EXPLORING THE RELATIONSHIP BETWEEN SOCIO-ECONOMIC INEQUALITY, POLITICAL INSTABILITY AND ECONOMIC GROWTH:

WHY DO WE KNOW SO LITTLE?

Gunhild Gram Giskemo

Master Thesis

Department of Comparative Politics
University of Bergen

May 2008

EXPLORING THE RELATIONSHIP BETWEEN SOCIO-ECONOMIC INEQUALITY, POLITICAL INSTABILITY AND ECONOMIC GROWTH:

WHY DO WE KNOW SO LITTLE?

Gunhild Gram Giskemo

Master Thesis

Department of Comparative Politics
University of Bergen

May 2008

Good people, things cannot go right in England and never will, until goods are held in common and there are no more villeins and gentlefolk, but we are all one and the same […]. They are clad in velvet and camlet lined with squirrel and ermine, while we go dressed in coarse cloth. They have the wines, the spices and the good bread: we have the rye, the husks and the straw, and we drink water. They have shelter and ease in their fine manors, and we have hardship and toil, the wind and the rain in the fields […]. [W]e want things to be changed, or else we will change them ourselves.*

We must not lose sight of the fact that the greater challenge is to address the issues of injustice, inequality and unfairness, which clearly gave rise to the sorry situation we find ourselves in.**

ABSTRACT

This thesis investigates the contention that socio-economic inequality has a detrimental effect on economic growth by breeding political instability. This hypothesised causal pattern has been subject to academic interest and empirical investigation for centuries, but still no agreement has been reached as to what the relationships really look like. During the 1990s the number of quantitative studies in this field increased dramatically, but they reached highly different conclusions. An important objective of this thesis is to explore what lies behind these inconsistent empirical findings.

The hypothesised causal pattern is investigated both theoretically and empirically. Socio-economic inequality is thought to produce political instability by breeding relative deprivation and discontent. Political instability is, in turn, likely to decrease growth because it constitutes a disincentive to invest. These hypotheses are explored through cross-national time-series regression analyses of a total of 188 countries from 1950 to 2004. This study is the first to use the recently updated and expanded dataset on socio-economic inequality provided by the UN-based World Institute for Development Economics Research in May 2007. Previously, lack of comparable and extensive data on this variable has made it difficult to conduct large-scale quantitative analysis of the political and economic effects of socio-economic instability.

The analytical results reveal the following: while instability does seem to affect growth negatively, the hypothesis that inequality breeds political unrest is not unanimously supported. This is because a series of robustness tests show that the results are dependent on how inequality and instability are measured. The use of the largest and most recently updated data source on socio-economic inequality has thus not contributed to clarifying whether inequality reduces economic growth by breeding political instability. It has contributed in a different way, however, because the analysis offers a possible explanation of why previous empirical studies have reported such diverging findings: namely that they measure socio-economic inequality and political instability in different ways. This is an important finding because it shows that the reliability of the conclusions of existing studies can be questioned.

[*] The priest John Ball, observed preaching to the masses in public squares by Jean Froissart in 1318 (Froissart and Brereton 1968: 211-212).

[**] Editorial in the Kenyan newspaper Daily Nation (Daily Nation 2008, January 10th)

ACKNOWLEDGEMENTS

First and foremost, I want to thank my supervisor, Michael Alvarez, whose encouragement, devotion and extraordinary generosity have been decisive to the writing of this thesis. Thank you for all your patience, flexibility and firm assistance from the beginning to the very end, and for your efforts to help me solve all the statistical tangles that have arisen along the way.

Thanks to the Christian Michelsen Institute (CMI) for having provided me with office facilities and an amazingly inspiring research environment. Being part of the social and scholarly activity at the CMI has been of great value to the writing of my thesis.

I would also like to thank the Methods Group at the Department of Comparative Politics, where I have been offered the opportunity to present parts of my thesis. The comments from the scholars and students in the group have been very useful.

Many people have helped me by commenting on my work, and with methodological and technical issues. For this I am extremely grateful. Thanks to Inga Brautaset, Ragnhild Hvoslef Krüger, Magnus Hatlebakk (CMI), Kyrre Moe (NSD), Elin Skaar (CMI), Inge Amundsen (CMI), Arild Een, Erik Aarebrot, Tor Midtbø (UiB), Anne Katrine Mortensen, Frode Løvlie and the IT staff at the CMI. Thanks also to my contact person at the CMI, Espen Villanger, and to Ingvild Hestad and Chris Jacob for their efforts to arrange all the practicalities related to my stay at the CMI.

I would like to thank all my fellow students, Inga, Katrine, Frode, Torgunn, Kari, Chandra and Carmeliza, and my other coffee-drinking, quiz-loving colleagues at the CMI – your company has kept my spirits up in this period.

Thanks to my aunt Liv Inger and uncle Finn for inviting me to Sunday dinners. Thanks to my dear cohabitants Ingrid and Gunvor for serving me champagne and apple pie. Thanks to the University of Bergen, the Norwegian State, HBO and the Sopranos, Timian the cat, and Kvamme.

Finally, I leave this thesis indebted to my mother, father and brother, and last but not least, to Joachim.

Bergen, May 2008
Gunhild Gram Giskemo

CONTENTS

TABLES AND FIGURES

1 INTRODUCTION

This thesis focuses on the political and economic consequences of socio-economic inequality and takes as its point of departure the contention that socio-economic inequality reduces the rate of economic development by breeding political instability. This alleged causal pattern has been subject to academic focus and empirical investigations since the times of the Ancient Greeks, and distributional issues and ideals of equality have played a central role in many of the world's conflicts and revolutions. As stated by Lichbach (1989: 433), the three great ideologies of the last three centuries – nationalism, liberalism and socialism – all spawned revolutionary movements based on ideas of inequality, although different ones. For example, in the French Revolution people called for "Liberté, égalité, fraternité"; the propaganda of the Russian Revolution was "peace, land and bread"; a wartime slogan of the Chinese Revolution was "those who have much give much; those who have little give little"; and the rhetoric of the American Revolution was "all men are created equal". Indeed, the timelessness and universality of this topic is demonstrated by the two opening citations. Several centuries and widely different contexts separate them, but their message remains the same. The first citation comes from one of Jean Froissart's *Chronicles*, "The Peasants' Revolt in England" (1318). It describes the great uprisings and rebellion of the peasantry and common people in England that came as a response, according to Froissart, to the stark contrast between serf and noble, peasant and gentry, in material well-being and freedom from subjection. The second citation is taken from an editorial of the Kenyan newspaper Daily Nation, commenting on the political violence that has marked Kenyan reality since the elections in December 2007.

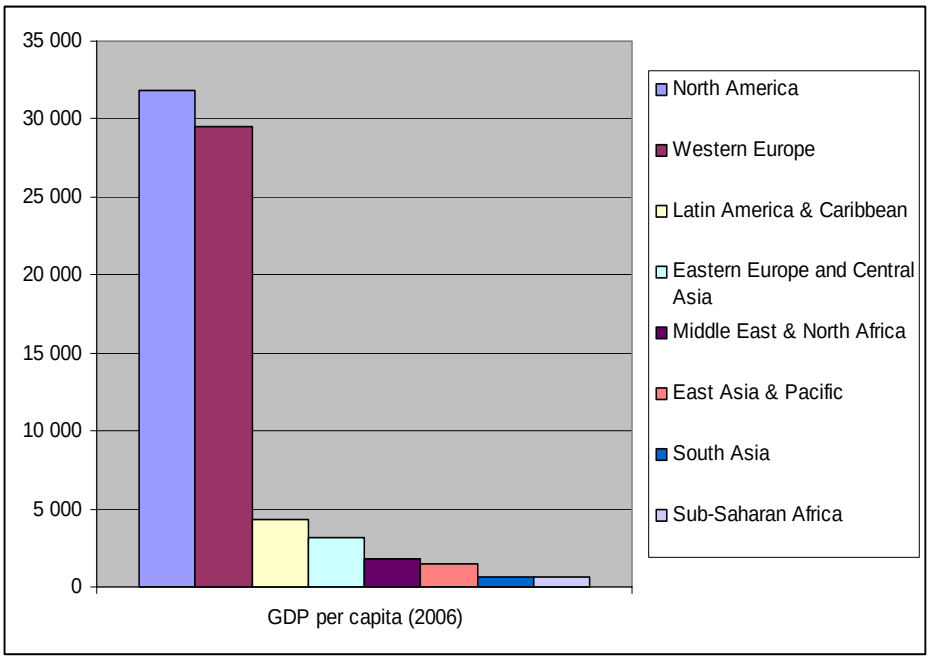
1.1 WHY STUDY THE EFFECTS OF SOCIO-ECONOMIC INEQUALITY

The aim of this thesis is to answer the following questions: 1) does socio-economic inequality reduce economic growth by increasing the level of political instability, and 2) why have previous studies on this subject reached such diverging conclusions? The theoretical basis on which to expect such a relationship is the following: First, a highly unequal, polarised distribution of resources produces relative deprivation and is thus an important source of discontent. In that way, inequality creates strong incentives to engage in violent protests, assassinations, coups or other politically destabilising activities. In turn, political instability discourages investment for at least two classes of reasons: first, it creates uncertainty

regarding the political and legal environment, and secondly, it disrupts market activities and labour relations, with a direct adverse effect on productivity (Perotti 1996a: 151).

Several circumstances contribute to the persisting academic interest in this subject. First, as emphasised by the World Development Report from 2006, and clearly shown by Figure 1 below, the differences in GDP per capita among Western countries and the rest of the world are still severe. According to the most recently available data from the World Bank, Western Europe and North America, comprising 13 percent of the world's population, held 72 percent of the world's total income in 2006 (World Development Indicators 2006). Being able to shed light on some of the underlying causes of this empirical pattern is a goal in itself.

Figure 1: Regional differences in GDP per capita (the world)[1]

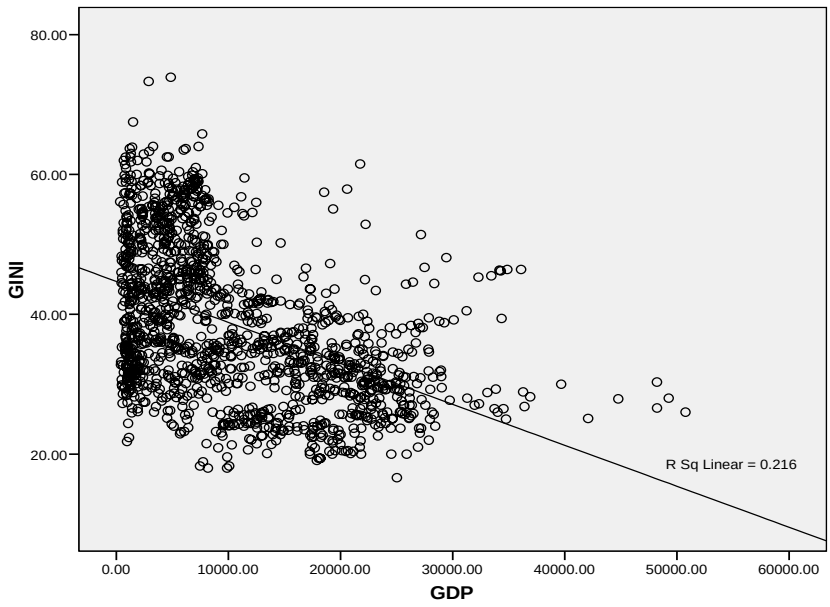


Secondly, as depicted by Figure 2, there is a negative correlation between income distribution, as measured by the Gini coefficient, and level of economic development (GDP per capita), something that calls for an explanation. This relationship was also shown by Easterly (2002), and according to him, "if this link is causal from inequality to income, it provides further strong evidence that there is a long-run negative association between growth (of which

[1] The graph is constructed by the author on the basis of data from the World Development Indicators (World Bank 2006).

income is of course the cumulative sum) and inequality" (Easterly 2002: 2).[2] This correlation has been an important motivation behind the 1990s increase in empirical studies of the relationship between economic inequality and growth (Galor and Zeira 1993: 35).

Figure 2: Scatterplot of GDP per capita and Gini[3]



Thirdly, despite the long-lasting academic and public interest in the political and economic consequences of socio-economic inequality, still no agreement has been reached as to what these really are, and the divergence in both theoretical expectations and empirical findings has spurred a continued interest in the subject (MacCulloch 2005: 93). A major problem associated with applying a quantitative approach to studying socio-economic inequality has been the limited data availability and comparability: until the late 1990s most studies were cross-sectional analyses with very few observations. In 2007 the UN based World Institute for Development Economics Research published an updated and expanded version of their inequality dataset, which is the most comprehensive international data collection on within-country economic inequality. This dataset enables cross-sectional time series analyses that include socio-economic inequality as a variable in a scale that has not previously been possible.[4] The dataset is unique, not because of its large expansion, but also in its detailed

[2] Easterly's term *income* can be interpreted here as the level of GDP per capita.
[3] The scatterplot is produced by the author with the computer software SPSS version 14. The variable definitions of GINI and GDP and their data sources are specified in Appendix A.
[4] This dataset has existed since 1997, and was first updated in 2004.

information on each data point, making it possible to conduct a differentiated analysis that takes into account different ways of measuring inequality and contextual specifics that have made cross-sectional comparison difficult. The hope is that with this new data collection on economic inequality it will be possible to answer the research questions of this thesis, which have remained unsettled for so long.

Finally, any conclusions regarding the effects of socio-economic inequality on political instability and growth can have significant implications for what kind of policies are recommended and implemented around the world to enhance growth (Muller and Seligson 1987). If inequality is thought to be irrelevant, or as having a positive effect upon economic growth, then a government might not be compelled to introduce redistributive programs, and it can pursue a model of development that promotes rapid wealth accumulation at the expense of distributional equality. If, on the other hand, great inequality is expected to cause mass protest and collective violence, then a government should be compelled to implement policies that reduce the effective inequality in society. As Barro puts it, redistribution can have a positive effect upon growth if it reduces crimes and riots caused by income inequality, and even in a dictatorship, self-interested leaders would favour income-equalising transfers if the net effect were a decrease in the tendency for social unrest and political instability (Barro 2000: 7).

It is imperative to note that political stability not only is important in relation to growth, it is also a goal *per se*: "The question of why nations differ in rates of domestic political violence is of intrinsic interest because the maintenance of political stability is a goal of all governments" (Muller and Weede 1990: 624). Since the end of the Cold War most conflicts have been intra-national, non-ideological disputes that take place in less developed countries. Accounting for these conflicts has become a high-priority global task, and theories on the determinants of conflict can provide practical tools for preventive diplomacy (Auvinen 1997: 177). It is also important to note that political instability and the incidence of political upsurge are sometimes a necessary means to achieve a "greater good". When people protest it is sometimes a reaction to an inacceptable situation such as in the presence of a repressive, non-democratic regime, and collective violence has thus sometimes led to the creation of new and more satisfying political communities (Gurr 1970: 3).

1.2 STRUCTURE OF THE STUDY

The hypotheses subjected to analysis in this thesis are examined empirically through a cross-section time-series multivariate regression analysis of a panel of 188 countries from 1950 to 2004. The thesis is divided into three parts. The first part presents the theoretical expectations and previous empirical findings for the hypothesised relationship between socio-economic inequality and political instability, and subsequently for that between political instability and economic growth. The chapter starts out by explaining the reason why it is important to study what factors affect the rate and level of economic development. It then gives an overview of the different hypothesised paths of causation that link inequality to negative growth, with the objective of putting political instability as a path of causation into context. The hypothesis about the effects of inequality on political instability is far more complex and ambiguous both in its underlying theoretical assertions and related empirical findings than that of the economic effects of political instability. Therefore, the main focus of chapter 2 will be on the relationship between socio-economic inequality and political instability.

Chapter 3 discusses the methodological issues related to the approach adopted here, and how the various variables are measured and operationalised. Both socio-economic inequality and political instability are measured in various ways, the purpose of which is to test the robustness of the analytical findings. The analytical model consists of a two-equation system: in the first equation political instability is the dependent variable, and in the second the dependent variable is economic growth. Because causality between economic growth and political instability is expected to run in both directions, tests for simultaneity are introduced to avoid biased results. When simultaneity is found to be present, a simultaneous equation model is employed. When it is not, a recursive model is used. The third part of the thesis, chapter 4, presents and discusses the results of the analyses, followed by a conclusion that sums up the thesis and its main findings.

2 WHAT WE THINK WE KNOW: THEORETICAL EXPECTATIONS

The aim of this chapter is to show how the hypotheses of this thesis are justified. The chapter is organised in the following way: First, I elaborate on why the underlying goal of both this thesis and an infinite amount of academic work, ultimately is contributing to knowledge about what can increase or reduce the economic development of a country. Secondly, I present an overview of the literature on the relationship between inequality and growth, with the objective of putting the path of causation explored here in context. Thirdly, I discuss the ways in which inequality is expected to affect political instability. Finally, I elaborate on the relationship between instability and growth, and show how the former is thought to affect the latter. These sections are accompanied by examples of previous empirical findings, which, together with the insights offered by the various theoretical approaches, affect our expectations about causality between these variables. A summary of the theoretical discussions concludes the chapter.

2.1 THE IMPORTANCE OF ECONOMIC DEVELOPMENT

The variable ultimately to be explained in this thesis is the rate of economic development, more specifically the growth rate of GDP per capita. An important objective of the thesis is thus to achieve knowledge about some of the factors that cause variation in levels of economic development across time and space. But what makes economic development so important that explaining its variation constitutes a basic purpose, not only of this thesis, but of an infinite amount of literature since the very beginning of academic enterprise? The scope of this thesis does not allow me to enter into the philosophical complexities associated with this fundamental question. Nevertheless, avoiding the discussion altogether is too simplistic because its conclusions are decisive for the very grounds on which research is justified. If economic development does not increase quality of life, then an important reason for explaining its variation disappears. Most studies concerned with economic development treat it as a given good. However, it is being questioned whether, and when, material well-being increase *actual* well-being and personal happiness. For example, several studies have shown that the societal upheavals, extensive demographic changes, and environmental degradation often associated with rapid economic development can create more problems than

improvements to the lives of those affected (Kenny 2005: 204), and others demonstrate that in some cases the fruits of economic growth are so unevenly distributed that the well-being of the poor majority does not improve (Easterly 1999: 240; Todaro 1997: 15).

On the other hand, few would disagree that in a modern world economic well-being is strongly associated with such basic aspects of human life as survival, comfort, and ultimately, self-realisation. With widespread misery, poverty and unfulfilled material needs as a historical – and in many places, current – backdrop, the pursuit of economic growth has naturally been central to human life. But are the marginal returns of continued development constant, or are there limits beyond which further growth in material well-being does not produce more *actual* well-being? Studies of the relationship between subjective and objective measures of well-being commonly find that the level of economic well-being is associated with happiness only up to a certain point (Kenny 2005; Seghieri et al. 2006). Therefore, even though both developed and developing countries are included in this analysis, I do not contend that economic growth is equally important independently of its starting point. Rather, there are obvious difficulties related to determining at what level of development to draw a line, if such exists, above which further development does not generate increased well-being. Such an evaluation would be highly subjective. This, added to methodological considerations, such as the advantage of having a large dataset, commends me to abstain from discriminating among different levels of development. The implications of this choice are further discussed in relation to the analysis in chapter 4 (see section 4.2.5).

The common practice of equating economic development with the level of GDP per capita is criticised for missing important aspects of economic development such as measures of life expectancy, literacy rates, political liberties and legal justice, poverty rates, occupational patterns, and similar demographic characteristics. But increases in GDP per capita have been shown to affect these aspects of development positively. Indeed, a range of studies have found that quality-of-life indicators generally are higher in richer countries (e.g. Barro 1996; Mauro 1995; Pritchett and Summers 1996) – stated in Huntington's terms: "all good things go together" (Huntington 1968, cited in Easterly 1999: 240). The correlation between these indicators and level of GDP per capita provides the use of the latter as a measure of economic development with validity, and it indicates that, although growth processes might produce unwanted effects, higher levels of GDP per capita in general is associated with improved life conditions. In addition, data on GDP per capita is far more easily attainable for a wide range

of countries and years than alternative measures such as those mentioned above (this will be discussed further in chapter 3), and using the growth rate of GDP ensures comparability with most of the existing studies on the subject. Therefore, in line with most studies on economic growth, I make the basic assertion that economic development is a good – in general facilitating the improvement of conditions of human life, independently of context. The crude simplification of complex matters that this necessarily entails, and the important caveats of the abovementioned literature, will be taken into account in the discussion of the analytical results presented in chapter 4.

2.2 SOCIO-ECONOMIC INEQUALITY AND ECONOMIC GROWTH: DIFFERENT PATHS OF CAUSATION

There is no universally agreed-upon answer to how socio-economic inequality within a population of a social unit should be defined. Theoretically, it can be said to refer to the degree to which the economic abilities of the rich are greater than those of the poor. For more practical ends, it can be defined as the number of people living in either extreme poverty or affluence relative to the total population.[5] For centuries economic inequality has been an important subject of analysis. One has studied both the causes and effects of inequality from different perspectives: economic, political, sociological, psychological and philosophical (Thorbecke and Charumilind 2002: 1477). However, modern macroeconomic thought and the theoretical study of the relationship between economic inequality and economic development have gone through many different phases (Galor and Zeira 1993), and as an explanatory factor of economic growth, the distribution of income "has been very much out in the cold" (Atkinson 1997: 297). Only at the end of the last century economic inequality started reappearing in economic and political studies, and during the 1990s it received renewed theoretical interest and empirical attention. While during the 1950s and 1960s the main focus tended to be on the determinants of inequality, inspired by the influential work by Kuznets,[6] inequality was now re-established as one of the main factors in explaining economic growth (Atkinson 1997).

[5] In the methodological discussion in the next chapter I will elaborate further on how inequality is conceptualised and measured in this analysis.
[6] Simon Kuznets: "Economic Growth and Income Inequality" (1955).

Several so-called growth puzzles – typically, when two quite similar countries display very different growth records – have partly spurred the renewed focus on socio-economic inequality. A common example of such a growth puzzle is described by Lucas in his article "Making a Miracle" (1993). In the early 1960s South Korea and the Philippines were similar with regard to all major economic aggregates: GDP per capita, investment per capita, average saving rates, population, urbanisation, and primary and secondary school enrolment rates. But over the following 30 year period fast growth in South Korea resulted in a fivefold increase of the output level, while that of the Philippines barely doubled. In an attempt to explain this puzzle, Benabou (1996: 11-12) and Aghion, Caroli and Garcia-Peñalosa (1999: 1615-1616) point to the fact that the two countries were actually very different on one dimension, namely, the degree of socio-economic inequality. For example, the Gini index for the Philippines was 18 percentage points higher than for South Korea, and the ratio of the income share of the 20% richest compared to that of the bottom 20%, or even to the bottom 40%, was about twice as large in the Philippines, and similar disparities characterised land ownership.

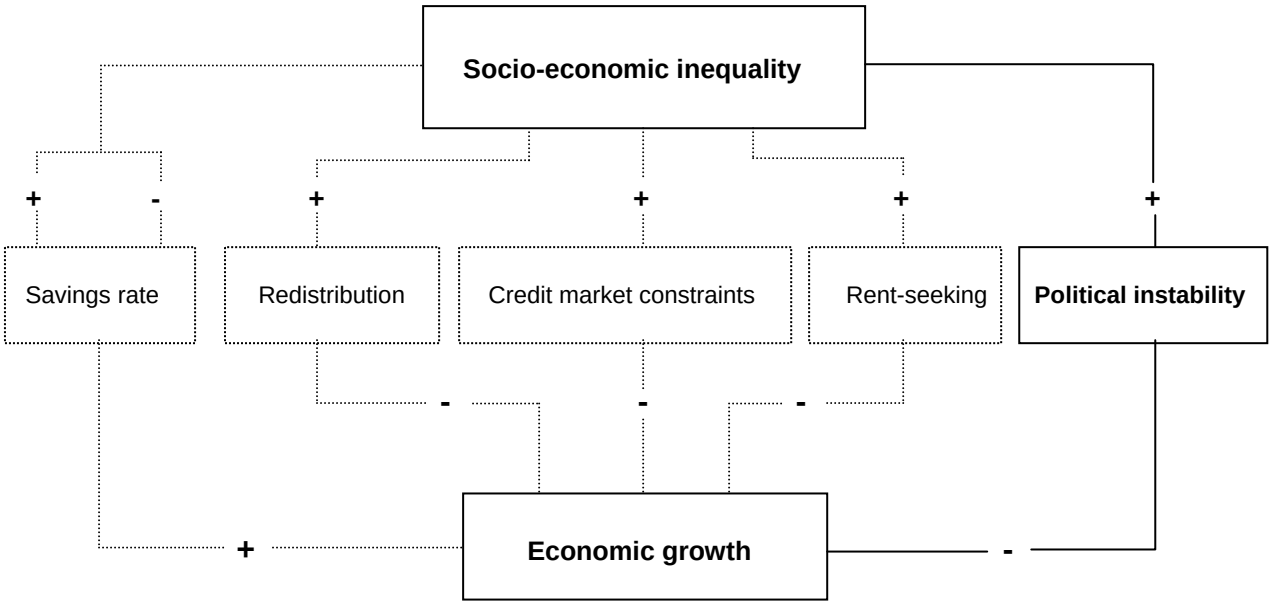
Contrary to the conventional wisdom that socio-economic inequality is good for growth,[7] the vast majority of the studies pertaining to the so-called "new growth wave" of the 1990s[8] thus claim to show that the relationship is actually negative. In his detailed overview of the literature in the mid-1990s, Benabou (1996: 13) concludes: "These regressions, run over a variety of data sets and periods with many different measures of income distribution, deliver a consistent message: initial inequality is detrimental to long-run growth". This contention, as we shall see, was not completely warranted at the time, and has been both confirmed and challenged since (Nel 2003). For example, Tanninen (1999: 1115) found "a clear negative reduced-form relationship between several "predetermined" income inequality measures and the average long-run per-capita-growth rate". Similarly, Clarke's (1995: 422) finding is that inequality is directly and negatively correlated with growth, and that, although substantially fairly small, this relationship is statistically significant. Forbes (2000), on the other hand, claims to show through her analysis that inequality is actually positively related to growth. While an enormous literature investigates the effect of inequality on growth, less attention has

[7] This assumption has been made based on the expectation that a concentration of assets will enable large-scale investments necessary for economic development (see section 2.2.1).

[8] This term refers to the renewed academic interest in examining the determinants of economic growth, especially in such factors that are not strictly economic but rather political and social, largely due to the fact that the stark differences in levels of development between poor and rich states seemed to persist (Aghion et al. 1999; Castello and Domenech 2002; Fielding 2003b).

been given to exploring the specific determinants through which inequality affects growth (Bandyopadhyay and Basu 2005: 1273). Reduced form analyses[9] such as those above-mentioned are not very enlightening in this regard. In fact, and as pointed out by Barro (2000: 8), a problem is that the different theories on the relationship between inequality and growth tend to have off-setting effects and that the net effect of inequality on investment and growth therefore are ambiguous. It is therefore much more useful to look at that part of the literature where several different paths of causation have been explored either directly or indirectly. The main ones are the savings rate, redistributive policies, credit market constraints, rent seeking and political instability. These are depicted in Figure 3.

Figure 3: Overview of hypothesised causal links between socio-economic inequality and growth



While the literature focusing on the savings rate hypothesise a positive effect of inequality on growth, the other four mechanisms imply a negative effect of inequality on growth (Castello and Domenech 2002: 187). Table 1 below gives an overview of the main existing studies on the effect of socio-economic inequality on economic growth, the paths of causation they have explored and the effect that they have reported. Before moving on to the path of causation hypothesised in this thesis, namely that of political instability, I will briefly present the alternative linkages.

[9] Reduced-form analyses investigate the direct effect of socio-economic inequality on growth and do not specify or analyse how the former is supposed to affect growth. They skip a causal step, one could say.

Table 1: Recent quantitative studies on the effect of socio-economic inequality on growth[10]

Study	***Presented effect***	***Path of causation***
Hardy (1979)	No relationship	Political instability
Venieris and Gupta (1986)	Negative	Savings rate and political instability
Cukierman et al. (1992)	Negative	Political economy
Bertola (1993)	Positive	Savings rate/ political economy
Galor and Zeira (1993)	Negative	Credit market constraints (H.D.)
Alesina and Rodrik (1994)	Negative	Political economy
Persson and Tabellini (1994)	Negative	Political economy
Birdsall et al. (1995)	Negative	Credit market constraints
Clarke (1995)	Negative	Not specified/ reduced form
Alesina and Perotti (1996)	Negative	Political instability
Benhabib and Rustichini (1996)	Negative	Political economy/ political instability
Perotti (1996)	Negative	Political instability/ credit m. constraints
Torstensson (1996)	No clear relationship	Political economy
Knack and Keefer (1997, 2002)	Negative	Rent-seeking
Deininger and Squire (1998)	Negative	Credit market constraints (H.D.)
Li and Zou (1998)	Positive	Political economy
Temple (1998)	Negative	Not specified/ reduced form
Aghion et al. (1999)	Negative	Credit market constraints
Tanninen (1999)	Negative	Not specified/ reduced form
Barro (2000)	No clear relationship/ Inverted U-curve	Not specified/ reduced form
Chang and Ram (2000)	Negative	Credit market constraints (H.D.)
Forbes (2000)	Positive	Savings rate/ not specified
Sylwester (2000)	Negative	Political economy
Easterly (2001)	Negative	Reduced form (rent-seeking)
Landa and Kapstein (2001)	Negative	Political economy
Castelló and Domenéch (2002)	Negative	Credit market constraints (H.D.)
Easterly (2002)	Negative	Rent-seeking/ credit market constraints
Panizza (2002)	Negative (weak, not robust)	Not specified/ reduced form
Banerjee and Duflo (2003)	Inverted U-curve	Not specified/ reduced form
Nel (2003)	Negative	Risk perceptions of investors
Odedokun and Round (2004)	Negative	Political instability; credit market constraints (H.D.) and the fertility rate
Bandyophyay and Basu (2005)	Positive/negative*	Not specified/ reduced form
Knowles (2005)	Negative for developing countries	Not specified/ reduced form
Frazer (2006)	No clear relationship/ country specific	Not specified/ reduced form
García-Peñalosa and Turnovsky (2006)	Positive	Not specified/ reduced form

** Positive effect in developed countries, negative in developing countries*

2.2.1 THE SAVINGS RATE

Most of the literature that studied the economic effects of inequality during the 1950s and 1960s, focused on its effect on saving. It was claimed by neo-classical economic theorists that individual savings rates rise with the level of income. That is, the marginal propensity to save is higher for the rich than the poor. Because a high level of investments is a prerequisite of rapid growth, then transfers from rich to poor reduce capital accumulation and investment,

[10] The table is constructed by the author for the purpose of this thesis and based on a review of the literature in the field.

and hence growth (Barro 2000: 8; Chang and Ram 2000: 788; Adelmann and Robinson, in Chenery and Srinivasan 1989: 951; Clarke 1995: 404) – a line of reasoning that lends legitimacy to refuting redistribution of wealth from the rich to the poor in a society (Kaldor 1978, in Birdsall et al. 1995). This conventional wisdom has been challenged from several theoretical stands during the course of the last decades. For example, Venieris and Gupta (1986) rejected the classical approach by claiming to have demonstrated that the bulk of savings is in fact produced by the middle income class and not the upper class. Others maintain that consumption is more important for growth than savings. This is in part what has led Landes (1998: 217-221) to conclude that an ideal society for development and growth would be one with a large middle class, i.e., a relatively equal distribution of economic wealth. Even though some have maintained the classical argument and found empirical support for it, most prominently Bertola (1993) and Forbes (2000), the vast majority of empirical studies have hypothesised a trade-off between inequality and growth.

2.2.2 REDISTRIBUTIVE POLICIES

One of the most studied paths of causation linking inequality to growth is that of political demands for redistributions (Temple 1998: 318). During the 1990s several path-breaking studies took as their theoretical approach, and found empirical evidence for, that skewed income distribution encourages higher taxes and redistributive policies, and in that way reduces growth (Alesina and Rodrik 1994; Benhabib and Rustichini 1996; Cukierman et al. 1992; Persson and Tabellini 1994). The logic of this approach is the following: If the mean income in an economy exceeds the median income, then a system of majority voting tends to favour redistribution of resources from rich to poor.[11] These transfer payments, such as taxes, distort economic decisions and thus lower growth (Barro 2000: 6). The political economy approach has been criticised on theoretical grounds, most notably by Dagdeviren et al. (2001) who argue that it is difficult to understand how the median voter (and those below) in developing countries can effectively overcome the powerful differential that gave rise to inequality in the first place, and force through redistributive policies (cited in Nel 2003: 625). Indeed, empirical evidence has been found that inequality can actually lead to *lower* income taxation and thus higher growth (Li and Zou 1998). More generally, the political economy approach has been criticised by several authors that have found evidence that reject the

[11] According to Acemoglu and Robinson (2000: 1191-1193), the democratising reforms in Britain, France, Germany and Sweden that took place during the nineteenth century led to a fall in inequality due to subsequent tax reforms that redistributed wealth to the poor.

median voter hypothesis. E.g., as pointed out by Deininger and Squire (1998) it is logically derived from the so-called median-voter theorem that tax rates are democratically determined, and that one would not observe the relationship in non-democratic settings. By splitting their dataset into democratic and authoritarian regimes, it is shown that this assertion does not hold, and that democratic voting does not appear to be the root of the inequality-growth link (Clarke 1995; Knack and Keefer 1997; Levine and Renelt 1992).[12]

2.2.3 CREDIT MARKET CONSTRAINTS

Credit constraints and market imperfections constitute a path through which inequality can reduce growth. Investment in human and physical capital constitutes both a source of growth at the country level and an important way out of poverty for individuals and families. Especially, the accumulation of human capital, as measured by the educational attainment of the population, has consistently emerged as an important cause of economic growth and development (Birdsall et al. 1995: 483; Castello and Domenech 2002). As large segments of the population in poor countries do not possess initial wealth, investment has to be financed through credit. This is not unproblematic, however, because of constraints in the credit market. Credit market imperfections arise as the interest rate for individual borrowers is higher than that for lenders. This is due to the lenders' lack of information about the borrowers, which implies that there are costs connected to acquiring such information and preventing default (Galor and Zeira 1993: 39). Many poor people can thus not afford to borrow, and they often lack collateral. The moral hazard associated with lending increases as the borrowing amount increases, because the incentive to default rises at higher borrowing levels, and hence tracking costs rise. Consequently, as education represent high initial costs which only pays off in the long run, limitations in the access to credit makes poor households forego human-capital investments, which would offer relatively high rates of return (Barro 2000: 6). Thus, greater inequality *increases the credit restrained share of the population* and in this way affects growth negatively (Aghion et al. 1999; Deininger and Squire 1998; Galor and Zeira 1993). On the aggregate level, countries with high inequality thus invest less in human capital and are less able to benefit from technological innovations, resulting in that they grow more slowly and remain poor (Galor and Zeira 1993).

[12] From a different point of view, Milanovic (2000) finds evidence that redistribution is higher in more unequal societies, but still rejects the median voter hypothesis because his analysis reveals that the middle class (including the median voter) does not benefit from this redistribution.

