	0.01 [0.29]
Const	0.0072 [0.61]	-0.0144 [-0.63]	0.0010 [0.06]	-0.0132 [-0.55]	0.0112 [0.89]	-0.0006 [-0.04]	0.0029 [0.21]	-0.0034 [-0.25]	0.0023 [0.16]	-0.0026 [-0.2]	0.0088 [0.75]	0.007 [0.61]
R2	0.4528	0.4741	0.4584	0.4811	0.4947	0.5403	0.5630	0.5280	0.5405	0.5249	0.4938	0.4526
Obs	34	34	34	34	33	35	35	35	35	35	34	34

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity. #, + and * implies significant at 99%, 95% and 90% respectively.

Source: Authors calculations.

Table 7
Linear regression results of *rpcy1* on government expenditure variables.
Dep. Var.: ***rpcy1***

Variable	1	2
rpcyn1	1.0988# [5.03]	1.0963# [5.04]
govex1	0.0005* [1.93]	.
sw	.	0.0006 [2.36]+
Cons	-0.0318 [-1.38]	-0.0223 [-1.13]
R2	0.5049	0.5444
Obs	27	27

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity. #, + and * implies significant at 99%, 95% and 90% respectively.
Source: Authors calculations.

Table 8
Linear regression results of *rpcy1* on political variables variables.

Dep. Var.: ***rpcy1***

Variable	1	2
rpcyn1	0.8622# [5.83]	0.9268# [5.87]
rulelaw	0.0045 [1.21]	.
voice	.	0.0014 [0.31]
Cons	0.0065 [0.5]	0.0028 [0.18]
R2	0.5799	0.5594
Obs	33	33

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity. #, + and * implies significant at 99%, 95% and 90% respectively.
Source: Authors calculations.

Table 9

Linear regression results of *rpcy1* on asset inequality variable.

Dep. Var.: ***rpcy1***

Variable	1
rpcyn1	0.9785# [6.86]
resgini	-0.0001 [-0.3]
Cons	-0.0027 [-0.21]
R2	0.5348
Obs	34

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity. #, + and * implies significant at 99%, 95% and 90% respectively.
Source: Authors calculations.

Table 10

Linear regression results of *rpcy1* on initial income inequality variables.

Dep. Var.: ***rpcy1***

Variable	1	2	3
rpcyn1	0.9331# [7.1]	1.0931# [8.43]	1.0969# [8.5]
gini	-0.0003 [-0.83]	.	.
rpnp	.	-0.2699+ [-1.96]	.
q1	.	.	-0.3093+ [-2]
Cons	0.0150 [0.72]	0.0050 [0.38]	0.0059 [0.44]
R2	0.5393	0.5689	0.5705
Obs	35	35	35

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity. #, + and * implies significant at 99%, 95% and 90% respectively.
Source: Authors calculations.

3.II Regressions across factors

The next set of regressions presented in Table 11 are the ones where *rpcy1* have been regressed on *rpcyn1* and any two of the variables representing the four factors-health, education, gender inequality and initial income inequality. The variables used as factors are the ones which have turned out to be significant in the regression analysis of section 3.I. implying that these are the ones which best represents the factors. In the next set of regressions the variables *resgini, rulelaw, voice and govex1* have been dropped out. The reason for dropping the first three variables is that they were not found to be significant and hence not representative of the factors they were initially thought to represent. However such a consideration does not hold in the case of the variable *govex1* as it was found to be quite significant in the analysis of section 3.1. This variable had to dropped out due to scarcity of data on this variable whereby the number of observations for the regression when it is considered individually is merely 28(refer to regression 1, 7). Therefore addition of other variables to *govex1* in a regression would lead to a reduction in the degrees of freedom to a level below 30, which is undesirable from the point of view of the accuracy of the regression results which would get adversely affected.

