Economic inequality, democratization and civil conflict

Measuring the long-run effects

Ranveig Drolsum Flaten

Master Thesis at the Department of Political Science

UNIVERSITY of OSLO

Spring 2012

Economic inequality, democratization and civil conflict: Measuring the long-run effects

II

2012

Economic inequality, democratization and civil conflict: Measuring the long-run effects

Ranveig Drolsum Flaten

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Abstract

The democratization literature has suggested the possibility of a dynamic interplay between democratization and civil conflict (see Boix 2003; Acemoglu and Robinson 2006). In this thesis I argue that regime changes and civil conflict in many instances are so closely associated, that they should be studied in conjunction. I develop an analytical framework where economic inequality is seen as an important determinant for the speed and ease of democratization, and for the amount of civil conflict observed during its course. High inequality is hypothesized to both cause regime instability and violence.

Using a cross-sectional time-series dataset consisting of 164 countries observed between 1960–2010, I estimate a multinomial logit model of changes to countries' regime type and conflict level. Further, I use this model to simulate the long-run development of political regimes and civil conflict over a time-span of 40 years, taking the level of economic inequality into account. I extend a simulation routine developed in Hegre et al. (forthcoming), where civil conflict and the political systems in the world are endogenous, and evolve in accordance with the estimates of the multinomial logit model. The explanatory variables are assumed exogenous and kept at a constant level. Doing this, I am able to measure the overall effects of inequality on democratization and on conflict incidence, by capturing the reciprocal causality between the two events, as well as the impact from previous regime and conflict history.

The analysis provides support for the hypothesis that high economic inequality increases the amount of conflict in the long run, as well as the proportion of years with partially democratic institutions. High inequality seems to make complete democratization more difficult to achieve.

Acknowledgements

I will start this thesis by thanking my supervisor Håvard Hegre. His advice have been extremely helpful and constructive, and I highly appreciate his involvement in my project, and for guiding me along the way.

The thesis has been written in conjunction with the project *Conceptualization and Measurement of Democracy*(RCN project 204454/V10, see http://www.sv.uio.no/*isv*/ english/research/projects/ conceptualization-and-measurement/index.html. Thanks to Håvard Hegre, Carl Henrik Knutsen and Håvard M. Nygård for organizing the seminar group in conjunction with the project. Moreover, I would like to thank them for sharing their knowledge, and literature, and answering all of my questions. I would also like to thank the master students who participated in the group for useful comments and inspiring discussions.

I would like to thank Joakim Karlsen for invaluable technical support. Moreover, I would like to thank the people who attended PRIO's Brownbag on the 21th of March for helpful comments.

Two years at Blindern have come to an end, and during this time I have gotten to know many wonderful people. You have made my time here so much more pleasant, and for that I am extremely grateful. Last, but not least, I would like to thank Nils for reading every page of my thesis and providing me with help along the way, but most importantly for his constant moral support and encouragement.

Remaining errors are my responsibility alone.

Wordcount: 58 885 (all included)

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The number of democratic regimes has increased over the course of the last decades. Alongside the wave of democratization, 'democracy promotion' has become an important part of many Western governments' foreign policy agendas; democracy is both considered to have an inherent value, and to promote peace and prosperity (McFaul 2004). However, the road towards democracy, i.e. the process of democratization, may not necessarily be straightforward. Research has shown that transitions from an autocratic to a democratic system of government are associated with an increased risk of civil conflict (see e.g. Cederman, Hug and Krebs 2010). Moreover, democratization does not always result in the introduction of a Western-style democracy, but rather in the construction of partially democratic regimes (Diamond 2002). Why are some democratization processes more violent than other, and why are some incomplete? Moreover, what are the implications of these differences for countries that have not yet started to democratize, or are currently undergoing institutional changes?

The demise of the autocratic regimes in Tunisia, Libya, Egypt in 2011, as well as recent events in Syria, have underlined the importance of these issues. The events have become known as 'the Arab Revolutions', and they represent the largest political upheaval in the Arab world for decades (see Filiu 2011). Furthermore, they serve to illustrate how different countries' democratization processes follow different dynamics. In Tunisia, President Zine El Abidine Ben Ali withdrew from power after only a short period of protest, whereas in Syria, the regime has proven persistent and President Bashar al-Assad refuses to step down. As neither the regime, nor the opposition have been willing to make concessions, the situation has evolved into a massive bloodshed (News 2012).

The relationship between democratization and civil conflict is at the heart of this thesis. Drawing on the theories of Boix (2003) and Acemoglu and Robinson (2006),

who consider the dynamic interplay between violence and institutional change, I develop an analytical framework that tries to explain the logic behind democratization *and* its relationship to civil conflict. The basis for inquiry is the notion that countries experience different types of democratization processes; some are swift and peaceful, whereas others are long-drawn and violent. The main argument is that the level of economic inequality may determine the speed and ease of democratization. By increasing the demand for democratization, the elites resistance towards it, and the population's willingness to use violence to achieve it, high inequality may both induce regime changes and provoke civil conflict.[1]

When investigating the relationship between inequality, democratization and civil conflict I take a different point of departure than previous studies have done. Regime changes and civil conflict are in this thesis studied in conjunction, as two closely interrelated phenomena. Using a cross-sectional time-series dataset which includes information from 164 countries between 1960–2010, I construct a multinomial logit model that is able to estimate transitions between political regimes, civil conflict and peace simultaneously. Intrigued by previous research that has found a short-run increase in the risk of civil conflict from democratization, I focus on the long-run effects. By extending a simulation routine developed for the project 'Predicting armed conflict' (see Hegre et al. forthcoming), I evaluate the consequences of economic inequality on the prospects of democratization and civil conflict incidence over a period of 40 years.

1.1 Motivation

There are three main issues that motivates this thesis, and I address each of them in the following.

Firstly, there is a discrepancy between the theoretical understanding of the relationship between democratization and violent conflict, and the way in which these two phenomena are investigated within quantitative research: The dynamic relationship between violence and democratization is, as highlighted above, present in some theories of democratization Moreover, empirical evidence, for example recent events during the 'Arab Revolutions', indicates that regime changes and violence tend to occur in conjunction. Although the effects of institutional change on the risk of civil conflict are previously investigated (see e.g. Cederman, Hug and Krebs 2010; Hegre et al. 2001; Gleditsch, Hegre and Strand

[1] In the introduction I use institutional change and regime change, as well as inequality and economic inequality, interchangeably. Moreover, civil conflict refers to internal armed conflict. All of these concepts are more thoroughly explained in chapter 2.

2009), there has not been conducted any research where the two events are actually studied simultaneously. If we are to take seriously the the theoretical contributions, and the historical evidence, it should be mirrored in the way in which the phenomena are investigated empirically. Also, if violence is part of a process of institutional change the onset of conflict may manifest itself both before, as well as after, the political institutions are altered. As the causal sequence is ambiguous the statistical measurement should be able to take this into account. By constructing the model of political regime and civil conflict transitions I am able to capture the endogenous relationship between regime changes and civil conflict.

Secondly, inequality has been suggested as a trigger of civil conflict, as well as an impediment for democratization: Gurr (1970) argued that inequality can provoke rebellion. Inequality can induce 'relative deprivation', which is linked to violence through a psychological 'frustration-aggression'-mechanism. Within the literature on democratization, Boix (2003) and Acemoglu and Robinson (2006) argue that high inequality is a hindrance for democratization. When inequality is high the pressure on the elites to redistribute some of their wealth is great. Pressure for redistribution is equivalent to a pressure for democratization. Thus, inequality is considered a proxy of the potential welfare-loss of the elites under democracy. Despite inequality having a place within both literatures, there has, to my knowledge, not been conducted any statistical studies that scrutinize the link between inequality, regime changes, and civil conflict. This study attempts to remedy the gap.

There has been scarce evidence linking inequality to either civil conflict or democratization (see Collier and Hoeffler 2004; Fearon and Laitin 2003; Barro 1999; Houle 2009). One may therefore ask why it is necessary to once again examine its effect. I argue that measurement-problems, at least partly, explain the previous lack of findings. Former quantitative studies have often relied on the Gini coefficient (see Deininger and Squire 1996), although it has serious weaknesses. The coverage is poor, both over time and across countries, which has also been acknowledged in previous studies (see e.g. Fearon and Laitin 2003; Barro 1999). Moreover, the Gini coefficient is not very suitable for cross-country comparison, as it is constructed from different types of data in different countries (Houle 2009). The search for more valid measures of inequality is therefore important.

In this thesis, I use *capital share* to measure inequality. The variable has a straightforward coding – it measures the value-added that accrues to the owners of capital in the industrial sector – and better coverage (Houle 2009). Furthermore, as it measures the relative income between capital owners and laborers it is more directly related to the

theoretical contributions I draw on in this thesis (see Boix 2003; Acemoglu and Robinson 2006).

Thirdly, although the short-run effects of democratization on civil conflict are previously investigated (e.g. Hegre et al. 2001; Cederman, Hug and Krebs 2010), it has not been focused on capturing the impact of institutional changes on the conflict level *over time*. Hegre et al. (2001, 44) state that countries eventually are likely to end up as stable democracies and attain civil peace. However, they also emphasize that how long time a country spends in a semidemocratic category is decisive for how fast violence can be reduced. They acknowledge that in order to enhance our understanding of the mechanisms at work the long-run patterns of regime changes and civil conflict must be studied. I follow up on their proposal.

When scrutinizing the long-run effects there are two important issues of measurement that should be regarded. First, conflict history and previous regime instability is found to increase the risk of future violence and regime changes (see e.g. Clague et al. 2003; Strand et al. 2012; Collier, Hoeffler and Söderbom 2008; Hegre et al. forthcoming). Ignoring this 'shadow of the past' may underestimate the long-run effects. Second, accepting that there is an endogenous relationship between regime changes and conflict also has implications for aggregate effects over time. In order to measure the total consequences of economic inequality on democratization and civil conflict both the repercussions from recent conflict and regime history, as well as the reciprocal causality between regime changes and civil conflict must be taken into account. With the simulation procedure applied in this thesis I am able to capture both of these mechanisms, and how they affect each other over time, by updating the probabilities of regime and conflict transition during the course of the simulation.

Knowledge about all of the above issues; the reciprocal causality between regime changes and conflict, the impact of inequality on democratization and violence, and the overall, long-run effects of their interrelationship, is necessary to increase our understanding of the peace-generating potential of democratization, and for the possibilities of consolidating democracy.

In light of the issues discussed above, I formulate two research questions:

1) *Does economic inequality increase the risk of regime changes and being in a state of civil conflict?*

2) *How does economic inequality relate to democratization and civil conflict in the long run?*

Apart from presenting both theoretical and empirical contributions to the scholarly

community, the issues of inequality, democratization, and civil conflict are also highly relevant for policy makers. If the underlying economic structure matters for the amount of conflict observed over time and for the opportunities to establish democratic institutions, a certain amount of redistribution may be necessary to achieve durable peace and consolidate democracy. Ignoring the impact of inequality on the dynamics of these processes may not only increase the danger of failed policy, but more seriously, it may amplify long-run political instability and violence.

After conducting the analysis the results indicate that economic inequality may have an effect on the course and outcome of democratization, as well as its amount of violence. Higher inequality coincide with more civil conflict, as well as having partially democratic institutions in the long run.

1.2 Structure of the thesis

The thesis is divided into 8 chapters. Chapters 2–4 are theoretical, and chapters 5–7 present the methodology, the data, and the empirical analysis. Chapter 8 concludes the study.

In the thesis, I refer to many concepts which in the literature are disputed and/or used ambiguously, and chapter 2 clarifies some of the most relevant. Chapter 3 provides a review of the literature. The first part of the chapter deals with the literature on regime change and democratization, whereas the second part presents some theoretical contributions and prominent empirical studies within the civil conflict literature. The focus is directed at studies that have dealt specifically with the relationship between political regimes, institutional changes and civil conflict.

Chapter 4 develops the analytical framework, which is used as a point of departure for the subsequent empirical analysis. I discuss the relationship between economic inequality, relative deprivation and violence, as well as the link between economic inequality and democratization. The main argument is built around the notion that democratization processes differ, both in their length, course and outcome. Three 'ideal types' of democratization is used as a point of departure for an investigation of the short- and long-run effects of inequality on conflict incidence and political regimes. The analytical framework is the basis for the hypotheses, which I present at the end of chapter 5, i.e. after the presentation of the research design. The hypotheses are derived as observable effects, or manifestations, of the 'ideal types' and focuses on the effects of economic inequality on democratization and civil conflict, in the short and long run.

The research design is presented in chapter 5. I describe the multinomial logistic model, the transition probability matrix, and the methodology of simulations. I justify my choice of methods, and highlight the value-added of my specific research design. In chapter 6 the data and indicators used in the empirical analysis are presented. Chapter 7 reports the results from the analysis. First, the short-run analysis is presented. It consists of the results from the multinomial logit estimation. Next, I move over to the long-run results, which is divided into two sections. The first section presents global simulation results of the long-run relationship between economic inequality, conflict incidence and political regimes. In the second part I try to isolate the effect of inequality, by conducting a number of counterfactual 'experiments' where I change the level of inequality, while keeping income level and political and conflict history constant. At the end of the chapter, I discuss the uncertainty of the results, and some important issues concerning the validity of capital share. Chapter 8 concludes the study, by summarizing and discussing the main findings. I point at some caveats, and highlight this thesis' contribution to our understanding of the relationship between economic inequality, democratization and civil conflict.

Chapter 2

Conceptualization

In this chapter I present and define the most relevant concepts of this thesis. I find it useful to clarify their meaning initially, to lay down the basis for the subsequent theoretical and empirical investigation, and avert misconceptions. I start off with the definition of political regimes. The main dimension used to classify regimes is degree of democracy, and I therefore provide a brief review of democracy definitions. This thesis apply a tripartite classification of political regimes, and I argue for its relevance in section 2.1.2. Civil conflict is the second phenomenon of interest. I choose to look upon civil conflict as a political tool, and as a constituent part in a cycle of violence. The interrelationship between civil conflict and regime change, the crux of this theses, is conceptualized in section 2.3. In section 2.4 I touch upon the various understandings of inequality, and argue why I focus on economic inequality. At the end of the chapter I account for the aspect of time, separating between the short and the long run.

2.1 Political regimes

Central to any theory of democratization is the notion of political regimes. There is a continuous debate on how to characterize and categorize political regimes. A particular contended issue concerns which dimensions should be given prominence. The most common dimension used to differentiate between political regimes is *democracy*. As my second research question directly relates to to democratization, to think about political regimes in terms of their degree of democracy is relevant. Secondly, I emphasize the importance of political institutions:

Political institutions are considered to be manifestations of *patterns of authority*; they are a "set of asymmetric interactions among hierarchically ordered members of social units

that involve the direction of the unit" (Eckstein 1973, 1153). Authority patterns include both written rules and practices about the way in which power is organized and executed in a system of governance, as well as the invisible and unspoken rules about who holds power over whom (Ibid.). My emphasis is on the formal rules and practices. Political regimes are classified in the following, but first, I briefly present two prominent categories of democracy definitions.

2.1.1 Defining democracy

There is no consistent understanding of democracy in the literature. I leave aside normative discussions of the most 'correct' way of conceptualizing democracy, and focus on how the theoretical definition has consequences for the possibilities for valid measurement.

Knutsen (2011, 46) provides a very thorough review of democracy definitions. He differentiates between institutional and substantive definitions of democracy, whereof, he claims, the first is functional and narrow, whereas the latter offers more room for a broader understanding of democracy.

Institutional definitions of democracy tend to be minimalist in character, and focus on the functioning of some specified institutions considered to represent the core in a democratic system of governance (Knutsen 2011). Schumpeter (1976, 269) is a much-cited advocate of a minimalist approach. He views democracy as the "institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people's vote". Thus, a main feature of institutional definitions is a focus on contested elections; they are so-called 'electoral definitions' of democracy. As electoral definitions are narrow they have been criticized for placing too much emphasis on what happens at election day, ignoring the preliminary election period, as well as the post-election implementation of policies (Knutsen 2011, 56).

Substantive definitions of democracy, on the other hand, point to some core principles which are viewed as the essence of what democracy is. One core principle is "popular rule of popular control over collective decision making" (Beetham 1999, 90), and an additional core principle is *political equality*. Political exclusion of for example minorities or women would count as political inequality, and therefore disqualify a country as democratic.[1] Moreover, the people, rather than specific institutions, are the point of departure. Institutions are only considered to be instruments that underpin democracy (Knutsen 2011, 59).

[1] An extreme form of political inequality is dictatorship, where everyone but a small elite is excluded from political decision-making.

A further division can be made between dichotomous and continuous democracy classifications. Substantive definitions are considered to be more suitable for those who argue that democracy is a continuous concept, because an encompassing, multidimensional definition, *in general*, is incompatible with a sharp division between autocracy and democracy. In this sense, substantive definitions secure better face validity (Knutsen 2011, 64).[2] However, including several, and arguably more normative, aspects of democracy creates challenges. The more encompassing, the higher the propensity for vagueness, which may provide definitions which are difficult to operationalize (Knutsen 2011, 61). This is where minimalist definitions have their strength. As they are arguably easier to operationalize they also have the potential of securing better content validity.[3] This is one reason why I have chosen an institutional definition of democracy.