2.2.4 RENT-SEEKING

According to Hall and Jones (1999, in Easterly 2002: 6) institutions can have both a productive and a so-called predatory equilibrium. In the predatory equilibrium, resources are diverted towards seeking to seize others' assets or protecting one's own instead of spending resources on asset creation. Many authors have pointed out that a higher gap between rich and poor would tend to raise the returns to predation relative to production. As Barro (2000: 7) puts it, the participation in activities such as collective violence and crime represents a direct waste of resources because the time and energy of these participants are not devoted to productive efforts. Moreover, the threats to property rights deter investment. Through these mechanisms, more inequality tend to reduce the productivity of an economy (Alesina and Perotti 1996: 1214). A related approach is the one that focuses on the effect of social polarisation on growth. Here it is hypothesised that societies that are polarised with respect to income, ethnicity, religion, and/ or other dimensions, exhibit lower growth rates than un-polarised societies. This is because polarisation reduces social trust and increases possible gains from predatory actions by chief executives, which in turn affect growth rates detrimentally (Easterly and Levine 1997; Keefer and Knack 2002; Knack and Keefer 1997; Svensson 1998; Woo 2003; 2005).[13] Thus, polarised societies, e.g. in terms of socio-economic inequality, will tend to have higher levels of rent-seeking and hence lower economic growth.[14]

2.2.5 POLITICAL INSTABILITY

All of the above-mentioned ways in which inequality can affect economic growth have been dealt with in a wide array of studies. The path of causation that will be in focus in this thesis, however, is one that has also received a large amount of attention both theoretically and empirically, namely, political instability. When comparing different paths of causation, Perotti (1996a) found that the mechanism linking inequality to growth that received the strongest result from empirical investigation was that of political instability. Illustrating this path of causation, Stewart and O'Sullivan (1998: 28) argue that establishing an economic and social system which spread the fruits of progress widely and to all significant regions/ ethnicities/ religious groups, was central in mitigating conflict in Uganda, Sri Lanka and Kenya.

[13] During recent years, substantial empirical evidence has been found that support the notion that social trust is an important factor for economic development (e.g., Zak and Knack 2001).

[14] The approach is related to the hypothesis that political instability reduces the time-horizon of politicians in power, making it more likely that they will engage in short-term policies and rent-seeking at the expense of macroeconomic stability and long-term development policies (Alesina and Perotti 1996: 1206; Woo 2003: 394).

Similarly, Acemoglu and Robinson (2000) argue that many Western countries democratised during the 19th century as a response to the threat of revolution, which constituted a political and economic menace to the elites' position. This threat had been intensified by increasing socio-economic inequality in the population and redistributive demands. Partly induced by industrialisation, large and pronounced differences in the standards of living between the elites and the people created political unrest and made revolution more attractive to the distributional losers.

Such observations have led many to argue that "[…] redistribution of income must be ranked as one of the more meaningful reforms that a modernizing government can undertake in the interest of achieving political stability" (Muller and Seligson 1987: 444). Formulated in a synthesised way, the idea is that political instability, produced by stark socio-economic inequalities, creates an environment unfriendly to both savings and investment that, in turn, lowers growth rates (Landa and Kapstein 2001: 282). Important contributions in fields such as psychology, sociology and political science form the theoretical basis for the expectations we have about the relationship between socio-economic inequality and political instability, and the effects of political instability on economic growth has long been an important research field in economic theory and political economy. In what follows I will present the theoretical foundations for the expected causal relationships between inequality and political instability, and between the latter and economic growth, and give examples of central empirical findings.

2.3 THE EFFECT OF INEQUALITY ON POLITICAL INSTABILITY

What do we mean by political instability? As stated by Perotti (1996b: 80), the concept of political instability is rather nebulous and does not have an immediately obvious and objective counterpart. Consequently, empirical studies display a wide range of different operationalisations. As stressed by Russett (1964), it is therefore necessary to clarify what is meant by political instability. According to Alesina and Perotti (1996), political instability can be viewed in two ways: 1) as executive instability such as the propensity to observe government changes, or other kinds of regime-related political instability such as coups d'état, purges and governmental and constitutional crises; or 2) as social unrest and political violence, that is, civil society-induced manifestations of political instability. Many studies focus on *either* the relationship between socio-economic inequality and political instability *or* on the relationship between the latter and economic growth. Studies of the first type focus

primarily on political instability induced by civil society, due to the fact that what is analysed is how individuals respond to socio-economic inequality. Studies focusing on the effect of political instability on economic growth often focus in addition to regime-related political instability. This is because the channel through which political instability is thought to affect growth is primarily investment. Thus, what is relevant in these studies is the kind of instability that affects investment, which not necessarily involve violent acts – or even civil society at all.[15] In this thesis, political instability enters into the hypothesis as a mediating link between socio-economic inequality and economic growth. Hence, regime-related political instability is irrelevant in this setting as it is a kind of instability that inequality is not likely to affect. Therefore, political instability is defined here as collective unrest that arises from civil society and that has political objects as its targets. Given that such unrest often involves violent action, and in accordance with the literature on the subject, the terms *political unrest, political violence* and *political instability* will be used interchangeably in what follows.

There are of course many potential sources of conflict and political instability in a nation, such as ethnic, religious and regional disputes, discrimination and lack of political democracy. Nevertheless, the principal political contest and debate in a nation often involve a polarisation of social groups around distributional issues (Lichbach 1989: 432), and the above-mentioned sources of conflict often become precarious only when followed by such polarisation. Revolution was defined by Marx as a class struggle, and by Pareto as a circulation of elites, both placing the role of economic inequality at the centre of the field. As stated by Sen (1973: 1): "The relation between inequality and rebellion is indeed a close one". Many would argue that the general issue of inequality has been involved in all major episodes of conflict. For example, Verba and Orren state that "[t]he demand for equality has lain at the epicentre of the major upheavals that have erupted on the American political scene: the Revolution, the Jacksonian era, the Civil War and Reconstruction, the Populist-Progressive period, the New Deal and the tumultuous 1960's and 1970's" (Verba and Orren 1985: 21, cited in Lichbach 1989: 433).

A remarkably diverse literature, ancient and modern, ideological and theoretical, has coalesced on the assertion that political violence is a function of economic inequality

[15] Elite struggles for power and alternations in office in a political setting where basic rules are lacking and unpredictability prevails can produce uncertainty around property rights and economic policies, and thus disincentives for investment.

(Sigelman and Simpson 1977: 105). Indeed, income distribution was a subject of central importance to the classical economists. Aristotle (cited in Linehan 1980: 193) identified inequality as the "universal and chief cause" of instability: "Inferiors revolt in order that they may be equal, and equals that they may be superior". He asserted further that "where the middle class is large, there are least likely to be factions and dissension". Centuries later, in *The Federalist No. 10,* Madison (1787-1788) described inequality in the distribution of property as the "most common and durable" source of political faction (Easterly 2001: 317). Later still, Engels argued that political violence occurs when political structures are not synchronised with socio-economic conditions (Sigelman and Simpson 1977: 106). These contentions have remained central, and economic inequality is still regarded as a crucial factor leading to social conflict (Horowitz 2000; Schock 1996; Thorbecke and Charumilind 2002). This is forcefully expressed by Hibbs (1973: 196-198), who contends that all major cross-national quantitative studies of dissent that do not include economic inequality as an independent variable, must acknowledge specification error.

Are these seemingly strong and uniform theoretical expectations supported empirically? Many studies do find that inequality is associated with higher political instability (see Table 2). However, the finding has not remained unchallenged, and the expected strong and positive relationship between inequality and political dissent sometimes conflicts with the data (Lichbach 1989: 432). As Midlarsky (1988: 491) states: "[…] rarely is there a robust relationship discovered between the two variables. Equally rarely does the relationship plunge into the depths of the black hole of non-significance". Macculloch (2005: 93) similarly concludes that two decades of empirical research and over 3 dozen studies on the relation between inequality and conflict has produced a diverse and contradictory array of findings, and thus that the impact of inequality on conflict is still being debated.[16] Below, Table 2 lists central empirical studies and their findings, an overview that confirms Schock's (1996: 101) observation that "an uncomfortable ambiguity prevails with regard to the relationship between income inequality and political violence". The causes of this empirical inconclusiveness will be a main topic in chapter 4. Another, related problem is the simplistic manner in which most studies deal with the theoretical foundations for their hypothesised relationship between inequality and instability, and more generally for that between inequality

[16] Interestingly, Lichbach wrote almost exactly the same in 1989, indicating that the conventional hypothesis has been challenged almost for fifty years, but still without any solution: "two decades of empirical research in conflict studies have challenged the conventionally accepted view that a strong, positive relationship exists between inequality and political conflict" (Lichbach 1989: 440).

and economic growth. At the centre of the controversy lies the question of what determines each individual's support for radical change through violent means, yet the determinants of e.g. revolutionary preferences remain largely unstudied in the literature (MacCulloch 2005: 94): "The reasoning behind various EI-PC propositions – how and why economic inequality breeds political conflict – has typically been neglected" (Lichbach 1989: 436).

Table 2: Findings in the literature on the effect of inequality on political instability[17]

Study	***Presented effect***	***Type of study***
Kling 1956	Positive	Small-N, qualitative study
Russett 1964	Positive	Quantitative, cross-section
Feierabend & Feierabend 1966	Positive	Theoretical work and cross-section analysis
Runciman 1966	No relationship	Theoretical work and empirical analyses
Huntington 1968	Positive	Theoretical work
Mitchell 1968	Negative	Case study (South Vietnam)
Gurr 1970	Positive	Theoretical work
Russo 1972	No relationship	Quantitative, cross-section
Hibbs 1973	No relationship	Quantitative, cross-section
Parvin 1973	Negative (weak)	Quantitative, cross-section
Nagel 1974	Positive (curvilinear)	Case-study (South Vietnam) and quantitative, cross-section
Sigelman & Simpson 1977	Positive (weak)	Quantitative, cross-section
Hardy 1979	No relationship	Quantitative, cross-section
Weede 1981	No relationship	Quantitative, cross-section
Muller and Jukam 1983	No relationship	Quantitative, survey analysis
Panning 1983	Positive	Theoretical model
Muller 1985	Positive	Quantitative, cross-section
Muller and Seligson 1987	Positive	Quantitative, cross-section
Midlarsky 1988	Positive	Quantitative, cross-section
Muller and Weede 1990	No relationship	Quantitative, cross-section
Moaddel 1994	Positive	Cross-national structural modelling
Alesina and Perotti 1996	Positive	Quantitative, cross-section
Perotti 1996	Positive	Quantitative, cross-section
Schock 1996	Positive*	Quantitative, cross-section
Temple 1998	Positive	Quantitative, cross-section
Fearon and Laitin 2003	No relationship	Quantitative, panel
Nel 2003	No relationship/ Negative	Quantitative, panel** (Sub-Saharan Africa only)
Collier and Hoeffler 2004	No relationship	Quantitative, panel
Maccullock 2005	Positive	Quantitative, survey analysis (*taste* for revolt)

** Contingent upon regime repressiveness*
*** A significant positive relationship is found only when subjective measures of political instability are used*

How, then, is socio-economic inequality thought to cause political unrest? As Dahl (1966, in Sigelman and Simpson 1977: 125) argued, the causal chain connecting objective socio-economic conditions with actual political behaviour is long and tenuous. There are three stages in the process connecting socio-economic inequality with the outbreak of political

[17] The table is constructed by the author for the purpose of this thesis and based on a review of the literature in the field.

violence: 1) that in which discontent is generated, 2) when it is politicised, and 3) that when it is actualised in political violence. In what follows I will elaborate on the mechanisms that impel the progression of these processes. By doing so, I intend to show how socio-economic inequality can cause political instability.

A large body of cross-national research has been undertaken based on the assumption that socio-economic inequality produces relative deprivation, and according to Gurr (1970: 62), most of the literature on collective violence assumes a causal relationship between relative deprivation (or some equivalent concept) and the occurrence of violence. As stated by Alesina and Perotti (1996: 1214) in their seminal work "Income distribution, political instability and investment":

> A large group of impoverished citizens, facing a small and very rich group of well-off individuals is likely to become dissatisfied with the existing socio-economic status quo and demand radical changes, so that mass violence and illegal seizure of power are more likely than, when income distribution is more equitable.

Relative deprivation can be defined as a perceived discrepancy between a person's value expectations and his or her value capabilities (Gurr 1970: 13). Value expectations are the goods and conditions of life to which people believe they are rightfully entitled, and value capabilities are the goods and conditions they think they are capable of attaining or maintaining, given the social means available to them. Thus, by being an important source of relative deprivation which in turn produces discontent, socio-economic inequality is thought to be an indirect cause of political violence and instability. Important contributions to the theory of relative deprivation are the influential works of Davies (1962), Feierabend and Feierabend (1966), Huntington (1968), Gurr (1970) and Runciman (1966; 1972). They show that relative deprivation is decisive to the first stage of the process of political unrest, the generation of discontent, thus its theoretical implication and argumentative logic will be central in the following sections.

2.3.1 THE GENERATION OF DISCONTENT

A general contention is that revolution is not very likely in a country with a high degree of socio-economic equality and well-being (Acemoglu and Robinson 2000). Why exactly is this? In Muller and Jukam's (1983: 159) terms: "People who take part in acts of civil disobedience or political violence are discontented about something. That is a truism". The discrepancy that

constitutes the definitional core of relative deprivation, between the goods and conditions in life that people perceive that they should and could have, and the goods and conditions that they actually have or think are within reach, spurs the incipience of a disposition to aggressive action (Gurr 1970: 319; Snyder 1978: 502). This disposition is expressed here as *discontent*, a term that is central in explanations of collective violence. Discontent refers to the socio-psychological mechanisms that make people resort to violent behaviour.[18] Due to the importance of material well-being in human life, this discontent will often be closely connected to one's economic situation, and the notion of being relatively deprived is based upon an evaluation of this situation. According to Maslow's hierarchy of needs, material needs and physical well-being is fundamental to all life. As discussed in the beginning of this chapter, in the context of a modern world, this translates into assigning primacy to one's economic situation relative to less fundamental needs in Maslow's hierarchy. Referring to Cantril's cross-national survey of human concerns, Gurr (1970: 69) concludes that "[m]aterial values are clearly of greatest concern to the people of the world".[19] The survey revealed no significant difference between countries of various developmental levels.

This might lead one to think that absolute deprivation, that is, poverty, makes a person discontent. But an important postulate of the theory of relative deprivation is that discontent is not produced by poverty *per se*. Poor people might be unaware that they are deprived, or they might think that a better life is unattainable, or even undeserved. As pointed out by Weede (1981: 641), if e.g. an untouchable street-sweeper in Calcutta accepts his low status and extremely meagre income as deserved, he will not feel deprived relative to his standards of justice and expectations, however absolutely deprived he may seem by Western standards. More importantly, extremely poor people lack available resources to protest. As argued by Huntington (1968: 53), poverty itself is a barrier to instability, because those who are concerned about the immediate goal of the next meal are not in a position to worry about the great transformations of society. Supportive of this argument are the studies of Nel (2003), and Krueger and Maleckova (2003), the latter showing that poverty was not a decisive factor in explaining who participated in Hezbollah's militant wing in the late 1980s and early 1990s. Thus, according to relative deprivation theorists, it is relative – not absolute – deprivation that

[18] A term that is also frequently used, is *systemic frustration* (e.g. Feierabend and Feierabend 1966). This concept is the equivalent of discontent.

[19] When answering questions about one's hopes and fears about the future, the majority of the respondents mentioned materially related issues such as standards of living, health, technological advances, economic stability, and owning a house or land. Political participation and security were also important categories.

it is though to produce discontent and in turn aggressive political behaviour, and a prominent source of relative deprivation is the extent of socio-economic inequality in a society (Muller and Jukam 1983: 160; Schock 1996: 101). However, as resources are scarce, when some people are very rich this necessarily is at the expense of someone else (Horowitz 1985), and one could thus say that socio-economic inequality necessarily involves some degree of absolute deprivation. This has led some to believe that inequality is curvilinearly associated with instability. For example, when finding that a high degree of inequality (i.e., many poor people) was actually associated with less violence in South Vietnam, Mitchell (1968: 423) ascribed this partly to what he calls the "docility and low aspirations" of poorer peasants.

The fact that relative deprivation refers to a person's subjective notion of his own situation makes it clear that socio-economic inequality, referring to the actual distribution of economic means in a society, does not necessarily cause discontent. According to the theory of relative deprivation, it is the gap between an individual's *expected and achieved well-being*, rather than structural conditions *per se*, that ultimately produce political violence (Thorbecke and Charumilind 2002: 1486). Some theorists go so far as to claiming that objective measures of inequality are irrelevant for outbreaks of political violence unless they are accompanied by measures of subjective notions of this inequality (Sigelman and Simpson 1977; Zimmermann 1983). This might not be completely warranted, however, because *the likelihood* that people feel deprived is higher when the actual distribution of wealth in society is unequal. As expressed by Nafziger and Auvinen (2002: 156): through the demonstration effect of consumption levels of the relatively well off, high income concentration increases the perception of relative deprivation by large population segments. In addition, if measures of the subjective notion of deprivation are to be included in the analysis, serious problems of data availability and cross-national comparability would arise that could stall any attempts at testing the relationship empirically.

Social comparison and modernisation

A crucial question thus appears: when does socio-economic inequality translate into a notion of being relatively deprived? Decisive here is the potential that lies in socio-economic inequality of increasing the expectations in the population of a higher level of material well-being when such levels are exposed through the life-style of the relatively better-off. This implies that if socio-economic inequality is to translate into relative deprivation, and in turn cause discontent, it must be accompanied by individual comparison between people, because

if one is not aware of the situation of others, there is no basis on which to feel deprived. In Panning's (1983: 323) terms, "inequalities in a society affect the behaviour of its members *only if those individuals compare their own lot with that of others* and as a result become gratified or dissatisfied" (emphasis added). It therefore becomes pivotal to answer the questions: when do people compare their own situation with that of others, and whom do they compare with? Festinger's (1954) path-breaking theory of social comparison processes states that evaluating one's own abilities lies in the nature of the human being. Further, it shows that comparison with others is the main tool for such evaluation, as objective and *a priori* given standards hardly ever exist. According to Hobbes, one of the foremost political philosophers of rational political action, people seek relative advantage and not merely absolute advantage. As Taylor, (1987, cited in Lichbach 1989: 458) says: "In *Leviathan*, Hobbes seems to assume that each man seeks to maximize not merely his own payoff, but also his 'eminence', the difference between his own and other people's payoff". One could thus say that social comparison is an inherent part of human nature, and that unless people are extremely poor and spend all their time and energy on merely surviving, comparison will always take place (Popkin 1979 and Maslow 1954, in Besancon 2005: 395).

It has been argued that modernisation widens the scope and increases the intensity of comparison between people, and that it therefore facilitates the generation of discontent (Moaddel 1994: 295). Modernisation processes lead to increased social mobility and greater information flows across geographical and social borders, and in this way people are exposed to new ways of life and experience changes in the lives of individuals or groups around them. As a consequence, people become aware that these new ways of life are within reach, and that misery can be avoided (Panning 1983: 323). A prominent advocate of such mechanisms is Huntington. In his seminal book *Political Order in Changing Societies* (1968) he relates manifestations of instability with economic inequality through modernisation. Huntington (1968: 57) asserts that as a society modernises, social mobilisation makes people aware of, and resentful towards the fact that they are relatively deprived: "The influx of new ideas calls into question the legitimacy of the old distribution and suggests the feasibility and the desirability of a more equitable distribution of income". This "diffusion" of equity and social justice in modern times, Toqueville's "dread (of) the insistent and immediate demand for equality in our lives", implies a strong relationship between inequality and instability because the spread of norms of equality makes invidious comparisons universal and all forms of subordination illegitimate (Lichbach 1989: 437).

It is commonly assumed that people compare their situation with others that are relatively close to themselves on the point of comparison. For example, a poor farmer will not compare himself with the rich, urban elite, but with people that belong to his own, or approximate, socio-economic strata (Festinger 1954: 120; Nagel 1974: 455; Runciman 1966: chap. 10). This has led many (e.g. Nagel 1974) to assume that the shape of the relationship between inequality and instability resembles and inverted U. Even though inequality is expected to cause grievances, collective political violence is assumed to occur at intermediate levels of inequality. This is because if the distribution in a society is close to equal, people are not discontented, and if the distribution is very unequal, they do not rebel because comparison with economic strata too far from their own does not take place.[20] I contend, however, that the social mobility associated with modernisation processes incites comparison across such strata by making it evident that it is possible to improve one's life conditions. Indeed, Festinger (1954: 137) emphasises that comparison across highly deviant groups can take place if the area of comparison is sufficiently important and central in a person's life. In light of the central place of material needs and economic well-being, it follows that the most decisive point of comparison between individuals and groups is the economic conditions under which they live.

In addition to social comparison, people are also likely to compare their present situation with how they have fared previously. As modernisation is associated with increases in material well-being (at least temporarily) a situation can arise where many people have experienced some improvement to their own situation, and have "gained an appetite for more" (Mitchell 1968: 437). In instances when they are worse off, or expect to be worse off, than they have been, this can lead to discontent and in Davies' (1962) terms, *a revolutionary mood*. This process is prominently expressed by Soule (1935, cited in Gurr 1970: 114, emphasis added):

> When the people are in their most desperate and miserable condition, they are often least inclined to revolt, for then they are hopeless... Only after their position is somewhat improved and they have *sensed the possibility of change,* do they revolt effectively against oppression and injustice. What touches off insurrection is hope, not the lack of it, rising confidence, not bleak suffering.

To sum up, comparison with others is natural to human beings and its scope and intensity is enhanced by modernisation processes. Due to the universal primacy of material well-being and the social consequences of modernisation, comparison across socio-economic strata is

[20] Alternatively, as mentioned earlier, that they are very poor and thus too occupied with ensuring the next meal. A U-shaped relationship, that is, that political violence will be most frequent at either high or low levels of inequality has also been hypothesised (e.g. Davies 1962).

likely to take place. This tendency to compare one's own situation with that of others exposes to people the degree of socio-economic inequality in society, and where this inequality is wide-spread, this becomes a prominent source of relative deprivation. Relative deprivation, defined as a discrepancy between people's expectations and how they really fare, in turn produces discontent. But as Muller (1985: 52) states, discontent is only a necessary condition for collective violence, discontent must be mobilised.

2.3.2 THE MOBILISATION OF DISCONTENT

Recall that there are three stages in the process that transforms socio-economic inequality to the actual outbreak of political unrest. Following the generation of discontent are the second and third stages in the process of political violence, that is, the politisation and concrete manifestation of discontent in acts of protest behaviour. Together these stages can be referred to as the mobilisation of discontent. This section will present two different approaches to explaining incidents of political unrest. While one emphasises the importance of relative deprivation-induced discontent, the other focuses on each individual's participation decision, contending that discontent is not sufficient to produce unrest. This is the so-called rational actor approach. The section concludes that despite the objections put forth this latter perspective, socio-economic inequality is likely to be positively related to the outbreak of political unrest.

When discontent is politicised, it means that people blame the political system and its actors for the state of things and that discontent therefore is directed at political targets. While the general term *collective violence* refers to a situation where members of a collectivity take violent action against others, political violence can be defined as "all collective attacks within a political community against the political regime, its actors or its policies" (Gurr 1970: 157). It has been argued that insurgency is more likely if the less advantaged can identify the perpetuators of their poverty and suffering, and that the probability of manifest conflict increases when discontent is focused on governments or related political targets (Nafziger and Auvinen 2002: 156; Snyder 1978: 503). According to Gurr (1970: 159), discontent is a necessary precondition for political violence, but as he states: "the relationship is not exact. […] Depending on peoples' perspectives on violence and politics, relative deprivation-induced discontent may be either focused on, or deflected from, the political system".

How relevant is the distinction between collective and political violence? Discontent is likely to be politicised to the extent that a regime affects, or is thought to have the potential to affect, the lives of most or all of its citizens (Gurr 1970: 181). As there are strong and widespread doctrines in the modern world, stating that the political system is and should be responsible for social problems, discontent tends to be politicised. Indeed, a glance at the internal conflicts that have taken place during the contemporary history, confirms Gurr's (1970: 178-179) assertion that nearly all large-scale outbreaks of collective violence are politically motivated and directed towards political targets, and the distinction between collective and political violence is therefore more theoretical than real. These terms will therefore be used interchangeably in what follows. The literature on explanations of the actual outbreak of civil unrest can be divided into two strands, and I will now present the first approach to explaining how discontent is mobilised.

Relative deprivation and the mobilisation of discontent

What determines whether discontent actually materialises as violent action? Lichbach (1989: 459) points out that the synaptic link transforming discontent into behavioural dissent has received too little attention. One approach focuses on relative deprivation-induced discontent, and emphasises the importance of the grievances brought about by economic conditions. This perspective treats the outbreak of political unrest as a function of discontent: if people are discontented enough, they will engage in acts of protest. Due to the importance of economic concerns in people's lives, it is assumed that these to a high degree influence individuals' participation decision in political violent activity. Supportive of this notion is a study of Nel (2003) on the effect of inequality on political instability and growth in sub-Saharan Africa. He finds that the more a government is committed to welfare spending, the less prone to instability a polity would be. According to relative deprivation theorists, the disposition to collective violence thus depends on how badly societies violate socially derived expectations about the means and ends of human action. As stated by Festinger (1954: 124), if there is a discrepancy between this person's evaluation of his own "performance" and the performance of those with whom he compares himself with, then he will act to reduce this discrepancy.

In his article "Toward a Theory of Power and Political Instability in Latin America" Kling (1956: 33) describes a situation where inequality is so severe that violence is seen as unavoidable: land ownership is so heavily concentrated that no individual not already possessing great tracts of agricultural land can reasonably hope to achieve wealth through

farming, and of the possible sources of enrichment, only government is open to competition. Political office provides such a unique source of gain that "large segments of the population are prepared to take the ultimate risk of life, in a revolt, in a coup d'état, to perpetuate a characteristic feature of Latin American politics – chronic political instability". History provides us with plenty examples of other groups experiencing such situations: the Parisian workers in 1848, Mexican peasants in 1910, the German *Kleinburgertüm* in the 1920s, Hungarians in 1956 and black South Africans in the 1960s (Gurr 1970: 92). Therefore, relative deprivation theorists hypothesise that the potential for collective violence is greatest when people feel sharply deprived with respect to their most deeply valued goals.

The critique of the relative deprivation approach

The results from empirical studies testing the relative deprivation hypothesis, however, are highly mixed (Snyder 1978: 504). This could be due to the fact that there are several qualifying conditions that affect whether discontent actually translates into collective violence. As Sen (1973: 1) points out: "[…] that a perceived sense of inequity is a common ingredient of rebellion in societies is clear enough, but it is also important to recognize that the perception of inequity […] depend[s] substantially on possibilities of actual rebellion". Formal modellers have thus shown that a relationship between economic inequality and discontent can be derived logically, but have not linked discontent with violent behaviour (Lichbach 1989: 454). The most extensive critique of the explanations of political violence that are founded on relative deprivation and discontent, comes from an alternative perspective often referred to as the *rational actor perspective*, or the *political opportunity approach*. The critics belonging to this stance hold that much of the relative deprivation literature tends to overlook factors that explain the varying degree to which discontent is organised and resources mobilised for sustaining political insurgency (Muller 1985: 48; Snyder 1978: 503).

According to Festinger (1954: 117), "[a] person's cognition (his opinions and beliefs) about the situation in which he exists and his appraisals of what he is capable of doing (his evaluation of his abilities) will together have bearing on his behaviour". Muller and Weede (1990: 625) similarly remind us that individuals do not rebel in a vacuum. "If they are rational actors, they will take the costs of rebellion into account". According to Blainey's (1975, in Besancon 2005: 396) theory of victory, there must be a belief in the ability to win in order to choose violent rebellion. While the relative deprivation perspective assumes that actors will compare their returns from economic activities with the economic returns of others, the

rational actor perspective assumes that actors will compare their economic returns with *their returns from dissident activities* (Lichbach 1989: 462). This theoretical stand thus focuses on each person's decision to participate, characterising it as being based on an expected utility calculation that balances potential personal gains and losses (MacCulloch 2005: 94). It therefore hypotheses that opportunity costs influence the relationship between relative deprivation and collective violence (Collier and Hoeffler 2004; Skocpol 1979; Tilly 1978).

A basic assumption made by the rational actor approach is that the sources of discontent that lead to political violence are inherent in all societies and that the occurrence of such violence depends on the political opportunities and constraints of the immediate political environment, rather than variations in levels of economic inequality or the intensity of economic discontent (Schock 1996: 104). For example, Collier and Hoeffler (2004) find that socio-economic conditions are unrelated to politically motivated violence, and that it is the opportunities for rebellion that decide whether such violence occurs. They assert that an important motivation for participation in collective violent action is *greed*, that is, each individual's conscious evaluation of the potential gain he or she will have from participation. Here the so-called free-rider problem emphasised by Olson's collective action theory is illustrative (Lichbach 1989: 463). Since everyone will get the fruits of collective dissent regardless of their participation, why should anyone bear the personal costs of participating in dissent? The participation decision is thus affected by a considerate evaluation by rational actors of the costs and opportunities associated with dissident activities. The following question thus arises: when does this evaluation prevent discontent from materialising in political violence?

Resource mobilisation and regime repressiveness

It has been claimed that it is not discontent per se that causes collective violence but the organisation of discontent (Feierabend and Feierabend 1966: 251; Muller and Seligson 1987: 426; Snyder 1978: 505; Weede 1981: 640). That is, acts of protest behaviour depend on "[…] the extent to which dissident groups are able to acquire control of the resources necessary to develop strong and effective organizations for the purpose of obtaining collective goods" (Muller 1985: 48). Even though inequality creates *incentives* for the deprived to rise up, people do not necessarily have the resources, the ability to communicate, or the symbolic power generated by social capital that it takes to do so. These resources are sometimes controlled by the ruling elite, who will use them to suppress any attempted resistance (Rodriguez 2000, in Nel 2003: 629). Jenkins and Perrow (1977) claimed to have demonstrated

in their study "Insurgency of the Powerless" that farmers were unable to organise in political protest actions without a surrounding political environment that facilitated insurgency, as for example sponsorship by established organisations or the absence of collective controls and resources used to suppress it. According to Skocpol (1979: 112-157), the peasant revolts crucial in the French, Russian and Chinese revolutions did not take place because of the inequalities in these societies but rather because the peasants had sufficient autonomy from the local landlords to enable them to mobilise collectively. Some scholars have gone even further by proposing that inequality in fact can affect insurgency negatively. For example, in his study from 1968 of South Vietnam, Mitchell (1968) claimed to have shown that economic inequality decreased political instability, because high economic inequality is associated with powerful elites that have the resources and the incentives to use social, economic and political power to repress political dissent.[21]

The dissident groups' capacity to undertake collective action is thus conditioned by the context in which they find themselves, and a key determinant of this context is the repressiveness of the political system or regime (Eisinger 1973: 11; Snyder 1978: 505). A basic premise of rational action theory of political violence is that the costs of rebellion are a function of the degree of regime repressiveness and the behaviour of the government in response to protest (acts of coercion or negative sanctions) (Muller and Weede 1990: 625). As stated by MacCulloch (2005: 95), "[w]henever high levels of inequality is accompanied by a repressive military, tastes for revolt may not be manifested in terms of observable rebellions". As found by Blomberg (1996), increases in military defence counteract domestic collective violence. An important factor conditioning the balance between discontent and utility calculation is thus the degree to which the regime is expected to use repressive means to counteract violent protest behaviour (Feierabend and Feierabend 1966: 251; Schock 1996: 104; Tilly 1978: 49). Muller (1985) hypothesises that regime repressiveness is related to collective violence in a curvilinear, inverted U shaped fashion: an extremely repressive, closed regime will leave little opportunity for dissident groups to engage in collective protests and violence, while in an open democracy political violence should rarely occur, as most groups can effectively pursue their interests through peaceful channels of political participation. According to this logic then, it is under a regime structure of intermediate

[21] Paranzino (1972) took a second look at Mitchell's study, and questioned Mitchell's interpretations of what his land-tenure variables actually measure. In addition to showing a significant regional effect, Paranzino's findings reversed those of Mitchell for almost all measures of tenure inequality (Hardy 1979: 211).

repressiveness that collective political violence should be most likely (Huntington 1968: 275; Schock 1996: 105). This assertion has resulted in a wide array of studies of political violence incorporating a variable measuring regime repressiveness in their analysis, contending that repression increases conflict except when it is severe (Collier and Hoeffler 2004; Hegre et al. 2001; Muller and Seligson 1987).[22]

Concluding about the effect of inequality on political instability

What role then, should be ascribed to socio-economic inequality in explanations of the mobilisation of discontent and the outbreak of collective violence? The rational actor approach has been criticised for the tendency to treat economic discontent as ubiquitous and as a constant (Schock 1996: 105), while in fact it is not: Levels of socio-economic inequality do vary both across space and time. Further, there is reason to question the central contention of the rational actor perspective that individuals assume a rational weighing of different alternatives and ultimately chooses the one that is best for him or her (Lichbach 1989: 460). How credible is this assumption? Are people really calculating personal gains versus costs, and based on this, making a conscious choice how to behave, at the same time as enduring deep frustration and a notion of being unjustly deprived? Indeed, a weakness of the rational actor approach is that it is not able to account for events that are apparently spontaneous and that do not have an organisational base (Snyder 1978: 506).

According to Gupta (1990: 108), a pervasive view among experts and observers of incidents of collective violence is that a large part of participation decisions are not rational:

> The image of a violent, wild-eyed revolutionary, with his guns blazing, or blowing himself up in a suicidal attempt to accomplish his impossible political goals, provides us with a typical example of an emotive, if not an irrational, act.

Frustration, anger and desperation drive people to engage in protest and violent behaviour, not a calculative weighing of benefits against costs. This contributes to the case for the relevance of inequality, relative deprivation and discontent in explanations of violent behaviour. And even if one admits to the notion that humans generally are rational, this does not necessarily mean that they are selfish. According to Gupta, (1990: 115) "[w]e must be prepared to recognize the absurdity of the viewpoint that any goal other than the one of selfishness constitutes irrationality. Hence, we should broaden the definition of human rationality to

[22] As Muller (1985) points out, there seems to be truth in Machiavelli's dictum that a leader should either embrace or crush his opposition.

include the desire to maximize group interest". Thus, even if it means taking personal risks and bearing heavy costs, people might act on their discontent as much out of frustration on their own behalf as on their group's behalf, *and they might still be acting rationally*.

The following line of reasoning also lends support to the notion the socio-economic inequality is important in explanations of the outbreak of political violence: As inequality increases, the *pool of possible participants* in conflict behaviour increases Lichbach (1989: 436). As expressed by Muller (1985: 53): "It seems plausible to expect that in societies with high inequality, where the distribution or scope of discontent is presumably widespread, discontent is more likely to be mobilized somehow, ceteris paribus, than in societies with low inequality." Therefore, as the mobilisation of discontent is correlated with the extensiveness of inequality, such that when inequality is pervasive some mobilisation is almost bound to occur, then the relationship between inequality and political violence should be "*positively accelerated*". Hence, despite the criticisms, the idea that economic inequality is related to a high potential for violent challenges to the state or its policies remains a plausible assumption (Schock 1996: 99). On the basis of these insights I conclude that socio-economic inequality can be expected to increase the level of political instability, and the following hypothesis can be formulated: (H_1): *Socio-economic inequality produces political instability.*

I will now move on to discussing the second link in the relationship subjected to analysis in this thesis, namely that between political instability and economic growth, and answer the following question: Why is the presence of political instability expected to reduce the rate of economic development?