The variables which best explains the four factors-health, education, gender inequality and initial income inequality *le, csy, cgp and rpnp*. The variable *le* was chosen over *imr* because the former came out with a higher coefficient value than that of the latter (refer regressions 1 and 2 in Table 4). The choice of csy to represent education was simply because it appeared as the most significant variable (refer regression 4 in Table 5). Due to similar reasons *cgp* was selected to represent the factor gender inequality as it turned out to be the only significant variable (refer regression 7 in Table 6) among the others to represent this factor. In the case of the factor income inequality *rpnp* was chosen over another variable *q1,* (refer regressions 2 and 3, Table 10) although the latter was more significant, simply because of the fact that *rpnp* as a measure of inequality was more appealing than *q1* [1]

[1] (rpnp is defined as the ratio of the percentage share of income of the poorest quintile to that of the remaining eighty percent of the population and q1 represents the percentage share of income accruing to the poorest quintile).

The results of the six regressions are presented in Table 11. The variable *rpcyn1* enters all the regressions along with all the other variables. It is observed that in all the four regressions the variable *rpcyn1* enter significantly with a positive coefficient when added to any other variable. This proves that it is an important factor determining the growth rate of income of the poor.

The variable *rpnp* appears to be significant in all the regressions at 95 percent level (regressions 5and 6) and at 99 percent in one (regressions 3). The variable *le* appears to be significant at 99 percent in all the regressions except in regression 1 where it is not significant. In the case of *csy* is found to be significant in all the regressions 1,4 and 5, at 90 , 95 and 99 percent level of significance , respectively. For the variable *cgp* it is observed to be significant in regressions 2 and 6 at 95 percent.

Therefore from this analysis it can be concluded that the factors growth rate of per capita income of the remaining eighty percent of the population, health, education, gender inequality and initial income inequality turn out to be the most important determinants of long run growth rate of per capita income of the poorest quintiles.

Comparing with the R squared value of regression 1 in Table 1 in chapter 1, it is observed that the value improves from 0.52 to 0.60 (the lowest value) as shown in regression 4 of Table 11.

Hence from this analysis it is seen that the cross-country differences in the growth of per capita income of the poorest quintile is explained significantly by the factors: initial health, change in secondary schooling years, change in gender inequality and change in initial income inequality when the growth rate of the poor is regressed on the growth rate of the non-poor (representing the growth rate of per capita income of the economy) along with the above factors.

Table 11

Linear regression results of *rpcy1* on variables taking three at a time.

Dep. Var.: ***rpcy1***

Variable	1	2	3	4	5	6
rpcyn1	0.8682# [6.76]	0.6931# [4.12]	1.0405# [9.17]	0.7974# [4.78]	1.0061# [8.46]	0.9231# [5.99]
le	0.0007 [1.61]	0.0010# [2.77]	0.0012# [3.32]	.	.	.
csy	0.0068* [1.94]	.	.	0.0086+ [2.24]	0.0099# [3.57]	.
cgp	.	0.0442+ [2.67]	.	0.0245 [1.43]	.	0.0382+ [2.38]
rpnp	.	.	-0.3823# [-3.07]	.	-0.2625+ [-1.85]	-0.2818+ [-2.33]
Cons	-0.0451 [-1.53]	-0.0538+ [-2.02]	-0.0577+ [-2.39]	-0.0043 [-0.31]	-0.0015 [-0.11]	0.0113 [0.83]
R2	0.6241	0.6555	0.6831	0.6067	0.6326	0.6112
Obs	35	35	35	35	35	35

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity. In regressions 4 and 9 there are four regressors, as *rulelaw* and *voice* are taken together. #, + and * implies significant at 99%, 95% and 90% respectively.

Source: Authors calculations.

3.III Multivariate regressions across factors

The regressions presented in Table 12 are the ones where *rpcy1* is regressed on *rpcyn1* and any three of the regressors-*le, csy, cgp, and rpnp.* There are total five regressions in Table 12. Similar to that of Section 3.II the variable *rpcyn1* appears in all the four regressions and invariably it is found to be highly significant in all of them.

Observing the results of the regressions presented in Table 12 it seen that *rpnp* appears significant in one of the regressions at 95 percent level of significance and at 99 percent in the remaining. The variable *rpnp* maintains its significance in all the regressions no matter what regressors it is entered with. Therefore *rpnp* appears to be a very significant determinant of *rpcy1*. The result is same for the variable *le* which maintains its significance in all the regressions and with whatever regressors it is entered with. The variables *cgp* is significant in all the regressions except one. In the case of *csy* it appears to be significant at 95 percent in only one case. Looking at regression 5 where all the variables are taken together it is seen that except *csy* all other variables maintain their significance at the conventional levels.