2.1.2 Classifying political regimes

This thesis accentuates the importance of political institutions in a system of governance, and this is the point of departure for my classification of political regimes. I follow Gates et al. (2006), who utilize a modified version of Gurr (1974), and differentiate a regime along three institutional dimensions: (1) the *recruitment* of the executive; (2) *constraints* on the power of the executive; and (3) opportunities for political *participation* in regulated and competitive elections. The degree of democracy increases as the executive is recruited through competitive elections, as the level of participation increases, and as the power of the executive is restricted.[4]

Many prominent researchers adhere to a dichotomous classification, defining regimes as either democratic or autocratic (see e.g. Przeworski et al. 2000). I follow Gates et al. (2006) and choose a tripartition, where a regime is classified as either a democracy, an autocracy, or an inconsistent regime, based on the three institutional dimensions defined above. As the inconsistent regimes plays an important part in this thesis, a brief presentation of this third regime category is in place.

[2]Face validity is used in various ways by different authors, but what Knutsen (2011) refers to is whether the theoretically defined concept adequately captures the 'background concept', i.e. "the broader constellation of meanings and understandings associated with a given concept" (Adcock and Collier 2001, 531).

[3]Content validity assesses the degree to which an empirical indicator represents the the entire content entailed in the theoretically defined concept. For a discussion of measurement validity, see Adcock and Collier (2001).

[4]It is important to note that restrictions on the executive only increases democracy up to some point. For example, too many constraints on the executive is detrimental for the functioning of democracy, as it provides other institutions, such as the parliament or the judiciary, with too much power.

Institutional inconsistency

In recent years inconsistent regimes have received increasing attention in the literature. Historically they are far from a new phenomenon, and can be located in both Europe and Latin America as far back as the 19th century (Diamond 2002). Diamond (2002, 24) claims that "virtually all hybrid regimes in the world today are quite deliberately 'pseudodemocratic". What he means is a regime that has some formal democratic institutions, like multiparty electoral competition, but where open, free and fair electoral competition is inhibited. There is also a relatively new literature on authoritarian systems that introduce political institutions which were previously only present in democratic regimes (see e.g. Gandhi and Przeworski 2007).

There are several reasons why I consider a tripartition to be preferable to a dichotomous regime classification scheme. Firstly, there are few *ideal* democracies and autocracies in the world today, but rather a wide range of institutional constellations mixing both democratic and autocratic features. The number of hybrid regimes has increased dramatically after 1990, with the fall of the Soviet Union and the communist states of Eastern Europe and Central Asia (Diamond 2002). Secondly, in many instances the democratization process is incomplete (Strand 2007, 216). As noted by Epstein et al. (2006, 567), "[l]eaving autocracy is not the same as entering democracy." Hence, dichotomizing democracy in a study of democratization may potentially produce erroneous, or misleading, results.

Several terms are used to denote the inconsistent regimes. Some refer to them as 'anocracies' (Vreeland 2008), others as 'semidemocracies' (Hegre et al. 2001; Fearon and Laitin 2003), 'hybrid regimes' (Diamond 2002), or 'partial democracies' (Epstein et al. 2006). I prefer to call them *inconsistent*, in line with Gates et al. (2006).

2.1.3 Regime change

A regime change is a substantial, often abrupt, change in a country's institutions (Gurr 1974, 1483).[5] I define a *regime change* as a transition between two of the three regime categories defined above; autocracy, inconsistency and democracy.[6] Hence, a 'substan-

[5]I sometimes refer to a regime change as a political transition, and I also use institutional change and regime change interchangeably, although formally there is a clear distinction between the two (see below).

[6]It is important to stress that whenever I refer to a 'regime' or regime change, it refers back to the three regime categories defined in section 2.1.2, and the definition in this section. It does not imply a change of government or president within one of the categories. If for example the Syrian President

tial' change is a change that is large enough to induce a movement between two regime categories. This thesis is preoccupied with transitions in direction of democracy, which is what I refer to as democratization. *Democratization* is defined as a regime change away from autocracy, into either the inconsistent or democratic regime category.[7]

Regime change is closely related to *political instability*, which I define as a situation where the institutional composition is short-lived, and/or where there is incidence of civil conflict (Hurwitz 1973).[8] As this concept is beset with unclarity about its precise meaning and demarcation, I try to avoid using it and rather refer to the two events, regime change and civil conflict, separately.

2.2 Civil conflict

A single definition of civil conflict does not exist today, and several research projects and scholars have presented their own definition of the concept (see e.g. Fearon and Laitin 2003; Sambanis 2004*b*; Small and Singer 1992; Gleditsch et al. 2002).[9]

I follow UCDP/PRIO (2010, 3) who define an *armed conflict* as "a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths." An *internal armed conflict*, i.e. a civil conflict, "occurs between the government of a state and one or more internal opposition group(s) without intervention from other states."

There are three issues I want to highlight in the following. The first relates to whether one should differentiate between different types of violence and if civil conflict is a distinct phenomenon. I consider civil conflict to be a clearly defined and demarcated event, in line with most researchers within conflict studies (Sambanis 2004*b*). However, this does not necessarily imply that it should be studied in isolation from other types of violence. It is for example not uncommon that groups shift between various forms of violence, like

Bashar al-Assad is replaced by another authoritarian leader this does not count as a regime change.

[7]An institutional change and a regime change are considered to be two distinct events. An institutional change is a necessary, but not a sufficient, condition for a regime change. The empirical analysis deals exclusively with regime change.

[8](Hurwitz 1973) points to five common ways of defining political stability: as (1) the absence of violence; (2) governmental duration; (3) the existence of a legitimate constitutional regime; (4) the absence of structural change; and (5) a multifaceted societal attribute. I focus on the first two.

[9]In the following I use the terms *civil conflict*, *conflict*, and *civil war* interchangeably. A more nuanced usage of the terms is perhaps preferred by some readers, but when otherwise not specified, they all refer back to the definition of given in this section.

terrorism or organized crime, during the course of a conflict (Ibid.). I follow Sambanis (2004*a*, 259), who contends that civil conflict should be viewed as "one phase in a cycle of violence". Hence, in my theoretical argument I also place emphasis on the process leading up to the event of war.

Second, the literature differentiates between onset, incidence, duration, and severity or armed conflict. *Onset* is when a conflict breaks out, whereas *duration* is simply how long the conflict lasts. Conflict *incidence* is the proportion of country-years that has an ongoing armed conflict. Incidence is a function of the risk of onset and duration (Gleditsch, Hegre and Strand 2009, 159). Incidence is high if the risk of onset is high, or if the conflict is likely to last for a long time. My focus is on conflict incidence. In order to evaluate the overall effects of democratization on civil conflict the incidence of conflict is, in my opinion, the most appropriate metric.

Last, the thesis emphasizes civil conflict as a *political tool* (see Strand 2007, 13). The reason why the instrumental aspect of violence is underlined relates to the link between conflict and regime change. When organized groups use violence to overthrow the government, or more drastically, to change the entire political system, violence is used as a means to achieve some defined political goal.

2.3 Regime change and civil conflict

When violence is used as a political tool, regime changes and civil conflict become conceptually difficult to separate. Recent events in North Africa and the Middle East, during the Arab revolutions corroborate this. Both in Tunisia, Egypt, Libya and Syria, what started out as peaceful protests against authoritarian regimes intensified, and led to minor or major episodes of violence. These cases are only a few examples indicating that regime changes and violence tend to occur in conjunction. Moreover, they highlight the relevance of looking upon civil conflict as a constituent part in a cycle of violence, where the outbreak of war is the final stage.

2.4 Inequality

The notion of 'inequality' is fraught with conceptual vagueness. In order for it to be meaningful it has to be ascribed concrete characteristics, and discussed in relation to its many manifestations. The most important distinction is between vertical and horizontal inequality: *Vertical inequality* (VI) measures inequality at the individual (or household)

level, and concerns the ranking of individuals at the regional-, country- or global-level. This is also referred to as inter-individual inequality (Stewart 2002, 3). *Horizontal inequality* (HI), on the other hand, concerns inequalities between culturally defined groups, and relate to both economic, social or political dimensions, or cultural status (Stewart, Brown and Mancini 2005, 3).

Inequality has many dimensions, both political, social, cultural and economic (Stewart, Brown and Mancini 2005, 7). I focus on the latter, and argue that economic disparities are amongst the most notable and severe expressions of inequality, with a clear link to both regime changes and violence (see e.g. Gurr 1970; Boix 2003; Acemoglu and Robinson 2006). In the following, I define economic inequality as the unequal distribution of wealth between groups in society. More specifically, I look at economic inequality between the elites and the citizens.

Some may argue that given the increased attention that has been directed towards horizontal inequalities (e.g. Østby 2008; Cederman, Weidmann and Gleditsch 2011; Gurr and Moore 1997), my attention should be focused here. I have chosen differently because I, firstly, want to be in compliance with the understanding of inequality in the main theoretical works applied in this study (see Acemoglu and Robinson 2006; Boix 2003). Secondly, I consider economic inequality to be an appropriate point of departure for this thesis' inquiry. Unequal economic opportunities and persistent social cleavages was partly what motivated the young revolutionaries during the Arab uprisings (Filiu 2011, 35). However, my approach towards inequality does, to a certain degree, capture HI, as income inequality often coincide with social, political or cultural cleavages (Lichbach 1989, 432).[10]

2.5 The short run versus the long run

The second research question of this thesis concerns the measurement of democratization and conflict incidence in the long run. A specification of the 'long run', as distinct from the 'short run', is therefore required. The distinction is best understood in relation to democratization. A democratic transition is an event that takes place at a given moment in time, and it has an instantaneous effect, namely the formation of a new regime. This event is a *short run* effect. However, democratization is often not a linear and swift process. As noted by Strand (2007, 216), "[d]emocratization in particular tends to be

[10]In the following, I use the terms economic inequality and inequality interchangeably, but inequality always refers back to economic inequality.

partially successful, often resulting in a semi-democratic regime." Thus, an expectation is that some countries experience oscillation between regime types, or periods of violence, before they manage to consolidate democracy. All of these subsequent changes are part of long-run effects of democratization. I therefore consider it beneficial to differentiate between the effects of changes to a country's political institutions and its conflict level in the short run, and the effects in the *long run*, i.e. those that endure.

Chapter 3

Literature review

The following review of literature provides an overview of the most relevant contributions within the large and varied research on democratization and civil conflict. What is notable while reading these two literatures is that, although they deal with some of the same thematic they, to a little extent, communicate directly to each other. This acknowledgement is the point of departure for the subsequent analytical framework, in chapter 4.

The chapter is divided as follows. I first present the conditions that support democratization and regime stability. Focus is directed towards economic determinants, as it is most relevant for this thesis. Next, I discuss the relationship between economic inequality and democratization. The 'opportunity and motivation' framework is used as a starting point for a presentation of the literature on civil conflict. The final sections deal with the relationship between political regimes, institutional changes and civil conflict.

3.1 The conditions for democratization and democratic stability

There is a large theoretical literature that seeks to explain the rise and fall of democracies. In recent decades it has been supplemented with a vast array of empirical research, both qualitative and statistical. According to Boix (2003, 4), at least three main strands of research that can be identified. The first one is the modernization school; the second is situated within the sociology literature and focus on regime formation; the third, and most recent, present democracy as an institutional equilibrium. The first and third are particularly relevant for this thesis, and I highlight them in the following.

The first perspective gives precedence to economic conditions when explaining democratization and regime (in)stability. It has its roots in the modernization perspective dating

back to the 1950s, most notably with the work of Lipset (1959). Lipset discovers a positive correlation between economic development and democracy, and argues that economic development, measured as increases in GDP per capita, triggers democratic transition. The relationship between economic development and democracy has been at the core of the democratization debate ever since (see e.g. Bollen 1979; Burkhart and Lewis-Beck 1994; Diamond 1992; Przeworski and Limongi 1997; Przeworski et al. 2000; Boix and Stokes 2003; Hadenius and Teorell 2005*b*; Inglehart and Welzel 2006). In later years, many have questioned the link between economic development and democratization (see e.g. Przeworski et al. 2000; Przeworski and Limongi 1997). Przeworski and Limongi (1997) find evidence that the relationship between income and democracy is not a result of richer countries being more likely to democratize, but rather that democracies become more stable as they reach higher levels of economic development. The result is contested, and other studies have rejected the claims made by Przeworski and his colleagues (Hadenius and Teorell 2005*a*; Boix and Stokes 2003).

The most recent strand of research emphasizes democracy as an equilibrium condition. That democracy is an institutional equilibrium implies that it is a stable outcome, which can be reached through the strategic choices taken by the actors to maximize their welfare (Boix 2003, 8). Dahl (1971) formulates an early version of the democratic equilibrium argument. He claims that the choice of political regime can be evaluated through a cost-benefit analysis: The political actors in position consider the chance of losing elections and the costs of possible policies enforced on them by the other parties, against the costs of excluding the opposition permanently through authoritarianism. If the price of repressing the opposition increases, democracy becomes more viable, because all actors gain from it. Both economic, cultural and historical factors may induce democratization by changing the cost-benefit ratio. Most of the literature that defines democracy as an institutional equilibrium employs game-theoretical tools, and develop formal models of regime change (see e.g. Weingast 1997; Przeworski 1991; Boix 2003; Acemoglu and Robinson 2001, 2006). Contributions from this strand of research are the point of departure for the analytical framework I develop in the next chapter.

3.1.1 Economic inequality and democratization

Alongside the debate over economic development's relationship to democracy others have tried to reframe the argument away from the level of wealth, towards the *distribution* of wealth. The main argument has been that inequality has detrimental effects on the prospects of democratization (Boix 2003; Acemoglu and Robinson 2006, see e.g.). The

idea is far from new. Earlier scholars, like Lipset (1959) and Dahl (1971), maintain that an unequal distribution of wealth within autocratic regimes reduces the likelihood of a democratic transition. Lipset (1959) argues that economic inequality increases social tension and civil unrest, and inn addition inhibits widespread education. As education promotes democracy, and civil unrest is detrimental to regime stability, inequality indirectly impedes democratization.

From empirical studies the evidence of inequality's effect on democratization and regime stability is mixed. Muller (1988, 1995) finds that income inequality is associated with less democratization, whereas Bollen and Jackman (1995) do not find any clear effect of income inequality in democratization. Barro (1999) uses the Gini coefficient of income inequality from Deininger and Squire (1996), and finds a negative, albeit weak, statistical relationship between inequality and democracy. Of more recent studies, Houle (2009) does not find any link between inequality and democratization. However, he does find a positive, statistical relationship between low economic inequality and democratic consolidation.

3.2 The causes of civil conflict: Opportunity and motivation

The theoretical point of departure for much research on civil conflict is the opportunity and motivation for rebellion (see Collier and Hoeffler 2004). Collier and Hoeffler (2004, 564) note that the political science literature has developed theoretical accounts that focus on the motive behind rebellion, and the forces that unite people in rebel groups. A small economic conflict literature has, on the other hand, developed formal models of rebellion where economic opportunities are accentuated (see e.g. Grossman 1991).

The *Motive* behind rebellion can be understood in a negative sense, as grievance against the government concerning social conditions, the way in which state business is organized etc. However, it can also understood in a positive sense, as a desire to achieve wealth and acquire resources, often referred to as the motive of greed (Gleditsch, Hegre and Strand 2009).

The *Opportunity* to rebel includes all factors that enables a rebel group to join forces against the government. It relates to the rebel organizations' own resources and capacity, the repressive power of the government and the military forces they possess, as well as geography (Collier and Hoeffler 2004, 588). Opportunities may also arise when costs of rebellion decline, and/or when the opportunity-cost of rebellion is low. This is typically

the case in low-income countries where there are few jobs or educational opportunities, and where the prospects for future improvements in welfare are poor. The income lost by joining rebellious activities are in these cases perceived to be lower than the net present value of the rebellion.

3.2.1 Greed and grievance explanations of civil conflict

The 'greed and grievance debate' originates within the opportunity and motivation framework, and has become a well-known contribution in contemporary research on armed conflict. *Greed* relates to the economic motive for insurrections and the profit-seeking rebel. Civil conflicts are hypothesized to occur when there are atypical profitable opportunities for looting and other types of profit-generating activity (Collier and Hoeffler 2004; Collier 2000). Grievance-based rebellion stands in contrast to greed-based rebellion. *Grievances* are often associated with the term 'relative deprivation', which refers to expectations of material well-being failing to be realized (Gurr 1970). The logic behind a grievance-based argument is that broader social movements channel their deprivation through the use of violence. Objective grievance measures, such as political repression, ethnic or religious hatred, political and/or social exclusion, and economic inequality, are used as proxies within the quantitative literature.

Collier and Hoeffler (2004, 564) strongly argue against grievance-based explanations and claim that "misconceptions of grievances may be very common: all societies may have groups with exaggerated grievances. In this case, as with greed-rebellion, motive would not explain the incidence of rebellion".

One reason why grievance-based accounts have been rejected relates to the collective action problem. The collective action problem is relevant for all types of collective mobilization, but it is argued to be even harder to solve when motivation is based on grievance rather than greed. As the costs of securing a 'public good', for example democracy, only accrues to the rebels, but all members of society reap the benefits, recruitment is more difficult than in situations where private gains from for example looting are attainable (Collier 2000).

Another reason why grievances have been dismissed relates to the lack of statistical evidence. Previous statistical studies did for the most part focus on income inequality (e.g. Collier and Hoeffler 2004; Hegre, Gissinger and Gleditsch 2003; Fearon and Laitin 2003), but found scarce support of the hypothesis that income inequality causes more violent conflict.[1]

[1]Because previous research has found little evidence of a link between income inequality and civil

Although grievance-based explanations of civil conflict have received a lot of opposition there is a close relation between greed and grievance, as this passage from Collier (2000, 852) shows: "[...] greed may need to incite grievance. Thus, grievance and greed may be necessary for sustained rebellion: grievance may enable a rebel organization to grow to the point at which it is viable as a predator; greed may sustain the organization once it has reached this point."