2.4 THE ECONOMIC EFFECTS OF POLITICAL INSTABILITY

Both the theoretical expectations and empirical findings are much more complex and ambiguous regarding the relationship between inequality and instability compared to that between instability and growth. This section will therefore be briefer than the previous one. Let us begin by comparing two illustrative cases, namely, the records of Argentina and Japan on the dimensions of political instability and economic development. In the beginning of the previous century Argentina was one of the world's wealthiest countries, with an income per capita in the top twenty of the world in 1960. The same year Japan had a per capita income below Iraq, Ireland and Argentina and was not even among the top twenty-five. During the

course of the last forty years Japan has experienced one of the fastest growth rates of the world. Argentina, on the other hand, has often come close to economic collapse. While Argentina has had a history of political instability, with several coups d'etat and much political violence, Japan has been a model for political stability. It has not been marked by political violence, and it has had the same political party in office continuously from 1960 until 1993 (Alesina et al. 1996). Does this conjunction contain a pattern of causality? Kuznets (1966: 451) states that "[o]ne could hardly expect much economic growth under conditions of economic turmoil…and unpredictable changes in regimes". Likewise, Gupta (1990: 250) concludes his comprehensive study of the effect of political instability on economic growth by saying that "[t]he presence of political violence poses a serious threat to the pursuit of peace and prosperity of a nation". This strong assertion has been confirmed empirically by an almost infinite array of studies (see Table 3 below), giving support to the notion that political instability is harmful for growth.

Table 3: Findings in the literature on the effect of political instability on economic growth[23]

Study	***Finding***	***Type of study***
Hardy (1979)	Negative *and* positive	Quantitative, cross-section
Venieris and Gupta (1986)	Negative	Quantitative, cross-section
Londregan and Poole (1990)	No relationship	Quantitative, panel
Barro (1991)	Negative	Quantitative, panel
Fosu (1992)	Negative	Quantitative, cross-section
Mauro (1995)	Negative	Quantitative, cross-section
Alesina et al. (1996)	Negative	Quantitative, panel
Alesina and Perotti (1996)	Negative	Quantitative, cross-section
Benhabib and Rustichini (1996)	Negative	Formal model
Blomberg (1996)	Negative	Quantitative, panel
Perotti (1996a)	Negative	Quantitative, cross-section
Auvinen (1997)	Negative	Quantitative, panel
Svensson (1998)	Negative	Quantitative, cross-section
Persson and Tabellini (in Taylor and Woodford 1999 1c)	Negative	Formal model
Easterly (2001)	Negative	Quantitative, cross-section
Feng (2001)	Negative	Quantitative, cross-section
Campos and Nugent (2003)	Positive*	Quantitative, panel
Abadie and Gardeazabal (2003)	Negative	Case study (time series)
Fielding (2002; 2003a; 2003b; 2004)	Negative	Case studies (time series)
Nel (2003)	Negative	Quantitative, panel (Sub-Saharan Africa only)
Woo (2003)	Negative	Quantitative, panel
Gwartney et al. (2006)	Negative	Quantitative, panel

** Dependent variable: investment, not growth*

[23] The table is constructed by the author for the purpose of this thesis and based on a review of the literature in the field.

Casting a glance at the world scenery as of the beginning of 2008, several examples stand out that illustrate the underlying mechanisms of this relationship. Perhaps the most current is the situation that has followed the Kenyan presidential election on December 27th 2007.[24] In the wake of the alleged electoral victory of the incumbent president Kibaki and the subsequent outbreak of mass political violence, headlines such as "Experts warn of downturn in economic growth", "Violence too costly" and "Kenya's economic future at stake" have marked the Kenyan media. Professor Terry Ryan, consultant to the Treasury and a member of the Monetary Policy Advisory Committee, says to the newspaper Daily Nation on January 11th that the impact of the political violence "could reduce growth from seven to between two and 4.5 per cent" (Wachira 2008, Daily Nation, January 14th). What lies behind these poor prospects? The direct economic consequences of the violence include the loss of property and vandalising of utilities of both businesses and civilians. But conflict is not only immediately damaging to those directly involved. Conflict also results in large development costs: "We must realize that, under conditions of social unrest, political upheaval and wars, economic modernisation is impossible" (Fei 1997, cited in Stewart and O'Sullivan 1998: 2). As described in an editorial in the Daily Nation on January 14th, the tourism industry, decisive to Kenya's economy, was affected severely: "Tourists do not go into areas seen as unstable or dangerous. […] Empty hotel rooms mean workers laid off. Twenty thousand jobs have already been shed. The Kenya Tourism Board estimates that if the trend continues, 120.000 jobs will have been lost by March. And this does not include jobs that will be lost in the ripple effect from other industries that service the tourist sector" (Daily Nation 2008, January 14th).

INVESTMENT AND THE IMPORTANCE OF PROPERTY RIGHTS

The Kenyan example highlights a central point to the argument that instability is detrimental to growth, namely, its negative consequences for investment. Indeed, who would want to invest their money in a firm that produces hotel furniture or establish a restaurant in Nairobi these days? Both domestic and foreign investors are likely to prefer a stable political environment. As stated by Kuznets (1966: 451), "[…] clearly some minimum of political stability is necessary if members of the economic society are to plan ahead and be assured of a relatively stable relation between their contribution to economic activity and their rewards". In a similar vein, Levine and Renelt (1992: 958) contend that it is not surprising that "countries that experience a high number of revolutions and coups tend to be countries that

[24] Other current examples of highly politically unstable areas would be Afghanistan, Pakistan, Sri Lanka, Zimbabwe, Sudan, Palestine, Lebanon, Chechnya, Kosovo and Haiti.

invest less of their resources domestically than countries with stable political environments". Investment has long been established as a decisive driving force of growth (Barro 1996: 7), and by reducing the incentives to save and invest, political instability reduces growth (Alesina and Perotti 1994: 359).

The reason why political instability is thought to decrease investment incentives is because it creates *uncertainty*. Collective violence represents an immediate threat to property, and in addition the probability of the government being overthrown is higher when social unrest is widespread. This makes the course of future economic policy and the protection of property rights more uncertain (Alesina and Perotti 1996: 1214). It has also been pointed out that in a situation of political instability, the probability that the government will repudiate contracts increases (Thorbecke and Charumilind 2002: 1484). The uncertainty associated with such circumstances increases the risk of capital loss. Because a high risk discourages investment in physical capital, the availability of factors of production is reduced. Furthermore, a high risk also raises the uncertainty of future rates of return of investment projects, and thus the likelihood of loan defaults rises. In this way the cost of capital increases, resulting in reluctance among investors to take productive economic initiatives. They might even "exit" the economy altogether (Alesina et al. 1996: 191). As formulated by Feng (2001: 273): "An impending political crisis puts investors' decisions to send their money into the market on hold". Therefore, capital flight and luxury consumption among the local capital holders based on foreign import[25] might be characteristic of politically unstable situations (Fosu 1992: 830-831; Venieris and Gupta 1986).[26]

For the economic actors that choose to stay in the market, political instability may influence the timing and organisation of the production process negatively. The uncertainty associated with political instability can produce rampant stops and starts in investment projects – resulting in the inability of the economy of attaining the optimal growth path (Fosu 1992: 831). In addition, by reducing the planning time-scale for business leaders, uncertainty has negative consequences for projects that have longer-term growth goals. It can also prevent state expenditure in sectors that would have long-term growth-enhancing consequences, such

[25] Such luxury consumption takes place at the expense of the investments necessary to create economic development.
[26] Supportive of this expectation is Fielding's study of the economic effects of the Palestinian *Intifada*, where he found that violence lead to Israeli capital flight (Fielding 2004).

as investing in human capital and infrastructure (Nel 2003: 614) (cf. section 2.2.3). Due to mechanisms such as these, Zak (2000, in Feng 2001: 274) notes the following:

> In countries where regime change is common, investment tends towards the liquid and speculative, leaving these countries with low-investment and low-productivity industries (such as textiles and mining), rather than capital-intensive enterprises requiring investment in the plants and high-technology equipment that provide the foundation for advanced industrialisation.

Several studies find a robust negative correlation between investment rates and various measures of political instability (Barro 1991; Easterly and Rebelo 1993; Levine and Renelt 1992; Venieris and Gupta 1986). For example, through several studies on the economic consequences of the political instability associated with the Palestine-Israeli conflict, Fielding finds that violence of all kinds following the uprising (*Intifada*) among Palestinians between 1987 and 1993 depressed Israeli investment substantially (Fielding 2003a; 2003b; 2004). He also finds similar effects of the political instability associated with the Apartheid period in South Africa (Fielding 2002). Studying the terrorist activity in the Basque Country from the 1970s to the 1990s, Abadie and Gardeazabal (2003) find a strong negative effect of the terrorist activities on economic prosperity. What these examples depict, are situations of violence and threat to life, property and property-rights; a powerless political authority, incapable of establishing security; and a high degree of political uncertainty as to what the future might bring.

It becomes clear that the disincentive to investment that political instability represents ultimately rests on the fear of losing property (Barro 1991: 437). As formulated by MacCulloch (2005: 93): "A fundamental requirement of market economies is the security of ownership claims to property. Yet throughout history existing claims to property have been regularly challenged by revolts". The importance for investment of the rule of law and contractual certainty in general, and the protection of property rights specifically, is emphasised by several scholars (e.g., Barro 1996; Landes 1998; North 1981; Svensson 1998; Tornell and Velasco 1992) . The occurrence of attempted or successful revolutions or coups indicates a propensity to abandon the rule of law and therefore, in principle, a threat to established property rights (Alesina and Perotti 1996: 1214). Investors will be reluctant to risk their capital when property rights are weak and poorly protected, because they fear that their returns will be appropriated by others (Gwartney et al. 2006). As stated by Landes (1998: 32): Why should anyone invest capital and labour in creating or acquiring wealth that one risks

losing? When property rights are poorly enforced, a wedge is created between the marginal product of capital and the rate of return that can be privately appropriated by investors (Svensson 1998: 1318). Thus, by reducing the certainty that property rights will remain protected, and in that way creating disincentives to invest, political instability is negatively related to growth.

Despite the obvious logic and intuitiveness of this argumentation, the importance of political stability for growth has not always been recognised in economic science. But, as emphasised by several scholars, the importance of political variables for economic development should not be underestimated. Gradually economists have started to take seriously the possibility that political instability can have a large and lasting impact on the economy (Fielding 2004: 465). On the importance of political factors Feng (2001: 288) writes:

> I am tempted to argue that their negative effects on private investment far exceed the unwholesome impact of monetary or fiscal policies on private investment. Political determinants do matter in private capital formation, and their effects are fundamental and far-reaching.

But is instability always bad? As noted earlier, sometimes collective, political violence can bring an authoritarian regime to collapse, and, as such, instability can have long-term positive consequences. Alesina et al. (1996: 192) state that government change might be viewed favourably by economic agents if the current government is incompetent or corrupt and its successors are seen as an improvement. Similarly, Mancur Olson (1982, in Fedderke et al. 2001) has suggested that where political instability disrupts rent-seeking activities, it may have a positive impact upon growth. However, in such cases any positive effects on growth will most likely not be immediate, but will rather set in after a new government has been established and its pro-growth policies have actually been realised. Thus it is not instability in itself that affects growth positively, but the establishment of policies by the new government that promote investment and economic activity. In sum, the following hypothesis is generated: (H_2): *Political instability affects economic development negatively.*

2.5 SUMMARY – THEORETICAL DISCUSSION

This chapter has attempted to show how socio-economic inequality can reduce economic growth by producing political instability. It started out by discussing the importance of studying what affects the rate of economic growth, and concluded on this point that in general

increases in the level of GDP per capita are indeed associated with an increased quality of life. An overview of the different hypothesised paths of causation between socio-economic inequality and economic growth was then presented, before attention was turned to the causal link subjected to analysis in this thesis, namely that of political instability. Central contributions in the literature dealing with the relationship between inequality and political instability were presented, followed by a theoretical discussion of the underlying mechanisms of this hypothesised causal pattern.

Socio-economic inequality is expected to produce political instability due to its central role in breeding relative deprivation – and as such, its potential to generate discontent. This potential becomes effective in the face of social comparison, a human characteristic that is magnified by various processes related to modernisation. According to the rational actor perspective, the actual manifestation of discontent in acts of collective unrest is conditioned by individual evaluations of costs versus gains from participating in such acts. However, people do not always act based on such a rationalistic evaluation of possible personal gains. They might also take into account the possible gains of their collectivity, in which case it is sometimes rational to participate in acts of collective protest. Further, by breeding discontent, socio-economic inequality *increases the pool of potential participants* in protest behaviour. Therefore, and independently of the objections put forth by rational action theorists, socio-economic inequality is hypothesised to affect the level of political instability positively.

The chapter closed by discussing the relationship between political instability and economic growth. By creating uncertainty around future protection of property rights and increasing the chances of capital loss, political instability discourages investment. On these ground, then, political instability is hypothesised to have a detrimental effect on growth. Put together, we thus see that socio-economic inequality, by breeding political instability, is thought to have a negative effect on economic growth.

3 VARIETIES OF KNOWING: METHODOLOGY AND MEASUREMENT

In this chapter I will discuss the methodological issues related to my analysis. The research questions are: 1) does socio-economic inequality affect economic development negatively by producing political instability, and 2) why do previous studies report such diverging results? The *methodological approach* adopted to answer these questions is the quantitative approach, and the *methodological tool* is regression analysis. In the first part of this chapter I discuss the methodological issues in detail and relate them to the sample selection and analytical design, and in the second part I describe how the variables are measured and operationalised.

3.1 METHODOLOGY

3.1.1 A QUANTITATIVE APPROACH

In studying the current research question I adopt a quantitative approach. While qualitative analyses are decisive in generating theory and constituting the basis on which hypotheses are made, quantitative studies in turn test these hypotheses with the objective of either modifying them or providing them with empirical support (Ragin 1987: 55). Qualitative studies are *case-oriented* in that they focus one or a few particular social units or historical events and emphasise their complexity and uniqueness (Ragin and Zaret 1983). Quantitative, large-N analyses, on the other hand, aim at an over-arching scientific goal: the ability to generalise - making statements about the relationship between phenomena that are independent of time and space. Quantitative analyses are thus *variable-oriented*: they seek replacing the proper names of social units or historical events with variables. The inclusion of a large number of units enables statistical control, which is necessary to make statements about the relationships *with a sufficient degree of certainty* (Landman 2003: 26-27). However, large-N studies face a trade-off. The contextual uniqueness of each case forces the renouncement of the accuracy and complexity associated with the qualitative approach, in Sartori's (1970) terms, it implies moving up the *ladder of abstraction*. This is important because it touches upon what a researcher can hope, and claim, to achieve when applying a quantitative approach. One cannot make accurate statements about specific cases on the basis of statistically established relationships between variables. Awareness of this will mark the interpretation of the results of the analysis in the next chapter.

3.1.2 MULTIPLE REGRESSION AS THE ANALYTICAL TOOL

Compared to simple correlation analysis, regression analysis offers additional possibilities that are decisive for the two main goals of quantitative analyses. The first goal is related to determining causation. When employed correctly, regression analysis enables estimation of the *direction* and *strength* of a causal relationship. As social phenomena are determined by a range of causes, *multiple* regression analysis will be used to examine the relationships between the variables in this analysis. This technique incorporates several explanatory variables, and enables estimation of the isolated effect of each explanatory variable in addition to their relative strength, by holding constant the effect of the other variables (Lewis-Beck 1980: 47). As opposed to the physical sciences, however, where relationships are law-like in nature, the relationships in the social sciences can only be expressed in probabilistic terms (Lijphart 1971: 684). This means that an error term (e), representing the variation in the dependent variable that is due to causes that are unknown (or not explicitly determined), has to be included. In my analysis this leads to the following mathematical expressions:

$$Y\,(\textit{instability}) = a_0 + b_1X_1(\textit{inequality}) + b_2X_2 + b_3X_3 + \ldots + b_kX_k + e \qquad \text{(equation 1)}$$

$$Y\,(\textit{growth}) = a_0 + b_1X_1(\textit{instability}) + b_2X_2 + b_3X_3 + \ldots + b_kX_k + e \qquad \text{(equation 2)}$$

We commonly assume a linear relationship between the dependent and independent variables, even though this is not always correct.[27] Therefore, regression analysis implies producing a line in a scatterplot that in the best possible way describes the statistical tendency in the data we are analysing (Skog 2004: 220). This is done using the principle of Ordinary Least Squares (OLS), which produces the line with the smallest possible sum of distances, between the line and the data points. In other words, OLS minimises the sum of squared errors (SSE).[28] Regression analysis helps achieving a second goal in quantitative studies, namely, making generalisations, and thus also predictions, about real world relationships. Prediction rests upon the premise that the sample subjected to analysis is drawn from the population of interest in a random way,[29] and based on this sample we can make inferences about the population through

[27] This assumption can generally be made in the face of the following four circumstances: 1) there are empirical grounds to assume linearity; 2) the linear specification is most parsimonious; 3) theory might be too weak to suggest what an alternative specification should be; and 4) empirical investigation of the data might fail to suggest a clear alternative to the straight line model (Lewis-Beck 1980: 13). However, when analysing the data, the linearity assumption will be examined in relation to the specific model adopted here.

[28] The sums are squared to avoid positive errors cancelling out negative errors (Lewis-Beck 1980: 14).

[29] In studies where aggregates such as countries (and perhaps at different points in time) are the analytical units, one could argue that it is difficult to randomly draw the sample, as the population is necessarily limited to a fixed

significance testing. Following the norm on this field (Lewis-Beck 1980: 31), I base the significance testing in my analysis on the t-statistic. A two-tailed 95 % confidence interval is constructed around the slope estimate,[30] and if the value of zero does not fall within this interval (above 1.96/ below -1.96 on the t-distribution), the null-hypotheses is rejected at the .05 level of significance, i.e. with a confidence of 95%. When determining whether the relationships are significantly different from zero, I apply this level of significance.[31]

When conducting a regression analysis, the following five assumptions are made: 1) there is no specification error. That is, the relationship between the dependent and independent variables is linear and no relevant explanatory variables have been left out and no irrelevant variables included; 2) There is no measurement error, meaning that the variables are accurately measured; 3) The error term is homoskedastically and normally distributed and is not auto-correlated; and 4) There is no correlation between the independent variables and the error term (Lewis-Beck 1980: 26; Skog 2004: 236-257). In addition comes a fifth assumption, namely the absence of multicollinearity, which we have when two or more independent variables are approximately linearly correlated (Kennedy 2003: 48-49). In connection with the analysis in the next chapter the different assumptions will be discussed and any possible violation will be dealt with. However, some violations are expected *a priori* due to 1) the panel design of the data and 2) the endogenous nature of the analytical structure. These violations will therefore be discussed in the following sections.

3.1.3 PANEL DESIGN

SAMPLE SELECTION

The sample subjected to analysis in this study, that is, the choice of countries and years, is to a large degree determined by the data availability for the inequality variable, as this is the measure with the most limited data extension. To enable the inclusion of as many observations as possible on inequality, the time range in my dataset extends from 1950 to 2004, and the number of countries included is 188. All regions of the world are well

number of countries (and years). However, the events preceding the values that the different countries exhibit on variables such as inequality, instability and level of GDP have arguably taken place in a stochastic manner and *could have* resulted in different outcomes (Midtbø 2000: 59-60).

[30] A two-tailed test is applied to factor in the possibility that the direction of causality of some variables on the dependent variable(s) might not be as expected.

[31] However, as there is no given answer to what level of significance is appropriate, I do not treat the 5% level as an absolute, but rather an approximate indicator of statistical significance. Variables that are significant at the 10 percent level of significant will therefore also be reported in the analysis.

represented. This results in an unbalanced panel, as the coverage on the different variables varies in terms of which countries and years are included. Selecting a sample on the basis of data-availability raises the question of representativeness, due to the fact that data availability is known to be correlated to level of national development. Some scholars have even suggested that an index of data availability could serve as a *measure* of societal modernity (Sigelman and Simpson 1977: 115).

Another problem of studies based on country-level data is that countries are so different that their comparability can be questioned (Abadie and Gardeazabal 2003). Many methodological theorists warn against "comparing the incomparable" and measuring and analysing a phenomenon across inherently different cases (Russett 1964: 444). The inclusion of countries from all world regions and different levels of development creates problems of measuring variables in a consistent and accurate way both across countries and over time. It thus constitutes a risk of violating the principle of measure validity (Muller and Jukam 1983: 161).[32] Further, it increases the risk of disregarding relevant explanatory variables at the intra-country level (Dogan and Kazancigil 1994: 41-42). On the other hand, one could argue that the comparison of inherently different cases is necessary to establish *exactly what it is about these cases* that cause the differences they display, and that the insight that studies of different cases yields is necessary for generalisation. For example, Barro claims that it is impossible to get an accurate understanding of long-term effects of different factors on growth studying the experience of one or a few countries (Barro 2000). These problems of comparability will be taken into account when interpreting the results in the analysis in the next chapter.

THE PANEL DESIGN'S ADVANTAGES – AND PROBLEMS

As opposed to cross-sectional and time series analyses, the panel design incorporates both dimensions. This has two important advantages. Through a panel design one is able to test whether a relationship is truly general, that is, whether it holds across space and time. Combining the time and space dimensions also facilitates the identification and measurement of effects that perhaps would not appear in pure cross-sectional or time series analyses. Further, to be able to estimate the statistical significance of a causal relationship, one depends on a sufficient number of freedom degrees.[33] When only one dimension is included the

[32] E.g., developing countries tend to have large measurement errors in national accounts (Barro 2000: 10-11).
[33] The concept of freedom degrees refers to the number of observations relative to the number of variables (Skog 2004: 163).

number of observation is necessarily limited, and restricted data availability for certain countries and early time periods accentuates this problem. Over-determination of explanatory systems is therefore a danger to one-dimensional designs: as the number of societies, cultures and political units is highly limited and the number of relevant variables is very high, the number of freedom degrees becomes too low to allow consideration of all relevant factors (Przeworski and Teune 1970: 30). In Bartolini's (1993: 160) terms:

> [I]n order to control for the validity of empirical generalizations the safest way to proceed is in a contrast between cross-sectional and cross-temporal results. It is therefore not simply a question of one or the other; rather, research designs should always try to observe *both* types of variance.

The panel design also has the ability to analyse dynamic behaviour, which the cross-sectional design can tell us nothing about, and by creating more variability, a panel design alleviates the multicollinearity problem (Baltagi 2005: 4-6; Kennedy 2003: 302).

A common objection to the panel design is that the inclusion of both the space and time dimension unavoidably increases the number of missing values in the sample. However, using an unbalanced panel counteracts this problem, as it allows for the inclusion of values that are only available for some countries and/or for some years. The inclusion of the space dimension increases the risk of observing a heteroskedastically distributed error term, which means that the variance in the error term is dependent upon the values on the X-axis, that is, the independent variable (Midtbø 2000; Skog 2004: 246). As heteroskedasticity renders the significance testing unreliable, White's heteroskedasticity corrected covariance matrix is included in the final models to correct for this. Further, non-stationarity is common in time series and panel data analyses, and leads to spurious results. In variables that are non-stationary, the mean and variance depend on time, which invalidates the t-statistic and the estimation of the explanatory power of the model, the R^2 (Gujarati 2003: 797). The Phillips-Pearon unit root test will be applied in the analysis to examine whether non-stationarity is present. The most serious problem associated with panel data structure, however, is violation of the assumption that the error term is not auto-correlated (Kennedy 2003: 140).

Autocorrelation in the error term is due to the fact that many variables display little variation across time on each unit in the analysis, which leads to a situation where the error term related to one observation is correlated to the error terms of other observations, and thus violating the fourth regression assumption (Lewis-Beck 1980: 28). There is disagreement on how much

autocorrelation can be present without constituting a problem, but it is usual to draw the limit at .30, or 30 %. Common ways of dealing with autocorrelation is to control for its effect by including an autoregressive parameter or the lagged dependent variable at the right hand side. Here, the latter will be done when the estimated autocorrelation is above .30. Including the lagged dependent variable is not an uncontroversial practice. It has been criticised for "eating up" all the variation in the dependent variable (Wilson and Butler 2003, in Gudbrandsen 2005: 39), and for leaving any remaining autocorrelation un-dealt with (Plümper et al. 2005). Nevertheless, other scholars such as Beck and Katz (2004) recommend the inclusion of a lagged dependent variable, and point out that one can estimate any remaining autocorrelation after the lagged dependent variable has removed its part. I will therefore report estimated autocorrelation when presenting the regression results in the next chapter, and whenever this is below .30 it will be considered unproblematic.

The second major advantage of the panel design is that it facilitates estimation of the effect of unobserved variables. With repeated observations on each unit one is in a better position to account for unmeasured effects (Petersen, in Hardy and Bryman 2004: 331). In any cross-section there are innumerable unmeasured explanatory variables that affect the "behaviour" of the units in the analysis. Also, there may be time-series variables that affect the units uniformly but differently in each time period. In a pure cross-sectional analysis these omitted effects cannot be accounted for. Many would argue that this cross-sectional heterogeneity is the normal state of affairs, as there are so many unknown variables determining *Y*. Nevertheless, their influence results in a different intercept for each unit, and their omission thus causes biased estimation. Therefore, applying the standard OLS technique is inappropriate "unless the influence of these omitted variables (embodied in the different intercepts) is uncorrelated with the included explanatory variables" (Kennedy 2003: 303). With a panel design there is a way of improving estimation that allows for different intercepts, that is, applying OLS to either the Fixed Effects Model (FEM) or the Random Effects Model (REM) (Petersen, in Hardy and Bryman 2004: 331; Kennedy 2003: 303-305). I will now discuss their differences and explain my choice of model.

CHOICE OF MODEL: FIXED OR RANDOM EFFECTS?

The difference between the FEM and REM lies in their assumptions about the error term. Contrary to the REM, the FEM assumes that the error term is not randomly related to the independent variables or to country-specific effects. Thus, it assumes that the effect of the

omitted variables on the dependent variable is non-random. If true, this violates a premise of regression analysis and renders the statistical output unreliable. Thus, the Fixed Effects Model includes a dummy variable for each group (country) to take into account the group-specific effects. In this way one is able to control for the non-random effects of omitted group-specific, time-invariant variables in the analysis (Gujarati 2003: 636-652). There are, however, important drawbacks to the FEM, some of which the Random Effects Model is designed to overcome (Kennedy 2003: 304). First, there is the "degrees of freedom"-problem. The introduction of a large number of dummy variables reduces the degrees of freedom, and the more groups (countries) in the analysis, the more dummy variables. (Gujarati 2003; Kennedy 2003: 303-305). The principal drawback, however, is that the FEM complicates the inclusion of other time-invariant variables because it will not be able to estimate the impact of these variables (Gujarati 2003: 646; Petersen in Hardy and Bryman 2004: 338). Plümper et al. (2005: 333) point out that if a level effect of at least one variable or a time-invariant variable of theoretical interest exists, then the inclusion of country dummies is problematic because it suppresses level effects.

The debate about which model is preferable is still going on, and thus there is no universally agreed-upon answer. However, despite its attraction, there is a major qualification of the Random Effects Model that makes it applicable only in certain circumstances (Kennedy 2003: 305). The assumption of the REM that the error term is randomly related to the independent variables is a strong assumption, and if it does not hold, the REM produces biased estimates. In my analysis the chance that there are unspecified variables captured by the error term which are non-randomly related to specific countries or to the explanatory variables, is relatively high, and this makes the use of the REM inappropriate. There is also a statistical tool that can be used to determine whether the Random Effects Model or the Fixed Effects Model should be applied, the Hausman Test.[34] The null-hypothesis of the Hausman Test is that the results obtained by the two models are not significantly different. In that case, the REM is more efficient. The alternative hypothesis is that the efficient model is biased, implying that the FEM is preferred. The values produced by the Hausman test when applied to the multivariate models here, indicate strongly that the assumption of randomness is unfounded. Therefore, due to both substantial and econometric objections, the Fixed Effects Model will be applied in this analysis.

[34] This test was developed by J. A. Hausman in 1978 and presented in "*Specification Tests in Econometrics*" in volume 46 of *Econometrica* the same year.

When it comes to the above-mentioned drawback to the FEM, only one is relevant. In my dataset the freedom-degrees problem is not pronounced since the extensive panel structure including a sample of 188 countries implies a modest loss of freedom degrees. The second problem, on the other hand, may be more serious. There are theoretical reasons to assume that several variables that are consistent across time, and only vary between countries, are relevant explanatory variables in this analysis. The inclusion of these variables is rendered impossible by the Fixed Effects Model, and consequently one of the regression assumptions might be violated. In the evaluation of the analytical model in the next chapter, I will discuss whether relevant explanatory variables have been excluded and the consequences this might have.

3.1.4 THE ENDOGENEITY PROBLEM

The two-equation system of my analysis brings additional challenges into the thesis. In the first equation political instability is the dependent variable, and in the second equation the dependent variable is economic development, measured by the rate of economic growth.[35] At this point a serious problem arises: Causality in the relationship between political instability and growth can run in both directions (Abadie and Gardeazabal 2003: 114; Alesina and Perotti 1994: 355). That is, the dependent variable in the second equation, economic growth, is at the same time theoretically assumed to be an explanatory variable in the first equation. It has conventionally been argued that economic growth promotes political stability (Abadie and Gardeazabal 2003; Barro 1991), and intuitively this is not far-fetched. Economic hardship causes discontent in the population, which can lead to an increased threat of violence (Acemoglu and Robinson 2000: 1183). This assertion is held by relative deprivation theorists such as Gurr (1970) and Hibbs (1973) (cf. section 2.3.1), in addition to theories of systemic frustration and Marxist theories of rebellion (Schock 1996:112-113). On the other hand, it has also been argued that growth may have a destabilising effect, as it can produce social dislocation and increased socio-economic disparities (Olson 1963). Independently of whether the effect of growth is positive or negative, there are thus reasons to assume that growth should be included in analyses explaining political instability.

However, including growth in the first equation as an explanatory variable creates a problem of endogeneity, which violates the regression assumption that there is no correlation between

[35] In addition, an alternative measure, the yearly rate of change in the level of investment (*INVESTC*), is introduced to test the robustness of the results. This will be explained further in section 3.2.3.

the explanatory variables and the error term (Gujarati 2003: 719; Kennedy 2003: 180). To deal with this, many studies apply a simultaneous equations model using a so-called Two Stage Least Squares system (2SLS). In a simultaneous equations model instrumental variables are used to create predicted values which replace the right-hand side endogenous variables in the equations.[36] To investigate whether simultaneity is present in the different models I introduce a simultaneity test (details on this test are presented in Appendix C). This makes it possible to decide in which cases it is appropriate to use a simultaneous equations model. If the test reports simultaneity, a simultaneous equations model as described above is used. When it does not, a recursive model is employed.[37]

When specifying the recursive models, some considerations must be dealt with. First, due to the theoretical expectation of causality running in both directions between instability and growth, the safest thing to do to ensure that the right-hand side variables are truly exogenous would be to lag these variables before including them in the recursive models. In that way one would be completely sure that the estimation is not biased due to any simultaneity that the test was not able to capture. Further, in the case of the instability equation, there are reasons to believe that in addition to having an immediate effect on political instability economic growth is likely to have *a delayed* effect on instability. For example, some of the consequences of a lower rate of economic growth will not be manifested and felt by people until some time has elapsed, for example massive job lay-offs and higher living costs. On the basis of this, a lagged version of *GROWTH* is used in the recursive models where instability is the dependent variable. As for the equation where economic development is the dependent variable,[38] however, the effects of instability are expected to be immediate. As explained in the previous chapter, incidents of collective unrest and violence disrupt investment decisions and consumption patterns as they take place. The effect on the growth rate will thus appear within weeks or months, not years. Therefore, emphasising these theoretical considerations, instability is not lagged in the growth equation.[39]

---

[36] When applying a simultaneous equations model, the particular equation must be *identified*. An equation is identified when "the numerical estimates of the parameters of a structural equation can be obtained from the estimated reduced form coefficients" (Gujarati 2003: 739). To achieve this, the so-called order and rank conditions must be fulfilled - which they are in this analysis (see Appendix C).

[37] A recursive model simply indicates that the explanatory variables are *de facto* exogenous.

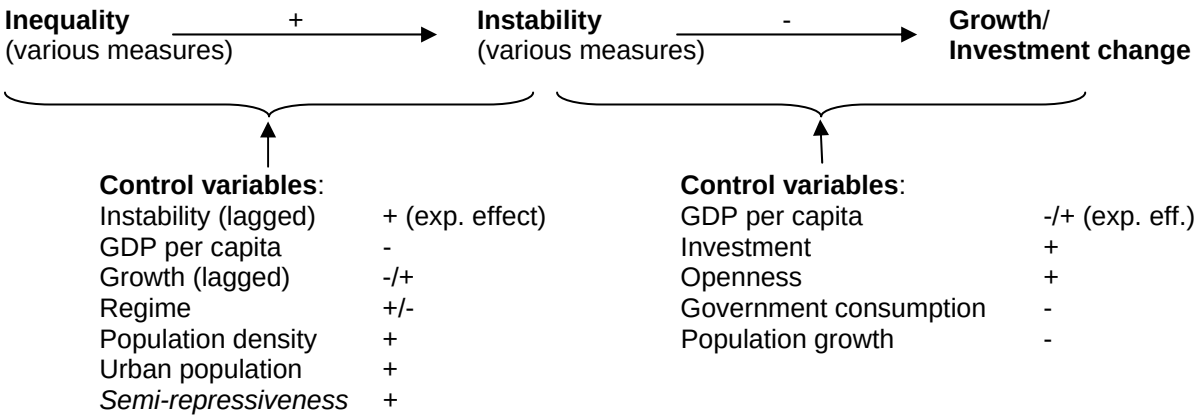
[38] For reasons of simplicity, this equation will hereafter be referred to as "the growth equation" even though it includes the models where the change in the investment share of GDP (*INVESTC*) is the dependent variable.

[39] To make sure that this is the right choice, the lagged effects of instability and the non-lagged effects of growth were explored. In the case of instability, it appeared that even though the l-year lagged effect is occasionally significant in the growth equation, it is the immediate effect of instability that is most important. This effect is not affected by the introduction of the lagged variable. As for the growth variable, it is the lagged version that is

3.2 MEASUREMENT

I will now elaborate on how the different variables are measured and operationalised. Below, Figure 4 sums up the variable relationships subjected to analysis in this thesis, the control variables introduced in each equation, and the hypothesised directions of causality. As was explained in section 2.3 and 2.4, inequality is thought to have a positive effect upon political instability, and this, in turn, is hypothesised to have a detrimental effect upon growth. The background for including the control variables listed in Figure 4 will be given in section 3.2.4.[40]

Figure 4: Overview of the hypothesised variable relationships and the control variables included:



3.2.1 MEASURING SOCIO-ECONOMIC INEQUALITY

As touched upon in the previous chapter, there are no agreed-upon answers to how inequality should be defined and the complexity and vagueness of this concept makes measuring it a challenge. Consequently, inequality has been measured differently across time, space, and scientific branches, but also within each branch (Lichbach 1989: 441-442). In the following section I will therefore elaborate on how inequality is measured in this analysis. Some scholars have employed the share of the middle class of the total income of the population as a measure of inequality (Alesina and Perotti 1996; Easterly 2002), others have looked at the ratio of the income shares of the first decile to the tenth decile of the population, and yet others compare the share of the richest 20 percent to that of the poorest 40 percent (Nel 2003).[41] The most common measure, however, is the so-called Gini coefficient.[42] It is an

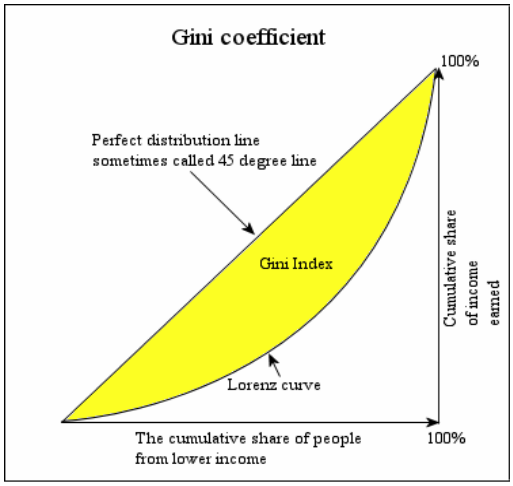
most significant in the recursive instability model: the non-lagged effect of *GROWTH/ INVESTC* is actually never significant at the 5 per cent level.