Hence the variables that emerge as the most significant even in the presence of other factors are *rpcyn1*, *le, cgp, csy* and *rpnp*. Therefore much in line with what has been found in the literature is the fact that growth rate of the economy plays an important role in the growth rate of the poor. This is in conformity with the results of Dollar and Kraay (2002) and Birdsall and Londono (1997). In this study initial health also appears to be a major factor determining the growth rate of income of the poor. This result is at variance with the findings of Birdsall and Londono and Dollar and Kraay analyses. The issue of education based gender inequality appears to be a significant determinant which is once again not taken up in the two above mentioned studies. The factor education is however common to both this study and the other two and there is agreement in the result that education is an important factor affecting the growth rate of income of the poor, though of course it is the secondary education which emerges to be of significance in this study. Finally coming to the factor of income inequality it is seen that it is not only important here but has also been found to be important in the analysis of Birdsall and Londono (1997).

Table 12.
Linear regression results of *rpcy1* on variables taking four at a time.

Dep. Var.: ***rpcy1***

Variable	1	2	3	4	5
rpcyn1	0.7017# [4.21]	1.0052# [8.72]	0.8158# [6.14]	0.9015# [6.08]	0.8162# [6.05]
le	0.0009+ [2.15]	0.0010+ [2.42]	0.0013# [3.79]	.	0.0013# [3.28]
csy	0.0033 [0.82]	0.0050 [1.62]	.	0.0083+ [2.37]	0.0004 [0.12]
cgp	0.0387+ [2.15]	.	0.0491# [4.16]	0.0268 [1.54]	0.0483# [3.41]
rpnp	.	-0.3598# [-2.84]	-0.4113# [-4.21]	-0.2721+ [-2.14]	-0.4088# [-4.08]
Cons	-0.0493* [-1.74]	-0.0505+ [-1.89]	-0.0572+ [-2.48]	0.0040 [0.29]	-0.0565+ [-2.27]
R2	0.6603	0.6962	0.7511	0.6516	0.7512
Obs	35	32	35	35	35

Note: The values in the parentheses give t values. The results have been checked for heteroscedasticity and multicollinearity #, + and * implies significant at 99%, 95% and 90% respectively.
Source: Authors calculations.

CHAPTER 4

Conclusion

The issue of poverty has received a lot of attention from researchers and academicians. The subjects of absolute and relative poverty have been widely researched. In this study the area of interest is *relative* poverty whereby a person is poor on a comparative scale. It is the kind of poverty that is common to all societies and nations at large. The "poor" in this study have been those who belong to the poorest twenty percent of the population of a country.

The aim in this piece of analysis has been to examine the factors which can help explain the cross-country differences in the growth rate of per capita incomes of the poorest quintiles. In so doing two papers Dollar and Kraay (2000) and Birdsall and Londono (1997) have been influential in guiding this course of study.

Birdsall and Londono (1997) undertake a cross-country analysis with factors that can affect the growth rate of income of the poorest quintile, over a time period of at least five years. They arrive at the conclusion that factors like capital accumulation, educational inequality and initial asset inequality are crucial to the growth in income of the poor.

Dollar and Kraay (2000) look at almost similar issues. They too consider a five year period and inspect the factors that affect the per capita income of the poorest quintile. The paper finds that the per capita income level of the poor have a one to one correspondence with that of the overall economy. They test for certain other factors like primary educational attainment, social spending, agricultural productivity and formal democratic institutions which are known to directly affect the incomes of the poor but find none of these to be statistically significant at the conventional levels.

This study poses a slightly different question. It looks at the factors that directly affect the growth rate of per capita income of the poorest for a long enough period of at least twenty years. This study looks at the long term factors which have implications for sustained improvement in the welfare of the poor. It has been pointed out in the literature that the growth rate of per capita income of the economy is a significant determinant of the growth rate of per capita income of the poorest. A preliminary analysis carried out in

chapter 1 reveals that the growth rate of per capita income is a highly significant factor affecting the growth rate of per capita income of the poor. This finding is in consonance with the findings of Birdsall and Londono (1997) and Dollar and Kraay (2002). However it is important to note that the explanatory power of the growth rate of per capita income of the economy for that of the poorest quintile is relatively less than that for the other quintiles in an economy.