I end the preliminary discussion of inequality's effect on violence on this note. In the final section of this chapter the center of attention is the relationship between regime changes and civil conflict.

3.2.2 Political regimes and civil conflict

The relationship between political regimes and civil conflict has received considerable attention within conflict research over the years (see e.g. Muller and Weede 1990, 1998; Hegre et al. 2001; Gleditsch, Hegre and Strand 2009; Strand 2007). It is often differentiated between whether it is the *process of change*, or particular *types of political institutions*, that heightens the risk of conflict. Different arguments have been put forward, and in the following I present some of them. As I turn to political regimes, the inconsistent regimes are in focus, as they play a prominent part in this thesis.

Institutional change and civil conflict

Democratization involves a process of mass mobilization. Failing to accommodate the high levels of participation that tend to follow in the wake of a democratization is thought to increase the risk civil unrest and violence (Huntington 1968, 83). Moreover, when old institutions are replaced with new ones, institutional weakness may threaten the ability of the government to utilize its monopoly of violence. This 'institutional deficit' can, amongst other things, lead to hasty and hostile political strategies, like playing various factions against each other (Mansfield and Snyder 2002, 302). Institutional changes also create a 'window of opportunity' for leaders to mobilize popular support for collective action, often based on ideological and/or nationalistic sentiments, so-called 'belligerent nationalism' (Ibid.).[2] Moreover, the period after a regime change is often marked by

conflict, more recent studies have direct attention away from vertical differences in the distribution of wealth, towards horizontal, or group-level, inequalities (see e.g. Østby 2008; Cederman, Weidmann and Gleditsch 2011; Gurr 2000).

[2]Mansfield and Snyder (2002) focus on the effects of democratization on institutional strength and the prospects for interstate war. However, their arguments have been applied by researchers who focus

uncertainty about the potential of further institutional changes, and as to how the new institutions will influence the balance of power. It is not unlikely that this situation triggers a tug-of-war between different actors over the future distribution of power and resources (Ibid.).

Elections, which are often held in close proximity to a regime change, are also potential triggers of violence. Elections can provide motivation for rebellion *before* an election, if opposition groups feel that they have no real chance of winning, or *after* an election, if the opposition consider the incumbent to have won unlawfully (Gleditsch, Hegre and Strand 2009, 165).

There has been conducted a vast array of quantitative studies on the relationship between regime changes and civil conflict. One notable finding is that democratization is not only a process of institutional change, it is also a manifestation of general regime instability. Both of these factors are found to increase the likelihood of civil conflict (Hegre et al. 2001). Cederman, Hug and Krebs (2010) examine the relationship between democratization and civil war, and assert that the dynamic effects of democratization and autocratization differ. Whereas democratization increases the probability of civil war onset over several years, autocratization increases the probability of civil war instantly. They find evidence of both effects, although to a lesser extent for autocratization. Fearon and Laitin (2003) have been occupied with the way in which institutional changes challenge the state's capacity, and the state's opportunity to repress rebellious activity. They find evidence that a weakened state capacity increases the risk of civil conflict. The final stage of an autocratic regime is also found to increase violence, as the repressiveness of the regime often increases when it acknowledges its demise (Zanger 2000).

Instability and violence in inconsistent regimes

Although the process of democratization can lead to violence, it is thought that when democratic institutions are in place, the likelihood of rebellion is reduced. Gleditsch, Hegre and Strand (160 2009) state that "[d]emocratic governance in itself can be seen as a conflict management system where different interests meet and are resolved peacefully." The motivation for rebellion is reduced because first, grievances are channeled through legitimate institutions, and second, the oppressiveness of previous autocratic institutions disappear (Muller and Weede 1990). Although democracies offer new possibilities for organization, and hence, increase the opportunities of rebellion, it is accentuated that the freedom to organize is most often used peacefully (Gleditsch, Hegre and Strand 2009,

on intrastate, i.e. civil war.

161) Moreover, the opportunities for political leaders to repress and coerce are heavily constrained in a democratic system of governance (Ibid.). However, it may not always be the case that a country transitions directly from autocracy to democracy (Hegre et al. 2001).

Strand (2007, 216) states that "democratization in particular tends to be partially successful, often resulting in semi-democratic regime." The fact that democratization is not always swift and linear has consequences for the opportunities and motivation to rebel. Hegre et al. (2001) find that semidemocracies are more prone to civil war, and argue that "semidemocracies are partly open yet somewhat repressive, a combination that invites protest, rebellion, and other forms of violence" (Hegre et al. 2001, 33). This is often referred to as the 'inverted U-shaped' relationship between democracy and conflict, and it has been confirmed by other studies (e.g. Reynal-Querol 2002; Fearon and Laitin 2003; Hegre and Sambanis 2006). Moreover, Epstein et al. (2006) highlight the importance of partial democracies for understanding democratic transitions. They claim that partial democracies are more volatile and less well understood, and prevent a consistent understanding of democratic transitions. Further, Gates et al. (2006) find that inconsistent regimes are less stable than both democracies and autocracies.

Despite considerable evidence, the link between inconsistent regimes and the risk of conflict has been scrutinized by other scholars: Testing the findings of Hegre et al. (2001) with an improved method, and a new dataset, Strand (2007) finds that it is a particular type of semidemocracies, which he refers to as *illiberal* democracies, that have the highest risk of conflict onset. Illiberal democracies have a high level of participation, through competitive elections, combined with limited or no constraints on the power of the executive (or ineffective constraints). These institutional features are thought to create an 'explosive' mix (Strand 2007, 310). Vreeland (2008) claims that the posited inverted U-shaped relationship relates to measurement errors, rather than a factual empirical connection. He attributes the error to the use of the Polity IV Participation Index for construction of the democracy measure.[3] Gleditsch, Hegre and Strand (2009) exclude the troublesome components from the Polity Index, and is still able to find the contended U-shape.

Evidently, the debate over the instability and conflict-proneness of inconsistent regime is not concluded. Despite the level of attention the causal connection between inconsistent regimes, violence, and regime stability has received, there tends to be less focus on why inconsistent regimes come into being and what maintains them. This is one of the issues

[3]The Polity participation indicator includes a category denoted as 'factionalism'. Factionalism refers to a situation where political competition is intense, hostile and frequently violent.

I address in the next chapter.

3.3 Moving further

A realization provided by this review of literature is that there is potential for more interdisciplinary communication and research between scholars of democratization and conflict studies, for several reasons. First, many contributions within the literature on civil conflict focus on the link between violence and institutional change. However, the reciprocal causality between the phenomena is seldom highlighted. Second, both literatures have had continuous debate over the importance of inequality and greivances, as a trigger of violence, or as a break on democratization. The empirical evidence so far has not been convincing. Researchers within both camps disagree as to whether this is due to an in fact non-causal link, or whether it is an artifact of erroneous, or inadequate, measurement, or focus on the wrong *types* of inequality. In the following, I argue that economic inequality can prove relevant for the link between institutional changes and violence. Third, some contributions within conflict research highlight that democratization is a process that increases the risk of conflict over several years (see e.g. Cederman, Hug and Krebs 2010). No one has, to my knowledge, focused first-and-foremost on the long-run effects of democratization, and on capturing them empirically.

In the next chapters democratization and civil conflict are treated as two closely interrelated phenomena. Further, focus is directed towards the process of democratization, and economic inequality is argued to shape its course and outcome. I start off with the theoretical foundation and try to develop an analytical framework, which is used as a point of departure in the subsequent empirical analysis.

Chapter 4

Theory

In the literature review I tried to highlight that although the theoretical perspectives and empirical evidence on the causes of regime change and of civil conflict are diverse, the two literatures share some common features. It is those I seek to unite. Inspired by the work of Boix (2003) and Acemoglu and Robinson (2006), this chapter places the dynamic relationship between democratization and civil conflict at the forefront.

In the following, I develop an analytical framework that explains the logic behind democratic regime changes *and* its relationship to civil conflict. In contradiction to previous studies, I emphasize the long-run perspective: The main argument is that differences in the level of economic inequality partly explain why democratization processes unfold so differently in different countries. In line with Boix (2003), Acemoglu and Robinson (2006), and Gurr (1970), I argue that by increasing the demand for democratization, and the willingness to use violence, inequality can both induce regime changes and provoke civil conflict. The outcome of the discussion is the following propositions:

1) Democratization is relatively swift and peaceful in equal societies: The elites have little to loose by democratizing, and neither the citizens nor the elites have incentives to fight.

2) Democratization is more long-drawn, and conflict prone, in highly unequal, than equal, societies: The citizens' demand for democracy increases, but the elites strongly reject its introduction. Both sides thus have an incentive to fight.

The propositions are explicitly or implicitly suggested within the literature, although they have not, to my knowledge, been systematically tested.

The argument of this chapter is built in several stages: First, I present the basic assumptions of the analytical framework. The three subsequent sections serve as a basis for the final argument. I discuss the relationship between inequality, group organization and

rebellion. Next, I explore the relationship between economic inequality and democratization, inspired by the works of Boix (2003) and Acemoglu and Robinson (2006). Last, I account for the institutional equilibrium model. In the final part of the chapter, from section 4.5 onwards, I draw the acquired knowledge together. I formulate three 'ideal types' of democratization, which I use as a point of departure for a discussion on the relationship between democratization and civil conflict, in the short and long run. At the end of the chapter the main points of the analytical framework are summarized.

4.1 The basic assumptions

As a simplification, I consider the world to consist of two groups, or types of actors; the elites and the citizens, where the citizens are more numerous than the elites. Although each individual belongs to one of these groups, they act to maximize their individual welfare. The most important determinant for individual behavior is considered to be economic welfare, and individuals seek to preserve, or introduce, the type of political institutions that enhances their economic welfare.[1] When assessing the choice between democracy and autocracy, the actors consider the economic and social consequences of these political institutions. The approach taken above is often described as 'economic', or rational choice, and it is in line with previous work on democratization (see e.g. Boix 2003; Acemoglu and Robinson 2006).

Second, I assume that the citizens always prefer democratic institutions. The citizens expect democracy to enhance their economic welfare through the redistribution of wealth. Further, it provides them with the power to influence decision-making, both now and in the future. As a consequence there is an underlying pressure for democratization in all non-democracies.[2]

Third, violence is in this framework considered to be a political tool, as described in section 2.2. But, under which conditions are violent means applied, and hence, civil conflict a possible outcome? I follow Boix (2008, 398), who states that "the excluded majority may resort to violence whenever the expected gain of revolting is larger than the value of accepting an authoritarian regime." The costs and benefits of using violence

[1] Welfare is understood in a broad sense to include power, income, peace, security, and other 'goods' that a person desire or need, in order to create the best possible living environment. I acknowledge these 'goods', but assume that economic welfare is the most important, as it often determines a person's ability to require food, shelter and other basic necessities.

[2] The notion of an underlying pressure for democracy is concurrent with historical evidence, which shows that the world has become increasingly more democratic over the last century.

depend on a number of factors, for example the perceived probability of victory, the resources and technology available, the gains from victory etc.

The last assumption concerns the path towards democracy. Karl (1990, 8–9) points at four different modes of democratic transition: revolution, reform, imposition and pact. Reform and revolution are most often driven from below, by mass actors, and entail the control and subversion of traditional ruling elites. Elite pacts are 'transitions from above' and involve continued elite control of power, but with gradual, often limited, democratic reforms. The last mode is imposition, which often imply that external actors' use coercion in order to impose democratic institutions. It also happens that parts of the elite, often the military, decide to force through a democratic transitions. I am chiefly interested in the first two modes of transition, where the citizens serve as the main driver of democratization, and the focus in the following is on transitions from below.

With the basic assumptions in place, I now examine how economic inequality can trigger 'men to rebel'. A main point is that the grievance-based accounts, often discredited in studies of civil conflict, may be reconcilable with a rational actor perspective.

4.2 Economic inequality, mobilization, and rebellion

> *"The institutions, persons, and policies of rulers have inspired the violent wrath of their nominal subjects throughout the history of organized political life"* (Gurr 1970, 3)

In his seminal work *Why Men Rebel*, T.R Gurr provides a thorough investigation into the causes of political violence. Gurr (1970, 3–4) defines political violence as "all collective attacks within a political community against the political regime, its actors [...] or its policies." Political violence is thought to vary in its magnitude and form. Magnitude concerns how many who participate in the activities, how intense or destructive the violence is, and its duration. The forms of political violence relate to its different manifestations; e.g. unorganized violence (turmoil), coups d'etat, or civil war. These are not uniform variables, but should rather be understood in terms of *degree* or *quantity* of violence (Gurr 1970, 9–10). Thus, the occurrence of civil conflict is only a manifestation of an evolving process of violence.

Gurr's theory explains how social and economic inequalities increase the risk of violence, through frustrated expectations. The theory's main building block is the psychological phenomenon *relative deprivation* (henceforth, RD), which is defined as "[...]

a perceived discrepancy between men's value expectations and their value capabilities. Value expectations are the goods and conditions of life to which people believe they are rightfully entitled. Value capabilities are the goods and conditions they think they are capable of attaining or maintaining, given the social means available to them" (Gurr 1970, 13).

The causal link between RD and political violence works through a 'frustration-aggression mechanism'. It contends that deep frustration is a precondition for aggressive behavior. When the discrepancy between what people feel they are entitled to, and what they actually are able to attain, persists over time, or become intensified, it triggers aggression (Gurr 1970, 9). In the causal chain that links RD to violence, both societal and psychological variables are relevant: First, the development of discontent takes place. Second, the individual actors come together and the discontent is politicized. Over time the group may decide to use of violence to achieve their political objectives (Gurr 1970, 12–13). It is in the first phase, the development of discontent, that economic inequality comes into play. The citizens form expectations about the economic welfare they *potentially* can achieve, given the resources available in society and the welfare of 'better off' elite groups. Thus, increasing inequality is hypothesized to cause RD, as it increases the gap between value expectations and value capabilities.

Gurr's theory is often cited in relation to the 'greed versus grievance' debate within conflict research. The discord over the significance of grievances in general, and economic inequality in particular, can be related to two distinct approaches within the political science literature; the 'Deprived Actor' (DA) scientific research program and the 'Rational Actor' scientific research program (RA) (Lichbach 1989). Gurr's theory, evidently, belongs under the heading of the Deprived Actor. DA theories emphasize demand-pull factors, such as preferences and attitudes, and how psychological processes activate grievances and convert them into rebellious action (Lichbach 1989, 456).

Collier and Hoeffler (2004) and their colleagues, whose research was presented in section 3.2, are more in accordance with the Rational Actor approach. RA theories are occupied with cost-push factors, such as opportunities, costs and benefits. As people make rational choices, grievances cannot be turned into violent rebellion, because the mobilization process anticipated by DA theories fail to materialize: The preferences of a rational actor are always self-regarding, and the main goal is to maximize ones own income, irrespective of what others are able to receive (Lichbach 1989, 460). Hence, the 'relative' aspect of inequality becomes irrelevant. Moreover, an individual is only going to partake in rebellion if the private benefit is greater than the benefit from other economic

activities. Increasing inequality is only going to increase the risk of rebellion if absolute poverty also increases. Only then will the gains from other economic activity decrease, and the opportunity cost of rebellion decline (Lichbach 1989, 461).

One particularly contended issue for the RA and DA approaches is how, and if ever, people are able solve the collective action problem (see e.g. Olson 1965). The collective action problem states that rational actors will never mobilize and rebel, because the gains from rebellion are shared by all members of society, it is a so-called 'collective good', whereas the costs are private and accrues to each individual member of the rebel organization. The collective action problem is expected to be particularly hard to overcome when the government has a very strong repression apparatus or the rebels are highly unorganized. DA theories bypass the collective action problem by assuming that people act irrationally, or that they receive some psychological 'benefit' from rebellion (Lichbach 1989, 459).

In the following, I view rebellion as serving two distinct purposes. First, it is a rational act. The goal is to increase one's economic welfare, and the strategy is to influence the elites and change the political institutions. Second, rebellion satisfies the need to agitate when one's living conditions are seriously deprived. The recent events during the Arab revolutions supports this second notion (see Filiu 2011). Thus, in line with Collier (2000, 852), grievances can serve as an impetus for mobilization, but economic rationale is considered necessary to carry through an organized rebellion.

Having laid down the premise for the link between grievances and violence, I now examine two theories of democratization.

4.3 Economic inequality and democratization

Two prominent contributions to the literature on economic inequality and democratization are Boix (2003) and Acemoglu and Robinson (2006) (from now on referred to as A&R). Their work represent two schools of thought that provide quite different predictions about the effect of inequality on democratization, and democratic consolidation (Houle 2009, 589). In the following I first present the main features common to both theories, before I briefly review each of them, highlight their principal differences, and point at some of their weaknesses.[3]

Both Boix and A&R provide formal game-theoretical models, and 'prove' how different

[3]It is important to stress that I do not strictly follow any one of these theories, but rather use their main arguements as a point of departure for the analytical framework.

levels of inequality impact on the possibility of reaching the democratic equilibrium. Political regimes are ultimately seen as the product of the nature of economic assets, where the most important assets are land and capital, and the distribution of assets between the elites and the citizens.[4] Further, they operate with two representative actors; the citizens (the poor) and the elites (the rich). The citizens make up the majority of the population and they have, by definition, less endowments than the elites.[5] The share of resources that accrues to the elites and the citizens varies between societies; some societies are fairly equal, while others are highly unequal.