[40] Details on the variables are given in Appendix A.

[41] See Table 4 below.

expression of how the total income of a society is distributed among its members, and varies between 0 and 1 (alternatively 0 and 100). To explain how it is calculated one can draw upon a graphical representation of the Lorenz curve[43]:

Figure 5: The Lorenz curve and the Gini coefficient[44]



The horizontal axis measures the cumulative proportion of the population starting from the poorest and ending with the richest, the vertical axis measures the cumulative proportion of income (or consumption) accruing to the corresponding unit on the horizontal axis (UNU/WIDER 2007: 17). The Gini coefficient is the simply the area between the 45 degree line and the Lorenz curve (multiplied by two): the smaller, the more equal a distribution, the larger, the more unequal. In line with most empirical studies of income inequality, and because it is the measure that is most widely covered by the large-N data collections, I use the Gini coefficient as the main measure of socio-economic inequality.[45]

In 1997 The World Institute for Development Economics Research at the United Nations University (UNU-WIDER) initiated the project "Rising Income Inequality and Poverty Reduction: Are They Compatible?". In connection with this, data on inequality was compiled.

---

[42] The Gini coefficient was developed by the Italian statistician Corrado Gini and published in 1912.
[43] The Lorenz curve was developed by Max O. Lorenz in 1905 for representing income distribution.
[44] The figure is collected from http://www.singaporeangle.com/2007/01/non_sequitur_economics_ii_the.html.
[45] An important aspect of variation in the literature regarding how inequality is measured, concerns what economic sector is the basis for the inequality measure. As stated by Sen (1992: 12): "[i]nequality of what?" The most important distinction is the one between inequality of land or of income (Lichbach 1989). Over the years it has become more common to use the latter. The distribution of land holdings is becoming less relevant as modernity spreads and urban populations grow. Muller and Seligson claimed already in 1987 that it is income inequality that is of greatest importance in studies of political instability, as urban populations are crucial in the mobilisation of discontent (Muller and Seligson 1987: 427).

As the data set grew larger UNU-WIDER decided to make it publicly available in 2000 under the name World Income Inequality Database (WIID). Over the years updates have been made, the last, WIID2b, was published in May 2007. WIID2b contains 4982 Gini observations, and to a less extent data on quintile and decile income shares, gathered from altogether 228 different sources.[46] These in turn build upon different primary sources and surveys, making the WIID2b a so-called secondary dataset. Due to this, the data in WIID2b vary on several different aspects: a) what segments of the populations are covered; b) what the unit of analysis is; c) what weights are employed when the income sharing unit is an aggregate (family or household); d) what income definition has been used and e) what source the data are collected from. The use of secondary sources poses the problem of comparability and accentuates the need for a thorough and appropriate documentation of all aspects of the data: their primary source, what criteria have been used when gathering them and what exactly it is that they measure. The critics of the use of secondary sources point in particular to the problem of insufficient documentation (Atkinson and Brandolini 2001; Pyatt 2003; Székely and Hilgert 1999). Therefore, when updating the first version of the World Inequality Indicators Database (WIID1), one has taken this objection into account and aimed at minimising the problems of secondary datasets by offering extensive information on the different observations, enabling the researcher to assume an informed and sober use of the data.[47]

The WIID2b dataset contains not only different kinds of inequality data from a whole array of different sources, *the data also overlaps in many instances*. Of the complete WIID2b dataset of 4982 Gini observations, only 1560 remain when counting only single observations.[48] In an effort to maximise the number of observations on the Gini variable in my dataset, I have thus included 1560 Gini observations from the WIID2b dataset. The process of reducing the dataset to including only single observations was intricate and laborious, and only possible after an extensive and considerate evaluation of the three aspects of variation: 1) definition of income; 2) unit of analysis and different weighing systems applied to them; and 3) the different data sources. To illustrate what sort of challenge it represents to anyone who wishes to use the dataset, Table 24 in Appendix D shows how Argentina's Gini coefficient in 1961 is

[46] The coverage varies across different countries and years, making the dataset an unbalanced panel.
[47] In addition, the WIID datasets offer quality ratings covering all observations, based on criteria that are described in Appendix D.
[48] This is after having removed all observations that do not cover all areas, population segments and age groups. Some observations were based only on e.g. rural or urban areas, or only on the economically active part of the population, or on certain age groups. For the purpose of this thesis, these observations were deleted from the dataset.

reported in the WIID2b. As shown here, there are 14 different Gini observations for Argentina in 1961. Appendix D also contains a detailed description of how the selection scheme was constructed to guide the choice of which observation to include when several were available for the same country-year. Most importantly, disposable income is the preferred income definition,[49] and person-weighted household data are preferred as units of analysis. These choices are in line with the recommendations of the Canberra Group[50] and based on qualitative considerations (The Canberra Group 2001: 37-38).

The WIID dataset is by far the compilation of data on income inequality with the greatest coverage both in terms of time and space, for the first time making possible relatively large-N studies including developing countries (Nel 2003: 612). Most of the literature studying income inequality uses datasets that are both very old and much less extensive in coverage, some of which form part of WIID. Although comparability with the findings of these studies might be partly sacrificed as a consequence, the fact that this analysis uses the most recently updated and improved WIID2b nevertheless represents an important improvement compared to existing studies in that it increases the number of observations substantially and offers the information necessary to ensure awareness regarding how the data are treated.

As an endeavour to test the robustness of the findings, I include in addition another common measure of income inequality, namely the income share of the middle class, more specifically the share of the third and fourth quintiles of the population.[51] This variable, (*MIDCLASS*), also taken from the WIID2b dataset, contains 928 single observations after having removed duplicates using the same procedure and selection scheme as described for the Gini variable. If the analyses produce similar results independently of whether inequality is measured using the Gini coefficient or the income share of the middle class, the findings are strengthened. The

[49] In an effort to reduce the bias that arises due to the fact that substantial parts of the economy in many less developed countries does not primarily consist of incomes, inequality is measured on the basis of consumption in these countries when such data are available (see Appendix D for details on this point).
[50] The Canberra Group on Household Income Statistics was formed with the aim of improving national statistics on household income distribution and inequality, and the quality of international comparisons in this area. The group was organised in 1996 and completed their work in 2001 with a final report and recommendations. In addition to various countries, the following institutions participated: Economic Commission for Latin America and the Caribbean (ECLAC), Statistical Office of the European Union (Eurostat), Inter-American Development Bank (IADB), International Labour Organization (ILO), Organisation for Economic Cooperation and Development (OECD), Luxembourg Income Study Group at the Centre for Population, Poverty and Public Policy Studies / International Networks for Studies in Technology, Environment, Alternatives, Development (CEPS/INSTEAD), United Nations Statistics Division, World Bank and the Economic Commission for Europe.
[51] The quintile group shares express the share of total income going to each fifth of the population ordered according to the size of their incomes.

1560 and 928 observations that are included in *GINI* and *MIDCLASS* vary on several of the dimensions of variation in the WIID2b dataset. To make sure that the inclusion of observations that are measured differently does not affect the results in the analyses in any significant way, I subject these two inequality measures to several sensitivity tests. Thus, four alternative versions of *GINI* and *MIDCLASS* are constructed, which consist of a smaller data sample that is more uniform with regard to what kind of data is included.

First, *GINIINC* and *MIDINC* exclude gross income measures and non-adjusted units of analysis altogether. This reduces the number of observations from 1560 (*GINI*) and 928 (*MIDCLASS*) to 917 and 679, respectively. Further, the variables *GINICON* and *MIDCON* consist only of consumption and expenditure data. However, the exclusion of income data reduces the data set substantially (to 348 Gini and 181 middle class observations). I therefore constructed two alternative variables, *GINIIN* and *MIDIN*, which exclude consumption and expenditure data instead of income data. In this way I can test whether the *exclusion* of consumption/ expenditure data alters the results in any significant way with an N that is large enough to avoid the problem of too few freedom degrees (*GINIIN* has 1273 observations and *MIDIN* 778). Finally, I include two versions that combine the different exclusion criteria of the former versions: *GINISM* (664) and *MIDSM* (530). These exclude gross income measures, consumption and expenditure data and non-adjusted units of analysis.[52]

3.2.2 MEASURING POLITICAL INSTABILITY

The plan of this study is to test the hypothesis that economic inequality affects economic development negatively by causing an increase in political instability. I have explained how economic inequality is measured and will now proceed to outlining the measurement and operationalisation of political instability. In part due to their different research questions, the existing literature displays a wide range of different operationalisations: the frequency of coups d'etat; the frequency of revolutions; the extension of politically motivated violence and the existence of guerrilla groups; political demonstrations or strikes; the degree of weakness of the constitution (the frequency of changes made to it); the frequency of changes in the executive; and the *perceived* risk of such phenomena. Table 4 below lists what measures both of political instability and socio-economic inequality that central studies employ:

[52] I do not pretend here that the list of sub-samples is exhausted. I could admittedly have create even more versions of the *GINI* and *MIDCLASS* variables and run more sensitivity tests. However, I conclude that 10 different inequality measures are enough to serve the purposes that the scope of this paper allows.

Table 4: Measure differences across studies on the effect of inequality on instability[53]

Study	*Inequality measure*	*Instability variable*
Kling 1956	Relative shares in land holdings (UN - Department of economic affairs)	Government turnovers, revolutions, uprisings, violence, coups d'etat
Russett 1964	Gini index and relative shares in land holdings	Instability of personnel, internal group violence, internal war, stability of democracy
Feierabend and Feierabend 1966	An index of systemic frustration measuring exposure to modern lifestyles not available to one self	Index containing various measures of internal conflict
Mitchell 1968	Owner-operated land as a percent of all land, and the coefficient of variation of the distribution of land-holdings by size	Degree of government control
Russo 1972*	The Gini coefficient	Degree of government control
Hibbs 1973	Social structural imbalances: the ratio of educational level to economic development level, and the ratio of urbanisation to economic development level	Factor-analysed data on six types of mass anti-system events
Parvin 1973	The Gini coefficient (Kuznets 1963)	Number of deaths resulting from group domestic violence per 1 million
Nagel 1974	The Gini coefficient	Degree of government control
Sigelman and Simpson 1977	The Gini index (Paukert 1973)	Hibb's measure of political instability
Hardy 1979	The Gini index (Paukert 1973)	Riots, armed attacks, deaths from political violence, political strikes
Weede 1981	Top 20 % income share (Paukert 1973, Ahluwalia 1974)	Amed attacks and deaths from political violence
Muller and Jukam 1983	Variuos variables measuring subjective discontent	Index of aggressive political participation
Muller 1985	Income share of upper quintile	Deaths from domestic political violence (Jodice and Taylor 1983)
Muller and Seligson 1987	The Gini coefficient (land distribution), and the income share of the upper 20%	Deaths from political violence per 1 million
Midlarsky 1988	A systematic departure of the pattern of holdings of one societal sector relative to the pattern of holdings of another	Deaths from political violence
Muller and Weede 1990	Average life expectancy	Political violence death rate
Moaddel 1994	Income share of the top 20%	Index comprising riots and deaths, demonstrations, assassinations and government sanctions
Alesina and Perotti 1996	Share of the middle class (Jain 1975)	Index comprising assassinations; deaths; coups d'etat or coup attempts; and authoritarian regime
Perotti 1996	Share of the 3rd and 4th quintile, and share of 3rd quintile alone	Index comprising assassinations; deaths; coups d'etat or coup attempts; and authoritarian regime, and Gupta's instability index (1990)
Schock 1996	Ratio of the total value added in the manufacturing sector to salaries in the manufacturing sector (World Bank data)	Total deaths from demonstrations, political strikes, riots, armed attacks, and assassinations (Taylor and Jodice 1983)
Temple 1998	Income share of the middle class (Deininger and Squire 1996)	Assassinations; Perotti's sociopolitical instability index (1996)
Fearon and Laitin 2003	The Gini coefficient	Civil war
Nel 2003	Gini index; expenditure share of bottom 40%; expenditure share of top 20%; and expenditure share of 3rd and 4th quintile (WIID version 1 (2000))	Index that combines propensity towards coups d'etat with actual events of elite and civil instability; and a subjective measure of instability (survey)
Collier and Hoeffler 2004	Gini coefficient on income inequality and land inequality (Deininger and Squire 1996)	Civil war
Maccullock 2005	The Gini index (Deininger and Squire 1996); and the 90/10 ratio (Luxembourg Income Study)	Preference for revolt (survey results)

[53] The table is constructed by the author and based on a review of the literature in the field.

As discussed in section 2.3, the measures of political instability can roughly be divided in two broad categories: regime-related instability and instability induced by civil society. Due to the nature of the research question of this thesis, it is the latter kind of political instability that is relevant here.[54] Several different measures of this kind of political instability are included in this analysis, for a number of reasons: First, as argued by Moaddel (1994: 283), it is impossible to justify the choice of one single measure – which should we choose, and for what reasons? It is also evident from the great variety in the literature (cf. Table 4) that political instability can manifest itself in several different ways and cannot be captured by a single variable (Woo 2003: 395). As pointed out by Sigelman and Simpson: "[t]he primary problem we faced in measuring political violence was less one of locating suitable data than of choosing an apt measure from among the rich variety of candidates" (Sigelman and Simpson 1977: 113). Most studies therefore include several measures of political instability, and sometimes construct indices of political instability. Secondly, using different measures enables comparisons of this study with a greater number of existing empirical studies, and thirdly, by running the analysis with different measures of instability the robustness of the analysis is tested: similar results across the different measures strengthens the reliability of the results.

The Cross-National Time Series Archives dataset compiled by Arthur Banks and collaborators is a widely used data source on political instability (e.g. the seminal work of Barro 1991), which offers data on a wide range of measures that are extensive both in time and country coverage. Of the various measures of instability offered in Banks' dataset, I focus on the following: assassinations, general strikes, guerrilla warfare, riots, revolutions and anti-governmental demonstrations. [55] However, as Dogan points out in the article "The use and misuse of statistics in comparative research": "[s]ingle isolated indicators are often misleading. […] By compounding various indicators in an index, the sociological significance of statistical data could be enhanced"(Dogan and Kazancigil 1994: 48-49). In line with this argument, and with many existing empirical studies on the subject (e.g. Alesina and Perotti 1996; Nel 2003), I thus construct a conflict index, *CINDEX*, which consists of the sum of the

[54] It has also been argued that using objective measures of political instability is misleading, and that it is the subjective measure, that is, investors' perceptions of political uncertainty that determine the investment rate (Mauro 1995: 690). A study by Nel (2003) supports this notion empirically as he finds that perceived instability decreases growth but not "real" instability measures. However, as the existence of a link between inequality and *perceived* instability is rather dubious, such measures are not included.

[55] As discussed in the previous chapter, measures such as coups d'état, governmental crises, purges and cabinet changes, are not included in the analysis, as they refer to kinds of instability that is primarily *regime-induced*, while what this study focuses upon is instability induced by *civil society.*

above-mentioned instability measures. However, the use of indices is not unproblematic. It mixes measures that tap different kinds of conflict events, and their aggregation can complicate the theoretical interpretation of analytical results (Hardy 1979: 212). Therefore, I run in addition the analysis with all the individual components of *CINDEX*. Thus there are a total of 7 different measures of political instability that enters into the analysis.[56] The extension of the 7 variables measuring political instability in my dataset adds up to 8255 country-years each. The large coverage is especially important since the number of observations on my inequality variables is quite small and highly dispersed across countries and years.

3.2.3 MEASURING ECONOMIC DEVELOPMENT

To measure economic development I follow a traditional approach by focusing on the yearly change in percent of Gross Domestic Product (GDP), in other words, the growth rate (see section 2.1). This variable consists of data gathered from Penn World Table (PWT), version 6.2, a widely used and acknowledged data source with extensive coverage both in terms of countries and years, and which many of my studies of reference employ.[57] PWT reports data on three different measures of real gross domestic product (GDP) per capita. One is measured in current prices and two in constant prices. One of the latter, *RGDPCH*, uses a price chain index with the base year changed each year and it is adjusted both annually, to capture price changes and cross-nationally, to reflect purchasing-power parity. According to Feng (2003: 4), this is the preferable measure of GDP. It is also the basis for the growth measure provided by PWT.

As described in the previous chapter, the mechanism through which political instability is thought to lower growth is by affecting the investment decision. By creating uncertainty regarding the security of property and property rights, something that is decisive to both domestic and foreign investors, political instability constitutes a disincentive to invest (cf. section 2.4). Further, in economies that rely on exports of primary products the growth rate may sometimes be largely driven by increased international prices and consequently it will not reflect the actual economic development of these countries. Therefore, as a supplement to

[56] All the variables are measured in absolute numbers and are not adjusted for population size. This is based on the position held by Alesina and Perotti, who argue that assassinations and events that similarly rare are "just as disruptive of the social and political climate in a small country as of a large country" (Alesina and Perotti 1996: 1208). I regard this description as apt for all the variables used here.

[57] The Penn World Table is also referred to as the Summers and Heston dataset.

GROWTH, I test the hypothesis with an alternative measure of economic development, the yearly change in percent of the investment share of GDP per capita. PWT offers data on the investment share of GDP, a variable that is based on the other measure of GDP in constant prices, *RGDPL*. This measure does not use a price chain index but is obtained by adding up consumption, investment, government consumption and exports, and subtracting imports in any given year (Heston et al. 2006, Data Appendix). On the basis of this variable I constructed the change variable, *INVESTC*.

3.2.4 THE CONTROL VARIABLES

To be able to compare the results of one's analysis with the rest of the literature, substantial components have to be similar, such as variables and the country and time coverage (Campos and Nugent 2003: 533). I therefore take the existing literature on the field as a starting point when determining what variables to control for in the analysis. However, there are limits to how many control variables one should include. As Alesina et al. put it: different authors have their own "favourite" explanatory variables, from purely economic ones, to geographic, legal, political, cultural, religious and historical ones (Alesina et al. 2003: 182). For example, Levine and Renelt points out that over 50 variables have been found to be significantly correlated with growth in at least one regression (Levine and Renelt 1992: 942). At the other extreme, Achen claims that "[a] statistical specification with more than three explanatory variables is meaningless" (Achen 2002: 446). Combining the objectives of validity, parsimony, and minimising the risk of collinearity, I choose an approach in between and include a limited set of control variables in this analysis (6 in the first equation and 5 in the second). As the analysis includes sensitivity tests in which the variables are measured in different ways, the total number of models becomes very high. Therefore, the inclusion of the following control variables is fixed, that is, they will not be removed from the final models in those cases where they do not yield statistically significant results. [58]

EQUATION 1: POLITICAL INSTABILITY AS THE DEPENDENT VARIABLE

In the first equation measuring inequality's effect upon political instability, the literature suggests that controls for the following variables should be included:

[58] For details on the various variables and their sources, see Appendix A.

GDP per capita (natural logarithm)

Contrary to the central assertion made by relative deprivation theorists that it is relative well-being that matters for the generation of discontent (see section 2.3.1), others claim that absolute well-being is more decisive for political stability than relative well-being. For example, Parvin found that absolute economic well-being, measured as per capita income, was a far more influential factor than relative well-being, measured as economic inequality, which in fact turned out to have the opposite effect than what is usually assumed (Parvin 1973). The finding was partly supported by the results of Sigelman and Simpson's analysis. While they found that the relationship between inequality and political violence was moderate and linear in the expected direction, they too emphasised that the effect of income per capita was a more critical determinant than inequality. It is commonly assumed that level of economic development affects political stability positively, and a variable measuring this is often included in equations explaining political instability of various kinds (Hardy 1979; Huntington 1968; Muller 1985; Muller and Seligson 1987; Nagel 1974; Parvin 1973; Sigelman and Simpson 1977; Weede 1981; Zimmermann 1983). I thus include a variable measuring GDP per capita to control for level of economic development in the first equation. Due to the great dispersion in the levels of GDP across countries, the natural log of *GDP* is used.[59]

Growth

As discussed in detail in section 3.1.4, economic growth is though to be an important explanatory variable when studying political instability, and many such analyses include growth as an independent variable (e.g. Alesina and Perotti 1996; Auvinen 1997; Barro 1996; Londregan and Poole 1990; Nel 2003; Schock 1996). I thus include *GROWTH* as a control variable in this equation. As was discussed in section 3.1.4, in the recursive models a lagged version (1 year) of *GROWTH* is included. This is done to be sure that the estimations are not biased due to simultaneity that the test for simultaneity has not been able to capture, and because there are theoretical reasons to lag this variable.

Regime

The inclusion of regime type as a control variable stems from the possibility that its effect makes the estimation of the effect of inequality on instability biased. As democracies tend to

[59] The measure is the same as the GDP that forms the basis for *GROWTH*, as explained in section 3.2.3.

have both more equal income distributions and less political instability than authoritarian regimes (Gurr 1970; Rogowski 2004; Østby 2005) (cf. section 2.2.5), any established causality between inequality and instability might be spurious. It is assumed that political violence is most likely to occur in societies that do not provide non-violent patterns to value-satisfying action, that is, an open, democratic contest over political priorities and resource distribution (Feierabend and Feierabend 1966: 251). This is the reason why studies such as Keefer and Knack's (2002) "Polarization, politics and property rights: Links between inequality and growth" include a regime dummy variable. The opposite effect of regime is also expected in the literature, primarily stemming from the hypothesis that in authoritarian regimes, repressiveness will prevent people from engaging in political violence (e.g. Moaddel 1994).[60] Despite their different expectations, these two last-mentioned studies both hypothesise a linear effect, and in line with Moaddel's study I have operationalised regime as a dummy variable where the value of 1 is assigned to authoritarian countries. I use the ACLP dataset as my data source (see Appendix A).

Urban Population and Population Density

Finally I include two demographic variables, urban population and population density, that seem important when studying the determinants of political instability. Many political scientists, such as Huntington (1968) and Hibbs (1973), argue that more urbanised societies should be more politically unstable because political participation and social unrest are more likely to be higher in cities (Alesina and Perotti 1996: 1214). According to the relative deprivation hypothesis outlined in the previous chapter it also seems reasonable to suggest that when people are crammed more densely together, they become more aware of their situation relative to that of others. Scholars like Muller and Seligson and Moaddel similarly point out that it is easier to mobilise urban populations than rural ones (Moaddel 1994: 295; Muller and Seligson 1987: 427). These variables are included in empirical studies to a varying degree, but when included they are often found to be important (Annett 2001; Auvinen 1997; Collier and Hoeffler 2004; Nel 2003). *POPDEN* is entered into the analysis as its natural logarithm, and *URBPOP* is simply the percentage of urban population of the total population.

[60] See section 2.3.2.

Instability (lag)

I include in the equation a 1-year lag of the instability variable. This is simply because previous instability is thought to affect the present instability positively (Gurr 1970). In addition, in many of the models autocorrelation was above 0.3 when this lag was not included (see section 3.1.3, page 42). The lagged variables are specified as the original variable name plus *L*, e.g. *CINDEXL*.

Semi-Repressiveness

As discussed in section 2.3.2, theories of rational action, resource mobilisation and political opportunities, hypothesise that it is political, economic and organisational opportunities that are decisive factors explaining political violence, and that the occurrence of violence will be greatest at *intermediate* levels of regime repressiveness. As parsimony is a methodological goal, it makes sense to explore the simple general effects before moving on to more complex models (Weede 2002: 99). Therefore, a measure of regime repressiveness will only be included in those cases where inequality is significant to test whether this effect disappears when a control for regime repressiveness is included.[61] The variable *SEMI* is a dichotomous variable constructed on the basis of the democracy-autocracy variable from the POLITY IV dataset. The value 1 is given to the cases displaying values between -3 and +3, and the value 0 otherwise.

EQUATION 2: THE RATE OF ECONOMIC DEVELOPMENT AS THE DEPENDENT VARIABLE

In the second equation measuring the effect of political instability upon the rate of economic development, I introduce the following variables:

GDP per capita (natural logarithm)

It is common practice to control for the initial level of GDP when studying the determinants of growth, and in line with most studies on the field (e.g. Barro 1997; Easterly 2002; Knack and Keefer 1997; Krieckhaus 2004), I thus include GDP per capita in this equation. According to neo-classical growth theory, its basic version provided by Robert Solow, diminishing returns to capital makes poorer countries grow at faster rates than rich countries, leading to convergence between rich and poor countries in the long run (Mankiw 1995; Solow

[61] In these cases the variable *REGIME* is not included, as this is another measure of regime repressiveness, although measured in a different way. The inclusion of both would confuse the analytical interpretation of the semi-repressiveness variable.

1956). According to this viewpoint, the expected effect of initial GDP per capita is therefore negative. However, the convergence hypothesis has been disproved by several studies. For example, Benhabib and Rustichini observe that poor countries grow more slowly than rich countries due to the poorer countries lower investment rates in physical and human capital (Benhabib and Rustichini 1996), while Keefer and Knack ascribe the same empirical pattern to the lack of institutional quality in developing countries (Keefer and Knack 1997). The expected direction of the effect of this variable thus remains open.[62]

Investment

Including the level of investment as an explanatory variable in equations where growth is the dependent variable is in line with neo-classical growth theory, which states that a higher savings rate (i.e. investment rate) is an important determinant of growth, and further, numerous works have identified investment as a major vehicle for accelerated growth (Barro 1996: 9; Benhabib and Rustichini 1996: 125; Feng 2001: 288). I therefore include this variable in the second equation. Investment is here measured as the investment share of GDP in constant prices, as specified in section 3.2.3 and Appendix A.[63]

Trade openness

Trade openness is thought to affect growth positively, and many studies, also those focusing on the effect of instability on growth, find that the effect of trade openness is significant (Barro 2000; Easterly 2002; Sachs and Warner 1997). Here, it is measured as total trade as a percentage of GDP, and data are collected from PWT.

Government consumption

Government consumption, on the other hand, is assumed to have a negative effect on growth. Many scholars view this variable as particularly important, and find statistical support for this assertion (Alesina et al. 2002; Barro 2000; Sylwester 2000). Government consumption is measured as the government's share of GDP. Here too, data are collected from PWT.

Population growth

Many contributors to the growth literature point to the negative correlation that is often shown to exist between population growth, (alternatively, the fertility rate), and economic growth

[62] Here too, *GDP* is represented in the equation by its natural log.
[63] This is the variable from which *INVESTC* (see section 3.2.3) has been constructed.

(Barro 1996; Krieckhaus 2004; Perotti 1996a). Fertility theory states that as family size increases, parents diminish their average investment in human and physical capital per child (Becker et al. 1990). I therefore include a control for this in my analysis and expect its effect to be negative. *POPG* is operationalised as the annual percentage growth of total population, as defined by Global Development Network Growth Database.

3.3 SUMMARY – METHODOLOGY AND MEASUREMENT

To investigate whether socio-economic inequality affects economic growth negatively by increasing political instability, I apply a quantitative approach, and my analytical tool is the regression analysis. The sample subjected to analysis is a time-series cross-section panel of 188 countries from 1950 through 2004. To control for the possibly significant effect of omitted variables that is due to the panel design of the analysis, I apply the Fixed Effects Model. Another challenge associated with the structure of the analysis is the possibility of the presence of endogeneity: the causality between instability and growth might run in both directions. Therefore, a simultaneous equations model is used when simultaneity is found to be present, otherwise a recursive model is employed. To test whether the results are robust to different ways of measuring inequality and instability, these phenomena are measured in different ways. The growth variable is complemented with a variable measuring annual change in the investment level to examine whether this can supply the analysis with any additional knowledge that is more specific with regard to the research question. The choice of control variables is guided by the existing literature on the field and remains fixed throughout the analysis.

4 WHAT WE HAVE LEARNED AND WHAT WE STILL DO NOT KNOW: TENTATIVE FINDINGS

This chapter consists of three main sections. In section 4.1 the results of the analysis are presented. In section 4.2 the results are discussed elaborately and methodological issues are dealt with. Finally, section 4.3 sums up the analysis and its main findings, and offers some tentative conclusions. Tables 5 and 6 below list the variables in the two equations and the effects they are expected to have.

Table 5: Variables and expected effects – EQUATION 1*

Concept	***Variable names***	***Expected effect***
Socio-economic inequality (various)	*GINI; GINIIN; GINIINC; GINICON; GINISM; MIDCLASS;MIDIN; MIDINC; MIDCON; MIDSM*	+
Lag (1 year) of political instability (various)	*CINDEXL; ASSASSL; DEMSL; RIOTSL; GWARL; REVL; STRIKESL*	+
Growth rate of GDP per capita (1 year lag)	*GROWTHL/ INVESTCL*	-/+
Level of gross domestic product per capita	*GDP***	-
Urban population	*URBPOP*	+
Population density	*POPDEN***	+
Regime	*REGIME*	+/-
Semi-repressiveness	*SEMI*	+

** Dependent variable: POLITICAL INSTABILITY. Independent variables: inequality; instability lagged; growth of GDP/ change rate of investment (lagged); GDP per capita; urban population; population density; regime type; (semi-repressiveness of the regime)*
*** Natural logarithm*

Table 6: Variables and expected effects – EQUATION 2*

Concept	***Variable names***	***Expected effect***
Political instability *(various)*	*CINDEX; ASSASS; DEMS; RIOTS; GWAR; REV; STRIKES*	-
Level of gross domestic product per capita	*GDP***	-/+
Level of investment	*INVEST*	+
Government consumption	*GOVCON*	-
Trade openness	*OPEN*	+
Population growth	*POPG*	-

** Dependent variable: GROWTH/ INVESTC. Independent variables: Instability; instability lagged; GDP per capita; level of investment; government consumption; trade openness; population growth*
*** Natural logarithm*

4.1 PRESENTATION OF ANALYTICAL RESULTS

As described in the previous chapter, I have constructed several measures both of socio-economic inequality, political instability and economic development:[64] When combining all the different measures in separate constellations, one is left with not less than 140 model

[64] See variable overview in Appendix A for variable details and Appendix E for descriptive statistics.

versions: 70 models where the rate of economic development is measured as growth (10 inequality measures multiplied by 7 instability measures), and equally 70 where it is measured as the yearly rate of change in the investment level.[65] Tests for simultaneity have been conducted for all of the different models, and in those cases when simultaneity was present a simultaneous equations model was used, and when it was not, a recursive model was applied. The simultaneity test reported simultaneity in 25 of the *GROWTH* models; the remaining 45 were therefore made recursive. Of the *INVESTC* models, simultaneity was present in 14 of the models; the remaining 56 were run recursively. In the next section I will present three models (A, B, and C) in detail.[66] As the total amount of information from all of the models is very large, I will synthesise the remaining results so as to identify their general patterns.

4.1.1 THREE MODELS

Tables 7- 8: Model A (recursive): *GINI; CINDEX; GROWTH*

Dependent variable: *CINDEX*

Variable	***b***	***SE***	***t-stat***	***P***	***Mean of X***	***Adjusted R²***	***Est. AC***
CINDEXL	0,47	0,10	4,85	0,00	2,80		
GROWTHL	0,06	0,03	1,79	0,07	2,23		
POPDEN	-0,41	1,50	-0,28	0,78	7,27		
GDP	-1,94	0,78	-2,48	0,01	8,83		
REGIME	0,89	0,53	1,68	0,09	0,28		
URBPOP	0,08	0,05	1,59	0,11	59,43		
GINI	-0,01	0,02	-0,20	0,85	38,79		
						0,41	-0,00

b: regression coefficient (unstandardised); SE: Standard error; P: level of significance; AC: autocorrelation
Number of units: 1233

Dependent variable: *GROWTH*

Variable	***b***	***SE***	***t-stat***	***P***	***Mean of X***	***Adjusted R²***	***Est. AC***
CINDEX	-0,10	0,02	-4,23	0,00	1,67		
GDP	1,31	0,56	2,32	0,02	8,30		
INVEST	0,16	0,03	4,71	0,00	14,73		
GOVCON	-0,14	0,03	-4,39	0,00	22,40		
OPEN	-0,01	0,01	-0,97	0,33	72,11		
POPG	0,19	0,37	0,52	0,60	1,93		
						0,06	0,11

b: regression coefficient (unstandardised); SE: Standard error; P: level of significance; AC: autocorrelation
Number of units: 5830

[65] As specified in section 3.2.4, the set of control variables that are included remains fixed in all of the models.
[66] These three models are chosen because they reflect the spectre of different findings across the models. This will be explained and discussed further in what follows.

As shown in Tables 7 and 8, Model A examines the relationship between *GINI, CINDEX* and *GROWTH*. The simultaneity test produced an insignificant residual for Model A, so the analysis was run recursively. The most important variable in the first equation is the variable measuring socio-economic inequality, namely *GINI*. The effect of *GINI* is not statistically significant in model A, and the implication of this is that socio-economic inequality, measured as *GINI*, does not affect political instability, measured as the conflict index, *CINDEX*. The first hypothesis of this thesis, that socio-economic inequality breeds political instability, is therefore rejected by this model.

As for the control variables in equation 1, we can see that the lagged conflict index is positively and significantly related to *CINDEX*: an increase of 1 on this variable causes an increase of .47 on the conflict index. This is to be expected: previous conflicts affect the present conflict level positively. The only other variable that is significant at the 5-percent level is *GDP*. This regression coefficient goes in the expected direction: a higher level of GDP per capita is associated with lower levels of conflict. As this is a logarithmic variable it cannot be interpreted linearly. In real numbers its coefficient of -1.94 must be interpreted in the following way: moving from having the lowest GDP per capita (170 US$ - that of Liberia in 1995) to the average level of GDP per capita in this dataset (7092 US$ - approximately that of Brazil in 1988), implies a decrease of 3.14 on the conflict index, and moving from the average to the highest GDP per capita (84 408 US$ - that of Qatar in 1974) implies a decrease of 2.66 on the conflict index. The effect is thus greatest at the lower end of the scale.

Two variables are significant at the 10-per cent level, *GROWTHL* and *REGIME*. *GROWTHL* has a positive effect on conflict, something that indicates that the economic growth of the previous year causes a rise in the present conflict level, even though the magnitude is relatively small (a rise of 1 percentage point on the growth rate causes an increase of .06 on *CINDEX)* and the significance level is rather low. This goes against the expectation that growth will reduce political instability, and lends support to Olson's hypothesis that growth actually has a destabilising effect due to its potential to produce social dislocation and increased socio-economic disparities.

Further, we can see that *REGIME* has a positive effect on conflict. As this is a dummy variable where the value 1 is given to non-democratic regimes, this coefficient can be interpreted in the following way: being a non-democratic regime is associated with a value on

the conflict index that is .89 higher than that of a democratic regime. The variable *urban population* is almost significant at the 10-percent level and its sign is positive. Although the direction of this variable is as expected, indicating that countries with a large urban population are more prone to be politically unstable than countries with greater segments of the population living in rural areas, the weak significance of this variable makes it questionable whether this effect really exists. One control variable is completely insignificant in this model: *POPDEN*. This indicates that a more densely populated country is not more likely to experience political unrest than countries that are less densely populated.