Owing to such revelations the focus of the study was directed towards the factors which can substantially affect the growth rate of per capita income fo the poorest, *other* than the growth rate of the overall economy. The idea being to analyse the differential sources that stand to affect substantially the growth rate of income of the poor. In order to do so this analysis deals with the factors which have been frequently cited in the growth literature. These consist of health, education, gender inequality, government expenditure, asset inequality, political rights and initial income inequality.

The hypothesis posed in this study is examined by using simple linear regression analysis of the growth rate of per capita income of the poorest on the growth rate of per capita income of the economy as a whole and all the above mentioned factors. As it is the growth rate of income of the poorest quintile which was being looked at, the relevant growth rate was calculated using income inequality data. For this a dataset on income inequality was compiled from the World Income Inequality Database version 2a compiled by World Institute for Development Economics Research, United Nations (UN-WIDER) and the "high-quality" dataset by Deininger and Squire. For the other factors used in the study the data had been taken from various other sources.

The regression results corroborates with the findings in the literature that the overall growth rate of the economy is a significant determinant of the growth of income of the poor. Among the factors that emerge as the most significant are health (measured by life expectancy, *le*), secondary years of schooling (measured by change in the average years of secondary schooling, *csy*) gender inequality (as measured by change in the ratio of the primary years of primary schooling in females to males, *cgp*) and the initial income inequality (as measured by ratio of the percentage share of income accruing to poor to that of the non-poor, *rpnp*). Another factor which emerges to be significant is government expenditure targeted towards the poorest section.

The finding that education turns out to be significant is supported by the study of Birdsall and Londono also income inequality as a significant factor affecting the growth rate of the poor finds support from Birdsall and Londono. The result that the growth rate of per capita income of the poor is significantly affected by that of the overall growth rate of per capita income of the economy finds support from both the papers. However it is only in this analysis that gender inequality and government expenditure turns out to be significant determinants of the growth rate of per capita incomes of the poor. Dollar and Kraay in fact find government expenditure to be insignificant.

Hence from this study it is seen that although the growth rate of per capita income of the economy is a significant determinant of the growth rate of per capita income of the poorest quintile it is probably not a sufficient one because it can explain only a part of the cross-country variation in long-run growth rate of per capita income of the poorest. In addition for a sustained improvement in the condition of the poor, the initial conditions in the form of lesser income inequality, lesser gender inequality and better health levels bear significance.

Therefore growth enhancing policies in the economy may not be enough to secure long-run growth in the income of the poor. Public policy must be directed towards creating favourable initial conditions under which the poor can enjoy a sustained increase in their income.

Finally coming to the end of this chapter a note on the limitations of the study must be cited. As it is an empirical study the first set of limitations arise from the datasets used. It is common knowledge that income inequality data are hard to come by and it is specially true for quality data. This is because for income inequality data the survey methods and definitions differ across and within countries. Therefore this study also derived its share of limitations from the use of the income inequality dataset.

Another limitation of the study had been the small sample of countries (36) used in the analysis. Due to the requirement of maintaining a time period of at least twenty years for all the countries a trade-off between the quality of data and the number of observations in the dataset, had to be faced.

A final limitation of the study stems from the fact that the data on the factor *asset inequality* could not be obtained. As this is an important factor affecting the growth rate of per capita income of the poorest quintile, the absence of data on this important factor seriously constraints the scope of the analysis.

References

Alesina, A. and D. Rodrik, (1994), "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 109, 465-490.

Barro, R.J. and Lee, J.W. (1994), "Data set for a Panel of 138 Countries", available at http://www.worldbank.org/research/growth/ddbarle2.htm

Barro, R.J., (1991), "Economic Growth in a Cross Section of Countries", *Quarterly Journal of Economics,* 106, 407-440.

Barro, R.J., and J.W. Lee, (1993), "Losers and Winners in Economic Growth" in proceedings of the World Bank, Annual Conference on Development Economics.