That "social choices are inherently conflictual" (Acemoglu and Robinson 2006, 15) is explicitly or implicitly assumed in both theories. Policies favorable to one group is not necessarily, and most often not, favorable to another. As different political regimes are better at promoting the interests of particular groups, conflict over the political institutions are almost inevitable. It is assumed that the preference for distributional outcome is decisive for the actors' choice of political regime. The two choices that face the rich and the poor are either a democratic- or an autocratic regime. The elites obtain a higher income in an autocracy and the citizens obtain a higher income in a democracy (Acemoglu and Robinson 2006, 176). In an autocracy the elites control most of the resources, whereas in a democracy the elites are forced to redistribute some of their wealth to the poor. This is in line with Meltzer and Richard (1981). Their model states that in an economy where income is distributed unevenly, the median voter can dictate the tax rate, and redistribute in a way that maximizes his own income.

Violence is a constituent part of both Boix' and A&R's theories. Violence can be applied by both parties, and is inherently political. It can be used by the elites to repress, and by the citizens in form of a violent revolution. When inequality is high concessions can be considered non-feasible to either of the parties, and both sides may then apply violent means (Boix 2003; Acemoglu and Robinson 2006, 26–27).

The first school of thought is represented by Boix (2003) and his book *Democracy and Redistribution*. Boix investigates the effects of differences in the distribution of land and capital, and his main claim is that unequal societies are less likely to experience a democratic transition than are equal societies.[6] The starting point for Boix is the

[4]This is what Boix (2003, 29) refers to as an endogenous theory of democratization; it explains the introduction of democracy and democratic consolidation with variables that determine the long run stability of a regime.

[5]The middle class is introduced at a later stage, but does not play a dominant part in the theories. The middle class have more endowments than the poor, but less than the rich elites.

[6]Boix (2003, 44–45) also emphasizes the mobility of assets. Capital is considered to be very mobile, whereas land is highly immobile. When capital mobility rises, the likelihood of democratization increases

autocratic regime. Under authoritarian institutions the rich hold decision-making power, and repress the poor. The rich prefer zero redistribution, as it reduces their share of total income. In autocracies, there is therefore little or no taxation, and the elites incur small costs, ceteris paribus. The poor prefer extensive redistribution, because it provides them with a higher level of welfare. The poor therefore want to introduce democratic institutions where they can impose heavy taxation on the assets of the rich, by exploiting the fact that they control the median voter. The implication of this 'clash of interests' is that democratization is impossible when the level of inequality is high. In equal societies, on the other hand, democracy can be introduced peacefully, because the rich do not have to fear extensive redistribution.[7]

The other school of thought, represented by A&R, proclaims that the effect of inequality on democratization works through an inverted U-shaped curve. They argue that it is rather at *moderate* levels of economic inequality that a transition to democracy is most likely, because the elites and the citizens are able to compromise about the institutional constellation. The poor do not pose as great a threat to the welfare of the rich when inequality is moderate, because the demand for redistribution is modified. Nevertheless, the elites face some degree of pressure to democratize (redistribute), and prefer to establish democratic institutions instead of incurring the costs of repression. When inequality is low, on the other hand, autocratic leaders do not face any threat of revolution. The demand for redistribution is low, and so is the demand for democratic institutions. When economic inequality is high, A&R provide the same prediction as Boix.

Both theories are quite similar in terms of their initial 'world-view' and the factors they emphasize. Their logic is also fairly simple, which is perhaps what makes them so attractive. However, there are some elements of the theories that can be questioned. First, the conclusion about the effect of inequality seems to rely on a couple of initial assumptions. The assumptions concern which actor that is given an informational advantage, and which actor that is allowed to make the first 'move'.[8] Both models may therefore present reasonable predictions about the effect of economic inequality on democratization,

---

because the holders of capital can threaten to move their wealth abroad if it is taxed too heavily. Land, on the other hand, is immobile, and can therefore be taxed more easily. When a lot of wealth is held in land, Boix expect it to be hard for democracy to prevail.

[7]When inequality is moderate, Boix predicts either authoritarianism or democracy. What kind of political regime that prevails depends on the the repression costs of the rich and political resources available to the poor.

[8]In Boix' model (2003, 30–31) it is the rich who make the first move and that is given the informational advantage, whereas in Acemoglu and Robinson's (2006) model the citizens have knowledge about the repression costs of the rich.

given the underlying assumptions of the model. Consequently, they are not necessarily contradicting theories.

Second, the definition of political institutions is very narrow in both of these models. Boix (2003, 10), for example, defines a political regime as "a mechanism employed to aggregate individual preferences about the ideal distribution of assets among those individuals governed by this institutional mechanism". This represents a functionalistic view of politics that only places emphasis on one particular attribute of political institutions, namely redistribution through taxation. Eckstein (1973, 1144) is amongst the critiques of functionalist definitions of politics and claims that they "unduly restrict and and handicap the field".

Houle (2009) has also been critical of both Boix and A&R, and points at several weaknesses in their arguments. The most crucial is perhaps that economic inequality can pull in direction of both autocracy and democracy: Inequality increases the cost of democracy to the elite, by increasing taxation for redistributive purposes. Moreover, it increases the demand for democracy, by increasing the potential welfare gain for the poor and currently disenfranchised. These two effects work in opposite direction and, according to (Houle 2009, 593), "the net effect of inequality on democratization is ambiguous".

To sum up the acknowledgements made so far; to Boix, only low levels of inequality promotes democratization, whereas to A&R, both high and low levels of inequality is a hindrance for democratization. In the following, I argue that democratization can occur *both* at high and low levels of economic inequality. However, when inequality is high the demand for, as well as the resistance towards, democracy is high. This has implications for the way in which the democratization process develops, as well as its outcome. The intermediary regime category, the inconsistent regime, comes to play a decisive role. Boix and A&R operate with a dichotomous regime classification, and although they eventually introduce the possibility of a partial democracy, it does not play a dominant role in their theories. In contradiction, I argue that in order to better explain the relationship between inequality, violence and democratization, a third regime category is essential. I therefore now turn to the notion of the institutional equilibrium.

4.4 The institutional equilibrium

Formal models of democratization have often adhered to the notion of democracy as an institutional equilibrium (see e.g. Boix 2003; Acemoglu and Robinson 2006; Przeworski 1991; Weingast 1997). An institutional equilibrium is defined by Boix (2003, 8) as a

"stable outcome that results from the strategic choices that different individuals or parties in contention make to maximize their own welfare". It has also been described as a state in which the political institutions are 'consistent', i.e. as a situation where "a set of institutions are mutually reenforcing" (Gates et al. 2006, 894). Gates et al. (2006) operate with two stable equilibria; a consistent democracy and a consistent autocracy, and an unstable constellation, referred to as the inconsistent regime. The main idea is that consistent democracies and autocracies distribute authority in a way that makes them durable and stable, in line with Eckstein (1973) and Gurr (1974).

Consistent or 'ideal' democracies are self-enforcing because authority is fully dispersed through democratic institutions. As power is sufficiently dispersed, the actors gain more by preserving, rather than subverting the institutions. That democracy makes up a stable equilibrium is also supported by Przeworski (1991, 30–31), who claims that "democracy will evoke generalized compliance, it will be self-enforcing, when all the relevant political forces have some specific minimum probability of doing well under the particular system of institutions." Another explanation is articulated by Weingast (1997), who claims that when democratic institutions become an established part of civil society, the citizens develop routines for how to react against potential incumbents who try to subvert the existing institutions.

Consistent autocracies, on the other hand, are stable for an entirely different reason. In a consistent autocracy, the power–maximizing autocrat prevents competition between elites, by concentrating power in his own hands. By banning electoral participation, he keeps the costs of challenging him insurmountably high. The ascribed or designated executive has self–interest in deterring any potential opposition through coercion, repression and the like, because it serves to prolong his tenure. This, in turn, make the autocratic institutions durable (Gates et al. 2006, 894).

Inconsistent regimes combine both democratic and autocratic features, and as the name implies, are not considered to be self-enforcing. The main reason why, according to Gates et al. (2006, 895), is that authority is neither sufficiently diffuse, nor sufficiently concentrated. This, first of all, provides various actors with an incentive to subvert the institutions, but it also creates uncertainty about the balance of power.

The equilibrium framework of Gates et al. (2006) explain why some regime types are more stable than other. However, democratization involves the movement from one equilibrium to another. The framework cannot explain transitions away from a consistent regime, or when a political transitions is likely to result in an inconsistent regime. In order to say something about these issues focus needs to be directed towards the actors that

Table 4.1: Ideal types of democratization

Ideal type nr.	*Name*	*Description*
Ideal type 1	Swift and peaceful democratization	Democratic institutions are introduced in a relatively quick pace and without violent conflict.
Ideal type 2	Prolonged and conflict-prone democratization	Inconsistent institutions are enforced and/or reluctantly accepted. The process is often marked by minor or major episodes of violence.
Ideal type 3	Failed democratization	The authoritarian regime remains in power. There are shorter or longer periods of violent conflict.

intervene, and on which ground they form their preferences. Hence, the theories of Boix and A&R become relevant. In the following sections I draw together the acknowledgments from the above discussion, in an analytical framework that serves as a basis for the subsequent empirical analysis.

4.5 Ideal types of democratization

One way to think about the relationship between democratization and civil conflict is to formulate 'ideal type' processes of change. I consider three such ideal types, as shown in Table 4.1.[9]

The first ideal type is a swift and peaceful democratization, where a country moves more or less directly and peacefully, from an authoritarian regime to a democracy. An example of this first ideal type is Portugal. Portugal's authoritarian regime fell in 1976, and the country immediately established democratic institutions. It has ever since been a stable democracy and has not experienced civil conflict on its territory for the last 50 years. Argentina's swift and peaceful transition from an authoritarian regime to a democracy at the beginning of the 1980s is another.

The second ideal type is a violent, often prolonged, democratization process where an autocratic regime is replaced by inconsistent institutions. This ideal type can be exemplified by Thailand. Thailand's political history is marked by both democratic progress and regression to authoritarianism, with periods of violence along the way. Guatemala and Nicaragua from the 1960s and onwards are two other examples.

The third ideal type is a failed attempt at democratization. It involves a shorter, or longer, period of violence, where the authoritarian regime remains in power. One

[9]These ideal processes are stylized, and should not be interpreted too literally. For example, 'relatively quick' does not necessarily imply a complete democratic transition in one or two years. Thus, there is room for some interpretation. However, the 'ideal types' serve to illustrate some features I find important.

such persistent authoritarian system is Iran, which ever since the revolution in 1979 has experienced episodes of unrest and violent protest, but where strong authoritarian rulers who have managed to stagger the demands for democracy by repressive means.

The three ideal types are by no means exhaustive, but I them sufficient to illustrate the main points of my argument. In the following the ideal types are used as a point of departure in a discussion of the effects of inequality on democratization and civil conflict, in the short and long run. How I distinguished between the short and the long run was explained in section 2.5. The division deals with the distinction between 'immediate' consequences, the aggregate of effects observed over time, and changes that become durable.

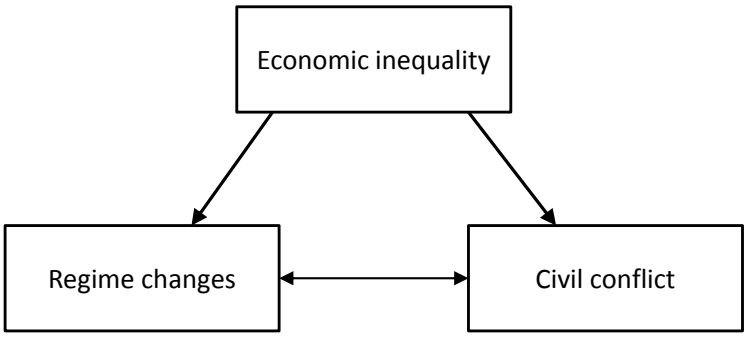
4.5.1 Democratization and civil conflict in the short run

The point of departure for a democratization processes is the authoritarian regime. In autocracies, power is concentrated in the hands of a powerful dictator, a military junta, a designated monarch etc. There are enormous differences between authoritarian systems, but in this framework these differences are subordinate. The most important feature is that autocracies are characterized by 'the rule of the few', rather than the 'rule of the many'. In line with Boix and A&R, authoritarian institutions are assumed to provide little or no redistribution to the citizens.[10] In an autocracy, the citizens are considered to have two options regarding the choice of political regime; to revolt, and try to impose democratic institutions, or acquiesce, and accept prolonged authoritarianism. As the second assumption in section 4.1 was the underlying pressure for democratization, to acquiesce is considered unsustainable over time. The elites either decide to make concessions, or they choose a strategy of repression. Whether the elites decide to repress or make concession depends on the costs associated with each strategy. The costs are here considered to be a function of the level of economic inequality.

In Figure 4.1 I have set up a simple causal structure illustrating the perceived the effects of economic inequality. High inequality is predicted to have a direct impact on the risk of regime changes, as well as on the risk of civil conflict. Moreover, the effect of inequality is expected to be exacerbated, as there is reciprocal causality between regime changes and violence. These causal effects are at the core of this thesis' analytical framework.

[10] In some cases the elites have strategic interest in rewarding specific groups in society with money, powerful positions etc., but this kind of favoritism or nepotism is not considered as redistribution.

Figure 4.1: Causal diagram of the relationship between economic inequality, regime changes and civil conflict



Peaceful transition to democracy

When are countries most likely to experience a swift and peaceful transition to democracy? The simple answer is that it will happen when there is consensus about democracy being a 'good'. That is, when both the citizens and the elites prefer democracy over autocracy. Low levels of inequality are expected to support a peaceful transition to democracy. When economic inequality is low, the redistributive demands of the poor are normally less pronounced. The elites do not fear extensive redistribution, and the costs of accepting democracy are correspondingly low. This is in line with Boix' (2003) model. Equal societies are predicted to have little conflict over the political institutions, and the regimes are therefore more stable. Moreover, in the absence of inequality-induced grievance, there is little social unrest and violence (Gurr 1970).

There is a rational for why the citizens prefer democracy even if the present level of inequality is low, namely expectations about the *future*. First, even if the elites abstain from widespread repression, there is little reason for expecting them to do so in the future. Second, even if inequality is low today, the tides may turn, and the economic structures can become less favorable. Thus, democracy secures political power, and hence ability to influence the distribution of wealth in the future.[11].

The reasoning above is supported by Boix, but not by A&R. According to Houle (2009, 594), the difference in conclusion hinges on an assumption made by A&R concerning the cost of repression in equal societies, which they expect to be close to zero. But, as Houle notes, the assumption is likely to be flawed, as all authoritarian regimes have to repress

[11]That political institutions secure the future political power is also highlighted by Acemoglu and Robinson (2006, 24). However, in their model it is not enough to induce the introduction of democracy in a low inequality country.

a little, or at least keep a repressive apparatus at standby. Hence, when inequality is low, the costs of maintaining authoritarian institutions through repression are likely to be greater than the costs of democratizing.

Violent democratization

In the same manner as certain economic structures make peaceful democratization possible, economic structures can also prove detrimental for the prospects of democratization. I predict that high levels of economic inequality prevent a country from swiftly, and peacefully, transforming from dictatorship to democracy. This prediction is in line with both Boix and A&R. The rationale is that high inequality creates more demand for democracy, while at the same time increasing the resistance of the elites to democratize. Hence, when inequality is high democracy is not an equilibrium solution. As the elites are unwilling to install democratic institutions the citizens must threaten with revolution. If credible, the threat has potential of forcing the elites to make concessions. The threat is a function of the level of inequality, and it is more credible the higher the level of inequality (Boix 2003).

Democratization is more likely to be violent when it takes place in an environment of high economic inequality. As elites fear extensive taxation of their wealth, the cost of democracy is larger than the initial costs of repression, and they often prefer to fight the poor rather than acquiesce. If the elites choose repression and the citizens refuse to acquiesce, the result is a violent conflict (see Boix 2008). Although high levels of inequality increase the risk of civil conflict, whether one actually breaks out depends on the uncertainty about the power-relationship, that is, about the political resources and organizational capacity available to the representative actors, as well as the repression costs of the elites (Boix 2008). The elites can decide to make concessions at an early stage in the process, or they are persistent, allowing the discontent to grow. This is where it becomes pertinent to consider the outcome of this institutional conflict.

I propose three different outcomes, in line with ideal types 2 and 3 described in Table 4.1.[12]

Consider ideal type 3, where there are no institutional changes in democratic direction,

[12]This framework is not exhaustive and focuses on the effects of inequality, as presented in Figure 4.1. Thus, I do not make accurate predictions about when the outcome is prolonged authoritarianism and when it is possible to introduce inconsistent institutions, only point to some factors that may increase the possibility of either of them. Important to note, these other factors are not part of the empirical framework. I acknowledge that it is a simplification, but for the purpose of this thesis it was necessary to restrict myself somewhat and not make the analysis too encompassing.

but where the country is driven into violence, and potentially a long-drawn civil conflict. I propose that prolonged autocracy and violence is more likely when inequality is very high, and when the previous regime has had a strong hold on power for a long time – an expression of their relative strength. The citizens will in this situation have a strong demand for democracy and willingness to use violence to achieve it.

There are several potential explanations why this situation may turn into a violent and failed democratization. A first explanation is that the elites may have crossed a certain threshold in terms of repressiveness. Because of current misdeeds the possibility to partake in the leadership of a new regime is obsolete, and the elites continue to repress. Syria is one good example. The violence escalated slowly over the course of 2011, and erupted in a war-like situation at the beginning of 2012. Despite massive international pressure the Assad-regime has, until recently, been unwilling to make any concessions, or restrict their massive use of violence against civilians (News 2012). A second explanation is plainly that the regime has a superior repression apparatus, in line with predictions from Boix and A&R. A third explanation is commitment problems and private information (Fearon 1995, 381). The elites may for example not consider it possible to commit to a deal that both the citizens and themselves prefer.