When it comes to the second hypothesis, that political instability reduces economic growth, we must look at the second table of Model A. Due to the fact that inequality does not enter into this equation, the sample size is much larger, namely, 5830 observations. As was hypothesised, the variable measuring political instability, *CINDEX*, is negative, and it is highly significant. An increase of one incidence of conflict (either an assassination, a riot, a revolution, an incidence of guerrilla warfare, a strike or an anti-government demonstration), decreases the growth rate in that same year by .10 percentage points.

Three of the control variables have significant and expected effects on the growth rate. *GDP* is positive with a t-statistic of 2.32, implying that the level of GDP affects the growth rate positively. Again, the logarithmic nature of this variable prompts the following interpretation: moving from having the lowest GDP per capita to the average (see the preceding interpretation of equation 1), is associated with an increase in the growth rate of 2.11 percentage points, while a move from the average to the highest level of GDP per capita results in an increase of 1.8 percentage points on the growth rate. The positive effect of the level of GDP per capita goes against the convergence hypothesis that claims that developing countries will grow faster than developed countries and that their levels of development will therefore converge over time. Rather, it supports the counter-expectation that cross-country differences in levels of GDP per capita will persist, or even grow, due to the less developed countries' lower investment rates in human and physical capital and lower institutional quality. Indeed, this is further indicated by the highly significant positive effect of the level of investment (*INVEST*) on the growth rate. Finally, *GOVCON* is negative and significant, supporting the expectation that the higher the share of government consumption of the total gross domestic income, the lower the growth rate. Two control variables, trade openness and population growth, did not have a significant impact on growth in this model.

The adjusted R^2 tells us how much of the variation in the dependent variable is explained by the independent variables in the model.[67] We can see from Model A that for the instability equation, adjusted R^2 is .41, and .06 for the growth equation. This means that for the population from which this sample is drawn, 41 per cent of the variation in *CINDEX* can be ascribed the set of explanatory variables in the model, while the explanatory variables included in equation 2 can only account for 6 per cent of the variation in *GROWTH*.[68] Further, we can see from the model that the estimated autocorrelation of both models is very low, -.04 and .11, respectively.[69] The results are therefore not biased due to autocorrelation.

According to Model A, then, inequality does not have an effect upon the level of political instability, but the latter has a statistically significant negative effect upon the rate of economic development. But are these results robust to alternative ways of measuring the variables? Let us take a look at one of the other models, that in which inequality as been measured as *MIDIN* (the income share of the middle class) and instability as *ASSASS* (number of politically motivated murders or murder attempts).[70]

Tables 9- 10: Model B (simultaneous equations): *MIDIN; ASSASS; GROWTH*

Dependent variable: *ASSASS*

Variable	*b*	*SE*	*t-stat*	*P*	*Mean of X*	*Adjusted R^2*	*Est. AC*
ASSASSL	0,29	0,10	2,96	0,00	0,31		
PRED	-0,01	0,02	-0,34	0,73	2,41		
POPDEN	0,34	0,31	1,09	0,28	7,39		
GDP	-0,46	0,19	-2,48	0,01	9,12		
REGIME	0,23	0,16	1,44	0,15	0,17		
URBPOP	0,03	0,01	2,49	0,01	65,60		
MIDIN	-0,09	0,04	-2,15	0,03	37,14		
						0,29	-0,07

b: regression coefficient (unstandardised); SE: Standard error; P: level of significance; AC: autocorrelation
Number of units: 646

[67] The R^2 is adjusted to remove any bias that is due to the number of explanatory variables included in the model.
[68] The following should be noted here: it has been argued that the importance of R^2 is limited and that a low R^2 does not necessarily imply a low degree of the variance has been explained. As Achen (1982: 59) points out, the problem with the R^2 in the social sciences is that the independent variables are not subject to experimental manipulation: in some samples, they vary widely and produce large variance, in others, the observations are grouped more tightly, thus producing little dispersion. Therefore, the variances are a function of the *sample*, not the underlying relationship. Although this objection is valid and should be kept in mind when interpreting the R^2 as reported here, the information that it gives is still useful.
[69] Durbin-Watson is a common measure of autocorrelation. It varies between 0 and 4 and values around 2 indicate non-presence of autocorrelation. An estimated autocorrelation of .10 gives a Durbin-Watson of 1.8. Assuming six explanatory variables and a large N, this is approximately within the range within which autocorrelation is not a problem (1.83 and 2.17) (Gujarati 2003: 481, 970).
[70] For Models B and C I will focus on the effect of main variables in the analysis. The control variables will be further dealt with in connection with the discussion of the general patterns.

Dependent variable: *GROWTH*

Variable	*b*	*SE*	*t-stat*	*P*	*Mean of X*	*Adjusted R²*	*Est. AC*
PRED	0,21	0,23	0,91	0,36	0,29		
GDP	-3,39	0,76	-4,47	0,00	9,12		
INVEST	0,11	0,04	2,74	0,01	18,43		
GOVCON	-0,05	0,06	-0,81	0,42	20,37		
OPEN	0,05	0,01	5,19	0,00	67,60		
POPG	0,24	0,39	0,62	0,54	1,03		
						0,30	0,09

b: regression coefficient (unstandardised); SE: Standard error; P: level of significance; AC: autocorrelation
Number of units: 646

When testing for simultaneity in Model B, the estimated simultaneity was significant. This implies that including *GROWTH* in the instability equation and *ASSASS* in the growth equation[71] would produce biased results due to the fact that these variables are not truly exogenous but are simultaneously affected by one another. Therefore, to avoid endogeneity and biased estimates, the analysis of this particular combination of variables was conducted using an S.E. (simultaneous equations) model. In Tables 9 and 10 the endogenous explanatory variables are thus replaced by the predicted variables and are called *PRED* in both equations.

The analytical results for the instability equation in Model B reveal that, compared to the results in Model A, the effects of some of the variables remain approximately the same, while others are very different. Most importantly in this setting is the fact that the variable measuring socio-economic inequality, *MIDIN*, is here strongly significant and has the expected sign: An increase of 1 on *MIDIN*, that means, an increase of one percentage point of the share of the middle class (the Third and Fourth quintiles) of the population's total income, causes a decrease on the dependent variable, *ASSASS*, of .09. This is quite a lot when one takes into account that the mean of *ASSASS* is .19 (see Appendix E). As for the second equation of Model B, the instability variable, appearing in Model B as *PRED,* has changed dramatically compared to Model A. While *CINDEX* was negative and statistically significant at the .01 percent level, *PRED* is not significant at all (and with a positive sign). Model B therefore does not support the hypothesised effect of instability upon growth as presented in section 2.4. To sum up the main message in model B: Socio-economic inequality increases the level of political instability, but political instability does not have an effect upon growth. Let us now look at a third model, and investigate whether either Model A or B is supported by

[71] Recall that "the growth equation" refers to the equation where economic development is the dependent variable, hereunder also the rate of change in investment (*INVESTC*).

the results in this model. Now, the measure of inequality is *MIDCON* (the income share of the middle class, measured in terms of consumption), *GWAR* (incidents of guerrilla warfare) is the measure of instability and *INVESTC* (the yearly rate of change in the level of investment) replaced *GROWTH* as the measure of economic development.

Tables 11- 12: Model C (recursive): *MIDCON; GWAR; INVESTC*

Dependent variable: *GWAR*

Variable	*b*	*SE*	*t-stat*	*P*	*Mean of X*	*Adjusted R²*	*Est. AC*
GWARL	0,07	0,20	0,34	0,74	0,19		
INVESTCL	0,02	0,01	1,83	0,07	-0,23		
POPDEN	-0,23	0,33	-0,68	0,50	7,19		
GDP	-0,27	0,26	-1,02	0,31	7,89		
REGIME	0,08	0,08	0,99	0,33	0,51		
URBPOP	0,00	0,02	-0,08	0,94	41,82		
MIDCON	0,01	0,01	1,96	0,05	35,21		
						0,42	-0,10

b: regression coefficient (unstandardised); SE: Standard error; P: level of significance; AC: autocorrelation
Number of units: 176

Dependent variable: *INVESTC*

Variable	*b*	*SE*	*t-stat*	*P*	*Mean of X*	*Adjusted R²*	*Est. AC*
GWAR	0,06	0,06	1,02	0,31	0,18		
GDP	-0,59	0,26	-2,27	0,02	8,30		
INVEST	0,14	0,02	6,08	0,00	14,76		
GOVCON	0,00	0,01	0,29	0,77	22,38		
OPEN	0,00	0,01	0,17	0,87	71,93		
POPG	-0,14	0,09	-1,50	0,13	1,93		
						0,06	0,01

b: regression coefficient (unstandardised); SE: Standard error; P: level of significance; AC: autocorrelation
Number of units: 5773

In this model simultaneity was not present, and a recursive system was thus applied. Beginning with the first equation in Model C, we can see that contrary to what was hypothesised, the measure of socio-economic inequality, *MIDCON*, has a positive and significant effect upon the level of instability. In other words, a higher level of socio-economic inequality leads to lower levels of political instability: an increase of 1 percentage point in the income share of the middle class is associated with an increase of .01 on *GWAR*. As for the second equation in Model C, we see that *GWAR* has a t-statistic of 1.02 and is thus insignificant as an explanatory variable of the rate of change in the level of investment. The central features of Model C are that socio-economic inequality, measured as *MIDCON*, decreases political instability, measured as *GWAR*, while the latter does not have any effect upon economic development, here represented by *INVESTC*.

4.1.2 GENERAL PATTERNS

The three models that have now been presented report entirely different results, and thus demonstrate that the hypotheses put forth in chapter 2 can both be rejected and supported, depending on what variables are used in the analysis. Model A, B and C display different extremes of the various explanatory variables in the two equations, but what is the general pattern for all of the 140 models that have been analysed? I will now present the "general behaviour" of the variables measuring socio-economic inequality, those measuring political instability and that of the different control variables in a synthesised manner, and investigate to what degree there can be found any general pattern across the models.

EQUATION 1: POLITICAL INSTABILITY AS THE DEPENDENT VARIABLE

As shown by the Models A, B and C, the effect of socio-economic inequality on political instability seems to depend on how the variables are measured, and the analysis of these models has thus not yielded any clear answer as to whether the hypothesis presented in chapter 2 about this relationship should be rejected or not. But these were just three models – does the broad picture that the 140 models constitute reveal any one salient effect, and are there certain combinations that give certain results?

Table 13: General effects of the inequality variables in the instability equation

Variable	*Total number*	*Insignificant effect*	*Significant positive effect*	*Significant negative effect*	*Expected Sign*
GINI	14	14	0	0	+
GINIINC	14	14	0	0	+
GINIIN	14	12	2*	0	+
GINICON	14	14	0	0	+
GINISM	14	14	0	0	+
MIDCLASS	14	12	0	2*	-
MIDINC	14	10	2*	2	-
MIDIN	14	12	0	2	-
MIDCON	14	11	1	2 (1*)	-
MIDSM	14	14	0	0	-

** Significant only at the 10 percent level*

We can see from Table 13 that the vast majority of inequality variables have a statistically insignificant effect in the instability models: Of a total of 140 entries, 127 are insignificant, 10 are significant with the expected sign (of these, 5 are significant only at the 10 percent level), and 3 are significant with the unexpected sign (of these, 2 are significant only at the 10 percent level). The general pattern in the models that have been analysed is thus that socio-

economic inequality does not affect the level of political instability. In other words: this is the result given 1) the methodological setup of this analysis with its cross-sectional, cross-temporal design; 2) the use of a fixed effects model and a simultaneous equations model when simultaneity appeared to be present; 3) the ten different ways of measuring socio-economic inequality; 4) the seven different ways of measuring political instability; 5) the two different ways of measuring economic development that have been used; and 6) in combination with the set of control variables that have been included in this analysis.

Let us look at the control variables that were included in this equation, (summed up in Table 14). The lagged instability variables are for the most part significant (in 84.3 % of the models), all of which have the expected sign. One can thus assert that how politically stable a country was in the previous year generally affects the present level of political stability. *GDP* has a less clear effect upon the level of political instability: in 59.3 % of the models this variable was insignificant. In those cases where *GDP* was significant, it had a negative sign in 55 out of 57 models, which is in accordance with the expectations. As for the *GROWTH/ INVESTC* variables (lagged in the recursive models), these were insignificant in 89.3 % of the models. The indecisiveness regarding the direction in which causality should run that can be found in the literature (see section 3.1.4) manifests itself in the fact that in those cases where these variables had a significant effect on instability, 7 had a positive effect and 8 had a negative effect. But in general, the rate of economic growth and the change in the investment level do not affect the level of political instability.[72]

Table 14: General effects of the control variables in the instability equation

Variable	*Total number*	*Insignificant effect*	*Significant positive effect*	*Significant negative effect*	*Expected Sign*
Instability lagged	140	22	118 (14*)	0	+
GDP	140	83	2	55 (13*)	-
GROWTH/ INVESTC	140	125	7 (4*)	8 (4*)	-/+
REGIME	140	87	34 (11*)	10 (4*)	+/-
URBPOP	140	111	29 (15*)	0	+
POPDEN	140	125	15 (9*)	2*	+

** Significant only at the 10 percent level*

[72] The recursive models were run with the un-lagged versions of *GROWTH/ INVESTC*, and the lagged versions were added to the S.E. models to see whether the time aspect affected the variables' effect. This appeared not to be the case.

The regime variable was insignificant in about the same number of cases as *GDP*, namely, in 62 % of the models. This might be an expression of an offsetting effect where its negative effect cancels out its positive effect. Alternatively, as suggested by Thorbecke and Charumilind (2002:1486), even non-democratic regimes can be forced to deal with public discontent to avoid social revolution and breakdown of the regime. For example, they describe how the dictator Suharto had to implement policies to reduce the economic inequality in the Indonesian population, mainly between the better-off Java and the less well-off periphery islands, to strengthen his own position. In the cases in which *REGIME* was significant, it displayed both positive and negative effects and the theoretical ambiguity related to the effect of regime type on the level of political instability (see section 3.2.4) is thus confirmed empirically in this analysis. However, since in 34 of the 44 cases where *REGIME* is significant it has a positive sign, more support is given to the view that manifestations of political instability are most likely to occur in authoritarian regimes, where peaceful channels of political participation are not available.[73]

When it comes to the two demographic variables in the analysis, *URBPOP* and *POPDEN*, the coefficients on both variables were for the most part insignificant. For example, *URBPOP* was insignificant in 111 out of the 140 models, that is, in 79.3 % of the cases. However, of the 29 models in which this variable had a significant effect upon instability, all were positive, and thus in line with the theoretical expectation that the more people live in urban areas, the easier it is to mobilise people to engage in acts of collective violence and unrest. Generally then, one can say that the amount of urban population that a country has does not affect the country's level of political instability, but that when it does have an effect, this is positive. As for *POPDEN*, this variable is insignificant in even more of the models than *URBPOP*. In 89.3 percent of the models this variable is insignificant, indicating that in general the density of the population does not affect the level of political instability in a country. As with the amount of urban population, when *POPDEN* was significant it was mostly positive (in 15 out of 17 cases), again in line with the theoretical expectation.

As for the explanatory power of the models where political instability is the dependent variable, the adjusted R^2 is mostly around .30, varying between .20 and .45. For some of the models, however, the adjusted R^2 is lower, sometimes almost zero, but these are very few

[73] Recall that the alternative hypothesis is that authoritarian regimes are least prone to political instability, as the repressiveness of these regimes prevents the outbreak of collective violence and unrest.

cases. Therefore, one can say that, to the degree that the R^2 tells us something about the goodness of fit, the fit of these models is quite good, meaning that the explanatory variables included in equation 1 account for a relatively large portion of the variation in the variables measuring political instability. The estimated autocorrelation for the models is generally around .10, and except in 9 of the 140 models, where the autocorrelation is estimated to more than .30, it is always below this level. Therefore, one can conclude that, generally, autocorrelation is not a problem in these models.

EQUATION 2: ECONOMIC DEVELOPMENT AS THE DEPENDENT VARIABLE

The general patterns for the first equation of this analysis have now been described, and I will now move on to the second equation. Table 15 below lists the levels of significance of the different instability variables that enter into the growth equation. We can see that almost all of the variables are insignificant in about half of models, *DEMS* was insignificant in the fewest number of cases and *RIOTS* had the highest number of insignificant entries. In those cases where the instability variables were significant, they were almost exclusively negative. This means that in about half of the models instability had a statistically significant negative effect upon the rate of economic development, which is in accordance with the theoretical expectation. There was only model in which instability had a positive effect, in which case it was significant only at the 10 % level. Due to the even distribution of negative and insignificant effects, it is difficult to decide what the general pattern is. This will be discussed and investigated further in section 4.2.

Table 15: General effects of the instability variables in the growth equation

Variable	*Total number*	*Insignificant effect*	*Significant positive effect*	*Significant negative effect*	*Expected Sign*
ASSASS	6	3	0	3 (1*)	-
CINDEX	9	5	0	4	-
DEMS	3	1	0	2	-
GWAR	7	5	1	1	-
REVS	5	3	0	2 (1*)	-
RIOTS	18	15	0	3	-
STRIKES	5	3	0	2 (1*)	-

** Significant only at the 10 percent level*

As for the control variables that have been included in equation 2, Table 16 below sums up the general patters. As discussed in section 3.2.4, the theoretical expectations concerning the effect of GDP per capita are twofold. While neoclassical economic theory expects the growth rates of less developed countries to be higher than those of the developed countries, so that in

the long run the levels of development will converge across countries, it has become more common to assert that it is the other way around and that the differences between countries therefore will persist. This dataset shows that *GDP* generally has a negative effect upon economic growth, as this variable is negative and significant in 26 of the 53 models, compared to 8 of 53 which are the cases in which *GDP* is positive and significant. Convergence theory thus receives more support from these analyses than the assertion that rich countries grow faster than poor countries. However, the high number of instances in which this variable is insignificantly related to growth, (19 of 53 models), indicates that this relationship is not sensitive to different variable operationalisations.

If we look at *INVEST*, the picture is clearer: only in 8 of the 53 models is this variable insignificant, and when it is significant it always has a positive effect upon growth. This gives support to the vastly agreed upon assertion that the level of investment is important for economic growth (cf. section 3.2.4). The contention that the level of government consumption is negative for growth is not supported in this analysis: in 44 out of 53 models *GOVCON* was insignificant. However, in all the 9 cases in which it had a significant effect upon the growth rate, this was always negative. Further, *OPEN* is positive and statistically significant in the majority of the models (27 out of 53),[74] and it never has a negative effect upon growth. It should be noted, nonetheless, that in about half of the models this variable was insignificant, so its effect is not very robust to different model specifications. Finally, population growth is generally unimportant as an explanatory variable of economic growth, as *POPG* is insignificant in 83 % of the models. Curiously, however, in those instances where this variable does have a significant effect upon economic growth this is positive in 8 of 9 cases. This is contrary to the theoretical expectation, and might be due to spurious effects between the variables.

Table 16: General effects of the control variables in the growth equation

Variable	*Total number*	*Insignificant effect*	*Significant positive effect*	*Significant negative effect*	*Expected Sign*
GDP	53	19	8	26	-/+
INVEST	53	8	45	0	+
GOVCON	53	44	0	9 (2*)	-
OPEN	53	26	27 (1*)	0	+
POPG	53	44	8 (4*)	1	-

[74] In the vast majority of these cases, *OPEN* was significant at the 1 % level or lower, with t-statistics around 4.

As for the explanatory power of the models where economic development is the dependent variable, the adjusted R^2 is below .12 in 26 models; between .20 and .30 in 7 models; and above .30 in 20 of the models. There is in other words great variation, and in the following sections I will explore what might lie behind these differences. The estimated autocorrelation of these models is never higher than .42, and around or below .10 in 48 of the 53 models. Generally, then, autocorrelation is thus not a problem in the models of equation 2.

4.1.3 SUMMARY – ANALYTICAL RESULTS

Summing up the main variable relationships in this analysis, we have seen that inequality does not seem to affect the level of political instability, but that this finding is not robust to alternative ways of measuring inequality and instability: In some models inequality did have a statistically significant effect upon political instability, and of these cases, inequality had both positive and negative effects. Castelló and Doménech's (2002: 198) finding that the effects of economic inequality is largely independent of how it is measured is thus not supported by this analysis. As for the relationship between political instability and the rate of economic development, the analytical results provide stronger evidence of the existence of a negative effect. However, the political instability variables were statistically insignificant in about half of the models, and therefore one cannot at this point conclude one way or the other. Taken together, then, this analysis has given little support to the hypothesis that socio-economic inequality affects growth negatively by producing political instability. In the following section I will explore the analytical results in detail and investigate what may lie behind the general patterns just presented. [75]

4.2 DISCUSSION OF ANALYTICAL RESULTS

In his study of the effect of inequality on growth across American states, Panizza (2002) shows that small differences in estimation techniques, in the method of measuring inequality or in the source of the data used to measure inequality can make a big difference in the observed relationship between inequality and growth. In this section I will explore to what degree the results are due to the various methodological specifications of this analysis. I will

[75] Due to the limited scope of this thesis I have not been able to run the analysis testing for the effect of time. The Fixed Effects Model used here allows for such a test, and it would have been interesting to see whether the effects between the variables in the analysis are in any way different across time periods.

begin by investigating whether there are systematic differences in the results according to what variables have been used to measure the different phenomena.

4.2.1 THE ROBUSTNESS OF THE DIFFERENT MEASURES

To what degree are the analytical results dependent on how socio-economic inequality, political instability and economic development have been measured? This subsection aims at answering this question, and will begin by exploring the variable relationships in equation 1.

EQUATION 1: POLITICAL INSTABILITY AS THE DEPENDENT VARIABLE

In ten of the 140 models the inequality variables are significant or close to significant but have the unexpected sign. In another ten models they have a significant or close to significant effect on political instability and with the expected sign.

Table 17 Regressions with the 10 most significant inequality variables with the *unexpected* sign*

Economic development measure	***Instability variable***	***Inequality variable***	***Model*** **	***t-stat***	***Expected sign***	***P***
GROWTH	ASSASS	GINICON	R	-1,45	+	0,15
	GWAR	GINICON	S	-1,54	+	0,12
	GWAR	MIDCON	R	1,57	-	0,12
	REVS	MIDCON	R	1,52	-	0,13
	REVS	MIDINC	R	1,74	-	0,08
INVESTC	ASSASS	GINICON	R	-1,60	+	0,11
	GWAR	GINICON	R	-1,49	+	0,14
	GWAR	MIDCON	R	1,96	-	0,05
	REVS	MIDCON	R	1,60	-	0,11
	REVS	MIDINC	R	1,67	-	0,10

** Dependent variable: POLITICAL INSTABILITY*
*** R = recursive model; S = simultaneous equations model*

Table 17 shows that there is a pattern among the inequality variables with significant or close to significant coefficients with the unexpected sign. It seems that *ASSASS*, *GWAR* and *REVS*, combined primarily with the inequality variables that are constituted only of consumption data, (*GINICON* and *MIDCON)*, and two cases of *MIDINC*, are variable combinations that produce unexpected effects, namely a positive effect of inequality upon the level of political instability. However, it is important to note that only one of these variables is significant at the 5 percent level, namely, *MIDCON* in the model where *GWAR* is the dependent variable – the variable combination that was presented in Model C.

Table 18: Regressions with the 10 most significant inequality variables with the *expected* sign*

Economic development measure	*Instability variable*	*Inequality variable*	*Model***	*t-stat*	*Expected sign*	*P*
GROWTH	CINDEX	GINIIN	R	1,68	+	0,09
	ASSASS	MIDCLASS	R	-1,76	-	0,08
	ASSASS	MIDIN	S	-2,15	-	0,03
	RIOTS	MIDINC	S	-2,17	-	0,03
	RIOTS	MIDCON	S	-1,68	-	0,09
INVESTC	CINDEX	GINIIN	R	1,71	+	0,09
	ASSASS	MIDCLASS	R	-1,76	-	0,08
	ASSASS	MIDIN	R	-2,11	-	0,03
	RIOTS	MIDINC	S	-2,21	-	0,03
	RIOTS	MIDCON	R	-2,36	-	0,02

** Dependent variable: POLITICAL INSTABILITY*
*** R = recursive model; S = simultaneous equations model*

Table 18 reports the models where the inequality variable is significant, or close to significant, and has the expected sign. As was the case in the former table, a pattern can be detected among these models. First of all, there are mostly middle class variables on this list. Further, both the Gini and middle class variables that only include observations where the income definition is either *income*; *disposable income*; *monetary income*; or *disposable monetary income* (*GINIIN* and *MIDIN*, see Appendix A and D), in combination with *CINDEX*, yield significant coefficients (at the 10 percent level) with the expected sign. In addition to *CINDEX*, *ASSASS* and *RIOTS* are instability variables that appear several times in this table.

When comparing the cases in which inequality as a positive effect upon instability with the cases in which this effect is negative, one finds that only in one case of the latter is the effect significant at the 5 percent level, and in two cases at the 10 percent level. Of the cases in which the effect of inequality is negative, on the other hand, its effect is significant at the 5 percent level in five instances and at the 10 percent level in another five. And in most of these instances is inequality measured as a version of the middle class measure. Further, in all the recursive models reported in Table 18 the effect of instability on growth is negative, while it is non-significant in the S.E. models.[76] *This implies that if one uses a recursive model; measures inequality as the income share of the middle class relative to the income of the total population (more specifically, as MIDCLASS, MIDINC, MIDIN or MIDCON; and instability*

[76] I will discuss the differences between the recursive and the S.E. models separately in section 4.2.2.

as either the number of riots or assassinations during a year, one will find that inequality, by producing political instability, has a negative effect on economic growth.[77]

Interestingly, in both Tables 17 and 18 we can see that the pattern is the same in the models where *GROWTH* is the dependent variable as when *INVESTC* is the dependent variable.[78] *This implies that these results are robust to whether economic development has been measured as the growth rate of GDP per capita or as the rate of change of the investment level, and it strengthens the observation of statistical significance of the specific combination of inequality and instability variables that appear in the two tables.*[79] Exactly why inequality, when it is measured as the relative share of the middle class, affects riots and assassinations more strongly than the other measures of political instability included here, is hard to tell. The immediate interpretation is that when the relative share of the total income of a country pertaining to the middle class is small, people more often engage in riots or attempts to assassinate politicians or government officials than e.g. engage in guerrilla warfare, strikes, anti-government demonstrations, or revolutions. As the intuitive logic of behind this a pattern seems rather vague, however, such an interpretation must be drawn with caution, and a thorough theoretical investigation is needed to establish a justification for such an expectation. At this point I can merely report the empirical existence of this pattern.

The control variables

Are the effects of the control variables in equation 1 dependent upon how instability was measured? For three of the variables, this seems to be the case. First, the effect of *REGIME* is positive and more significant both substantively and in terms of number when *GWAR* is the dependent variable than what is the case with any other of the instability variables, actually in as much as 18 out of 20 models. In other words, authoritarianism is an important explanatory variable for the presence of guerrilla warfare. In the models where *STRIKES* is the dependent variable, on the other hand, *REGIME* is sometimes significant with a *negative* sign, and as such these models stand out from the models where other political instability variables are the dependent variable. In other words, "any strike of 1000 or more industrial or service workers that involves more than one employer and that is aimed at national government policies or

[77] If one measures instability as the number of incidents of guerrilla warfare, and inequality as *MIDCON*, one will find that inequality actually *decreases* the level of political instability.
[78] Of course, this is only relevant in the simultaneous equations models, where *INVESTC* and *GROWTH* were included to produce the predicted values.
[79] As for the control variables in this equation, there were no clear systematic differences between the models stemming from different ways of measuring economic development.

authority" (see Appendix A), as opposed to other kinds of political instability, is more likely to take place in a democracy than in an authoritarian regime, which is not unreasonable, given that strikes are a common way of trying to alter the government's policies in established democracies.

Secondly, the effect of *URBPOP* appears to be significant and positive in more of the models where *ASSASS* and *DEMS* are the dependent variables than in the other models. This means that that the amount of urban population in a country is a more important variable for explaining the number of assassinations, or attempts of such, of a high government official or politician and the number of anti-government demonstrations than for explaining other kinds of political instability arising from civil society. Thirdly, *GDP* is negative and significant in about half of the models or more for all of the political instability variables except *REVS*. When this is the dependent variable, *GDP* is never negative and in some instances even positive and significant. This might be expressing that a revolution is more likely to take place in countries above a certain level of development, and suggests that carrying out a revolution is a resource demanding activity.[80]

The adjusted R^2 and estimated autocorrelation

Finally, let us see whether there are systematic differences in the explanatory power and the presence of autocorrelation across the models that stem from the different ways of measuring socio-economic inequality and political instability. As previously discussed, the adjusted R^2 in the models where political instability is the dependent variable are generally around .30. In some instances, however, this is below .20, and even below .10. A closer look at the models reveals that the adjusted R^2 is generally lower when *DEMS*, *RIOTS* and *STRIKES* are the dependent variables, indicating that the variables included in equation 1 have less explanatory power for the variation in these variables than for the variation in the other political instability variables. This suggests that these are kinds of political instability that are affected to a larger extent by other factors than those included at the right hand side in equation 1.

Further, although this is not an absolute trend, it seems to be the case that the adjusted R^2 is different in those models where either *GINICON* or *MIDCON* is the variable measuring socio-economic inequality. That is, for the political instability variables whose models have a

[80] However, the fact that *GDP* is insignificant in most of the models where *REVS* is the dependent variable, must be interpreted to indicate that *GDP* generally does not affect whether a revolution takes place or not.

generally high adjusted R^2, the few cases when it is low it is usually a model where either *GINICON* or *MIDCON* is included. And for the three variables mentioned above, where the adjusted R^2 generally is low, the exceptions where it is high also tend to involve these inequality variables. In addition, whereas the estimated autocorrelation generally is very low, it appears that in the models where socio-economic inequality is measured as either *GINICON* or *MIDCON*, there is a problem with autocorrelation.[81] In these models, autocorrelation is often estimated to be .30 or higher, up to as .64 at the most. This indicates that the observations in *GINICON* and *MIDCON* are correlated across time. Due to this the size of the standard error will be higher than what is estimated, and the significance testing will be therefore be unreliable (Skog 2004: 252). One reason why this is the case with these specific variables might be that they are based only on consumption data, which may not vary much across time, and in addition the number of observations on these variables is quite a lot lower than the other inequality variables.[82]

EQUATION 2: ECONOMIC DEVELOPMENT AS THE DEPENDENT VARIABLE

The effect of political instability upon economic development appeared to be generally independent of how political instability was measured. As was shown in Table 15, most of the instability variables had a negative and statistically significant effect upon the rate of economic development in about half of the models. One exception is, however, the case of *RIOTS*: this variable had an insignificant effect in 15 out of 18 models. As I will show in section 4.2.2, this must be seen in connection with the fact that most of these models were simultaneous equations models. Further, the effects of the political instability variables on the rate of economic development did not seem to depend on whether the latter was measured as the growth rate of GDP per capita or of the rate of change in the level of investment. One can thus say that the analytical results of the effect of political instability on economic development are robust to different ways of measuring the two variables.[83]

The control variables

A close look at the different models reveals that for some control variables there are systematic differences between the models where *GROWTH* is the dependent variable

[81] This is independent of whether the model is recursive or not, and what measures of political instability and economic development (in the case of a simultaneous equations model) have been used.
[82] *GINICON* has 348 observations and *MIDCON* has 181 (see Appendix A) – samples that are further reduced in combination with the other variables.
[83] In the simultaneous equations models, the effect of instability upon growth did not appear to depend on how inequality was measured.

compared to those where *INVESTC* is the dependent variable. This is the case for three of the variables in equation 2. First, *OPEN* is has a significantly positive effect upon the rate of economic development in about half of the models, and after a closer investigation it appears that this is mostly in the *GROWTH* models. When *INVESTC* is the dependent variable, *OPEN* is very often insignificant. This is possibly the case because trade openness is not as decisive for increases in the level of investment as it is for the growth rate of the economy in general, and probably affects consumption to a greater extent than investment. Secondly, *GDP* is insignificant in about half of the models where *GROWTH* is the dependent variable, and when it is significant its effect on growth is both positive and negative. In the *INVESTC* models, however, *GDP* is insignificant in less than one fourth of the models, and when it is significant (16 out of 21 models), its effect is always negative. Thus, one can assume that the level of GDP of a country is more important for investment specifically than for economic growth in general, and that less developed countries have generally higher rates of change in the level of investment than more developed countries.

To a less degree than *OPEN* and *GDP*, *GOVCON* displays a certain pattern in this regard: While the variable is insignificant in 20 of the 21 models where *INVESTC* is the dependent variable, in 8 of the 32 models with *GROWTH* is *GOVCON* statistically significant and negative. This might indicate that the level of government consumption is irrelevant for the rate of change in the investment level, but that it sometimes has a detrimental effect upon the growth rate of the economy. Again, this is probably due to the fact that the growth rate comprises more than investments (such as consumption and technological progress).

The adjusted R^2 and estimated autocorrelation

While the adjusted R^2 in the second equation is robust to different measures of political instability, it appears to be systematically different in the models where *GROWTH* is the dependent variable compared to those where *INVESTC* is the dependent variable. Of the 21 *INVESTC* models, the adjusted R^2 is around .10 or lower in 19, while of the 32 *GROWTH* models, the adjusted R^2 is around .30 in 25, and around .10 or lower only in 7 models.[84] The only way to interpret this is that the explanatory variables included in the system produce a better fit to the models where *GROWTH* is the dependent variable compared to those where the rate of economic development is measured as *INVESTC*. In other words, the explanatory

[84] The difference between the *GROWTH* models and the *INVESTC* models in this regard is not due to different sample sizes: that is, *GROWTH* and *INVESTC* have an approximately equal number of observations.

variables account for a greater proportion of the variation in the growth rate of GDP per capita than the rate of change in the level of investment. The estimated autocorrelation is approximately equal across the models, and is above .12 only in five cases (of these, the highest is .42, and all are simultaneous equations models where either *GINICON* or *MIDCON* is the variable measuring socio-economic inequality).

SUMMARY – MEASURE DIFFERENCES

This investigation has given us a very important insight: the ways in which inequality and political instability have been measured actually affects the analytical results in the first equation: There are specific combinations of inequality and instability variables that produce the finding that socio-economic inequality affects instability negatively, others that it affects it positively, and other still that it does not have any effect upon political instability. We have also seen that the effect of some of the control variables on the level of political instability depends on how the latter is measured, and that the explanatory power of the models depends on what measures both of instability and inequality are used. Further, the inequality measures to a varying degree affect the amount of autocorrelation that is present in the model. As for equation 2, we saw that the effect of instability upon growth did not vary systematically according to what measure of either instability or economic development was used, but that the effect of the control variables in some cases depended on how economic development was measured. And while the adjusted R^2 was generally higher in the *GROWTH* models than the *INVESTC* models, autocorrelation did not differ according to what measures were used. The fact that the analytical results are not robust to different ways of measuring the variables constitutes a possible explanation of why previous studies have concluded so differently.

THE QUALITY OF THE VARIABLES: ARE THERE ANY PREFERRED MEASURES?

At this point it becomes natural to pose the following question: among all the different measures that have been included in this analysis – are some measures better than the others in terms of validity and reliability? As for the measures of the rate of economic development, there are no substantial differences among the measures in this regard. The analysis revealed that the instability variables affected the variables measuring economic development approximately equally. This tells us that to the extent that political instability affects the rate of economic development, it affects both how the level of investment changes from one year to another and the growth rate of the economy in general. Therefore, political instability is

relevant not only for investments but for other aspects of economic activity as well, and the two measures included here thus appear to be equally relevant.