Birdsall, Nancy and Juan Luis Londono, (1997), "Asset Inequality Does Matter: Lessons from Latin America", Office of the Chief Economist working paper, Inter American Development Bank.

Deaton, Angus (2004), "Measuring Poverty", mimeo, Princeton University, available at http://www.wws.princeton.edu/~rpds/downloads/deaton_povertymeasured.pdf

Deininger, K. and Lyn Squire, (1996) "Measuring Inequality: A New Data Base", *World Bank Economic Review*, 10:3, 565-91.

Deininger, K. and Lyn Squire (1998), "New ways of looking at old issues: inequality and growth", *Journal of Development Economics*, 57, 259-87.

Deininger, Klaus and Lyn Squire, "Economic Growth and Income Inequality: Reexamining the Links", available at www.imf.org/external/pubs/ft/fandd/1997/03/pdf/deininge.pdf

Dollar, David and Aart Kraay, (2002), "Growth is Good for the Poor", *Journal of Economic Growth*, 7(3): 195-225.

Durlauf, S.N., P.A. Johnson and J. R. Temple, (2004): Growth Econometrics, handbook of growth economics (?) unpublished manuscript, available at www.ssc.wisc.edu/econ/Durlauf/

Easterly William, (1999), "Life During Growth," *Journal of Economic Growth*, 4(3), 239-275, available at www.cgdev.org/doc/expert%20pages/easterly/easterly_lifeduringgrowth.pdf

Galbraith James K. and Hyunsub Kum, (2005), "Estimating the Inequality of Household Incomes: A Statistical Approach to the Creation of a Dense and Consistent Global Data Set", *Review of Income and Wealth*, Vol. 51, No. 1, 115-145.

Galbraith James K. and Hyunsub Kum, (2005), "Inequality and Economic Growth: Data Comparisons and Econometric Tests", The University of Texas Inequality Project, UTIP Working Paper Number 21, available at http://utip.gov.utexas.edu/papers/utip_21rv.pdf

Government Finance Statistics Yearbook, 1979 to 1998, I M F.

Greene, W.H., (2000), *Econometric Analysis*, Fourth Edition, Prentice - Hall, Upper Saddle River, New Jersey.

Gujrati, N. Damodar, (2003), *Basic Econometrics,* Fourth Edition, McGraw – Hill, Singapore.

International database. available at http://www.census.gov/ipc/www/idbnew.html

Kuznets, Simon, (1955), "Economic Growth and Income Inequality", *American Economic Review*, 45, 1-28.

Lundberg, Mattias and Lyn Squire, (2000), "The Simultaneous Evolution of Growth and Inequality", *Economic Journal*, 113, 326-344.

Milanovic, Branko, (2002), "True world income distribution, 1988 and 1993: First calculation based on household surveys alone," *Economic Journal*, 112, 51-92.

Milanovic, Branko, (2002), "The Ricardian Vice: Why Sala-i-Martin's calculation of world income inequality cannot be right," Mimeo, September.

Patnaik, Utsa (2005), 'Theorizing Food Security and Poverty in the era of Economic Reforms', Social Scientist, 7-8 July-Aug, Vol. 33.

Penn World Table, version 6.1. available at http://pwt.econ.upenn.edu/

Sala-i-Martin, Xavier, (2002), "The World Distribution of Income", NBER Working paper No. 8905, available at www.nber.org.

Sala-i-Martin, Xavier, (2002), "The Disturbing "Rise" of Global Income Inequality", NBER Working paper No. 8905, available at www.nber.org.

Sen, Amartya and Jean Dreze, (1996), "India: Economic development and Social Opportunity", Oxford University Press.

Sen, Amartya, (1973), *On Economic Inequality*, New York, Norton.

UNU/WIDER-UNDP World Income Inequality Database, 2000, available at http://www.wider.unu.edu/wiid/wiid.htm

Wooldridge, J. M., (2000), *Introductory Econometrics: A Modern Approach*, South Western.

Fields, Gary. S., "The State of Development Economics Progress and Perspectives". Ed. by Gustav Ranis and T.P. Schultz.