Ideal type 2 involves the introduction of inconsistent institutions, often combined with a shorter or longer period of violence and civil unrest. When inequality is high conflict over the preferred institutional constellation is intense. Both the elites and the citizens have much to gain by preserving or subverting, respectively, the current authoritarian institutions. It therefore becomes very difficult to preserve a stable institutional equilibrium that both parties accept. I propose that an inconsistent regime can constitute a 'compromise' solution in this situation.[13] As the demand for democracy is high, but the resistance within the elites is strong, both groups make concessions in terms of accepting institutions that both have authoritarian and democratic features.

The degree and duration of violence will most likely depend on *how* unequal society is, i.e. how much the elites risk loosing by introducing democratic dimensions into the political institutions. When inequality is very high, violent conflict is expected to be almost inevitable, because neither the elites, nor the citizens are willing to make concessions initially. As inequality falls, the probability that an inconsistent regime can be introduced peacefully rises, in line with the argument of Boix.

[13] Although Boix and A&R predominantly operate with two regime types, Boix (2003, 50–53) introduces what he calls a 'limited democracy', and A&R introduces a 'partial democracy (2006, 262). Its main characteristic is limited suffrage for the rich *and* the middle class. This third regime type does not receive much attention in their theories, though, and I wish to highlight their importance.

Democratization from below or from above?

The democratization processes described above are driven by pressure from below, but involve both social classes, as they explicitly or implicitly require the consent of the elite. This contrasts to transitions from democracy to dictatorship, which do not require consent from the population (Regan and Bell 2010).

Although transitions from below are the main focus of this framework, intraclass conflict within the elites can also provoke democratization. According to Boix (2008, 400) intraclass conflict may emerge when "it does not jeopardize the dominant position of the elite". The specific conditions that satisfy this criteria is either that the poor have little political resources so that their capacity to revolt is modest, or that repression costs are low, which constrain them from trying (Ibid.). That democratization is driven from above is emphasized by Houle (2009) and Little (1997). Little (1997) draws on examples from Latin America where several democratic transitions have been initiated from above, although the masses have often been included in the process at a later stage. However, intraelite conflict is unlikely to lead to a complete democratic regime change, as the elites often seek to retain decision-making power, and are less interested in restricting the power of the executive or promoting universal franchise (Gandhi and Przeworski 2007).

4.5.2 Democratization and civil conflict in the long run

In the above sections I argued how economic inequality relates to the risk of regime change and violent conflict in the short run. The next section explores the long-run implications of this causal relationship.

Houle (2009, 595) states that "one striking observation is that most stable democracies that are poor turn out to be very equal." He uses e.g. Costa Rica, Uruguay, Jamaica, Papua New Guinea, and Mongolia after 1990 as examples. Based on the discussion in the last section, a first expectation is that democracy is more likely to become durable quicker when economic inequality is low, because it signifies that that the underlying economic structures are compatible with a democratic institutional equilibrium. When inequality is low democracy represents the institutional composition that maximize the welfare of both the citizens *and* the elites. In this situation the regime is stable, as neither the elites nor the citizens have any incentive to alter the political institutions. This is also in accordance with the 'consistency-inconsistency' argument (Gates et al. 2006), which states that when the political institutions are in consistency, the likelihood of political upheaval is low. Another implication is that a transition to democracy in equal societies

is expected to cause little or no violence, as the conflict over the political institutions are resolved peacefully. This is in line with ideal type 1.

Contrasting the democratization in Western Europe and Latin America supports the above notion (see Collier 1999; Acemoglu and Robinson 2006). It is emphasized that Western European countries are more equal than are countries in Latin America, and that this difference partly explains why it has been so hard to move to a stable democracy in many Latin American countries. Acemoglu and Robinson (2001, 938) also acknowledge the destabilizing effect of inequality, and write that "highly unequal societies are less likely to consolidate democracy, and may end up oscillating between regimes [...]". Drawing on the Latin American case again; social unrest, civil war, and a constant battle over the governing institutions have been some of the main characteristics of political and social life over the course of the last century (Little 1997). A reiterative theme has been that of populist politicians installing redistributive policies, which have resulted in authoritarian coups when the elites felt too threatened by the political development (Acemoglu and Robinson 2006, 38).

The second expectation is that democratization in unequal societies causes more violent conflict in the long run. This type of democratization tends to be partial, as highlighted in section 4.5.1. Although an inconsistent regime can be a 'compromise' outcome, establishing inconsistent institutions does not solve the root causes of the initial conflict – the unequal distribution of wealth and resources. Moreover, as power-diffusion is incomplete, the various actors have incentive to subvert the political institutions (Gates et al. 2006). The on-going struggle, where the elites opts for authoritarianism, while the citizens prefer complete democratization, is expected to cause instability, and more frequent regime changes. In this tug-of-war, the likelihood of civil conflict is higher. Hence, economic inequality is a driving-force of institutional change, but it may also explain why the inconsistent regimes often are more unstable than the consistent regimes.

To reiterate, the amount of violence induced by democratization is expected to be contingent upon how long, and if ever, democratization is completed. Thus, when inequality is low a durable democracy emerges quickly, and the overall amount of conflict in the long run is expected to be low. If, on the other hand, inequality is high, and an inconsistent regime is introduced, the more civil conflict, and the more long-drawn is democratization expected to be.[14]

What are the long-run prospects for the inconsistent regimes, and especially those

[14]Take note that I only refer to civil conflicts that arise over the political institutions, and as part of the democratization process. Other reasons why regimes experience more or less civil conflict are not addressed in this framework.

suffering from a highly skewed distribution of income? I expect there to be several determining factors. First, it may depend on the power relations between the citizens and the ruling elite, and between the elites and other power bases, such as the military (see Little 1997). The relative strength of the actors impact on how, and whether, it is possible to form a stable alliance, and move towards democracy. An example is Egypt, where the military seemingly resists the political changes that are taking place. They try to temporize and intervene into civilian politics, for example by postponing the writing of a new constitution, and maintaining control over the interim government until the election process is duly over (Kirkpatrick 2011).

Second, according to Boix (2003), handling power relations is insufficient. He claims that if a democratic equilibrium is to be attainable the citizens have to lower their redistributive demands. The only way such a commitment can be made credible, is when inequality falls to a sufficiently low level. Third, Acemoglu and Robinson (2006, 40) emphasize the role of the middle class in consolidating democracy. With a large middle class it is possible to have only some redistribution toward the poor, and still equalize the distribution of wealth. Hence, the middle class works as a restriction on redistribution and therefore as a security for the elites. They point to Colombia and Costa Rica as examples of countries who have comparatively large and influential middle classes, and which are also surprisingly stable democracies, in a Latin American context.

In the final section I summarize the main arguments from the discussion above, concerning the relationship between economic inequality, democratization, and civil conflict, in the short and in the long run.

4.6 Summary of the argument

The overarching question, which spurred the above discussion, is *why are some democratization processes swift and peaceful, while others are long-drawn and violent?* I argue that economic inequality is decisive for the progress and outcome of democratization, and the amount of violence observed over its course. Second, compatibility between political institutions and economic structures is decisive regime stability and peace. In line with the theories of Boix (2003) and Acemoglu and Robinson (2006), I assumed that the world consists of two groups, or actors; the citizens and the elites. All individuals prefer the type of political regime that maximizes their welfare. For the citizens this equates to democratic institutions, whereas the elites favor autocracy. Conflict over the political institutions is therefore inevitable.

I presented three ideal type processes, and used them as a point of departure for a discussion of democratization and civil conflict in the short and long run: *In the short run* democratization can occur in both equal and highly unequal societies, but the dynamic of the processes differ.[15] When inequality is low, democracy can be introduced peacefully, because the costs that accrues to the elites under democracy is lower than the costs of preserving autocracy through widespread repression. When inequality is high, on the other hand, the demand for democracy is strong, and the elites resistance towards it is high, because it implies a high cost connected to the redistribution of wealth. The incompatibility turns violent if the elites are unwilling to make concessions. The outcome of the conflict it either the introduction of inconsistent institutions, as compromise solution, or that the authoritarian system prevails. High inequality is thus considered to both provoke regime changes, and civil conflict.

In the long run economic inequality has two main effects. First, it makes completion of democratization more difficult, thus potentially preventing a country from reaching a stable democratic equilibrium. Second, the more swift democratization is, the less oscillation between regimes, and the less conflict incidence is observed over time. An implication is that in order for stability and peace to prevail, and for democratic institutions to become self-enforcing, a certain amount of redistribution may have to take place.

Having constructed the analytical framework it is now time for the empirical investigation. The approach relates to tracing the short- and long-run observable effects of the ideal type processes. I have so far not derived precise hypotheses. The reason why, is that I find it useful to wait until having presented the research design of this thesis, and the methodology of simulations. This is the focus of the next chapter.

[15]The reason why democratization eventually takes place in all societies relates to the assumption about an underlying pressure for democratization, as referred to in section 4.1.

Chapter 5

Research Design

> *The content of "science" is primarily the methods and and rules, and not the subject matter, since we can use these methods to study virtually anything"*
> (King, Keohane and Verba 1994, 9)

This chapter presents the research design of the thesis, and the methods I apply to come closer to an understanding of the relationship between economic inequality, democratization and civil conflict. It also discusses some important methodological issues. I begin with a justification for why a quantitative research design is the better way of dealing with my specific research questions. Next, I present the model of regime and conflict transitions, before I introduce the transition probability matrix. The multinomial logistic model forms the basis of the simulation procedure. I account for the model and its benefits, before I proceed with an argument for applying simulations to measure the long-run effects of regime changes and conflict incidence. Further, I thoroughly explain the simulation setup.[1] A particularly challenging issue for my analysis is data coverage on the inequality variable. I therefore devote an entire section to the issue of missingness in statistical analysis, and the technique I have chosen for dealing with it; multiple imputation. At the end of the chapter I formulate hypotheses related to the analytical framework presented in chapter 4. How I operationalize the dependent and independent variables are described in chapter 6.

[1] The sources of error and uncertainty that arise when I use simulations as part of my research design are discussed in section 7.5.

5.1 Why use a quantitative design?

Quantitative research designs are one of numerous approaches to explaining social science phenomena. Research on democratization and on the causes of civil conflict reflects this methodological diversity, being covered by both in-depth case studies, comparative designs, game theory, as well as large-N statistical studies. The rational is that qualitative and quantitative research designs are complementary, rather than competing approaches in social science, and should be valued by their individual strengths (King, Keohane and Verba 1994, 3).

A general rule to guide the choice of research design is to find the method that best enables one to answer the research question or shed light on the object under inquiry (King, Keohane and Verba 1994). The main reason why I have chosen a quantitative design is that I have interest in explaining *global* patterns of institutional change and violence in a *long-run* perspective. I want to investigate more closely how the two events – regime change and conflict incidence – relate to each other over time. In-depth case studies of individual countries are then, in my opinion, not the best choice of method. Also, as I want to be able to make general statements about these phenomena, investigating a large number of cases is the better way of approaching the research problem. Statistical techniques are considered a good way of generating results that allow for generalization (Bryman 2004, 76). The subsequent analysis is therefore a statistical analysis on a cross-sectional time-series (panel) dataset, where 164 countries are observed in the period between 1960–2010.

The aspect of *time* plays an important role in this thesis, and my research question specifically refers to the relationship between short- and long-run effects. Methodologically this forward-looking perspective relates to the use of simulations as a way in which to generate aggregated long-run effects. Before the simulation procedure is described, some of its main components have to be explained; the transition probability matrix and the multinomial logit model.

5.2 A model of regime and conflict transitions

I develop a model of regime and conflict transitions that relates the probabilities of having one particular regime type and being in a state of either conflict or peace, to the predictor variables. The same kind of model has also been used to gauge the relationship between regime transitions *or* conflict status by, amongst others, Przeworski et al. (2000), Epstein et al. (2006), Hegre et al. (forthcoming), and Hegre et al. (2011).

I distinguish between three regime types and two conflict states: autocracy, inconsistency and democracy, no conflict (peace) and conflict. This leaves me with 6 possible 'states' that a country can find itself in:

AP: 'Autocracy and Peace'

IP: 'Inconsistency and Peace'

DP: 'Democracy and Peace'

AC: 'Autocracy and Conflict'

IC: 'Inconsistency and Conflict'

DC: 'Democracy and Conflict'

The inspiration for constructing this model comes from the analytical framework in chapter 4. As the dynamic interplay between regime changes and civil conflict are assumed to be an important part of the democratization process, they should also empirically be studied in conjunction. No one has previously, to my knowledge, modeled transitions between *both* regime and conflict states simultaneously.[2] In order to illustrate all the possible regime and conflict transitions, and the distribution of countries in the various states, the next section accounts for the transition probability matrix.

5.2.1 The transition probability matrix

The regime and conflict constellations presented above can be treated as a first order "markov chain". Markov chains, or markov systems, are systems that can be in one of several so-called 'states', and move from one state at time t-1 to another state at time t, and at all subsequent time steps. A Markov chain is a random process characterized by being memoryless. This means that the next state depends only on the current state and not on events that have happened at previous points in time. This is also referred to as the 'Markov property' (Waner 2004).

The probability of a country passing from e.g. 'autocracy and peace' this year to 'democracy and conflict' next year is called a *transition probability*. The system of transition probabilities can be presented in a transition probability matrix, which I give an example of in Table 5.1.[3] The matrix specifies all the (annual) probabilities for transition between the various states. The rows in the matrix represent the state at *t-1*, the time

[2] In the following I often refer to the six different regime and conflict constellations as 'states'.

[3] Although I do not present the dataset and the operationalization of the variables until chapter 6, I include the transition matrix in order to better describe the simulation procedure, and explain the intuition behind it. I use the Uppsala/PRIO conflict dataset, and the coding of regime types is from (Gates et al. 2006). The cross-sectional dataset consists of 164 countries observed between 1960–2010.

Table 5.1: Transition matrix: regime type and conflict status at t vs. at t-1, 1960–2010

	Regime Type and Conflict Status at t						
	AP: Autocracy Peace	IC: Inconsistent Peace	DP: Democracy Peace	AC: Autocracy Conflict	IC: Inconsistent Conflict	DC: Democracy Conflict	All Countries
Reg. Type Conf. Status at $t-1$							
Autocracy	2,327	77	22	90	5	2	2,523
Peace	92.23%	3.05%	0.87%	3.57%	0.08%	0.08%	100.00%
Inconsistent	57	1,083	45	4	51	2	1,242
Peace	4.59%	87.20%	3.62%	0.32%	4.11%	0.16%	100.00%
Democracy	14	20	2,161	3	3	39	2,240
Peace	0.63%	0.89%	94.47%	0.17%	0.17%	1.74%	100.00%
Autocracy	80	4	0	439	25	4	552
Conflict	14.49%	0.72%	0.00%	79.53%	4.53%	0.72%	100.00%
Inconsistent	6	54	3	11	221	6	301
Conflict	1.99%	17.94%	1.00%	3.65%	73.42%	1.99%	100.00%
Democracy	1	5	35	5	2	293	341
Conflict	0.29%	1.47%	10.26%	1.47%	0.59%	85.92%	100.00%
All	2,485	1,243	2,266	552	307	346	7,199
Countries	34.52%	17.27%	31.48%	7.67%	4.26%	4.81%	100.00%

step before the observation, and the columns represent the state at the time of observation, t. The probabilities in each cell in every row add up to 1, or 100%. The transition probability matrix (cf. Taylor and Karlin 1998; Waner 2004) represents the core of the simulations, and thus the thesis' long-run analysis.

Table 5.1 reports the observed transition matrix for all countries included in the dataset, between 1960–2010. The rows represent the regime type and conflict status at time t-1 and the columns represent the regime type and conflict status at time t, the year of observation. The row proportions are equal to the annual transition probabilities. I also report the number of country-years in each constellation. With 6 different states there are as many as 36 possible transitions, and I can use the transition matrix to investigate the stability of each of the regime and conflict constellation.

The distribution between the regime types shows that most country-years are autocratic (3037, or 42.19%), as shown by adding the number of country-years in AP and AC in the last row. 2,612 (36.29%) country-years are democratic, and only 1,550 (21.53%) are inconsistent country-years. Out of a total of 7,199 observations, only 1,205 country-years are marked by conflict incidence at time t, as shown by adding the numbers of AC, IC and DC in the last row. The first row shows what happened to countries that were autocracies in peace (AP) at time t-1. Of 2,485 country-years of this type, as much as

2,327, or 92.23%, remained autocracies in peace at time t. Of the autocratic country-years that changed, most of them either changed into conflict (3.57%), or they changed into peaceful inconsistent regimes (3.05%). Only 22 times (0.87% of the country-years) a country transitioned directly from an autocracy in peace to a democracy in peace. Thus, a complete democratic transition in one year is a relatively rare event. The second row shows the transition probabilities for inconsistent countries in peace (IP), the third row for democracies in peace (DP), etc.[4]

The different constellations are relatively stable, as shown by the numbers in the diagonal cells from the upper-left hand side of the table, to the lower right-hand side. As expected, countries in peace have a higher probability of remaining in peace in the subsequent year, than countries in conflict have for remaining in conflict. Amongst the countries in conflict, democracies have the highest probability of remaining stable (85.92%). This is in line with previous research which has found that conflicts in democracies tend to last longer (see e.g. Gleditsch, Hegre and Strand 2009). The inconsistent regimes are overall more unstable than both democracies and autocracies, which is in line with previous research (see e.g. Gates et al. 2006).