As for the variables measuring political instability, one could argue that *CINDEX* is a preferred measure of political instability as it incorporates all the other measures and thus captures different aspects of political instability. However, as mentioned in chapter 3, the use of indices is not unproblematic. First, if the variables included do not measure the same phenomena to a sufficient degree, combining them in an index is not appropriate. To test whether this is the case, that is: how well *CINDEX* measures the general concept of political instability, I scrutinised the Cronbach's Alpha coefficient of the variable combination. It appeared to be .58, and thus relatively close to, but not above, the level which is generally seen as necessary in this regard (.70) (Ringdal 2001: 168).[85] Secondly, *CINDEX* is not a weighted index but merely the sum of the other six measures of instability in the analysis. Consequently, all the components have been ascribed equal importance. Whether this is appropriate is something that has not been discussed in this setting, as the evaluation of how the different expressions of instability should be weighted relative to each other would require a study of its own.

Further, this analysis has included several measures of political instability mainly for two reasons: The fact that the literature displays many different ways of measuring political instability demonstrates both the difficulty and the inappropriateness of choosing one expression of political instability, such as political assassinations, riots, revolutions, anti-governmental demonstrations or the like, and this is in part the reason why many studies construct indices to measure political instability. Political instability is an abstract term with several concrete counterparts which are all relevant in a setting where the effects of socio-economic inequality or the determinants of economic development are explored. In addition, the use of several measures enables testing the robustness of political instability and investigating whether some measures are empirically more important than others.

When it comes to the different measures of socio-economic inequality, the matter becomes more complicated. As discussed in chapter 3, a major challenge to all quantitative studies of socio-economic inequality is the limited data availability. This analysis has incorporated

[85] One could have used factor analysis to examine what instability variables measure different phenomena, and thus how an efficient index could be constructed. This is, unfortunately, outside the scope of this thesis.

several measures of inequality both to investigate whether the contrasting results of existing studies is due to the fact that different measures of inequality have been used, and because there is no agreement as to how socio-economic inequality should be measured.[86] The construction of the inequality variables faces a familiar trade-off, namely that between internal comparability and validity on the one hand, and sample size and generalisability on the other (Muller and Jukam 1983: 161). As was discussed in section 3.2.1, the inequality measures with the largest number of observations consist of different income definitions and units of analysis: *GINI* and *MIDCLASS* include all income definitions and all units of analysis; *GINIINC* and *MIDINC* include both income and consumption data, and *GINIIN* and *MIDIN* include both gross and net income and all units of analysis. The mixing of observations that are based on different income definitions and units of analysis might not be appropriate and may constitute a validity problem, as several scholars have argued (Deininger and Squire 1996; Keefer and Knack 2002; Nel 2003). Nonetheless, this is perhaps a price necessary to pay to have a sample that is large and varied enough to enable generalisation.

The measures that distinguish between different income definitions and analytical units, on the other hand, have smaller samples, and in some cases, biased samples. For example, *MIDCON* and *GINICON* include only observations that are based on consumption data. This implies not only that the sample size for these variables is much lower than for the other inequality variables, but also that they cover only developing countries. Consequently, the findings in the models where these measures of socio-economic inequality have been used cannot be generalised to developed countries. In addition, the quality rating of the data differs between developing and developed countries in a non-random way, and one could thus say that the data for developed countries are more reliable than those for developed countries.

This implies that the reliability of *GINICON* and *MIDCON* is arguably lower than of the other inequality measures, and the results of these models should not be emphasised. The analytical results revealed that the models where *GINICON* or *MIDCON* were included, differed from the other models with regard to the effect of the different variables in the two equations, and in addition the explanatory power was lower and the amount of autocorrelation was higher in these models. Thus, even though *GINICON* and *MIDCON* are internally comparable with

[86] For example, while Temple (1998: 319) argues that using the income share of the middle class is a more appropriate measure of socio-economic inequality than the Gini coefficient, the latter is the most commonly used.

regard to the way income has been measured, there is reason to doubt the reliability of the analytical results in the models including these measures of inequality. As was demonstrated earlier, most of the cases (8 out of 10) where inequality had a significant (or close to significant) and *positive* effect on political instability, inequality was measured either as *GINICON* or *MIDCON*, as opposed to only 2 of the 10 cases where it had a significant and *negative* effect. Therefore, one could say that to the extent that socio-economic inequality affects political instability, this analysis supports the hypothesis that it has a negative and not a positive effect, as some have claimed (cf. section 2.3, pages 17-19).

GINISM and *MIDSM* consist only of data that are comparable both regarding units of analysis and income definitions, and are qualitatively the best measures of inequality. The drawback, however, is the fact that many developing countries are not included in these samples due to the fact that they only report data on socio-economic inequality based on consumption or expenditure.[87] In addition, the samples are substantially smaller than for most of the other inequality variables, and their generalisability might thus be questioned. As we see there are drawbacks to all measures, and except in the cases of *GINICON* and *MIDCON*, it is hard to say that some measures are better than others. In addition, it has previously been shown that different measures of inequality are highly correlated (Clarke 1995: 405), something that is the case here as well (see Table 19 below). Therefore, it might be the variables' different samples that make them produce such different results, and not the way the variables have been constructed. This adds to the difficulty of distinguishing between the measures in terms of quality, and doing so thus might not be warranted.

Table 19: Correlation between the inequality measures

	GINI	***GINIINC***	***GINIIN***	***GINICON***	***GINISM***	***MIDCLASS***	***MIDINC***	***MIDIN***	***MIDCON***	***MIDSM***
GINI	1									
GINIINC	1	1								
GINIIN	1	1	1							
GINICON	0,97	0,97	0,90	1						
GINISM	1,00	1	1	0,91	1					
MIDCLASS	-0,91	-0,91	-0,92	-0,86	-0,92	1				
MIDINC	-0,92	-0,92	-0,92	-0,88	-0,92	1	1			
MIDIN	-0,91	-0,91	-0,92	-0,84	-0,92	1	1	1		
MIDCON	-0,89	-0,88	-0,92	-0,92	-0,92	0,96	0,96	0,89	1	
MIDSM	-0,92	-0,92	-0,92	-0,86	-0,92	1	1	1	0,89	1

[87] This is the case also for *GINIIN* and *MIDIN*.

4.2.2 DIFFERENCES BETWEEN THE RECURSIVE AND THE S.E. MODELS

In this section I will investigate to what extent the analytical results depend on whether a recursive model has been used or whether the analysis has been run using a simultaneous equations model. In equation 1, there is no systematic difference in the behaviour of the various inequality variables depending on what model has been used. As for the control variables, *POPDEN* is somewhat different in the two kinds of models: while in the S.E. models this variable is never significant, it is significant and with the expected sign in 13 cases of the recursive models. However, the fact that *POPDEN* is insignificant in about 86 of the 101 recursive models, suggests that it is not a very important variable for explaining the variation in levels of instability across countries.

Similarly, *REGIME* seems to be more decisive in the recursive models, where it has a significant effect upon the various measures of political instability in about half of the 101 models. And when it is significant, its effect is for the most part positive indicating that authoritarianism is associated with higher levels of political instability. In the simultaneous equations models, *REGIME* is insignificant in 32 of the 39 models, and in the remaining models its effect goes in both directions. Finally, the effect of *GDP* is more often significant in the recursive models, where it is significant in about half of the models, than the S.E. models, where it is insignificant in 29 of the 39 models. The difference is not as large as for *POPDEN* and *REGIME*, however, and the significant entries of *GDP* are equally distributed between positive and negative effects in the two models types.

In equation 2, it appears that there are systematic and sometimes large differences in some of the variable effects across the model types. As for the control variables, while *OPEN* is insignificant in all of the 14 recursive models, it has a significant and positive effect upon the rate of economic development in 27 of the 39 simultaneous equations models. *GDP*, on the other hand, is statistically significant in all of the 14 recursive models while it is significant only in half of the S.E. models. And in the cases in which *GDP* is significant in the recursive models, its effect is positive in 7 and negative in 7, while in the S.E. models the effect of *GDP* is almost always negative (19 of 20 significant entries).

Most important, however, is the difference in the effect of the political instability variables. As we can see from Table 20, the instability variables have a significant effect upon the rate

of economic development in almost all of the recursive models: only in 2 out of 14 models was the variable measuring political instability insignificant, and its effect was always negative. For the simultaneous equations models, on the other hand, as much as in 33 of the 39 models was political instability insignificant, and in one case it was actually significantly positive. Further, it appears that for equation 2 the adjusted R^2 is systematically different in the recursive models compared to the S.E. models when *GROWTH* is the dependent variable.[88] While the adjusted R^2 in the *GROWTH* models is generally around .30, it is approximately .06 in seven models, and it appears that these are all the recursive models.

Table 20: The effect of instability – differences between the model types

Variable	*Model*	*Total number*	*Insignificant effect*		*Significant positive effect*		*Significant negative effect*		*Expected Sign*
			N	*%*	N	*%*	N	*%*	
Instability (various)	Recursive	14	2	*14,3 %*	0	*0 %*	12 (2*)	*85,7 %*	-
PRED (various)	Simultaneous equations	39	33	*84,6 %*	1	*2,6 %*	5 (1*)	*12,8 %*	-

The common denominator seems to be the differences in the sample sizes. The recursive models in the second equation do not include the inequality variables, but merely the explanatory variables of economic development as specified in section 3.2.4, and the sample size in these models is therefore about 5830. In the simultaneous equations models, however, all of the right hand side variables, including the inequality variables, have been used to produce the predicted values. Since the inequality variables are the variables in the dataset with the fewest number of observations, the sample size in the S.E. models will be conditioned by these variables. This can explain why the adjusted R^2 is so dramatically different in the recursive *GROWTH* models than in the S.E. models, and it can also account for the different effect of the instability variables and some of the control variables on the rate of economic development in the recursive models as compared to the simultaneous equations models.[89] In this regard, it is tempting to view the results in the recursive models as more reliable, as their samples are so much larger than in the S.E. models. *Viewing these models as more reliable, I therefore conclude that political instability has a statistically significant*

[88] For the models where political instability is the dependent variable (equation 1), the adjusted R^2 does not vary systematically according to whether a recursive or an S.E. model has been employed. Further, the estimated autocorrelation does not depend on model type in any of the equations.

[89] As previously discussed, the fact that the adjusted R^2 was almost always very low for the *INVESTC* models indicates that the explanatory variables included in this analysis do not account for much of the variation in *INVESTC*.

negative effect upon economic growth. Choosing to give more weight to the models with the largest samples implies, however, also that the explanatory power of the set of variables included at in equation 2 is relatively low.[90]

An alternative explanation for the differences between the recursive and S.E. models, and one that might account for the fact that some of the control variables in equation 1 are different in the recursive models than in the simultaneous equations models, despite that there is no difference in sample size across the two kinds of models for equation 1, is the following: One instability variable and one inequality variable appear more often in simultaneous equations models than others: *RIOTS* in 16 out of 20 models and *GINISM* in 9 out of 14 models. It might thus be, in the case of equation 1, that *POPDEN*, *REGIME* and *GDP* have different effects on *RIOTS* than on the other instability variables, and in the case of equation 2, that the effect of *RIOTS* on economic development is different than the effect of the other instability variables. Alternatively, that the sample that is analysed in the presence of *GINISM* give different results than the samples produced by the other inequality variables, something that also might account for the differences in the explanatory power of the models.

4.2.3 IS THE EFFECT OF INEQUALITY ALTERED BY THE INTRODUCTION OF *SEMI*?

As described in the theoretical discussion it is a viewpoint in parts of the literature that it is not socio-economic inequality or relative deprivation that makes people engage in collective protest activity, but rather something that people might choose to do after having weighted its cost against its benefits (see section 2.3.2). To investigate whether the semi-repressiveness of the regime has a higher explanatory power of political instability than socio-economic inequality, I will include the variable *SEMI* (see section 3.2.4, and Appendix A for details on this variable) in the models where the measure of socio-economic inequality proved to have an incremental and statistically significant effect upon the level of political instability. If *SEMI* removes the significance of socio-economic inequality in these models, and is itself significant, there is reason to assume that regime repressiveness is indeed more important than socio-economic inequality in explaining political violence. Recall from the presentation of the analytical results that there were 10 models in which socio-economic inequality had a significant, or close to significant, and positive effect upon political instability.

[90] Recall the objection that the relevance of the R^2 is disputable and that it should no be given too much weight. Thus, the low adjusted R^2 does not necessarily imply a low explanatory power of these models.

Table 21: The effect of intermediate levels of regime repressiveness

Economic development measure	***Instability variable***	***Inequality variable***	***Inequality t-stat***		***SEMI t-stat***	***Adjusted R^2***		***Model****
			Without SEMI	**With *SEMI***		*Without SEMI*	**With *SEMI***	
GROWTH	CINDEX	GINIIN	*1,68*	**1,41**	**0,76**	0,40	0,40	R
	ASSASS	MIDCLASS	*-1,76*	**-1,76**	**0,58**	0,28	0,38	R
	ASSASS	MIDIN	*-2,15*	**-1,97**	**0,24**	0,29	0,29	R
	RIOTS	MIDINC	*-2,17*	**-2,15**	**-0,13**	0,05	0,06	S
	RIOTS	MIDCON	*-1,68*	**-1,96**	**-2,03**	0,12	0,12	R
INVESTC	CINDEX	GINIIN	*1,71*	**1,45**	**0,60**	0,39	0,40	R
	ASSASS	MIDCLASS	*-1,76*	**-1,75**	**0,62**	0,28	0,27	R
	ASSASS	MIDIN	*-2,11*	**-1,96**	**0,20**	0,29	0,29	R
	RIOTS	MIDINC	*-2,21*	**-2,29**	**-0,02**	0,12	0,06	S
	RIOTS	MIDCON	*-2,36*	**-2,28**	**-1,53**	0,13	0,13	R

** R = recursive model; S = simultaneous equations model*

We can see from Table 21 that the effect of socio-economic inequality was generally not affected by the inclusion of *SEMI*. As for the effect of *SEMI*, this was significant only in one model, and then with the unexpected sign. This implies that in this model, intermediate levels of regime repressiveness had a *decreasing* effect on the level of political instability. Further, we see that the fit of the model, measured as the adjusted R^2, remained largely the same in the models where *SEMI* was included as compared to those where it was not. One can thus conclude that 1) generally, the semi-repressiveness of the regime does not remove any explanatory power from the variables measuring socio-economic inequality; 2) the semi-repressiveness of the regime does not increase the level of political instability; and 3) the inclusion of *SEMI* does not add to the explanatory power of the model as a whole.[91] In other words, the hypothesis that political instability will be highest at intermediate levels of regime repressiveness is not supported in this analysis.[92]

4.2.4 EXCLUSION OF RELEVANT EXPLANATORY VARIABLES

It is commonly claimed that results from regression analyses are sensitive to changes in the set of control variables included, and it has been shown that this also concerns studies

[91] The estimated autocorrelation was around .10 or lower in all of these models, except when *MIDCON* was the inequality variable. Here, estimated autocorrelation was .38 and .37.

[92] The viewpoint that the decision to participate in acts of political protest and collective violence is a result of weighing costs against benefits might still be valid, although the notion that the repressiveness of the regime is an important conditioner of this evaluation does not seem to hold empirically. For example, it has been suggested that other variables are better measures of costs and political opportunities, such as state strength and degree of political institutionalisation (Schock 1996: 107-108).

involving income inequality (Clarke 1995: 404; Levine and Renelt 1992; Torstensson 1996). An important goal of this analysis has been to investigate why previous studies have concluded so differently with regard to the relationship between socio-economic inequality, political instability and economic growth, and the focus has been set on testing various ways of measuring these phenomena. In face of the limited scope of this thesis, the resulting large number of different models has made it necessary to have a fixed set of control variables. The analytical results could therefore have been different with other control variables, and one cannot rule out the possibility that the discrepancy in existing studies' conclusions about these relationships might partly be due to their different set of control variables. One can always debate which variables should be included, but one would probably never agree. More serious, however, is the fact that some variables that the literature has found empirically to be important explanatory variables of political instability and the rate of economic development have been impossible to include in this analysis. The reasons for this will now be presented and the consequences of their exclusion will be explored.

EXCLUSION DUE TO LIMITED VARIABLE COVERAGE

Controlling for human capital, i.e., the level of education in the population, is suggested by a wide variety of endogenous growth theories and human capital theories - Barro being perhaps the most prominent advocate for its inclusion (e.g. Barro 1991; 1997). The underlying logic is that human capital affects growth positively, by increasing the productivity of labour, elevating technological progress and facilitating technological transfers from richer countries (Krieckhaus 2004: 640). However, the limited coverage of variables measuring educational attainment in terms of number of observations, in the face of the panel design of this analysis and the varying coverage of the different variables that are included would reduce the sample size substantially and create a serious problem of missing data. In particular, the limited coverage of the inequality variables makes it problematic to include other variables with many missing values as the combination of these can wipe out a major part of the observations. In addition, measures of human capital are highly correlated with level of GDP and including both variables would therefore create multi-collinearity and violate one of the regression assumptions.

EXCLUSION DUE TO TIME INVARIANCE

The Fixed Effects Model used in this analysis renders impossible the inclusion of time-invariant country-specific variables. In equation 1, measures of ethnic, religious, linguistic

and regional fractionalisation have been excluded due to time invariance. Many studies of political instability focus on ethnic fractionalisation as one of the main driving forces behind various kinds of political instability, and several find empirical support for its causal effect (e.g. Alesina et al. 2003; Alesina and La Ferrara 2005; Easterly and Levine 1997; Fearon and Laitin 2003; Keefer and Knack 2002; Mauro 1995; Reynal-Querol 2002a; VanHanen 1999). As for equation 2, there are mainly three variables that could have been included. First, being an exporter of primary products is expected to be negatively correlated with economic development, something that is also evident in empirical investigations on this matter. Secondly, the inclusion of tropical climate and landlocked location is also suggested by Sachs and Warner (1995; 1997), who argue that agricultural productivity and health is lower in tropical climates and that these factors directly inhibit development or growth. Subsequent studies have found that these variables significantly influences growth (Bleaney and Nishiyama 2004; Easterly 2001; Easterly 2002; Krieckhaus 2006).

CONSEQUENCES

The exclusion of a relevant explanatory variable can have serious consequences for the regression estimates: if it is correlated with some of the included explanatory variables in addition to the dependent variable, the parameter estimates of these will be biased. Further, as the standard error is proportional with the variation in the dependent variable that remains unaccounted for, the standard error of the parameter estimates will be higher if relevant explanatory variables have been excluded, independently of whether it is correlated with another right hand side variable or not. This affects the significance testing, which depends on the ratio between the parameter estimate and the standard error. Thus the parameter estimates and significance testing of the included explanatory variables are not reliable, and any detected effects might be spurious. In addition, it can have consequences for the explanatory power of the model: the R^2 will be lower if a relevant explanatory variable has been excluded (Skog 2004: 271-290).

This being said, it should be emphasised that it is impossible to include all relevant variables. In some cases we are not even aware of their existence, and in other cases the correlation between the explanatory variables creates problems with multi-collinearity. In addition, parsimony is a methodological goal, and the inclusion of a high number of independent variables violates this objective. Further, as already shown, the inclusion of one of the abovementioned variables would create serious problems with losing too many data points,

and the inclusion of the others is impossible given their time-invariant characteristics.[93] Due to these circumstances, I have to accept any limitations to the analytical approach and results that might follow. At the same time, any quantitative study, especially those with a panel design using a Fixed Effects Model, face similar limitations due to excluded relevant variables, and it is doubtful whether the problem is larger in this analysis than elsewhere.[94]

4.2.5 THE EFFECT OF LEVEL OF DEVELOPMENT

Many analysts have claimed that the effect of inequality on economic growth depends on the level of development. For example, Perotti (1993) and García-Peñalosa (1995) both found that inequality is beneficial to growth in developing countries but has the opposite effect in developed ones (Temple 1998: 319). Other studies have found evidence of the contrary, namely that inequality is bad for growth in developing countries and that it has the adverse effect in developed countries (Bandyopadhyay and Basu 2005; Barro 2000; Chang and Ram 2000). Their disagreement notwithstanding, what these examples tell us is that the variables, whether in the reduced-form relationship or in a two-equation analysis with a mediating variable, may have different effects in developed as compared to developing countries.[95]

In his comment on Benabou's article "Inequality and Growth" (1996) Perotti (1996b: 78) asserts that the difference in the effect of inequality on growth in developing as compared to developed countries is due to measurement error and poor data quality in developing countries. Similarly, Knack and Keefer (1997: 327) state that "[i]f autocracies tend to be poor, and poverty and closed political systems generate less reliable data, we are less likely to detect a "true" relationship between inequality and growth among non-democracies than among democracies". As previously discussed, the quality ratings of the data in the WIID dataset do indeed differ systematically between developed and developing countries.

Whether an interaction effect between inequality and level of development exists has not been analysed specifically in this study. Although this might constitute a weakness, it is not

[93] In the latter case, one could divide the sample into subsamples and run separate analyses, but this is outside the scope of this thesis.
[94] Indeed, far from all studies of the effect of inequality on political instability, or of the effect of instability on economic growth, include all the variables that previously have been hypothesised or found to be relevant.
[95] Russett (1964: 452) contends that "extreme inequality of land distribution leads to political instability only in those poor, predominantly agricultural societies where limitation to a small plot of land almost unavoidably condemns one to poverty".

necessarily so that such an interaction effect really exists. Some studies have found that the effect of socio-economic inequality on economic growth is independent of level of development (e.g. Torstensson 1996), and the inclusion of level of GDP per capita as a control variable in all models at least captures the effect that the level of development might have. As described in the previous section, *GDP* was insignificant in about half of the models for both equations, and when it did have a significant effect it was generally negative. This might indicate that to the extent that the level of development matters, poor countries tend to be less politically stable but have higher rates of economic growth than richer countries.

A related issue is that there might be regional differences in the variable relationships. As stated by Fielding (2003b: 160), no one seriously claims that the causal link between political and economic performance is homogenous throughout the world. Further, inequality is thought to be greatest in sub-Saharan Africa and Latin-America and lowest in Northern Europe, an assertion that is empirically supported by Temple (1998: 320). The Fixed Effects Model that has been applied in this thesis implies that one has controlled for country-specific effects. That means that the effects that are detected are not biased due to country-specific circumstances. However, as discussed above, using a Fixed Effects Model also implies that one cannot include time-invariant control variables such as regional belonging. One possibility would be to split the sample and analyse these subsamples, but the number of observations would then become critically low due to the low degree of coverage of the inequality variables, and further, the scope of this analysis in terms of space does not allow it. Tests of regional effects have therefore not been conducted in this study.[96]

4.2.6 OTHER METHODOLOGICAL ISSUES

In section 3.1.2 the five regression assumptions were listed. Most of these have already been discussed, such as how well the models have been specified with regard to what explanatory variables have been included. Further, any bias due to heteroskedasticity has been removed by the inclusion of White's heteroskedasticity corrected covariance matrix, and all the analyses have estimated auto-correlation. Except in a few cases, the amount of auto-correlation present

[96] Some would argue that controlling for regions, or even countries, is flawed since it is just a way of avoiding the fact that one has not been able to find all the relevant explanatory variables. As Przeworski & Teune (1970: 18) put it: "[s]cience is concerned with the explanation of specific events by means of statements that are invariably true from one set of circumstances to another".

in the models was too low to represent a problem.[97] One of the regressions assumption is that the explanatory variables included in the equation should not be perfectly correlated, in which case there would be a multicollinearity problem related to the model. The presence of multi-collinearity boosts the standard error of the regression because the additional explanatory power of a variable that is highly correlated with another explanatory variable in the system is marginal (Skog 2004: 286). It therefore reduces the t-statistic and makes it difficult to know how much of the variation in *Y* that each of the variables actually explains. To investigate the presence of multicollinearity, I scrutinised the correlation between all the independent variables in each equation. Except for a relatively high correlation between *GDP* and *URBPOP* (Pearson's R = .70), none of the variables are correlated at a critically high level (see Appendix F).[98]

Examining the correlation coefficients is not sufficient to determine whether multicollinearity is a problem, however, and *tolerance tests* were conducted to supplement these correlations.[99] As can be seen in Tables 32 and 33 in Appendix F, the tolerance tests report serious problems of multicollinearity in both equations. However, many econometricians argue that the importance of this regression assumption is disputable, because some degree of multicollinearity will always be present (Berry and Sanders 2000: 43). As stated by Blanchard, "multicollinearity is God's will, not a problem with OLS or statistical technique in general" (Blanchard 1967, cited in Gujarati 2003: 363). It has also been argued that when acknowledged data sources are used, a high degree of multicollinearity is acceptable due to the absence of measurement error in these data (Pennings et al. 2006: 163). Although the consequences of multicollinearity should be taken into account in the interpretation of the statistical outputs, its presence should thus not be given too much emphasis.

Another regression assumption is that there is no specification error due to a non-linear relationship between the dependent and independent variables. In chapter 2 I discussed the assertion some have made that the relationship between inequality and political instability is curvilinear, and both a U-shaped and an inverted U-shaped relationship has been

[97] These cases were those in which either *GINICON* or *MIDCON* were the inequality measure, and as previously discussed, these are models with several other problems as well, and should not be ascribed very much weight.
[98] It should be noted that there is no agreement as to what is a too high correlation.
[99] Tolerance tests are run in the following way: each explanatory variable is regressed upon all of the other explanatory variables in the equation, and the R^2 is subtracted from 1. If the R^2 is high, it means that one explanatory variable is to a large degree explained by the other *X*-variables. The lower the tolerance, then, the more multicollinearity is present.

hypothesised. I tested the linearity assumption by producing scatterplots of the bivariate relationships between each of the dependent variables and the independent variable in the two models, in addition to the whole models, where the standardised predicted values entered as the independent variable and the standardised residuals as the dependent variable. I was not able to distinguish any obvious curvilinear pattern in these scatterplots. This being said, determining whether curvilinearity is present on the basis of a subjective evaluation of scatterplots necessarily implies little accuracy. In addition, the nature of the instability variables, which have a skewed distribution of observations (most have the value zero), made the interpretation even more difficult. Therefore, although the scatterplots did not reveal any clear pattern of curvilinearity, I cannot rule out the possibility that such does exist.

The final regression assumption to be dealt with is that the error terms should be normally distributed. The histograms in Appendix G show that the error terms are relatively normally distributed. Therefore, this regression assumption is regarded as met. As discussed previously, non-stationarity is a common problem in panel data analyses. Here, the Phillips-Pearon unit root test has been used as a test of non-stationarity. According to this test, non-stationarity is not present in the models of this analysis, and thus the results are not biased due to non-stationarity.[100] The statistical significance of all the models has been determined using two-tailed tests. One could argue that for some of the variable relationships a one-tailed test could have been used because the expected direction of causality is sufficiently uniform, and in that way achieving more accurate significance estimations. However, I have chosen not to exclude any possibility regarding how the explanatory variables affect the left hand side variables in the two equations as none of them should be treated as given. The consequence might thus be that some variables are in fact more significant than what appears from the models, but this is much less serious than the opposite, namely to ascribe more importance to a variable than what really is the case.

As for the generalisability and reliability of the analytical results, the following must be considered. First, there is reason to question the degree to which the results can be generalised to all countries and regions of the world, as data availability is not randomly determined (Temple 1998: 320). Appendix H lists the number of observations of each country on *GINI*, the most general inequality variable with the largest coverage. It reveals that Western

[100] Appendix I presents two examples of test output, obtained through Limdep.

countries are overrepresented in the sample as compared to other regions in terms of number of observations, especially Africa, the Middle East and Central Asia. The number of observations on inequality measures thus seems to be correlated with the level of development and to some extent also with the presence of democracy and its level of consolidation, a pattern that is not very surprising. However, and as others have concluded regarding this bias (Nel 2003: 631-632), the countries included cover a broad inequality spectrum, from highly equal to very unequal countries. In addition, there is a sufficiently large degree of variation in the sample to conclude that it is not skewed in the sense that it almost only covers countries at one end of the spectrum. Further, the fact that all regions are represented by a large number of countries, adds to the generalisability of the analytical results.

The second issue that must be discussed is the possibility that the degree of underreporting of incidents of civil disobedience, protests, violence and other manifestations of a politically unstable situation is non-randomly distributed across the sample. It is likely that such underreporting is most serious in those countries that are most politically unstable, have authoritarian regimes, or are poor (Weede 1981: 651). One obvious reason why this might be the case is the wish to downplay the presence of political dissent in the population in the face of possible international sanctions and pressure, another is the wish to avoid scaring off possible investors. Likewise, a similar bias is possibly present regarding the inequality variables as well: countries with severe income disparities might have incentives to downplay the existence of such disparities, especially if those who are controlling the financial information to statistical bureaus and research organisations come from segments of the population that want to preserve the status quo. These biases actually represent a kind of measurement error that reduces the reliability of the measures of instability and inequality. The consequence is that these variables might have different, and stronger, effects than what appears from the analytical results.

Finally, the appropriateness of using a panel design to analyse the relationships between the variables must be considered. Some argue that it is not necessarily appropriate to use panel methods with relatively high frequency data when the mechanisms being studied are quite stable over time and thus long term characteristics (Easterly 2002; Lindert and Williamson 2001). Indeed, it is the case that those that have approached the relationship between inequality and growth using panel data, have come up with different results than most analysts in the field, and have typically found a zero, non-linear, or positive relationship between

inequality and growth (Banerjee and Duflo 2003; Barro 2000; Forbes 2000). Others argue that economic inequality changes very slowly over time, while political instability changes erratically. It is therefore unlikely to observe a strong and direct relationship between inequality and political instability (Lichbach 1989: 438-439).

However, these objections are misplaced if the dataset reveals that there are relatively large variations in the data on the central variables of this analysis. For some of them this is the case, not only across sections but also across time. For example, Armenia's Gini coefficient is 28 in 1988 and 47.5 in 2003; in Bulgaria it varies between 15.6 and 36.8 from 1963 and 2003; in Taiwan between 27.7 and 50 in the same period; and in Zimbabwe the Gini coefficient was 46 in 1950 and 73.3 in 1995.[101] The variation on *GROWTH* is for most of the countries quite varied across time, but less so for OECD countries. For example, Albania's growth rate was -47 percent in 1991 and 26 in 1994, while that of Luxembourg varied between -8 and 10 percent from 1951 to 2004, although it stayed mostly between -1 and 5 percent. And as discussed in chapter 3, there are several advantages of using panel design that arguably outweigh the drawbacks related to any temporal invariance, especially regarding the instability variables, that might exist in the dataset. In particular, the fact that the data availability on socio-economic inequality is becoming greater and greater as time goes by and data is collected from a large number of countries, should be taken advantage of to measure its causes and effects more accurately. Nonetheless, Easterly's objection should be noted, and the possibility that the results could have been different if a plain cross-section had been employed, must remain open.

4.3 SUMMARY AND CONCLUSION OF ANALYSIS

The main finding of this analysis is, then, that socio-economic inequality does not seem to increase the level of political instability, and therefore it cannot be said to reduce the rate of economic development through this path of causation. The hypothesised causal pattern as described in chapter 2 has thus not been supported by this analysis. However, the analysis has also shown that this result is not robust to alternative ways of measuring inequality and instability. This finding, in turn, provides an answer to the second research question of this

[101] It should be mentioned here that the referred numbers are collected from the variable *GINI*, and the data are thus based on various income definitions and units of analysis. However, also the more internally comparable variables display similar variations. It must also be mentioned that in several countries the Gini coefficient is in fact relatively stable across time.

thesis: namely that one possible reason why previous studies of the inequality-instability nexus have concluded differently is that they diverge on this dimension.

The lack of measurement robustness notwithstanding, the general pattern in the analysis is that there is relatively strong evidence of a negative effect of political instability on the rate of economic growth, but that inequality in very few models had a positive effect upon the level of political instability. The literature offers some possible explanations of this. First, it might simply be the case that inequality does not breed political instability. For example, even though relative deprivation might still cause political violence, it is a possibility that socio-economic inequality does not produce feelings of relative deprivation in the population. As stated by Weede (2002: 98):

> Relative deprivation may result from different characteristics of nations or even be a purely psychic state largely independent of objective or measurable macro-characteristics of nations. Or, the impact of the size distribution of income on violence may depend on the presence or absence of some widely accepted ideological justifications of inequality.

It might also be that in the mobilisation of discontent other factors are more decisive than the presence of stark socio-economic inequalities or relative deprivation. Important here is the argument put forth by game theory: the role played by political entrepreneurs is decisive in the mobilisation of discontent and the presence of inequality or other societal cleavages alone are not sufficient to spur protest activity (Elster 2007; Munck 2001).

However, due to the strong theoretical grounds on which to expect a causal relationship between inequality and political instability, added to its intuitive logic, other possibilities should also be considered. Indeed, the fact that both this analysis and previous empirical studies have shown that the no-relationship finding is not robust, increases the likelihood that other explanations are more credible. The following suggestions should thus mark future research in this field. First, as previously shown, the possibility that the effect of socio-economic inequality is curvilinearly related to political instability has not been unambiguously ruled out in this analysis.

Secondly, some would argue that measuring inequality in terms of income or consumption is flawed because it does not capture "real" inequalities. Some measure inequality in terms of the distribution of land and land tenure, claiming that in developing countries this is the kind of inequality that is relevant in an instability and growth setting (Deininger and Squire 1998)

(cf. section 3.2.1). However, measuring inequality in terms of land tenure would rule out developed countries from the analysis, since this is not a relevant indicator of a person's economic ability in a modern economy. Other kinds of inequality have also been advocated for: For example, Castelló and Doménech, find evidence that human capital inequality measures (the amount of education obtained by the different income quintiles and across the Gini coefficient), provide more robust results than income inequality measures (Castello and Domenech 2002). However, it is likely that education and economic inequality is highly correlated. Generally, using these and other alternative measures of inequality imply serious problems of data availability. Consequently, despite their theoretical appeal, it is difficult to test such notions empirically.

Thirdly, and perhaps even more appealing, is the contention that the lack of a clear pattern in empirical findings is due to the fact that the relationship between inequality and instability is a conditional relationship (Lichbach 1989: 465). Within the relative deprivation framework, Gurr (1970: 230) argues that discontent, and its potential to breed violence, is tempered by the intensity and scope of normative and utilitarian justifications for political violence. That is, it depends on the degree to which overt aggression is culturally and sub-culturally sanctioned; on the degree of success of past violence; and on the legitimacy of the political system and the kind of responses it makes and has made to relative deprivation. This suggests that it would be necessary to control for the presence of such characteristics to find the "true" effect of socio-economic inequality on political instability.

It is also argued that other sources of conflict are more important, such as ethnic, religious and cultural cleavages (Sigelman and Simpson 1977: 126). It has been suggested that it is when inequality interacts with these other cleavages that it becomes an important determinant of conflict (Barrows 1976; Horowitz 2000; Lichbach 1989; Nafziger and Auvinen 2002; Reynal-Querol 2002b; Østby 2005). For example, Østby (2005: 7-11) investigates the existence of horizontal inequalities and finds that the interplay between socio-economic inequalities and identity cleavages (ethnic, religious, regional) enhances group grievances, which in turn facilitates conflict mobilisation. She finds that this relationship is further conditioned by regime type and level of political inclusiveness, that is, democracies and proportional electoral institutions are more vulnerable to horizontal inequalities because these political arrangements facilitate civil activism. Therefore, it might be the case that the link between inequality and instability does exist, but that it is dependent on the interaction with other kinds

of cleavages. The fact that such interaction effects have not been scrutinised here implies that there might be a pattern of causality between inequality and political instability that has not been captured by this analysis. Similarly, there is of course the viewpoint that the effect of inequality, or any other variable for that matter, is context-specific. But as argued earlier, context-specificity is just another way of admitting that one lacks knowledge about existing relevant variables. However, it might still be that context-specific variables have been left out that when analysed in interaction with inequality would disclose a clearer pattern of causality.

Finally, it is important to note the following: As discussed in section 3.1.1, the quantitative method applied to the research question with its variable-oriented and generalising focus, implies that the analytical results do not fit all cases and that many exceptions exist. The results of the analysis merely describe the *general patterns* in the sample, and do not capture all the nuances that exist in the real world. Further, this thesis has investigated one of several possible paths of causation between socio-economic inequality and economic development. Based on the results of this analysis one can neither say that socio-economic inequality decreases, increases or has no effect upon economic development – it has merely shown that the contention that socio-economic inequality decreases economic growth by producing political instability does not seem to hold, but again that this depends on how one measures the different phenomena. Hence, the thesis does not reject the possibility that socio-economic inequality reduces the rate of economic development through some of the other paths of causation presented in chapter 2.

5 CONCLUSION

This thesis has aimed at answering the following questions: 1) Does socio-economic inequality affect the rate of economic development negatively by producing political instability; and 2) why have previous studies reached so highly divergent conclusions? I started out by discussing the theoretical basis for the hypothesis that socio-economic inequality affects economic growth negatively by breeding political instability. The main focus in chapter 2 was given to the expected causal relationship between inequality and instability. Based on the theoretical contributions made by the literature in the field, this pattern of causality was explained in the following manner: Socio-economic inequality is an important cause of relative deprivation and as such it produces discontent. This discontent increases the number of potential participants in protest behaviour, and therefore is socio-economic inequality thought to increase the level of political instability. I then explained why political instability is expected to have a negative effect upon growth. By creating uncertainty around private property and the protection of this, political instability constitutes a disincentive to invest. As economic development depends on investments, political instability thus reduces the rate of economic growth. Through this path of causation then, socio-economic inequality was hypothesised to have a negative effect upon the rate of economic development.

In chapter 3 I presented the method used to analyse the relationships between these variables, namely a regression analysis of an unbalanced panel of 188 countries from 1950 through 2004. The hypothesis was approached using a two-equation system in which political instability was the dependent variable in the first equation, and economic development in the second. The chapter described how certain issues related to the panel design of the analysis, and the presence of simultaneity in the two-equation system, were dealt with. I then discussed the measurement and operationalisation of the variables in the analysis. An improvement relative to existing studies on this subject was the use of the most recently updated data source of socio-economic inequality, WIID2b, which is also the most extensive in coverage. While most of the previous studies have been limited to cross-section analyses, this new dataset enabled a large-N panel data analysis. Further, the detailed information about each observation that was offered by the data source made it possible to construct various measures of inequality, which together with different measures of political instability and economic

development were included in the analysis to test the robustness of the analytical findings. In chapter 4 the analytical results were presented and discussed.

FINDINGS

The analytical results did not support the hypothesis that inequality decreases the rate of economic development by producing political instability. While the analysis to a large degree established a negative effect of political stability on the rate of economic development, the vast majority of the models reported statistically insignificant effects of inequality upon political instability. However, the analysis also showed that these results are not robust to alternative ways of measuring socio-economic inequality and political instability. As the analytical results varied depending on what measures is used, one could in practice find support for whatever hypothesis one had put forth.

For example, when inequality was measured as the income share of the middle class (with income defined as disposable income (or disposable monetary income) and various income sharing units were included); when political instability was measured as the number of politically motivated murders or attempted murders during a year; and the rate of economic development as the yearly change in the amount of investment, I found that inequality increased political instability, and that instability decreased growth. In other words, *that specific constellation of variables provides support for the hypothesis that socio-economic inequality lowers growth by producing political instability.* If, one the other hand, the Gini coefficient was used to measure income inequality (with all definitions of income and units of analysis included); anti-government demonstrations was used as a measure of political instability; and the growth rate of GDP measured the rate of economic development, then I found that socio-economic inequality did not spur political instability. *Based on this specific variable combination, then, one would have to conclude that socio-economic inequality does not have a negative effect upon the rate of economic development by spurring political instability.* Another combination, still, would tell us that inequality actually affects economic development positively, because it decreased the level of political instability.

Despite the methodological improvements of this analysis with regard to data and sample size, one has thus not been able to answer the first research question of this thesis, that is, whether socio-economic inequality decreases economic growth by breeding political instability. However, using the largest and most recently updated data source on income

inequality has contributed in a different regard. It has given an answer to the second research question of the thesis, namely why previous studies on this alleged causal pattern have concluded so differently. The analysis has demonstrated that using different measures of socio-economic inequality and political instability produces different results. This can explain why the existing literature has reached such highly diverging conclusions, and it constitutes a reason why we still know very little about the relationship between socio-economic inequality, political instability and economic growth.

IMPLICATIONS OF THE FINDINGS

This finding highlights the importance of assuming a conscious and considerate use of quantitative data. The analysis has shown that the way we define and measure a phenomenon to a large extent affects the results we get, and awareness of this should mark the interpretation of the conclusions of previous studies. Probably, such "measurement dependency" is not only present in studies dealing with the political and economic effects of socio-economic inequality, but is most likely a challenge in all research fields in the social sciences. Another implication of the results of this analysis is that there is reason to doubt the validity of any policy recommendations concerning the distribution of income that is based on empirical analyses such as this. Whether or not socio-economic inequalities affect economic development negatively by breeding political instability remains unanswered, and policy-makers aiming at either reducing political instability or establishing a pro-growth environment cannot be sure of the effectiveness of adjusting the distribution of income in this regard.

However, as the empirical cases described in the first part of this thesis demonstrate, the discontent associated with stark socio-economic inequalities does sometimes result in collective protests and violence. As many observers of the violence that still marks the political scenery in Kenya have argued, promoting a more equal distribution of income thus appears as a necessary means to alleviate the tension among people and groups in the country and the human suffering that exists at the lower end of the distributional scale. In addition, it is important to note that reducing socio-economic inequality may be a goal *per se*, independently of its effects on political instability and economic growth. Indeed, even though people do not actively protest, their demands might be highly legitimate. After all, promoting human well-being and dignity is the ultimate goal of all scientific endeavours in this field, not least those that aim at explaining differences in rates of economic development.

BIBLIOGRAPHY

Abadie, A., and J. Gardeazabal. 2003. The economic costs of conflict: A case study of the Basque Country. *American Economic Review* 93 (1):113-132.

Acemoglu, D., and J. A. Robinson. 2000. Why did the west extend the franchise? Democracy, inequality, and growth in historical perspective. *Quarterly Journal of Economics* 115 (4):1167-1199.

Achen, C. H. 2002. Toward a new political methodology: Microfoundations and ART. *Annual Review of Political Science* 5:423-450.

Achen, Christopher H. 1982. *Interpreting and using regression*. Beverly Hills: Sage.

Aghion, P., E. Caroli, and C. Garcia-Penalosa. 1999. Inequality and economic growth: The perspective of the new growth theories. *Journal of Economic Literature* 37 (4):1615-1660.

Alesina, A., S. Ardagna, R. Perotti, and F. Schiantarelli. 2002. Fiscal policy, profits, and investment. *American Economic Review* 92 (3):571-589.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg. 2003. Fractionalization. *Journal of Economic Growth* 8 (2):155-194.

Alesina, A., and E. La Ferrara. 2005. Ethnic diversity and economic performance. *Journal of Economic Literature* 43 (3):762-800.

Alesina, A., and R. Perotti. 1994. The Political-Economy of Growth - a Critical Survey of the Recent Literature. *World Bank Economic Review* 8 (3):351-371.

Alesina, A., and R. Perotti. 1996. Income distribution, political instability, and investment. *European Economic Review* 40 (6):1203-1228.

Alesina, A., and D. Rodrik. 1994. Distributive Politics and Economic-Growth. *Quarterly Journal of Economics* 109 (2):465-490.

Alesina, A., S. Özler, N. Roubini, and P. Swagel. 1996. Political Instability and Economic Growth. *Journal of Economic Growth* 1 (June):189-211.

Annett, A. 2001. Social fractionalization, political instability, and the size of government. *Imf Staff Papers* 48 (3):561-592.

Atkinson, A. B. 1997. Bringing income distribution in from the cold. *Economic Journal* 107 (441):297-321.

Atkinson, A. B., and A. Brandolini. 2001. Promise and pitfalls in the use of "secondary" data-sets: Income inequality in OECD countries as a case study. *Journal of Economic Literature* 39 (3):771-799.

Auvinen, J. 1997. Political conflict in less developed countries 1981-89. *Journal of Peace Research* 34 (2):177-195.

Baltagi, Badi H. 2005. *Econometric analysis of panel data*. Chichester: Wiley.

Bandyopadhyay, D., and P. Basu. 2005. What drives the cross-country growth and inequality correlation? *Canadian Journal of Economics-Revue Canadienne D Economique* 38 (4):1272-1297.

Banerjee, A. V., and E. Duflo. 2003. Inequality and growth: What can the data say? *Journal of Economic Growth* 8 (3):267-299.

Barro, R. J. 1991. Economic-Growth in a Cross-Section of Countries. *Quarterly Journal of Economics* 106 (2):407-443.

Barro, R. J. 1996. Democracy and Growth. *Journal of Economic Growth* 1:1-27.

Barro, R. J. 2000. Inequality and growth in a panel of countries. *Journal of Economic Growth* 5 (1):5-32.

Barro, Robert J. 1997. *Determinants of economic growth : a cross-country empirical study*, *The Lionel Robbins lectures*. Cambridge, Mass.: MIT Press.

Barrows, W. L. 1976. Ethnic Diversity and Political Instability in Black-Africa. *Comparative Political Studies* 9 (2):139-170.
Bartolini, S. 1993. On Time and Comparative Research. *Journal of Theoretical Politics* 5 (2):131-167.
Beck, N., and J. N. Katz. 2004. Time-Series–Cross-Section Issues: Dynamics, 2004* Draft of July 24, 2004.
Becker, G. S., K. M. Murphy, and R. Tamura. 1990. Human-Capital, Fertility, and Economic-Growth. *Journal of Political Economy* 98 (5):S12-S37.
Benabou, R. 1996. Inequality and Growth. *NBER Macroeconomics Annual* 11:11-74.
Benhabib, J., and A. Rustichini. 1996. Social Conflict and Growth. *Journal of Economic Growth* 1 (1):125-142.
Berry, William D., and Mitchell S. Sanders. 2000. *Understanding multivariate research: a primer for beginning social scientists*. Boulder, Colo.: Westview Press.
Bertola, G. 1993. Factor Shares and Savings in Endogenous Growth. *American Economic Review* 83 (5):1184-1198.
Besancon, M. L. 2005. Relative resources: Inequality in ethnic wars, revolutions, and genocides. *Journal of Peace Research* 42 (4):393-415.
Birdsall, N., D. Ross, and R. Sabot. 1995. Inequality and Growth Reconsidered - Lessons from East-Asia. *World Bank Economic Review* 9 (3):477-508.
Bleaney, M., and A. Nishiyama. 2004. Income inequality and growth - does the relationship vary with the income level? *Economics Letters* 84 (3):349-355.
Blomberg, S. B. 1996. Growth, political instability and the defence burden. *Economica* 63 (252):649-672.
Campos, N. F., and J. B. Nugent. 2003. Aggregate investment and political instability: An econometric investigation. *Economica* 70 (279):533-549.
Castello, A., and R. Domenech. 2002. Human capital inequality and economic growth: Some new evidence. *Economic Journal* 112 (478):C187-C200.
Chang, J. Y., and R. Ram. 2000. Level of development, rate of economic growth, and income inequality. *Economic Development and Cultural Change* 48 (4):787-799.
Chenery, Hollis, and T. N. Srinivasan. 1989. *Handbook of development economics*. Amsterdam: North Holland.
Clarke, G. 1995. More Evidence on Income Distribution and Growth. *Journal of Development Economics* 47:403-427.
Collier, P., and A. Hoeffler. 2004. Greed and grievance in civil war. *Oxford Economic Papers-New Series* 56 (4):563-595.
Cukierman, A., S. Edwards, and G. Tabellini. 1992. Seigniorage and Political Instability. *American Economic Review* 82 (3):537-555.
Daily Nation, Editorial. 2008. "Violence too costly". January 14th.
Davies, J. C. 1962. Toward a Theory of Revolution. *American Sociological Review* 27 (1):5-19.
Deininger, K., and L. Squire. 1996. A new data set measuring income inequality. *World Bank Economic Review* 10 (3):565-591.
Deininger, K., and L. Squire. 1998. New ways of looking at old issues: inequality and growth. *Journal of Development Economics* 57 (2):259-287.
Dogan, Mattei, and Ali Kazancigil. 1994. *Comparing nations: concepts, strategies, substance*. Oxford: Blackwell.
Easterly, W. 1999. Life during growth. *Journal of Economic Growth* 4 (3):239-276.
Easterly, W. 2001. The middle class consensus and economic development. *Journal of Economic Growth* 6 (4):317-335.

Easterly, W. 2002. Inequality *does* cause underdevelopment: New evidence: *Center for Global Development, Institute for International Economics*.
Easterly, W., and R. Levine. 1997. Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics* 112 (4):1203-1250.
Easterly, W., and S. Rebelo. 1993. Fiscal Policy and Growth: An Empirical Investigation. *Journal of Monetary Economics* 32 (December):417-58.
Eisinger, P. K. 1973. Conditions of Protest Behavior in American Cities. *American Political Science Review* 67 (1):11-28.
Elster, Jon. 2007. *Explaining social behavior: more nuts and bolts for the social sciences*. Cambridge: Cambridge University Press.
Fearon, J. D., and D. D. Laitin. 2003. Ethnicity, insurgency, and civil war. *American Political Science Review* 97 (1):75-90.
Fedderke, J. W., R. H. J. de Kadt, and J. M. Luiz. 2001. Indicators of political liberty, property rights and political instability in South Africa: 1935-97. *International Review of Law and Economics* 21 (1):103-134.
Feierabend, I. K., and R. L. Feierabend. 1966. Aggressive Behaviors within Polities, 1948-1962 - Cross-National Study. *Journal of Conflict Resolution* 10 (3):249-271.
Feng, Y. 2001. Political freedom, political instability, and policy uncertainty: A study of political institutions and private investment in developing countries. *International Studies Quarterly* 45 (2):271-294.
Feng, Y. 2003. *Democracy, Governance, and Economic Performance: Theory and Evidence*.
Festinger, L. Katerina. 1954. A theory of social comparison processes. *Human relations* 7 (1):17.
Fielding, D. 2002. Human rights, political instability and investment in South Africa: a note. *Journal of Development Economics* 67 (1):173-180.
Fielding, D. 2003a. Counting the cost of the Intifada: Consumption, saving and political instability in Israel. *Public Choice* 116 (3-4):297-312.
Fielding, D. 2003b. Modelling political instability and economic performance: Israeli investment during the Intifada. *Economica* 70 (277):159-186.
Fielding, D. 2004. How does violent conflict affect investment location decisions? Evidence from Israel during the Intifada. *Journal of Peace Research* 41 (4):465-484.
Forbes, K. J. 2000. A reassessment of the relationship between inequality and growth. *American Economic Review* 90 (4):869-887.
Fosu, A. K. 1992. Political Instability and Economic-Growth - Evidence from Sub-Saharan Africa. *Economic Development and Cultural Change* 40 (4):829-841.
Frazer, G. 2006. Inequality and development across and within countries. *World Development* 34 (9):1459-1481.
Galor, O., and J. Zeira. 1993. Income-Distribution and Macroeconomics. *Review of Economic Studies* 60 (1):35-52.
Garcia-Penalosa, C., and S. Turnovsky. 2006. Growth and income inequality: a canonical model. *Economic Theory* 28 (1):25-49.
Gudbrandsen, Frøy. 2005. Exploring the effects of political regimes on international migration: a quantitative analysis, [F.Gudbrandsen], Bergen.
Gujarati, Damodar N. 2003. *Basic econometrics*. New York: McGraw-Hill.
Gupta, D. K. 1990. *The Economics of Political Violence: The Effect of Political Instability on Economic Growth*: Greenwood Publishing Group.
Gurr, Ted Robert. 1970. *Why men rebel*. Princeton, N.J.: Princeton University Press.
Gwartney, J. D., R. G. Holcombe, and R. A. Lawson. 2006. Institutions and the impact of investment on growth. *Kyklos* 59 (2):255-273.

Hardy, M. A. 1979. Economic-Growth, Distributional Inequality, and Political-Conflict in Industrial-Societies. *Journal of Political & Military Sociology* 7 (2):209-227.
Hardy, Melissa A., and Alan Bryman. 2004. *Handbook of data analysis*. London: Sage Publications.
Hegre, H., T. Ellingsen, S. Gates, and N. P. Gleditsch. 2001. Toward a democratic civil peace? Democracy, political change, and civil war, 1816-1992. *American Political Science Review* 95 (1):33-48.
Heston, A., R. Summers, and B. Aten. 2006. Penn World Table Version 6.2: Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
Hibbs, Douglas A. 1973. *Mass political violence: a cross-national causal analysis*. New York: Wiley.
Horowitz, Donald L. 1985. *Ethnic groups in conflict*. Berkeley, Calif.: University of California Press.
Horowitz, Donald L. 2000. *Ethnic groups in conflict*. Berkeley, Calif.: University of California Press.
Huntington, Samuel P. 1968. *Political order in changing societies*. New Haven, Conn.: Yale University Press.
Jenkins, J. C., and C. Perrow. 1977. Insurgency of the Powerless: Farm Worker Movements (1946-1972). *American sociological review* 42 (2):249.
Keefer, P., and S. Knack. 1997. Why don't poor countries catch up? A cross-national test of an institutional explanation. *Economic Inquiry* 35 (3):590-602.
Keefer, P., and S. Knack. 2002. Polarization, politics and property rights: Links between inequality and growth. *Public Choice* 111 (1-2):127-154.
Kennedy, Peter. 2003. *A guide to econometrics*. Oxford: Blackwell.
Kenny, C. 2005. Does development make you happy? Subjective wellbeing and economic growth in developing countries. *Social Indicators Research* 73 (2):199-219.
Kling, M. 1956. Towards a Theory of Power and Political Instability in Latin-America. *Western Political Quarterly* 9 (1):21-35.
Knack, S., and P. Keefer. 1997. Does inequality harm growth only in democracies? A replication and extension. *American Journal of Political Science* 41 (1):323-332.
Knowles, S. 2005. Inequality and economic growth: The empirical relationship reconsidered in the light of comparable data. *Journal of Development Studies* 41 (1):135-159.
Krieckhaus, J. 2004. The regime debate revisted: A sensitivity analysis of democracy's economic effect. *British Journal of Political Science* 34:635-655.
Krieckhaus, J. 2006. Democracy and economic growth: How regional context influences regime effects. *British Journal of Political Science* 36:317-340.
Krueger, A. B., and J. Maleckova. 2003. Education, poverty and terrorism: Is there a causal connection? *Journal of Economic Perspectives* 17 (4):119-144.
Kuznets, S. 1955. Economic Growth and Income Inequality. *The American economic review* 45 (1):1.
Kuznets, Simon. 1966. *Modern economic growth: rate, structure, and spread*. New Haven: Yale University Press.
Landa, D., and E. B. Kapstein. 2001. Inequality, growth, and democracy. *World Politics* 53 (2):264-+.
Landes, David S. 1998. *The wealth and poverty of nations: why some are so rich and some so poor*. New York: W.W. Norton.
Landman, T. Todd. 2003. *Issues And Methods In Comparative Politics*.
Levine, R., and D. Renelt. 1992. A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review* 82 (4):942-963.

Lewis-Beck, Michael S. 1980. *Applied regression: an introduction*. Newbury Park, Calif.: Sage Publications.
Li, H., and H. Zou. 1998. Income Inequality is not Harmful for Growth: Theory and Evidence. *Review of development economics* 2 (3):318.
Lichbach, M. I. 1989. An Evaluation of Does Economic-Inequality Breed Political-Conflict Studies. *World Politics* 41 (4):431-470.
Lijphart, A. 1971. Comparative Politics and Comparative Method. *American Political Science Review* 65 (3):682-&.
Lindert, P. H., and J. G. Williamson. 2001. Does Globalization make the World More Unequal?: National Bureau of Economic Research Working Paper Series.
Linehan, W. J. 1980. Political Instability and Economic-Inequality - Some Conceptual Clarifications. *Journal of Peace Science* 4 (2):187-198.
Londregan, J. B., and K. T. Poole. 1990. Poverty, the Coup Trap, and the Seizure of Executive Power. *World Politics* 42 (2):151-183.
Lucas, R. E. 1993. Making a Miracle. *Econometrica* 61 (2):251-272.
MacCulloch, R. 2005. Income inequality and the taste for revolution. *Journal of Law & Economics* 48 (1):93-123.
Mankiw, N. G. 1995. The Growth of Nations. *Brookings Papers on Economic Activity* (1):275-310.
Mauro, P. 1995. Corruption and Growth. *Quarterly Journal of Economics* 110 (3):681-712.
Midlarsky, M. I. 1988. Rulers and the Ruled - Patterned Inequality and the Onset of Mass Political Violence. *American Political Science Review* 82 (2):491-509.
Midtbø, T. 2000. Et spørsmål om tid: Tidsserieanalyse som et verktøy i samfunnsvitenskapen. *Tidsskrift for samfunnsforskning* 41 (4):58-84.
Milanovic, Branko. 2000. The median-voter hypothesis, income inequality, and income redistribution: an empirical test with the required data. *European Journal of Political Economy* 16 (3):367-410.
Mitchell, E. J. 1968. Inequality and Insurgency - Statistical Study of South-Vietnam. *World Politics* 20 (3):421-438.
Moaddel, M. 1994. Political-Conflict in the World-Economy - a Cross-National Analysis of Modernization and World-System Theories. *American Sociological Review* 59 (2):276-303.
Muller, E. N. 1985. Income Inequality, Regime Repressiveness, and Political Violence. *American Sociological Review* 50 (1):47-61.
Muller, E. N., and T. O. Jukam. 1983. Discontent and Aggressive Political-Participation. *British Journal of Political Science* 13 (APR):159-179.
Muller, E. N., and M. A. Seligson. 1987. Inequality and Insurgency. *American Political Science Review* 81 (2):425-451.
Muller, E. N., and E. Weede. 1990. Cross-National Variation in Political Violence - a Rational Action Approach. *Journal of Conflict Resolution* 34 (4):624-651.
Munck, G. L. 2001. Game theory and comparative politics - New perspectives and old concerns. *World Politics* 53 (2):173-+.
Nafziger, E. W., and J. Auvinen. 2002. Economic development, inequality, war, and state violence. *World Development* 30 (2):153-163.
Nagel, J. 1974. Inequality and Discontent - Nonlinear Hypothesis. *World Politics* 26 (4):453-472.
Nel, P. 2003. Income inequality, economic growth, and political instability in sub-Saharan Africa. *Journal of Modern African Studies* 41 (4):611-639.
North, Douglass C. 1981. *Structure and change in economic history*. New York: W. W. Norton.

Odedokun, M. O., and J. I. Round. 2004. Determinants of income inequality and its effects on economic growth: Evidence from African countries. *African Development Review-Revue Africaine De Developpement* 16 (2):287-327.
Olson, M. 1963. Rapid Growth as a Destabilizing Force. *Journal of Economic History* 23 (4):529-552.
Panizza, U. 2002. Income inequality and economic growth: Evidence from American data. *Journal of Economic Growth* 7 (1):25-41.
Panning, W. H. 1983. Inequality, Social-Comparison, and Relative Deprivation. *American Political Science Review* 77 (2):323-329.
Parvin, M. 1973. Economic Determinants of Political Unrest - Econometric Approach. *Journal of Conflict Resolution* 17 (2):271-296.
Pennings, Paul, Hans Keman, and Jan Kleinnijenhuis. 2006. *Doing research in political science*. London: Sage.
Perotti, R. 1996a. Growth, Income Distribution, and Democracy: What the Data Say. *Journal of Economic Growth* 1 (June):149-187.
Perotti, R. 1996b. [Inequality and Growth]: Comment. *NBER Macroeconomics Annual* 11:74-82.
Persson, T., and G. Tabellini. 1994. Is Inequality Harmful for Growth. *American Economic Review* 84 (3):600-621.
Plümper, T., V. E. Troeger, and P. Manow. 2005. Panel data analysis in comparative politics: Linking method to theory. *European Journal of Political Research* 44 (2):327-354.
Pritchett, L., and L. H. Summers. 1996. Wealthier is healthier. *Journal of Human Resources* 31 (4):841-868.
Przeworski, Adam, and Henry Teune. 1970. *The logic of comparative social inquiry*. New York: Wiley.
Pyatt, G. 2003. Development and the distribution of living standards: A critique of the evolving data base. *Review of Income and Wealth* (3):333-358.
Ragin, C., and D. Zaret. 1983. Theory and Method in Comparative Research - 2 Strategies. *Social Forces* 61 (3):731-754.
Ragin, Charles C. 1987. *The comparative method: moving beyond qualitative and quantitative strategies*. Berkeley, Calif.: University of California Press.
Reynal-Querol, M. 2002a. Ethnicity, political systems, and civil wars. *Journal of Conflict Resolution* 46 (1):29-54.
Reynal-Querol, M. 2002b. Political systems, stability and civil wars. *Defence and Peace Economics* 13 (6):465-483.
Ringdal, Kristen. 2001. *Enhet og mangfold: samfunnsvitenskapelig forskning og kvantitativ metode*. Bergen: Fagbokforl.
Rogowski, R and D. C. MacRae 2004. Inequality and Institutions: What Theory, History and (Some) Data Tell Us. In *APSA (American Political Science Association) Conference*. Chicago, IL.
Runciman, W. G. 1966. *Relative deprivation and social justice: a study of attitudes to social inequality in twentieth-century England*. London: Routledge & Keagan Paul.
Russett, B. M. 1964. Inequality and Instability - the Relation of Land-Tenure to Politics. *World Politics* 16 (3):442-454.
Sachs, J. D., and A. Warner. 1995. Natural Resource Abundance and Economic Growth: National Bureau of Economic Research (NBER) Working Paper Series.
Sachs, J. D., and A. M. Warner. 1997. Sources of slow growth in African economies. *Journal of African Economies* 6 (3):335-376.
Sartori, G. 1970. Concept Misformation in Comparative Politics. *American Political Science Review* 64 (4):1033-1053.

Schock, K. 1996. A conjunctural model of political conflict - The impact of political opportunities on the relationship between economic inequality and violent political conflict. *Journal of Conflict Resolution* 40 (1):98-133.
Seghieri, C., G. Desantis, and M. L. Tanturri. 2006. The richer, the happier? An empirical investigation in selected European countries. *Social Indicators Research* 79 (3):455-476.
Sen, Amartya. 1973. *On economic inequality*. Oxford: Clarendon Press.
Sen, Amartya. 1992. *Inequality reexamined*. New York: Russell Sage Foundation.
Sigelman, L., and M. Simpson. 1977. Cross-National Test of Linkage between Economic Inequality and Political Violence. *Journal of Conflict Resolution* 21 (1):105-128.
Skocpol, Theda. 1979. *States and social revolutions: a comparative analysis of France, Russia, and China*. Cambridge: Cambridge University Press.
Skog, Ole-Jørgen. 2004. *Å forklare sosiale fenomener: en regresjonsbasert tilnærming*. Oslo: Gyldendal akademisk.
Snyder, D. 1978. Collective Violence - Research Agenda and Some Strategic Considerations. *Journal of Conflict Resolution* 22 (3):499-534.
Solow, R. M. 1956. A Contribution to the Theory of Economic-Growth. *Quarterly Journal of Economics* 70 (1):65-94.
Stewart, F., and M. O'Sullivan. 1998. Democracy, Conflict and Development - Three Cases: Queen Elisabeth House Working Paper Number 15, University of Oxford.
Svensson, J. 1998. Investment, property rights and political instability: Theory and evidence. *European Economic Review* 42 (7):1317-1341.
Sylwester, K. 2000. Income inequality, education expenditures, and growth. *Journal of Development Economics* 63 (2):379-398.
Székely, M., and M Hilgert. 1999. What's Behind the Inequality We Measure: An Investigation Using Latin American Data Inter-American Development Bank Working Paper.
Tanninen, H. 1999. Income inequality, government expenditures and growth. *Applied Economics* 31 (9):1109-1117.
Taylor, John B., and Michael Woodford. 1999. *Handbook of macroeconomics*. Amsterdam: Elsevier.
Temple, J. 1998. Initial conditions, social capital and growth in Africa. *Journal of African Economies* 7 (3):309-347.
The Canberra Group. 2001. Expert Group on Household and Income Statistics: Final Report and Recommendations. Ottawa.
Thorbecke, E., and C. Charumilind. 2002. Economic inequality and its socio-economic impact. *World Development* 30 (9):1477-1495.
Tilly, C. *Collective violence in European Perspective* Center for Research on Social Organization - Working Paper Series (CRSO)
1978 [cited. Available from http://hdl.handle.net/2027.42/50953
Todaro, Michael P. 1997. *Economic development*. New York: Longman.
Tornell, A., and A. Velasco. 1992. The Tragedy of the Commons and Economic-Growth - Why Does Capital Flow from Poor to Rich Countries. *Journal of Political Economy* 100 (6):1208-1231.
Torstensson, R. 1996. Is equality really growth-promoting? *Applied Economics Letters* 3 (3):159-161.
UNU/WIDER. 2007. World Income Inequality Database: User Guide and Data Sources.
VanHanen, T. 1999. Domestic ethnic conflict and ethnic nepotism: A comparative analysis. *Journal of Peace Research* 36 (1):55-73.

Venieris, Y. P., and D. K. Gupta. 1986. Income-Distribution and Sociopolitical Instability as Determinants of Savings - a Cross-Sectional Model. *Journal of Political Economy* 94 (4):873-883.
Wachira, Nick. 2008. "A blip or a dip? Kenya's economic future at stake". *Daily Nation*, January 11th.
Weede, E. 1981. Income Inequality, Average Income, and Domestic Violence. *The Journal of conflict resolution* 25 (4):639.
Weede, E. Thomas. 2002. Some new evidence on correlates of political violence: income inequality, regime repressiveness, and economic development. *European Sociological Review* 3 (2):97.
Woo, J. 2003. Economic, political, and institutional determinants of public deficits. *Journal of Public Economics* 87 (3-4):387-426.
Woo, J. J. 2005. Social polarization, fiscal instability and growth. *European Economic Review* 49 (6):1451-1477.
World Bank. 2006. World Development Indicators (WDI).
Zak, P. J., and S. Knack. 2001. Trust and growth. *Economic Journal* 111 (470):295-321.
Zimmermann, Ekkart. 1983. *Political violence, crises, and revolutions: theories and research*. Boston: G. K. Hall.
Østby, Gudrun. 2005. Inequality, Institutions, and Instability. In *PIDDCP Workshop (Political Institutions, Development and Domestic Civil Peace)*. Oxford, UK: International Peace Research Institute, Oslo.