5.2.2 The multinomial logistic model

The transition probability matrix described in section 5.2.1 can be estimated by a multinomial logistic model (cf. Greene 1997, 914–917), sometimes referred to as a 'dynamic multinomial model'. In this model I include regime type and conflict status at t as the outcome variable, and the regime type and conflict status at t-1 as a set of dummy variables. Hence, the regime and conflict status in the previous year is used to predict the regime and conflict status this year.

In order to get to grips with what a (multinomial) logistic model is, it is useful to contrast it with the commonly used ordinary least squares (OLS). A standard OLS model assumes a linear relationship between the dependent and independent variables, where a unit increase in the independent variables increase or decrease the probability of some event with a fixed number. There are several problems with this assumption for my specific dependent variable. Firstly, it is not necessarily the case that the relationship between inequality and regime change and conflict onset is linear. At some levels of the independent variable the risk may increase more or less. Secondly, the indicators I use for statistical measurement of conflict and regimes are categorical variables with a

[4]In the subsequent transition matrices I present probabilities instead of percentages, as this is the convention in transition matrices.

finite sample space. A linear model can in this situation provide meaningless predictions outside the range of what is logically possible. When applying the multinomial logistic model these problem disappears, as described below.

The logistic model is based on the logistic distribution, rather than the normal distribution of OLS, and the unit of measurement is log odds or logit. In all types of logistic regressions, including the multinomial, the log odds of some event is estimated for different values of the independent variables. The odds of a regime change and/or conflict onset in country i in year t is the probability of change in status, P_{it}, divided by the probability of no change in status, 1 - P_{it}. The logarithm of the odds has the wanted property of not being restricted from 0 to 1, but extends from - ∞ to + ∞. In contrast to the probability model, the relationship between the independent variables and the log odds of conflict and regime change can be linear in the logistic model (Skog 2009, 354–357).

The multinomial logit model with six possible outcomes ($j = 0$: AP, $j = 1$: IP, $j = 2$: DP, $j = 3$: IC, $j = 4$: AC, $j = 5$: DC) is

$$p(Y_i = j) = \frac{e^{x\beta_j}}{\sum_{k=0}^{2} e^{x\beta_k}} \tag{5.1}$$

The β estimates can also be interpreted in terms of relative probabilities:

$$\frac{p\,(Y = IP)}{p\,(Y = AP)} = e^{\beta'_{IP}\mathbf{x}_i} \tag{5.2}$$

To identify the model I set 'autocracy and peace' (AP) as the base outcome. The estimate β_{IP} from equation 5.2 is interpreted as the impact of the explanatory variable on the probability of being in a state of 'inconsistency and peace' relative to 'autocracy and peace', holding all the other independent variables constant. The estimates for the lagged dependent variables and the constants, estimate the transition probability matrix when all explanatory variables are zero. This is also referred to as the 'underlying transition probability matrix' (Hegre et al. forthcoming, 6).

There are several reasons why the 'dynamic' multinomial logit model is suitable for this study. Firstly, using this model enables me to capture whether the probabilities of moving in and out of different regime and conflict states vary between regime types, and also whether various regime and conflict constellations are influenced by the same or different explanatory variables. Secondly, the 'dynamic model' allows capturing that variables may increase the risk of specific transitions, but not necessarily the duration dependence of these 'states'. It can be modeled by including interaction terms between the predictors and regime type and conflict status at t-1. Thirdly, the model can estimate both onset and termination of conflict and regimes simultaneously, which enables me to simulate the

long-run global and country-level incidence of specific conflict and regime constellations (Hegre et al. forthcoming, 6).[5] This is relevant for answering the second research question, concerning the long-run effects of economic inequality on democratization and conflict. The use of simulations to capture long-run effects is the main focus in the next part of this chapter.

5.3 Simulating long-run effects

In the following I present the methodology of statistical simulations. I extend a simulation routine developed in Hegre et al. (forthcoming), where conflict incidence and the political systems in the world evolve in accordance with the results of the multinomial logit model. The simulation routine allows me to incorporate prior knowledge about conflict and regime history into the model, and update this information over the course of the simulations. Doing this, I can investigate whether aggregated conflict incidence, regime stability and democratization is conditional on the level of economic inequality.

In the empirical analysis the *long run* relates to the forecasting period, which I have set to 2011–2050. Hence, the long run is defined as to be 40 years. It is important to note that I am not interested in making predictions about the level of democracy and conflict in the world, or in specific countries, in 2050. The notation in dates is therefore strictly for convenience, and I could have used the labels 0 (2011) to 40 (2050). 40 years is partly an arbitrary cutoff point, but I consider it a sufficient amount of time to capture the aggregated effects of for example a democratization processes that is initiated in 2010.[6]

5.3.1 Why use simulations?

In order to capture the long-run effects from a statistical model it is common to use predictions. Predictions are often associated with post-estimation in regression analysis, and is based on the parameter estimates from the regression model (cf. Greene 1997, 369–374). A deficiency of using this strategy, instead of simulations, to estimate long-run effects is the assumption of constant transition probabilities. With constant transition probabilities I would 'stretch' the effect of the transition matrix in Table 5.1 forward in

[5]Predicting conflict *incidence* has only been done in a few earlier studies (Hegre et al. forthcoming) and (Hegre et al. 2011). More commonly, prediction projects have dealt with conflict onset (cf. Goldstone et al. 2010; Hewitt 2008; Rost, Schneider and Kleibl 2009).

[6]The issue of forecasting period is further addressed in section 7.4.1, where I find that 40 years is a sufficient amount of time for the transition matrix to reach 'steady-state', i.e. approach a stable matrix (cf. Waner 2004).

time, without updating any of the explanatory variables based on previous year's predicted outcome.

Most crucial for this analysis is to update variables and capture the effect from previous regime and conflict history. When considering the probability that a democracy in peace will transform into a democracy in conflict next year, whether the country has been in conflict before t-1, is an important determinant for its future risk of conflict, especially in the first years after the last conflict ended (see Hegre et al. 2001; Collier and Hoeffler 2004; Hegre and Sambanis 2006). The same holds for regime changes. If a country has just recently experienced a regime change, the probability of new change increases, especially in close proximity to the recent change (see e.g. Strand et al. 2012). Without updating the transition probabilities over time based on new information about recent conflicts or regime changes, the overall effects of an initial change in regime and conflict status is likely to be underestimated in the long run.

The simulation procedure described below allows me to incorporating regime and conflict history further back in time, and update the transition probabilities on a yearly basis. This provides more accurate, and in my opinion, realistic results. In the model I use the effects of previous conflicts and regime changes are captured by variables that measure the time in each of the combined regime and conflict states, as presented in section 5.2, up to time t-2. This is further described in chapter 6.

When I use simulations I am also able to fully capture the long-run effects from the *endogenous* relationship between regime changes and civil conflict.[7] That a regime change at time t-1, may trigger a conflict onset at time t, which further increases the risk of new regime change and conflict, is the core of my analytical framework. By simulating, the reciprocal causality, and how it influences the history variables, are taken into account. Moreover, I am able to keep the other explanatory variables exogenous at a constant level, and thus isolate the effects of an initial change in the level of inequality on the dependent variable.

Although there are methodological challenges and areas of uncertainty, issues which I discuss in section 7.5, using simulation technique provides results that more accurately capture the long-run effects of inequality on democratization, and on conflict incidence. Thus, I am better able to model and measure the relationships that are outlined theoretically, but which have not yet been studied empirically. I now proceed with a description of the simulation procedure.

[7]By endogeneity I mean that there is a causal relationship between regime change and conflict, where the value on both variables at time t is dependent on their values at time *t-1*, as well as on other variables in the causal system.

5.3.2 The simulation procedure

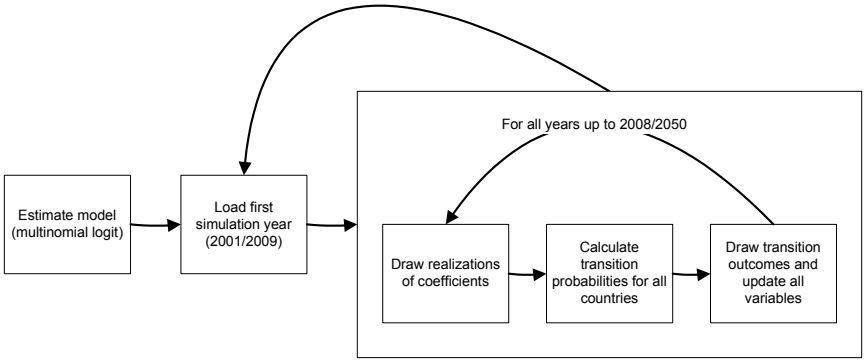
The simulation procedure is based on a statistical model where the probabilities of regime and conflict onset and termination depends on exogenous variables, such as economic inequality, GDP per capita and population size. In addition, I include regime- and conflict history as endogenous variables.[8] Figure 5.1 illustrates the simulation procedure. This figure, and parts of the description below, are based on the description of the simulation setup from Hegre et al. (forthcoming, 4–6), and Hegre et al. (2011, 17–18).

The first step (1) involves the specification and estimation of the multinomial logit model. I have explained the logic behind the model in the sections above, and the specification of it is given in chapter 6. In the next step (2) assumptions about the distribution of values for all exogenous predictor variables for the first year of simulation, and future changes to these, have to be made. Then the simulation can start: (3) The simulation routine starts at the first year of simulation, which in my case is 2011. First (4), a realization of the coefficients from the multinomial logit model is drawn. These are based on the estimated coefficients and the variance–covariance matrix of the estimates.[9] (5) Next, the program calculates 36 transition probabilities; between the three regime categories and peace and conflict status, before it (6) randomly draws an outcome amongst the six possible regime and conflict categories, based on the estimated transition probabilities. When the regime type and conflict status of a country in 2012 is 'decided', (7) the explanatory variables are updated based on this information. If for example Egypt was a peaceful autocracy in 2011, but the realization in 2012 defines it as an autocracy in conflict, this has consequences for the values ascribed to the endogenous history parameters, through the transmission mechanism described above. (8) Now steps (4)-(7) are repeated for each year in the forecasting period, which in this study is 2011–2050. The first simulation in then complete, but in order to reduce the impact of individual realizations of the coefficients

[8] All predictors could have been made endogenous to the model, but this is beyond the scope of my thesis. As stated by Hegre et al. (forthcoming, 7), at least some of the effect from regime changes and conflict on the predictors are captured by the history variables. Moreover, as I am interested in isolating the effects from the analysis' main variables, this is not preferable either.

[9] Variance and covariance are two statistical measures: The variance of a coefficient is a measure of dispersion. The true variance is always unknown, thus, has to be estimated. The covariance indicates how two coefficients change or vary together. If for example conflict and inequality vary together in the same direction relative to their expected values, the covariance of conflict and inequality is positive. If there is no independence between conflict and inequality the covariance is 0. The variance–covariance matrix presents data on both the variance and covariance for all the coefficients in an organized form; the variance is presented on the diagonal, while the covariances reside above or below the diagonal (cf. Greene 1997, 101–104).

Figure 5.1: Simulation flow chart



Source: Hegre et al. (forthcoming, 5)

from the multinomial logit model, as well as the individual draws from the probability distribution, the procedure (3)-(8) is repeated a large number of times. The confidence intervals, and the probability of random errors, should therefore be reduced (Hegre et al. forthcoming, 4–5).

To conduct the simulations I apply the statistical package STATA 12.1 and a simulation program developed by Joakim Karlsen, for Hegre et al. (forthcoming) and Hegre et al. (2011) called 'PrioSIM'. Its main input is a dataset with observed and projected data until 2050, and a parameter file with mathematical expressions that controls how the variables are updated in the simulation.[10] The main output is the simulation results.[11] The last step involves importing the results into Stata and summarizing them so that they can be presented in an comprehensible manner. I have modified a Stata do-file from Hegre et al. (2011) for this purpose. It extracts, amongst other things, the distribution of years in each regime and conflict state, by country, and globally. It also allows me to separate between countries with high and low levels of inequality.[12] In this manner, I attempt to identify the 'ideal type' processes presented in section 4.5, and test the hypotheses which I derive shortly, in section 5.5.

[10] An illustration of the parameter file used to conduct the analysis can be found in Appendix C.

[11] The simulation procedure is currently divided into three parts and the main components of the software are as follows: First, a Stata adofile reads the parameter file and performs the estimation of the multinomial logit model. It extracts and saves two .csv files; (1) draws of the beta-coefficients, and (2) the dataset, which includes all the relevant variables with observations for the entire time period up to 2050. Then, a C# class library imports the two files and the parameter file, runs the simulation and provides a result file, which is a .txt file that contains information from all the simulations. The parameter file is the interface that manages the simulation and MuParser, a general math parser, is used to evaluate the expressions in the parameter file. See http://muparser.sourceforge.net/ for details on the math parser.

[12] See Appendix C for a copy of the do-file.

5.4 Methodological challenges

Apart from validity problems related to the selection of indicators, something I address in chapter 6 and section 7.5, there are other challenges facing a quantitative design for civil conflict and regime change data: non-independent observations, endogeneity (Bryman 2004, 76), the rareness of conflict onset and regime change events (cf. King and Zeng 2001), and missingness, whereof the two latter are closely related in my study.[13]

The logistic regression model assumes that observations are independent of each other. Hence, the risk of a regime change or a conflict onset in Egypt next year is assumed to be independent of a regime change this year. But temporal independence is often a crude assumption (see Raknerud and Hegre 1997). In a 'dynamic multinomial model' dummy variables denoting the regime and conflict status of a country in the preceding year are included. These dummy variables can also account for temporal dependence (cf. Fearon and Laitin 2003, 82).

A second concern is endogeneity. In cross-sectional designs the direction of causal influence is often ambiguous (Bryman 2004, 76). In the simulations regime type and conflict are endogenous to the model, and hence, I am able to capture that there is a reciprocal causality between regime changes and conflict onset. The independent variables, on the other hand, are assumed to be exogenous. The ideal solution to endogeneity is to create an instrumental variable that is highly correlated with the independent variable, but only affects the dependent variable through the independent variable, and include it as a proxy for the independent variable. As instrumental variables are difficult to find, another solution is to reduce the threat of reverse causation by lagging the independent variables (Hegre and Sambanis 2006), which I do in the analysis. Moreover, I expect the endogeneity problem to be partly solved by the inclusion of information on regime and conflict history. The history variables capture some of the detrimental effect of, for example, a conflict onset on economic growth. In the proceeding part of this chapter I focus on of the most important challenges for this analysis, namely missing observations. Specifically, I discuss the technique I apply to solve the problem of missingness; multiple imputation.

[13] Another potential source of bias relates to the simulation procedure and the uncertainty of predictions. I address this in section 7.5.

5.4.1 Missingness and multiple imputation

Some of the indicators I use in the analysis originally contain missing values, in particular the variable I use to measure economic inequality. Missing data creates problems in statistical analysis for several reasons. I address them in the following, simultaneously as I discuss ways of solving the problem of missingness.

The most common way of handling missing values in panel data is plainly to discard the information by listwise deletion, also referred to as "Complete Case"-analysis (Honaker and King 2010). I am not satisfied with this approach to the missingness problem for several reasons. First, I would like an as complete dataset as possible. Missing values means that I cannot fully utilize all the information on conflict onsets and regime changes. The estimation of the 5-equation multinomial logit model requires an extensive amount of information, and I want to avoid problems with missing cells in the transition matrix in Table 5.1.[14] I already have one missing cell; the transition from AC at t-1 to DP at time t, and several cells with only one or two observations. Thus, removing country-years from the dataset because of missing values increases the probability of more missing cells.[15]

Second, in order for listwise deletion to be appropriate the observations should be Missing Completely At Random (MCAR), which means that the missingness is completely uncorrelated with the dependent variable. This is not always the case in panel data, and observations are then said to be Not Missing at Random (NMAR) (see Høyland and Nygård 2012). If data is NMAR the missingness depends either on predictors not included in the model, or on the missing value itself. Non-random missingness is potentially a problem in this study, as countries prone to civil conflict and regime changes may less-often have available data on both fiscal and social parameters which are represented in my dataset. Hence, there is a danger that 'high-risk' countries systematically become under-represented in the sample if I apply listwise deletion.

Missing at Random (MAR) is a third type of missingness. It implies that the probability of an observation being missing can be explained by other covariates in the data that is not missing (Høyland and Nygård 2012, 3). If the 'missing' predictors are included in the model, missingness can go from NMAR to MAR. But, as noted by Høyland and Nygård (2012, 5), listwise deletion may still induce bias if the deleted sample is not representative of the full sample.

[14] In addition, the simulation program cannot handle missing values. All country-years with missing values must therefore be removed from the dataset before the analysis.

[15] A missing cell provides extreme estimates in the analysis and the parameter estimate thus has to be restricted. For further description, see section 7.2.

Another method of handling missingness is qualified guesses, or replacing the missing value with a statistical estimate, for example the mean. In some instances this can be a satisfactory way of dealing with the problem at hand, but most times it is a subordinate technique. The most notable effect is over-confidence with the results, as standard errors and confidence intervals are underestimated (Høyland and Nygård 2012).[16]

A third way of dealing with missing data is applying multiple imputation techniques. Multiple imputation fills in the missing observations by using a predictive model that incorporates all available information in the observed data, along with any prior knowledge. The value that actually replaces the missing value is the mean of the imputed values across several imputed data sets. Uncertainty is accounted for by making each missing value a representation of multiple imputations for that specific value (Honaker and King 2010, 563). Where listwise deletion relies on the assumption of missing completely at random (MCAR), multiple imputation relies on the less stringent assumption of missing at random (MAR).

In the multiple imputation procedure one missing observation is replaced by an imputed value D. The variance of D is also the imputation uncertainty. An analysis is run on all the imputed datasets, and the parameters are then averaged over the D estimations. An alternative, which is the one I have used, is to combine the imputed datasets, average the observation, and then analyze them as one (Little and Rubin 2002, 86).[17]

I use the program Amelia II (Honaker and King 2010) to conduct the multiple imputation. Amelia II applies a bootstrapping algorithm (cf. Greene 1997, 184–185), which is uncommon in other imputation models. Normal imputation models assume that the missing values are linear functions of other variables' observed values, that observations are independent conditional on the remaining observed values, and that all the observations are exchangeable (Honaker and King 2010, 565). According to Honaker and King (2010, 566), this does not work for time-series cross-sectional (TSCS) datasets, because of the tendency of variables to move smoothly over time, and to jump sharply between cross-sections, like countries, and for time-series to differ across countries. Amelia II takes these issues into account.

As recommended by Honaker and King (2010, 567), I include lags of the independent variables, as well as leads, where the future is used to predict the past. Polynomials as a function of time are included in the model to let trends vary across countries (see

[16] As described in chapter 6, I have applied such 'guesstimates' on a couple of occasions, but always with a justification and thought-through argumentation for doing it.

[17] I could, in theory, have conducted simulations on all the ten datasets and then averaged the predictions. This would be very time-consuming, and I have therefore chosen not to do this.

Honaker, King and Blackwell 2011, 17). I also include empirical bounds on the variables, and set the empirical prior to 5%. The ridge prior places restrictions on the data in order to ease imputation in cases where the algorithm has problems converging, for example when there are many missing values. One way to look at it is picturing a ridge, which has a very restricted space. In those instances where there is poor data coverage the missing observation is replaced by a 'conditional mean'. A prior of 5% is quite large, but it is useful because of the extent of missingness in the economic inequality variable Houle (see 2009, 617). I impute 10 datasets altogether.[18]

After having presented the research design and the methodology of simulations, it is possible to be more precise in the formulation of hypotheses. The hypotheses relate to the short- and long-run effects of inequality on democratization and civil conflict, and are presented in the next section.

5.5 Arriving at hypotheses

This thesis operates with two sets of hypotheses, both based on the analytical framework presented in chapter 4. The short-run hypotheses deal with the immediate, direct, effects of economic inequality on the risk of having specific political institutions and being in a state of conflict in an given country-year. The long-run hypotheses take the form of empirical expectations from to the ideal types of democratization (see Table 4.1). I cannot observe the processes directly, but aggregated over 40 years they are expected to cause variation in conflict incidence and political institutions. As explained in chapter 4, I propose that this variation is contingent on differences in economic inequality between countries.

Short-run effects:

H_1: Economic inequality decreases the probability of being in a state of democracy and peace, relative to autocracy and peace.

H_2: Economic inequality increases the probability of being in a state of institutional inconsistency, relative to autocracy and peace.

H_{2a}: Economic inequality increases the probability of being in a state of inconsistency and peace, relative to autocracy and peace.

H_{2b}: Economic inequality increases the probability of being in a state of inconsistency and conflict, relative to autocracy and peace.

[18] I describe more details related to imputation under the indicator-heading for 'economic inequality', in section 6.2.

H_3: Economic inequality increases the probability of being in a state of autocracy and conflict, relative to autocracy and peace.

The hypotheses relate to the causal system described in Figure 4.1. Higher inequality directly increases the risk of conflict, but it also increases the risk of regime changes, and in particular, towards the inconsistent regime category. Moreover, I expect higher inequality to decrease the probability of moving peacefully from autocracy to democracy. The reason why I have formulated the short-run hypotheses in terms of 'states' and not 'transitions', which is the most intuitive given the theoretical argument, is for the sake of simplicity. I insert this simplification to decrease the number of parameters to be estimated – a model with many interaction terms would have gotten very complex –, but also to make it easier to interpret the results.[19] However, by estimating the probabilities for being in a given state I am implicitly studying transition probabilities, as the two are closely related. This is clear from inspecting the transitions probabilities in the diagonal of Table 5.1. The probabilities of remaining in the same state at time t-1 and time t is very high. Thus, if the probability of transition from autocracy to democracy is high, the probability of being in a state of democracy relative to autocracy is also expected to be high.

Long-run effects:

H_4: The total proportion of civil conflict incidence is higher in the long run if economic inequality is high.

H_5: The total proportion of regime changes is higher in the long run if economic inequality is high.

H_6: The total proportion of civil conflict incidence is higher in inconsistent regimes than democracies in the long run.

H_7: The total proportion of institutional inconsistency is higher in the long run if economic inequality is high.

Hypotheses H_4 and H_5 relate to the aggregated effects from the causal system in Figure 4.1, and are in line with 'ideal type' 2 in Table 4.1, as countries with high inequality are expected to have difficulty democratization and to more often end up in the unstable inconsistent regime category, and oscillation between regime types. Hypothesis H_4 also relates to 'ideal type' 3, where attempts at pushing through democratic changes are met with resistance from the government, causing minor or major episodes of conflict. Hypothesis H_6 contrasts ideal type 1, the peaceful and swift democratization, with ideal

[19] To model transitions I would have had to include interaction terms between economic inequality and the regime and conflict state at t-1.

type 2, the conflict prone and long-drawn democratization. The last hypothesis, H_7, also relates to ideal type 2, as countries with high inequality are expected to have difficulty transforming into a consistent democracy.

The short-run hypotheses will be investigated with the results from the multinomial logit model, whereas the long-run hypotheses relate to the simulation results. Before I can put the hypotheses to the test, the data and indicators used for measurement have to be presented.

Chapter 6

Indicators and data

In this chapter I account for the operationalization of the theoretical concepts, and the variables used in the empirical analysis. The dependent variable is 'regime type and conflict status'. The independent variable is economic inequality, which is measured by the variable *capital share*. The control variables are collected from the literatures on democratization and civil conflict. There are numerous potential control variables, but I focus on those I consider most relevant when studying regime change and civil conflict in conjunction. I expect the reliability, i.e. the notion of whether the data is measured properly (King, Keohane and Verba 1994, 255), to be high for the main variables, as they originate from sources which are widely applied and acknowledged. Hence, the validity of measurement receive more attention in the discussion below.

The main sources are the dataset developed by Hegre et al. (forthcoming), the dataset from Gates et al. (2006), and the dataset from Strand et al. (2012). Due to data availability the time-span of the analysis is limited to 1960–2010. Descriptive statistics for the variables is found in Appendix A.

In the following, I first present the dependent variable, 'Regime Type and Conflict Status', and its two components; regime type and civil conflict. Thereafter, I describe the independent variable, economic inequality, before I give an account of the control variables.

6.1 The dependent variable

This thesis is innovative in its choice of dependent variable, as I apply a combined regime and conflict variable. The choice is made to take into account the intimate relationship between changes to political institutions and violent conflict, and to underline that they

often are part of the same process, and therefore should be studied in conjunction. Before I describe the combined variable more intimately, I start with presenting its two individual components, regime type and civil conflict.

6.1.1 Political regimes

As discussed in section 2.1.1, it is common to separate between substantial and institutional definitions of democracy. I argued for using an institutional definition. The operationalization of political regimes should reflect this choice.

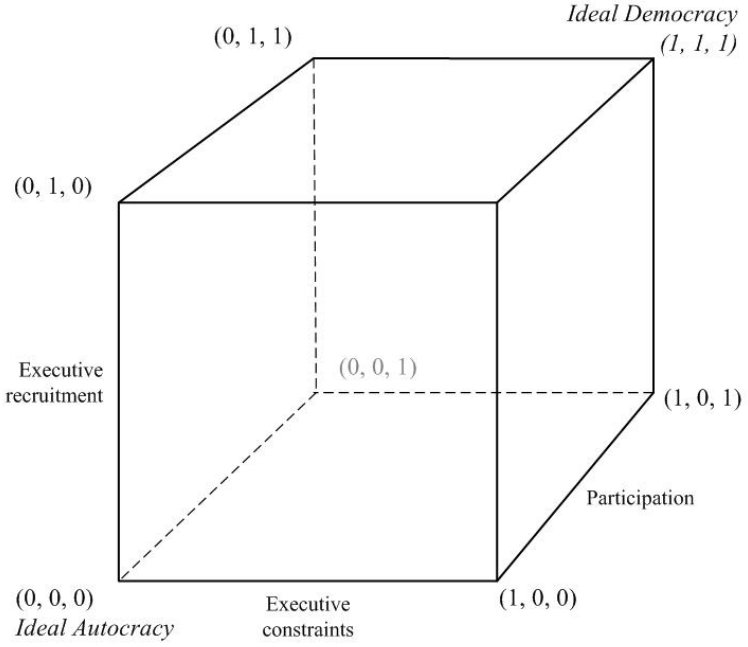
There are two well-known democracy measures; The Freedom House Index (FHI) and the Polity Index (PI). The FHI consists of two different indexes; the Political Rights index and the Civil Liberties index. Both indexes are aggregates of several sub-indicators. One of the main drawbacks with the FHI is that the scoring on the individual indicators are relative subjective. This increases the risk of bias and unsystematic measurement error. Moreover, Freedom House does not publish their original data, which reduces the transparency of the FHI (Knutsen 2011, 89–90). Despite these flaws, the FHI is a frequently used measure of democracy, and it is considered well-suited for operationalizing substantive definitions of democracy (Ibid.).

The Polity Index (PI), on the other hand, is well-suited for operationalization of narrower, institutional definitions of democracy. The PI is the basis for this thesis' operationalization of political regimes. A characteristic of the PI is that it only incorporates formal institutional structures, and excludes civil liberties and other elements that require subjective judgement in the coding process (Knutsen 2011, 91). Thus, it is better for securing content validity. However, it may create other validity problems. By ignoring the division between formal institutions and their actual use, face validity is threatened (Ibid.).[1] I have chosen to emphasize content validity, and this reflects in the operationalization of political regimes.

The basis for the operationalization of political regimes is the 'ideal polity' types, defined by Gates et al. (2006). The two 'ideal types' are the consistent democracy and the consistent autocracy, which are classified on the basis of three institutional dimensions: *Executive Recruitment*, *Executive Constraints*, and *Participation*. An illustration of the institutional dimensions, and the ideal regime types, adopted from Gates et al. (2006, 896), is found in Figure 6.1. The coding on the three dimensions provides a unique position for every polity on a cube with coordinates. The cube identifies the autocratic and democratic corners, along with the intermediary positioned 'inconsistent regimes'.

[1] See section 2.1.1 for a definition of face validity.

Figure 6.1: Three dimensions of authority structures



Source: Gates et al. (2006, 896)

The regime classification is based on calculating the distance from the point where a regime's coordinates are situated, to the eight corners as well as to the midpoint of the cube. If the regime is closer to either of the ideal corners than to the midpoint it is defined as a democracy (corner [1,1,1]) or an autocracy (corner [0,0,0]). Regimes that are not coded democratic or autocratic are defined as 'inconsistent' (Gates et al. 2006, 896).[2]

An inconsistent regime can have a democratically elected executive, but restricted participation. South Africa under apartheid is a good example of a so-called 'minority democracy'. Participation was restricted to only the Afrikaans citizens of European origin, and the native African population was systematically segregated, and excluded from political life. Another combination is an ascribed executive that allows extensive participation. A historical example is the United Kingdom in the 1800s, where there was a dual system with an ascribed monarch carrying much power, as well as an elected head of government.

The political regime variable is originally from Gates et al. (2006), but I have collected it from Strand et al. (2012). Tha basis for its constructions is the Polity dataset (Gurr 1974; Marshall and Jaggers 2002, 1485), where a set of six indicators characterize political systems. The Polity dataset has information on all countries with a population of more than 500,000 for the period between 1816–2010. Gates et al. (2006) group the indicators

[2]A country classified as an inconsistent regime either scores high on one dimension, and low(er) on the two others, or score in the middle on all of them (Gates et al. 2006, 896).

from the Polity dataset into variables that describe a regime along the three institutional dimensions:

Executive Recruitment consists of three indicators; 'Regulation of Chief Executive Recruitment' (XRREG), 'Competitiveness of Executive Recruitment' (XRCOMP), and 'Openness of Executive Recruitment' (XROPEN).[3] Countries with competitive elections are at the one extreme, and are coded 1, which signifies a high dispersion of power. At the other extreme are countries with succession by birthright, so-called 'Ascription', 'Designation' (informal competition within an elite), or a combination of the two. They are coded 0, recognizing the concentration of power (Gates et al. 2006, 897).

Executive constraints is the second dimension. It is based on a single indicator; 'Decision Constraints on the Chief Executive' (XCONST). The indicator goes from 1 to 7 (Marshall and Jaggers 2002, 21). A higher value means that there are more constraints on the executive, and the more constraints the more dispersed is power expected to be.

The third dimension is *political participation*. Vanhanen's (2000) Poyarchy dataset is the original source of this indicators, because of some inherent problems with the Polity participation index (see Gates et al. 2006, 897). The most decisive for this analysis, is the fact that the index includes systems which are characterized as 'factional', i.e. where "patterns of intense, often *violent competition between 'in' and 'out' factions* [author's emphasis] [...]." (Gurr 1974, 1486). This endogeneity problem introduced by the original Polity participation index is highly problematic, and two indicators from Vanhanen's (2000) Polyarchy dataset are used instead instead; 'Participation' and 'Competition' (Gates et al. 2006, 897). Participation is the percentage of the population that voted in the most recent election, and competition is the percentage of the valid vote won by all parties except of the plurality winner or winning electoral alliance.

The three institutional dimensions are normalized to range from 0 (maximum concentration of power) to 1 (minimum concentration of power), in line with Figure 6.1.[4]

The variable *regime type* consists of three regime categories; autocracy, inconsistent, and democracy. I have collected the variable from the dataset of Strand et al. (2012), and it covers 200 countries between 1816–2008. I have constructed a rule for updating the variable until 2010 (see below), in line with the original coding of a polity change from

[3]The three indicators are from the Polity dataset, and the abbreviation in brackets refers to the name used in the Polity dataset.

[4]For a more detailed description of the coding on the three institutional dimensions consult Appendix B of Strand (2007). For more information on the Polyarchy dataset, see http://www.prio.no/CSCW/Datasets/Governance/Vanhanens-index-of-democracy/Polyarchy-Dataset-Manuscript/.

Gates et al. (2006, 898):

A *regime change* is defined as a movement in any indicator, which results in one or more of the following three changes in the institutional dimensions: (1) a movement from one category to another on the 'executive dimension' (i.e., between ascription/designation, dual ascriptive/elective, and elective); (2) a change of at least two units in the 'executive constraints' dimension; or (3) a 100% increase or 50% decrease in the Participation dimension. If the indicators change according to the definition above, I code a new regime type in 2009 or 2010. Otherwise, I code the same regime type.[5] There are six cases where these criteria are met and I code a regime change; two in 2009 and 4 in 2010.[6] The dataset used in the analysis includes 330 regime changes between 1960–2010. 104 took place in autocratic regimes, 141 in inconsistent regimes, and 85 in democracies.

When a transition period is coded in the original Polity dataset, defined as either -66 (foreign interruption'), -77 ('interregnum or anarchy') or -88 (transition), Gates et al. (2006) have, consistent with the Polity project, coded the transitional regimes as 'missing'. I alter this coding rule, and let the most recent regime 'endure' through the transitional period, until the first year of the new regime. This is not entirely unproblematic, as it may inflate the number of autocracies and inconsistent regimes in the sample. However, I consider it a better solution than to exclude these country-years. This would lead to a loss of valuable information, since there are as many as 210 conflict country-years during these transitional periods. Moreover, I risk introducing bias as these country-years have proportionately more conflict than non-transitional country-years; 33% of the transitional country-years include a conflict, whereas only 15% of the non-transitional country-years. This has also been highlighted by Regan and Bell (2010, 7).[7] See Appendix A for an overview of the countries of concern.[8] After making the above changes, the distribution

[5]In lack of an updated version of Vanhanen's Polyarchy data, I inspect the Polity participation index, PARCOMP, and whether it has changed by more than two units between 2008–2010.

[6]The cases are Niger (from democracy to an inconsistent regime in 2009), Bangladesh (from inconsistency to democracy in 2009), Guinea (from autocracy to inconsistency in 2010), Madagascar (from democracy to inconsistency in 2010), Iraq (from autocracy to inconsistency in 2010), and Sri Lanka (from democracy to inconsistency in 2010).

[7]Regan and Bell (2010, 7) also included the transitional country-years, but chose the imputed version of the Polity data which interpolates the uncodeable cases.

[8]There are 11 countries in the dataset from Hegre et al. (forthcoming) with conflict data which are not coded with regime data. Most are missing because they have populations of less than 500 000 and are therefore not part of the Polity dataset, and a couple are 'new states' formed in the recent years. I assume the missingness to be uncorrelated with the conflict variable, thus, removing them from the sample is not likely to cause bias. In addition, there is one country coded in transition throughout. The country is Bosnia. I keep it in the dataset because it has available conflict data, and moreover, is not

over the three regime types between 1960–2010 is: Autocratic (3106), Inconsistent (2663), and Democratic (1594).

6.1.2 Civil conflict

The second component of the dependent variable is *civil conflict*. In the empirical analysis I focus on the *incidence of civil conflict*, as defined in section 2.2. Just as there is currently no precedent theoretical definition of civil conflict, there is no clear understanding of the best way to measure it in quantitative research (Sambanis 2004*b*).