APPENDIX A: Detailed variable description

Table 22: The variables and their sources

Variable:	***Description:***	***Source:***	***Coverage:***
GINI	Gini coefficient (WIDER). The variable includes different income definitions, units of analysis and data sources.	WIID2b (a)	1561 country-years
GINIINC	Equals *GINI* but excludes gross income; earnings; primary, market and factor income; plus undefined income data, and data on all other units of analysis than household or family that are person-weighted.	*Ibid.*	917 country-years
GINICON	Equals *GINI* but excludes all income data that are not consumption or expenditure.	*Ibid.*	348 country-years
GINIIN	Equals GINI but excludes data on all other income definitions than income, disposable income, monetary income and disposable monetary income.	*Ibid.*	1273 country-years
GINISMALL	Equals *GINIINC* but excludes data on all other income definitions than income, disposable income, monetary income and disposable monetary income.	*Ibid.*	664 country-years
MIDCLASS	The income share of the middle class (Third and Fourth quintile) of the total income of the population. The variable includes different income definitions, units of analysis and data sources.	*Ibid.*	928 country-years
MIDINC	Equals MIDCLASS but excludes gross income; earnings; primary, market and factor income; plus undefined income data, and data on all other units of analysis than household or family that are person-weighted.	*Ibid.*	679 country-years
MIDCON	Equals MIDCLASS but excludes all income data that are not consumption or expenditure.	*Ibid.*	181 country-years
MIDIN	Equals MIDCLASS but excludes data on all other income definitions than income, disposable income, monetary income and disposable monetary income.	*Ibid.*	778 country-years
MIDSMALL	Equals *MIDINC* but excludes data on all other income definitions than income, disposable income, monetary income and disposable monetary income.	*Ibid.*	530 country-years
ASSASS *(Assassi-nations)*	Any politically motivated murder or attempted murder of a high government official or politician.	Banks' Cross-National Time-Series Data Archive*	8255 country-years
STRIKES *(General strikes)*	Any strike of 1000 or more industrial or service workers that involves more than one employer and that is aimed at national government policies or authority.	*Ibid.*	*Ibid.*
GWAR *(Guerrilla warfare)*	Any armed activity, sabotage, or bombings carried on by independent bands of citizens or irregular forces and aimed at the overthrow of the present regime.	*Ibid.*	*Ibid.*
RIOTS	Any violent demonstration or clash of more than 100 citizens involving the use of physical force.	*Ibid.*	*Ibid.*
REVS *(Revolu-tions)*	Any illegal or forced change in the top governmental elite, any attempt at such a change, or any successful or unsuccessful armed rebellion whose aim is independence from the central government.	*Ibid.*	*Ibid.*

DEMS *(Anti-government demonstra-tions)*	Any peaceful public gathering of at least 100 people for the primary purpose of displaying or voicing their opposition to government policies or authority, excluding demonstrations of a distinctly anti-foreign nature.	*Ibid.*	*Ibid.*
CINDEX	An unweighed conflict index that equals the sum of *ASSASS, STRIKES, GWAR, RIOTS, REVS* and *DEMS*.	*Ibid.*	*Ibid.*
REGIME	Regime classification: 1 = Dictatorship; 0 = Democracy	ACLP(b)	8194 country-years
POPG	Annual percentage growth of total population.	Global Development Network Growth Database (c)	8899 country-years
URBPOP	Urban population as a percentage of total population.	*Ibid.*	8176 country-years
POPDEN	The natural logarithm of population density, calculated from an area in square miles (scaling: 1000) and population. Scaling: 0.1.	Banks' Cross-National Time-Series Data Archive	8403 country-years
GDP	The natural logarithm of real gross domestic product per capita (RGDPCH): a chain index obtained by first applying the component growth rates between each pair of consecutive years, t-l and t (t=1951 to 2000), to the current price component shares in year t-1 to obtain the DA growth rate for each year. This DA growth rate for each year t is then applied backwards and forwards from 1996, and summed to the constant price net foreign balance to obtain the Chain GDP series.	Penn World Table Version 6.2 (d)	7334 country-years
GROWTH	Growth rate of real GDP per capita in constant prices, Chain series, (RGDPCH). Unit: percent in 2000 constant prices.	*Ibid.*	7146 country-years
INVEST	Investment share of real GDP per capita. Unit: percent in 2000 constant prices (RGDPL)** RGDPL is obtained by adding up consumption, investment, government and exports, and subtracting imports in any given year.		
INVESTC	Annual change in the investment share of real GDP per capita.	*Constructed on the basis of INVEST*	7146 country-years
OPEN	Trade openness in constant prices. Unit: percent in 2000 constant prices. Exports plus Imports divided by real GDP: total trade as a percentage of GDP.	*Ibid.*	8329 country-years
GOVCON	Government consumption measured as government share of real GDP per capita. Unit: percentage in 2000 constant prices.**	*Ibid.*	8312 country-years
SEMI	A dichotomous variable based on the POLITY IV regime classification (values from -10 to 10: the higher the more democratic). Here, the value 1 is given to the cases displaying values between -3 and +3; the value 0 otherwise.	POLITY IV 2004 (e)	7001 country-years

****** *Political instability variables, Banks' Cross-National Time-Series Data Archive: All the variables are derived from the daily files of The New York Times. The eight variable definitions are adopted from Rudolph J. Rummel, "Dimensions of Conflict Behavior Within and Between Nations", General Systems Yearbook, VIII [19631, 1-50).*

******** *Since 1996 has been taken as the reference year for PWT 6.0, the real shares in constant prices are the same as the current shares in 1996. The components in international dollars are moved to another year by the national accounts growth rate for that component between 1996 and the given year. This includes exports and imports. INVEST and GOVCON are obtained by dividing each of them by real GDP per capita plus exports and minus imports in 1996 prices.*

a: UNU/WIDER World Income Inequality Database, Version 2.0b: http://www.wider.unu.edu/wiid/wiid.htm

b: Jose Antonio Cheibub and Jennifer Gandhi: "Classifying Political Regimes: An Extension and an Update", Yale University, 2004

c: Global Development Finance & World Development Indicators: New York University Development Research Institute (NYU-DRI): http://www.nyu.edu/fas/institute/dri/global%20development%20network%20growth%20database.htm

d: Alan Heston, Robert Summers and Bettina Aten, Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006: http://pwt.econ.upenn.edu/php_site/pwt62/pwt62_form.php

e: http://www.cidcm.umd.edu/polity/

APPENDIX B: Countries included in the dataset

Table 23: Overview of the total of 188 countries included in the dataset[102]

Afghanistan	Djibouti	Latvia	Samoa
Albania	Dominica	Lebanon	Sao Tome and Principe
Algeria	Dominican Republic	Lesotho	Saudi Arabia
Angola	Ecuador	Liberia	Senegal
Antigua	Egypt	Libya	Serbia and Montenegro
Argentina	El Salvador	Lithuania	Seychelles
Armenia	Equatorial Guinea	Luxembourg	Sierra Leone
Australia	Eritrea	Macao	Singapore
Austria	Estonia	Macedonia	Slovak Republic
Azerbaijan	Ethiopia	Madagascar	Slovenia
Bahamas	Fiji	Malawi	Solomon Islands
Bahrain	Finland	Malaysia	Somalia
Bangladesh	France	Maldives	South Africa
Barbados	Gabon	Mali	Spain
Belarus	Gambia, The	Malta	Sri Lanka
Belgium	Georgia	Mauritania	St. Kitts & Nevis
Belize	Germany	Mauritius	St. Lucia
Benin	Ghana	Mexico	St. Vincent & Grenadines
Bermuda	Greece	Micronesia, Fed. Sts.	Sudan
Bhutan	Grenada	Moldova	Suriname
Bolivia	Guatemala	Mongolia	Swaziland
Bosnia and Herzegovina	Guinea	Morocco	Sweden
Botswana	Guinea-Bissau	Mozambique	Switzerland
Brazil	Guyana	Namibia	Syria
Brunei	Haiti	Nepal	Taiwan
Bulgaria	Honduras	Netherlands	Tajikistan
Burkina Faso	Hong Kong	Netherlands Antilles	Tanzania
Burundi	Hungary	New Zealand	Thailand
Cambodia	Iceland	Nicaragua	Togo
Cameroon	India	Niger	Tonga
Canada	Indonesia	Nigeria	Trinidad &Tobago
Cape Verde	Iran	Norway	Tunisia
Central African Republic	Iraq	Oman	Turkey
Chad	Ireland	Pakistan	Turkmenistan
Chile	Israel	Palau	Uganda
China	Italy	Panama	Ukraine
Colombia	Jamaica	Papua New Guinea	United Arab Emirates
Comoros	Japan	Paraguay	United Kingdom
Congo, Dem. Rep.	Jordan	Peru	United States
Congo, Republic of	Kazakhstan	Philippines	Uruguay
Costa Rica	Kenya	Poland	Uzbekistan
Cote d`Ivoire	Kiribati	Portugal	Vanuatu
Croatia	Korea, Dem. Rep.	Puerto Rico	Venezuela
Cuba	Korea, Republic of	Qatar	Vietnam
Cyprus	Kuwait	Romania	Yemen
Czech Republic	Kyrgyzstan	Russia	Zambia
Denmark	Laos	Rwanda	Zimbabwe

[102] The number of time points for which there is data for each country is not included here as it varies depending on what measures of inequality and instability is used.

APPENDIX C: Issues regarding the presence of simultaneity

THE SIMULTANEITY TEST

To determine whether simultaneity was present in a model, a simultaneity test was employed. This implies that when creating the predicted values with which to replace the explanatory variables that we suspect are not truly exogenous, residuals were created and included in the equation along with the predicted value and the other explanatory variables.

EQUATION 1

REGRESS; LHS=GROWTH; RHS=ONE, LNGDP, INVEST, GOVCON, OPEN, POPG; PANEL; FIXED; HET; HET=CODE; STR=CODE; KEEP=PRED; RES=RES $

REGRESS; LHS=ASSASS; RHS=ONE, PRED, RES, ASSASSL, LNPOPDEN, LNGDP, REGIME, URBPOP, GINI; PANEL; FIXED; HET; HET=CODE; STR=CODE $

EQUATION 2

REGRESS; LHS=ASSASS; RHS=ONE, ASSASSL, LNPOPDEN, LNGDP, REGIME, URBPOP, GINI; PANEL; FIXED; HET; HET=CODE; STR=CODE; KEEP=PRED; RES=RES $

REGRESS; LHS=GROWTH; RHS=ONE, PRED, RES, LNGDP, INVEST, GOVCON, OPEN, POPG; PANEL; FIXED; HET; HET=CODE; STR=CODE $

If the variable representing the residuals turn out to be statistically significant, then simultaneity is present in the model, and a simultaneous equations model must be used to analyse that specific combination of variables. As a rule of thumb, I applied an S.E. model if *RES* in at least one of the equations was significant, but in practice it was significant in both or none of the equations.

THE RANK AND ORDER CONDITIONS

Below I present the structural equations of the system of simultaneous equations. (Hereafter, β represents estimated regression coefficients of endogenous variables, and γ represents estimated regression coefficients of exogenous variables).

(1) $REVS = \beta_{11}GROWTH + \gamma_{11}REVSL + \gamma_{12}POPDEN + \gamma_{13}GDP + \gamma_{14}REGIME + \gamma_{15}URBPOP + \gamma_{16}GINI + u_1$

(2) $GROWTH = \beta_{21}REVS + \gamma_{21}GDP + \gamma_{22}INVEST + \gamma_{23}GOVCON + \gamma_{24}OPEN + \gamma_{25}POPG + u_2$

Rearranging, the equations can be represented as:

$REVS - \beta_{11}GROWTH - \gamma_{11}REVSL - \gamma_{12}POPDEN - \gamma_{13}GDP - \gamma_{15}URBPOP - \gamma_{16}GINI = u_1$

$-\beta_{21}REVS + GROWTH - \gamma_{21}GDP - \gamma_{22}INVEST - \gamma_{23}GOVCON - \gamma_{24}OPEN - \gamma_{25}POPG = u_2$

Reduced form equations (endogenous variables expressed as functions of exogenous variables):

$REVS = \pi_{11}REVSL + \pi_{12}POPDEN + \pi_{13}GDP + \pi_{14}REGIME + \pi_{15}URBPOP + \pi_{16}GINI + \pi_{17}INVEST + \pi_{18}GOVCON + \pi_{19}OPEN + \pi_{110}POPG + \varepsilon_1$

$GROWTH = \pi_{21}REVS + \pi_{22}POPDEN + \pi_{23}GDP + \pi_{24}REGIME + \pi_{25}URBPOP + \pi_{26}GINI + \pi_{27}INVEST + \pi_{28}GOVCON + \pi_{29}OPEN + \pi_{210}POPG + \varepsilon_2$

(1) First, we check the order condition:

Given: G = # of equations = 2

K = # of variables (endogenous and exogenous) = 12

M = # of variables included in the equation

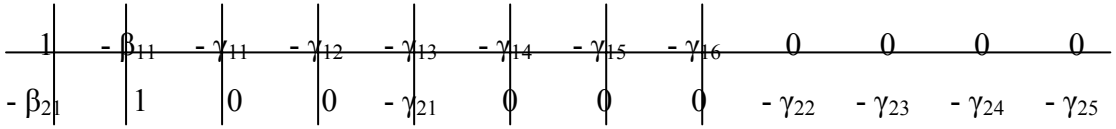
Order Equation: (K-M) must be equal to or greater than (G-1)

Equation 1: (K-M) = 12-7 = 5 > 1 = (G-1) ⟶ overidentified

Equation 2: (K-M) = 12-6 = 6 > 1 = (G-1) ⟶ overidentified

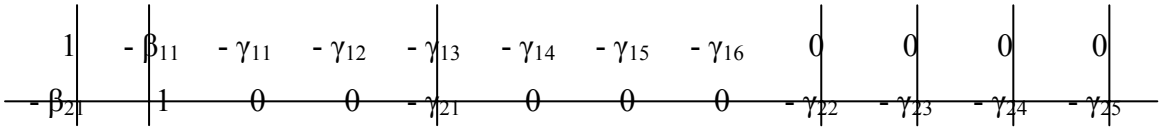
(2) Second, we check the rank condition:

Equation 1 (observe the coefficients of the matrix-form structured equations model presented above):



To check equation 1, we strike out row 1, and then strike out all the non-zero columns in row 1. If any of the - γ_{22}, - γ_{23}, - γ_{24}, - γ_{25} coefficients are non-zero then the equation is identified. We know from our estimation results that these coefficients are in fact non-zero, thus, equation 1 is identified.

Equation 2 (observe the coefficients of the matrix-form structured equations model presented above):



To check equation 2, we strike out row 2, and then strike out all the non-zero columns in row 2. If any of the – γ_{11}, - γ_{12}, - γ_{14}, - γ_{15}, - γ_{16} coefficients are non-zero then the equation is identified. We know from our estimation results that these coefficients are in fact non-zero, thus, equation 2 is identified.

APPENDIX D: Details on the inequality measures

Table 24: An example from the WIID2b dataset: Argentina 1961

Gini	*Reported Gini*	*Area Covr*	*Pop Covr*	*IncSharU*	*UofAnala*	*Equivsc*	*IncDefn*	*Source 1*	*Survey/ Source2*	*Quality*
43,4	41,90	All	All	Household	Household	No adjust-ment	Monetary Income, Disposable	Altimir 1986	Based on National Accounts	3
42,1	40,70	Urban	All	Household	Household	No adjust-ment	Monetary Income, Disposable	Altimir 1986	Based on National Accounts	3
42,3	42,30	All	All	Household	Household	No adjust-ment	Income, ..	Crom-well 1977	Weisskoff 1970	4
49,9	50,82	Agricult	All	Household	Household	No adjust-ment	Income, ..	Jain 1975	Cline 1972	4
51,3	53,11	Agricult	Income Reci-pient	Person	Person	No adjust-ment	Income, ..	Jain 1975	UN-ECLA 69	4
43,1	43,75	All	All	Household	Household	No adjust-ment	Income, ..	Jain 1975	UN-ECLA 69	4
47,6	48,95	All	Income Reci-pient	Person	Person	No adjust-ment	Income, ..	Jain 1975	UN-ECLA 69	4
41,3	42,04	Non-agricult	All	Household	Household	No adjust-ment	Income, ..	Jain 1975	Cline 1972	4
46,4	47,72	Non-agricult	All	.	Income Recipient	.	Income, ..	Jain 1975	Synthetic estimates (UN-ECLA 1970)	3
42,5	42,50	Urban	All	Household	Household	No adjust-ment	Primary Income	Lecaill-on et al. 1984	.	4
47,8	47,80	Urban	Econ. Active Pop.	.	Person	.	Primary Income	Lecaill-on et al. 1984	.	4
42,5	42,00	All	All	Household	Household/ Family	No adjust-ment	Income, Gross	Paukert 1973	Weisskoff 1970	4
43,3	43,40	All	All	Household	Household	No adjust-ment	Income, ..	Weiss-koff 1970	ACND	3
43,4	41,90	All	All	Household	Household	No adjust-ment	Monetary Income, Disposable	Altimir 1986	Based on National Accounts	3

The WIID2b dataset contains both "Reported GINI" – the Gini coefficient reported by the different sources, and "GINI", which is calculated by WIDER using methods developed by Tony Shorrocks and Guang Hua Wan to estimate the Gini coefficient from decile data "almost as accurately as if unit record data were used" (UNU/WIDER 2007: 9). When decile data were not available, "GINI" was equal to "Reported GINI". To enhance comparability, I use "GINI".

PRIORITISING INCOME DEFINITIONS

Both theoretical considerations and the coverage of the different definitions have guided the ranking of income definitions. There is disagreement in the literature whether inequality

measures should be based on income or consumption. Most industrialised countries and Latin American countries have traditionally collected data based on income while in most developing countries in Africa and Asia inequality has been assessed with reference to consumption. As societies are different, it is sometimes appropriate to use different expressions to measure the same phenomena (Przeworski and Teune 1970; Sartori 1970). In developing countries consumption is a better basis for measuring the economic situation of an individual than income. Hence, in an effort to ensure comparability of the actual inequality in different countries rather than in the means of measuring this inequality I have chosen to use income data for industrialised countries and Latin America, and consumption data for developing when available.[103]

Table 25: Prioritising the different income definitions in WIID2b

Priority	***Income category***	***Income definition***
1/2	Income	a) Disposable income b) Disposable monetary income c) Income,... d) Monetary income,…
2/1	Consumption/ Expenditure	a) Consumption b) Expenditure
3	Gross Income	a) Gross income b) Gross monetary income
4	Earnings	a) Net earnings b) Gross earnings
5	Market, factor and primary income	
6	Undefined	

Table 25 shows how the income definitions have been ranked. Income definitions such as gross income, gross monetary income, earnings and market, factor and primary earnings were only included when data with higher ranking were not available. This is because taxes and social assistance are not included in these data, while what we want for this purpose is data that reflect the *actual* economic ability of a person. Further, the coverage of most of these measures is quite small. Consumption and expenditure are grouped together as they are thought to be roughly expressing the same. However, when data on both definitions were available, consumption was prioritised before expenditure, as, to the degree that they differ,

[103] There are some exceptions from the list of countries that are defined as developing by the World Bank and the Global Development Network Growth Database. (These are: The Latin American countries, except Jamaica; Bulgaria; China; Croatia; Czech Republic; Estonia, the Republic of Korea; Latvia; Lithuania; Macedonia; Moldova, Poland, Romania, Slovakia). Income data have been prioritised before consumption/ expenditure data for these countries either because the quality of the income data is rated higher, or because they have much more data on inequality based on income than on consumption/ expenditure.

consumption is more precise (it measures what is actually consumed and not just merely purchased). Further, it is greater in coverage.

PRIORITISING UNITS OF ANALYSIS AND WEIGHTS

Next, the inequality data varied with regard to the unit of analysis. The following scheme guided the removal of duplicate country-years which varied on this dimension:

Table 26: Prioritising the different units of analysis and weights in WIID2b

Priority	***Income-sharing unit***	***Unit of analysis***	***Weight***
1	Household/family	Person	Household per capita
2			Household eq., (HBAI, national scale, OECD mod., OECD, square root, etc.)
3			Family eq. (square root, social assistance
4			Family per capita
5			Family unit per capita
6			-
7			No adjustment
8		Family/ household	Adjustment (undefined)
9			No adjustment
10		Undefined	Undefined

This priority scheme, as the former, is partly guided by conceptual relevance, partly by coverage. In line with the recommendations of WIDER and the Canberra Group for international comparisons of income distribution, the household/family is chosen as the basic statistical unit (UNU/WIDER 2007: 7). Further, "person" is prioritised before "family/household" as the unit of analysis both because the former exceeds the latter by far in coverage and because when the unit of analysis is person, it means that the needs of different sized households have been taken into account. Next, whether the observations are weighted or not constitutes an important point of variance in the dataset: when the income sharing unit is the household or family, it is not enough that the data takes into account the number of members in the household (which is the case when person is the unit of analysis). This adjustment ignores economies of scale in household consumption relating to size and other differences in needs among household members (e.g. age, sex, labour force status) (UNU/WIDER 2007: 17-18). Among the wide range of equivalence scales used in different surveys the priority scheme first differs between surveys that use such scales and those with no adjustment. It then ranks the different scales depending on coverage: the scale "household per capita" is prioritised because it is by far most widely used. This increases unity among the data used, which again promotes comparability.

PRIORITISING SOURCES

After these two selection rounds there still existed duplicates in the dataset, namely those with different sources. The total number of 228 sources implies that in some way one would have to rank them to decide which observation to include in the final dataset when two sources reported observations on the same country-year. In these situations the choices of which observation to include were guided both by the kind of source and their extension. Public and primary sources such as the World Bank (World Bank Poverty Monitoring Database 2002 and World Development Indicators, various years), the UN (1981 and 1985), international survey organisations such as Luxembourg Income Study and official national statistical units were deemed to have more credibility than the many different authors that refer to surveys conducted by these or other more or less acknowledged institutions. Of the official sources, the Deininger & Squire dataset from 1997, updated in 2004, was by far the most extensive one. As this dataset also constituted the basis for the first WIID dataset (1997), in addition to the fact that the people behind it are from the World Bank (Klaus Deiniger and Kihoon Lee) and the Global Development Network (Lyn Squire), this was the source that was first prioritised. In addition to these official data sources, some authors offer big compilations of inequality data which have been widely used in previous studies of income inequalities. These are Paukert (1973), Jain (1975), Cromwell (1977), Lecallion et al. (1984), Fields (1989) and Transmonee (2004).

This handful of sources covered most of the data, and in cases where country-years were covered by other sources, aspects such as coverage and recency guided the choice between them. I recognise that this can seem somewhat arbitrary. However, when it comes to which of these sources are most credible there is of course no agreement in the literature, and thus my guess is as good as anyone's. Further, the observations in question are quite few, they are not systematically distributed and in most instances have very similar values, so I have come to the conclusion that it will not affect the analysis in any significant way.

THE QUALITY RATINGS

Income distribution data suffer from several problems, one of them being the quality of the data (Castello and Domenech 2002: 188). And as Clarke (1995: 406) states, poor data quality can make it difficult to compare across countries. In the WIID2b dataset quality ratings are given to all the observations. It ranges from 1 (high quality) to 4 (low quality), and are based on the following criteria: 1) whether the concepts underlying the observations are known or

not; 2) the coverage of the income/ consumption concept; and 3) the survey quality, with regard to coverage issues, questionnaires and data collection methodology.

The WIID datasets builds upon the Deininger & Squire dataset from 1997 and employ the same quality rating system, however with a few improvements. Using this original dataset, Deininger and Squire (1998: 269) found in their important study "New ways of looking at old issues: inequality and growth" that the results of the analysis they conducted using their high quality dataset was not different from the findings of the existing empirical research at that time. This might suggest that the quality of the "lower quality" dataset might not be that low after all. In any case, it is arguably difficult to ensure that a subjective quality rating based upon criteria that are sometimes hard to judge by and most likely even evaluated by different people (e.g. area experts), are consistent and uniform. Further, Atkinson and Brandolini (1999, in Easterly 2002: 16) argue that many observations excluded from their high quality dataset have just as good reasons to be labelled high quality as the data included in this dataset. Thus, the quality ratings were not given very much weight when constructing the different inequality variables in this analysis. Only when all else was equal, that is, very rarely, I let the quality rating decide which observation to choose for the same country-year.

APPENDIX E: Descriptive statistics

Table 27: Descriptive statistics

Variable	*Mean*	*Std.Dev.*	*Minimum*	*Maximum*	*N*
GDP	7091,88	8049,66	170,55	84408,23	7334
GROWTH	2,04	7,58	-63,32	151,06	7146
INVEST	14,69	9,34	0,14	103,16	7334
INVESTC	0,01	2,93	-47,06	37,72	7146
OPEN	72,48	54,31	2,00	986,45	7344
GOVCON	21,93	11,17	1,53	93,72	7334
POPG	1,91	1,70	-44,41	21,76	7934
URBPOP	46,40	24,74	1,80	100,00	8175
POPDEN	3024,93	9398,82	14,00	175606,00	7429
REGIME	0,58	0,49	0	1	7222
SEMI	0,12	0,33	0	1	6813
ASSASS	0,19	0,92	0	25	7289
STRIKES	0,13	0,53	0	13	7290
GWAR	0,20	0,80	0	34	7284
RIOTS	0,45	1,83	0	55	7290
REVS	0,18	0,51	0	9	7290
DEMS	0,51	1,79	0	60	7290
CINDEX	1,66	4,08	0	87	7281
GINI	37,96	11,19	15,90	73,90	1560
GINIINC	36,59	11,36	18,00	73,90	916
GINICON	39,85	9,57	16,63	73,90	347
GINIIN	37,31	11,42	15,90	65,79	1273
GINISM	34,79	11,50	18,00	65,79	664
MIDCLASS	36,65	4,59	17,43	45,98	927
MIDINC	36,92	4,45	17,43	45,98	678
MIDCON	35,24	3,98	17,43	44,36	181
MIDIN	37,05	4,60	19,79	45,98	778
MIDSM	37,57	4,36	22,50	45,98	530
LNGDP	8,28	1,12	5,14	11,34	7334
LNPOPD	6,90	1,52	2,64	12,08	7429

APPENDIX F: Variable correlations and tolerance tests

Table 28: Correlation between the control variables in equation 1

	GROWTH	*INVESTC*	*GDP*	*URBPOP*	*POPDEN*	*REGIME*
GROWTH	1					
INVESTC	-0,08	1				
GDP	0,03	-0,01	1			
URBPOP	0,02	-0,02	0,70	1		
POPDEN	0,05	-0,02	0,15	0,21	1	
REGIME	-0,02	0,00	-0,35	-0,42	-0,03	1

Table 29: Correlation between the control variables in equation 1 and the inequality variables

	GROWTH	*INVESTC*	*GDP*	*URBPOP*	*POPDEN*	*REGIME*
GINI	-0,08	-0,04	-0,47	-0,25	-0,04	0,19
GINIINC	-0,09	-0,05	-0,52	-0,30	-0,14	0,23
GINIIN	-0,05	-0,03	-0,50	-0,23	-0,03	0,17
GINICON	-0,18	-0,07	-0,28	-0,21	-0,13	0,21
GINISM	-0,05	-0,03	-0,54	-0,22	-0,20	0,17
MIDCLASS	0,11	0,03	0,54	0,34	0,29	-0,28
MIDINC	0,12	0,05	0,54	0,33	0,28	-0,28
MIDIN	0,09	0,02	0,55	0,29	0,29	-0,28
MIDCON	0,13	0,12	0,33	0,32	0,24	-0,13
MIDSM	0,10	0,02	0,53	0,02	0,27	-0,25

Table 30: Correlation between the control variables in equation 2

	GDP	*INVEST*	*OPEN*	*GOVCON*	*POPG*
GDP	1				
INVEST	0,31	1			
OPEN	0,19	0,20	1		
GOVCON	-0,20	-0,10	0,19	1	
POPG	-0,09	-0,17	0,00	-0,05	1

Table 31: Correlation between the control variables in equation 2 and the instability variables

	GDP	*INVEST*	*OPEN*	*GOVCON*	*POPG*
ASSASS	-0,03	-0,04	-0,09	-0,04	-0,02
CINDEX	-0,06	0,01	-0,22	-0,04	-0,03
DEMS	0,03	0,06	-0,15	-0,04	-0,06
GWAR	-0,08	-0,04	-0,11	0,00	0,04
REVS	-0,17	-0,16	-0,10	0,03	0,06
RIOTS	-0,03	0,02	-0,16	-0,03	-0,01
STRIKES	0,00	0,01	-0,13	-0,03	-0,05

Table 32: Collinearity statistics: Tolerance test – equation 1 (example)*

Dependent Variable	***Tolerance***
GDP	0,04
URBPOP	0,02
POPDEN	0,01
REGIME	0,27
GINI	0,17

** Lagged variables have not been included*

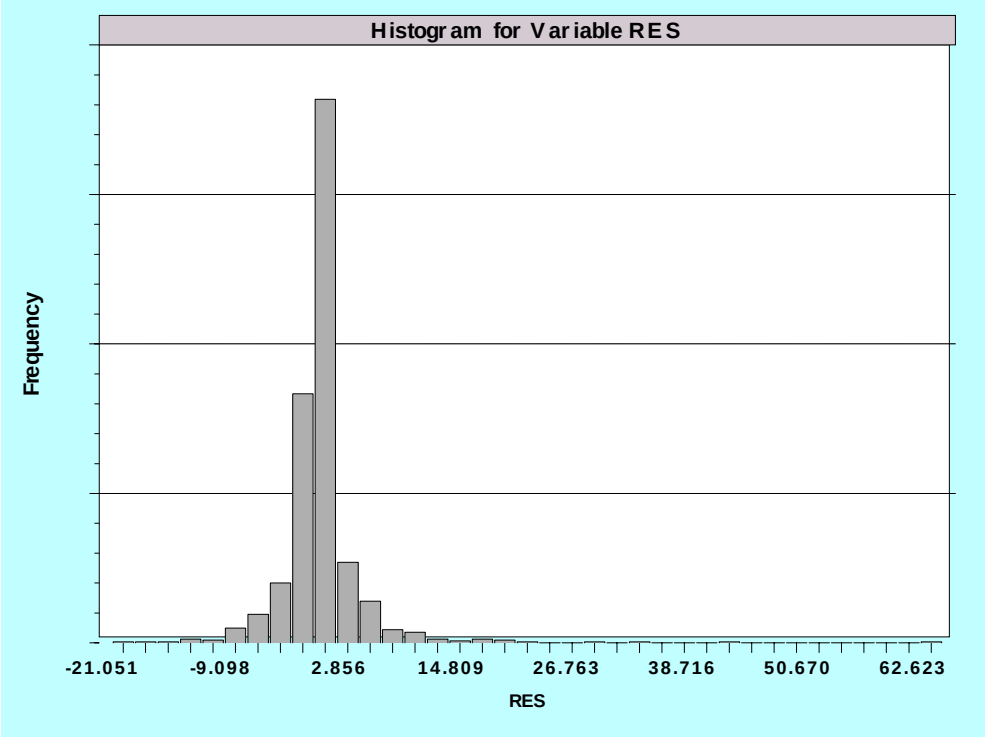
Table 33: Collinearity statistics: Tolerance test – equation 2 (example)*

Variable	***Tolerance***
GDP	0,07
INVEST	0,28
OPEN	0,23
GOVCON	0,20
POPG	0,52
CINDEX	0,76

** Lagged variables have not been included*

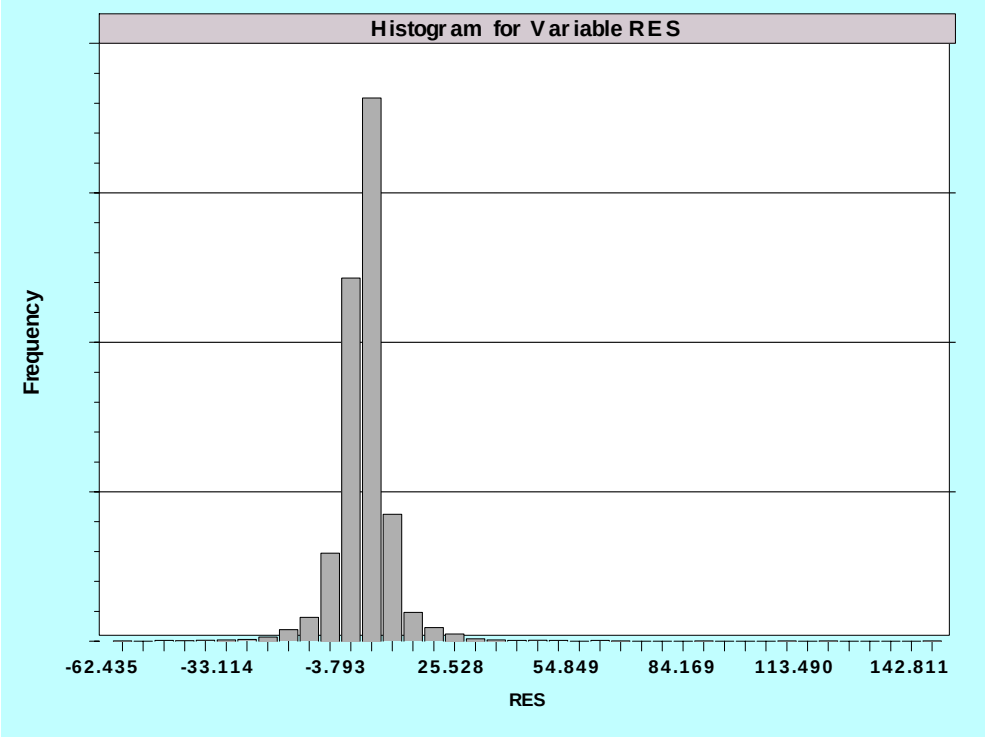
APPENDIX G: The distribution of the error terms

Figure 6: Frequency distribution of the error terms and the normality curve: Equation 1



(*Inequality variable: GINI)*

Figure 7: Frequency distribution of the error terms and the normality curve: Equation 2



(Instability variable: CINDEX)

APPENDIX H: Country coverage of the Gini variable

Table 34: Overview of countries covered by the Gini variable and their number of observations

Western countries		***Africa***		***Latin America***		***Asia***		***Middle East and Northern Africa***		***Central Asia and Eastern Europe***	
Country	*N*	*Country*	*N*	*Country*	*N*	*Country*	*N*	*Country*	*N*	*Country*	*N*
Albania	1										
Bosnia and Herzegovina	1										
Cyprus	2	Central African Republic	1								
Croatia	4	Congo, Republic of	1								
Serbia and Montenegro	5	Liberia	1								
Switzerland	6	Mozambique	1								
Latvia	9	Namibia	1								
Lithuania	9	Rwanda	1								
Portugal	9	Seychelles	1								
Macedonia	12	Somalia	1								
Austria	11	Swaziland	1	Cuba	1						
Belgium	13	Togo	1	Suriname	1						
Germany	13	Botswana	2	Uruguay	1						
Ireland	13	Burundi	2	Guyana	2						
Luxembourg	13	Djibouti	2	Barbados	3						
Slovenia	14	Guinea	2	Nicaragua	3						
Greece	15	Guinea-Bissau	2	Haiti	3	Papua New Guinea	1				
Romania	15	Uganda	2	Argentina	4	Laos	2				
Estonia	16	Burkina Faso	3	Bolivia	6	Mongolia	2				
France	16	Cameroon	3	Ecuador	7	Cambodia	3				
Poland	19	Zimbabwe	3	Paraguay	7	Fiji	3				
Spain	20	Gabon	4	Trinidad and Tobago	7	Vietnam	4				
Hungary	21	Gambia, The	4	Guatemala	8	Nepal	5	Benin	1		
Slovak Republic	22	Lesotho	5	Puerto Rico	8	Hong Kong	10	Lebanon	1		
Netherlands	23	Mauritania	5	Bahamas	14	Singapore	10	Jordan	2	Tajikistan	3
Czech Republic	24	Senegal	5	Honduras	14	Sri Lanka	10	Yemen	2	Azerbaijan	5
Finland	25	Sierra Leone	5	Dominican Republic	15	Indonesia	11	Iraq	3	Uzbekistan	6
Norway	26	Malawi	6	Jamaica	15	Philippines	12	Iran	6	Kazakhstan	7
Denmark	27	Mauritius	6	Peru	15	Malaysia	14	Israel	6	Belarus	12
Canada	30	Ghana	7	Panama	16	Bangladesh	15	Algeria	2	Kyrgyzstan	12
Italy	31	Madagascar	7	El Salvador	17	Thailand	17	Morocco	7	Russia	12
Sweden	33	Nigeria	8	Colombia	18	Korea, Republic of	20	Tunisia	7	Turkmenistan	5
Bulgaria	38	Zambia	8	Mexico	19	Pakistan	20	Egypt	9	Georgia	4
United Kingdom	44	Cote d`Ivoire	9	Chile	23	China	28	Sudan	2	Armenia	6
United States	54	Tanzania	9	Costa Rica	25	Japan	30	Chad	1	Moldova	6
Australia	17	Kenya	10	Brazil	29	India	33	Mali	2	Ukraine	9
New Zealand	18	South Africa	12	Venezuela	30	Taiwan	37	Niger	4	Turkey	9
Total:	**669**	**Total:**	**141**	**Total:**	**311**	**Total:**	**287**	**Total:**	**55**	**Total:**	**96**
Per country	**18**	**P. c.:**	**4**	**P. c.:**	**12**	**P. c.:**	**15**	**P. c.:**	**4**	**P. c.:**	**9**

APPENDIX I: Phillips-Pearon unit root test of non-stationarity

As can be seen from the two outputs, both Z (tau) and Z (rho) are well beyond the values indicated by all three confidence intervals, but perhaps more importantly, the slope coefficients are not significantly negative, but actually positive and statistically significant (.58 and .17, respectively). Therefore, non-stationarity does not represent a problem in these models.

EQUATION 1:

```
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant       .70512106      .04302947    16.387   .0000
 CINDEXL        .57560990      .00971882    59.226   .0000    1.67611793

+-------------------------------------------------------------------------+
| Phillips - Perron tests for unit root                                   |
| Sample size  =  7089, Number of regressors =    2                       |
| Sample statistics:        Z(tau) =   -42.4044 |Z(rho) = -2772.3866      |
|                           ------[Z(tau)]------+------[Z(rho)]------     |
| 3 cases (models)            99%    95%    90% | 99%    95%    90%      |
| 1. y(t)=ry(t-1)+u(t)       -2.58  -1.95  -1.62|-13.80  -8.10  -5.70      |
| 2. y(t)=a+ry(t-1)+u(t)     -3.43  -2.86  -2.57|-20.70 -14.10 -11.30      |
| 4. y(t)=a+ry(t-1)+dt+u(t)  -3.96  -3.41  -3.12|-29.50 -21.80 -18.30      |
+-------------------------------------------------------------------------+
```

EQUATION 2:

```
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant      1.73476653      .09197832    18.861   .0000
 GROWTHL        .16999058      .01174027    14.479   .0000    2.01230135

+-------------------------------------------------------------------------+
| Phillips - Perron tests for unit root                                   |
| Sample size  =  6962, Number of regressors =    2                       |
| Sample statistics:        Z(tau) =   -70.7590 |Z(rho) = -5815.2880      |
|                           ------[Z(tau)]------+------[Z(rho)]------     |
| 3 cases (models)            99%    95%    90% | 99%    95%    90%      |
| 1. y(t)=ry(t-1)+u(t)       -2.58  -1.95  -1.62|-13.80  -8.10  -5.70      |
| 2. y(t)=a+ry(t-1)+u(t)     -3.43  -2.86  -2.57|-20.70 -14.10 -11.30      |
| 4. y(t)=a+ry(t-1)+dt+u(t)  -3.96  -3.41  -3.12|-29.50 -21.80 -18.30      |
+-------------------------------------------------------------------------+
```