One disagreement concerns the issue of when to code a conflict onset. This is referred to as the 'threshold criterion'. Most conflict datasets, for example the well-known dataset from the Correlates of War project (COW) (Small and Singer 1992, 213–215) and the dataset from Fearon and Laitin (2003), apply a strict coding rule. COW only codes an onset after 1,000 battle-related deaths have been recorded, and require the same number of deaths in every subsequent year. Fearon and Laitin (2003, 76) use a somewhat less stringent coding rule, where, *after* an onset is coded (1,000 deaths necessary), the violence on *average* has to cause 100 deaths per year, including civilian casualties. A strict criterion has some advantages, but also considerable drawbacks. One advantage is that it is quite straightforward to code when the war starts and ends. However, main drawback is the exclusion of all warlike episodes, as well as low-key violence with only a few casualties (Sambanis 2004*b*). Another, frequently used, dataset, is the UCDP/PRIO Armed Conflict Dataset. It applies a more flexible coding rule, which is favorable to the analysis in this thesis.

One strength of the UCDP/PRIO dataset is a clear theoretical definition of armed conflict, which reflects in the coding criteria. It includes all conflicts that meet the following criteria: There has to be a (1) contested incompatibility that (2) concerns government and/or territory where the (3) use of armed force between two parties, of which (4) at least one is the government of a state. The violence has to result in at least (5) 25 battle-related deaths. An additional criteria for *internal* armed conflict, i.e. civil conflict, concerns the characteristics of the two opposing parties: The conflict has to occur (6) between the government of a state, and one or more internal opposition group(s). Further, there has to be (7) no outside intervention from other states (UCDP/PRIO 2010, 3).

The coding criteria focuses explicitly on an *incompatibility*, i.e. the issue which is

a 'microstate'. For Bosnia, I manually impute data on regime characteristic from a similar neighboring country, Croatia.

fought over.[9]

A drawback with the UCDP/PRIO dataset, is the lack of a threshold for battle-related deaths on both sides. Thus, it is impossible to establish with certainty whether there has indeed been effective resistance. A second weakness is that civilian deaths are excluded from the definition. However, this criterion can also be argued to be a strength, as it secures against one-sided violence being categorized as civil conflict. Overall, I consider the UCDP/PRIO Armed Conflict Dataset the better choice for this analysis.

I use data from the 2010 update of UCDP/PRIO Armed Conflict Dataset (Themnér and Wallensteen 2011; Gleditsch et al. 2002). The dataset differentiates between minor civil conflict, with 25 to 999 battle-related deaths in a given year, and major conflict, with at least 1,000 battle-related deaths. I have collected the conflict variable from Hegre et al. (forthcoming), but modifies it for the purpose of this analysis. I define a new variable, *conflict*, which has two distinct outcomes: 'non-conflict', coded 0, and 'conflict', coded 1. The 'non-conflict' category is referred to as 'peace' in the continuation.[10] I do not distinguish between minor and major conflict; all conflict incidence with at least 25 battle related deaths, in any given year, are coded as a conflict incidence. The conflict variable originally covers 169 countries between 1946–2010.

The reason why I combine the minor and major conflict relates to the research questions of this thesis. First, I am interested in the aggregated conflict incidence over time. Whether they are minor or major conflicts is less important. Second, the analytical framework emphasizes that civil conflict is but one of several manifestations of political violence. Hence, there is a well-founded reason for choosing a more lenient coding criterion for the onset of conflict, as it allows me to capture minor incidence of conflict, in addition to those that evolve into major civil wars. Gleditsch et al. (2002) states that "25 deaths in a single year - is high enough for the violence to represent a politically significant event, although the precise local and international impact may vary." A lower threshold results in more conflicts, and therefore makes statistical analysis more feasible and robust (Ibid.). Third, a practical reason for combining minor and major conflict is that a statistical model with

[9]The UCDP/PRIO Armed Conflict Dataset records the type of *incompatibility* the conflict is concerned with, and distinguish between an incompatibility over *government* or over *territory*. My interest is predominantly in conflicts over government, and thus, I should ideally distinguish between the two. I have chosen not to do this, because it in some cases is difficult to differentiate the incompatibilities; sometimes a conflict concerns both issues at the same time, other times the conflict may evolve from a conflict over territory, to a conflict over government. I would then miss out on this information, by omitting all conflicts over territory from my sample.

[10]For more information on the dataset, see http://prio.no/CSCW/Datasets/Armed-Conflict/UCDP-PRIO/.

a threefold conflict variable would have given me 81(!) possible transitions, instead of 36. The estimation of this multinomial model would be complicated. Moreover, to interpret the results from a multinomial logit model with 9 equations would be very demanding.

Incidence of civil conflict is a relatively rare event. There were a total of 1,219 country-years with conflict between 1960 and 2010. Conflict incidence thus make up 16.56% of all country-years. The number of conflict onsets is much lower; 201.

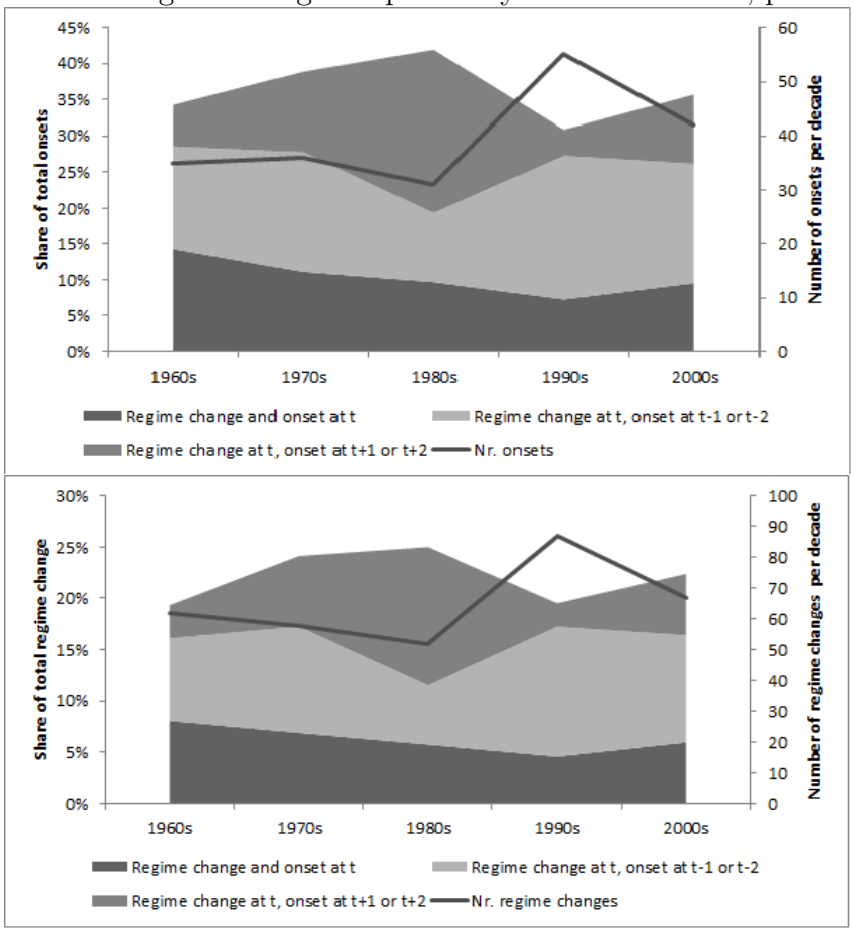
6.1.3 Regime type and conflict status

Based on the regime and conflict variables described above I construct a combined variable, *regime type and conflict status*. It consists of 6 categories: Autocracy and Peace (AP), Inconsistent and Peace (IP), Democracy and Peace (DP), Autocracy and Conflict (AC), Inconsistent and Conflict (IC) and Democracy and Conflict (DC). In the multinomial logistic model 'autocracy and peace' (AP) is set as the base outcome. As I am interested in democratic transitions, I judge this to be the most reasonable base outcome.

There are many different ways I could have coded the dependent variable, in particular in relation to different lag structures between conflict and regime change. I decided not to, for example, collapse two and two years, in order to take into account that a regime change and a conflict onset happen in close proximity to each other, but not manifesting themselves within the same year, thus emphasized the goal of having as many observations as possible, in order to avoid empty cells in the transition matrix in Table 5.1.

Figure 6.2 serves as an illustration that regime changes and conflict onsets often happen in proximity to each other. The upper panel reports the share of total conflict onsets, measured by decade, that coincided or took place in close proximity to a regime change. Analogous, the lower panel reports the share of regime changes that coincided with a conflict onset. There are 20 incidence in the sample where a regime change and a conflict onset took place within the same year. It amounts to approximately 10% of total onsets and 6% of total regime changes, reported as the darkest shaded area at the bottom of the two panels. It is reasonable to expect a time lag between the two events. For example, a regime change last year causes unrest and a civil war in the subsequent year. Or, political violence and protest culminates in a civil conflict, which only after a couple of years leads to the introduction of new political institutions. An example of the former is Haiti in 2006, which had an onset of conflict in 2004. An example of the latter is Peru, which transitioned to democracy in 1980 and had a conflict onset in 1982. Hence, I also include in the Figure 6.2 onsets that took place at time t-1 and t-2, and at time t+1 and t+2. Altogether, almost 40% of all conflict onsets took place in close proximity to a regime

Figure 6.2: The relationship between regime change and conflict onset, 1960–2010. Upper panel: Share of conflict onsets in proximity to regime change, per decade. Lower panel: Share of regime changes in proximity to conflict onset, per decade.



change. For regime changes the number is somewhat lower; approximately 23% of all regime changes took place in proximity to a conflict onset.

6.2 The independent variable: Economic inequality

Measuring economic inequality is far from easy. Lambert (2001) reviews some of the commonly used indicators of economic inequality, and finds that they rank countries very differently. It is therefore not unlikely that the indicator used for measurement, is determining for the results of a statistical analysis. Finding a valid indicator is therefore important.

One of the most well-known, and well-used, indicators of economic inequality is the Gini coefficient from Deininger and Squire (1996). It measures the equality in income distribution, on an index between 0 and 1, where 0 expresses perfect equality (every in-

dividual has exactly the same income), and 1 expresses maximal inequality (one person controls all income). I am reluctant to use the Gini coefficient, as it creates many problems when included in panel datasets. The most notable relates to poor coverage; the time-series are short, and many countries are left out of the sample. Coverage is especially poor on the African continent, where countries are disproportionately affected by civil conflict and detrimental systems of governance. However, perhaps the most serious weakness of the Gini coefficient is that the data used for its construction stems from various sources, which are difficult to compare; there are observations based on expenditure and income, on net and gross income, and on data from household and per capita surveys. Moreover, national household surveys are often based on different measurement techniques in different countries, and over time within the same country. According to Houle (2009, 598), the different data sources may significantly affect the inequality measure, and the results.

That there is little evidence linking inequality to democratization, or to civil conflict, from studies applying the Gini coefficient (see e.g. Collier and Hoeffler 2004; Fearon and Laitin 2003; Barro 1999), may therefore reflect inadequate measurement. The above concerns have made me search for another way of measuring economic inequality.

I use the variable *capital share*, which measures the value added in the industrial sector that accrues to the owners of capital, as opposed to the share that goes directly to the laborers, i.e. the wage share (Houle 2009; Ortega and Rodriguez 2006). *Capital share* is defined as 1 minus the wage shares, and measures the ratio of compensation of employees to the value added in production (Houle 2009, 602).[11] The variable is constructed from data collected by the United Nations Industrial Development Organization (UNIDO).[12] The closer the capital share is to one, the higher is the level of economic inequality expected to be, because owners of capital attain an increasingly larger proportion of the value added, at the expense of the laborers. Capital shares have previously been used to measure inequality by, amongst others, Przeworski et al. (2000), Acemoglu and Robinson (2006), and Dunning (2008). Houle (2009) uses this variable to explicitly test the theories of Boix (2003) and Acemoglu and Robinson (2006).

Using capital share as a proxy of inequality provides both empirical, as well as theoretical gains. First, it is probably the best cross-national indicator available in terms of coverage (Dunning 2008). Of the total possible number of observations in the Gini data

[11]The total value added comprises of value added in factor prices, producer prices, as well as a residual component of value added, which is undefined. The variable 'wages and salaries' is divided on total value added, to find the wage share of the value added in the industrial sector.

[12]I thank Carl Henrik Knutsen for lending me the UNIDO data. He constructed the wage share variable, which I converted to capital share.

from Deininger and Squire (1996), only 29% of the country-years between 1960 and 2005 are present. The UNIDO data, on the other hand, has originally 68% coverage between 1963 and 2008. Within my sample of countries the coverage is reduced to 47.9%.[13] Second, the UNIDO data is gathered by one source, and the same measurement method has been used in all countries. This safeguards against measurement error, and makes cross-country comparison more reliable (Houle 2009, 603). Third, the theoretical literature on democratization often focuses on intergroup inequality as a trigger of mobilization. Capital share captures intergroup inequality between the labor class, and the owners of capital. Houle (2009, 602) maintains that "capital shares measure the relative income of the elites." This focus is in accordance with Boix (2003) and Acemoglu and Robinson (2006). Using this variable, therefore, reduces the discrepancy between the theoretical concept and the empirical indicator, thus should increase measurement validity.

The variable *capital share* covers the years 1963 to 2008, and has 3533 observations, from 151 countries in the original dataset. As I wish to conduct the analysis on the years between 1960–2010, the timeseries has to be extended. In order to increase the number of observations I use the multiple imputation technique of Honaker and King (2010), as described in section 5.4.1. The sample is extended both in time, and across space, which is not uncontroversial. There are mainly two reasons why I have chosen to do so, although it is by no means a perfect solution. First, I have chosen to do so in an attempt to maximize the number of regime and conflict transitions in the sample. Removing the 31 countries without coverage, leads to a loss of 44 (22%) conflict onsets and 58 (18%) regime changes. Second, there is also a risk of introducing bias if their missingness is correlated with the dependent variable. There is 20% conflict country-years in the countries where capital share is missing completely, whereas the number is 15% in the countries with some or complete coverage. In the 'missing countries' autocracies are also disproportionately represented; 66% of the county-years are autocratic, whereas only 37% in the sample where these countries are excluded. If removed, the sample would be somewhat skewed somewhat in favor of more democratic and peaceful countries. An overview of the countries in question is provided in Appendix B, Table B.1. Descriptive statistics of capital share, before and after imputation, is found in Appendix A.[14] The

[13]The reason why I loose coverage is due to countries that are part of the UNIDO sample, but which are not coded with either conflict data, or regime type data.

[14]The minimum value on capital shares is -0.4. Although it is a *share*, and hence, should only have observations in the interval between 0 and 1, in practice, it is possible to obtain negative values. This happens for example during recessions, and in times of disinvestment, as the capital stock is reduced. Negative values only appear three times in the dataset; in Sierra Leone in 2005 and 2006, and in Syria

average capital share value is 0.66 in the imputed dataset, and somewhat higher – 0.71 – for the countries without coverage. Imputation diagnostics are performed below, in section 6.2.1.

Capital share is assumed exogenous in this analysis. Hence, the variable is not updated during the course of the simulation. Also, the level of inequality is assumed to be constant, at its 2010-level, from 2011–2050. I choose this strategy because I am not interested in the long-run *development* in economic inequality, but rather the long-run effects of having a certain level of inequality. This equates to 'stretching' the effect of the multinomial logit model beyond 2010. Doing this, I am able to isolate its impact on regime changes and conflict over time.

I operate with the following *'baseline scenario'* for the simulation period 2011–2050: I estimate the mean value of capital share in each country between 2006–2010, and use this value as a proxy for the level of inequality in 2010. The reason why I use this aggregate, and not the value for 2010, is that the point estimates vary slightly on a year-to-year basis. The capital share variable seems to be somewhat sensitive to changes in economic performance. However, inspecting the time-series for individual countries reveals that the trend in their capital share value is more or less constant. I expect to better capture this 'underlying trend' with a mean value, rather than a single point estimate. This volatility is noticeable both in individual countries, and in the sample as a whole.[15] I consider an average over the last five years to be sufficient to even out such busts and booms, while still maintaining the underlying inequality-level in 2010 intact. For an overview of the 20 most equal and unequal countries on the capital share variable, consult Appendix B.

6.2.1 Imputation diagnostics

According to Abayomi, Gelman and Levy (2008, 274) there has been a general belief that the assumptions of imputation models cannot be tested against the observed data, and therefore, that the quality of the imputed data cannot be checked. This has made many researchers sceptical towards using imputation techniques. Abayomi and his colleagues respond by suggesting a range of diagnostics techniques.

In order to perform unbiased multiple imputation one has to assume that the missing observations are MAR, i.e. that they are missing at random. The missingness has to be uncorrelated with the dependent variable, since the model conditions on the observed

in 2000. Capital shares jump back up almost immediately afterwards.

[15]For example, in 2008 and 2009 there was a relatively large decline in the mean level of capital share in the world (-1.52% and -1.33%, respectively), which was equalized by an increase in 2010.

data when drawing values for the missing observations. This assumption can be tested with a Kolmogorov–Smirnov test, as advised by Abayomi, Gelman and Levy (2008, 280). The KS-test compares the observed and imputed density distributions, and whether there are statistically significant differences between them.

A KS-test for capital share confirm that there is a statistical difference, at the 1% level, between the distributions with and without imputed data. This raises a first red flag, but it may not present a problem, as the imputed distribution can deviate from the observed distribution and still be missing at random (Abayomi, Gelman and Levy 2008, 280). There are reasons to be cautious, though, and I therefore continue with some additional diagnostics.

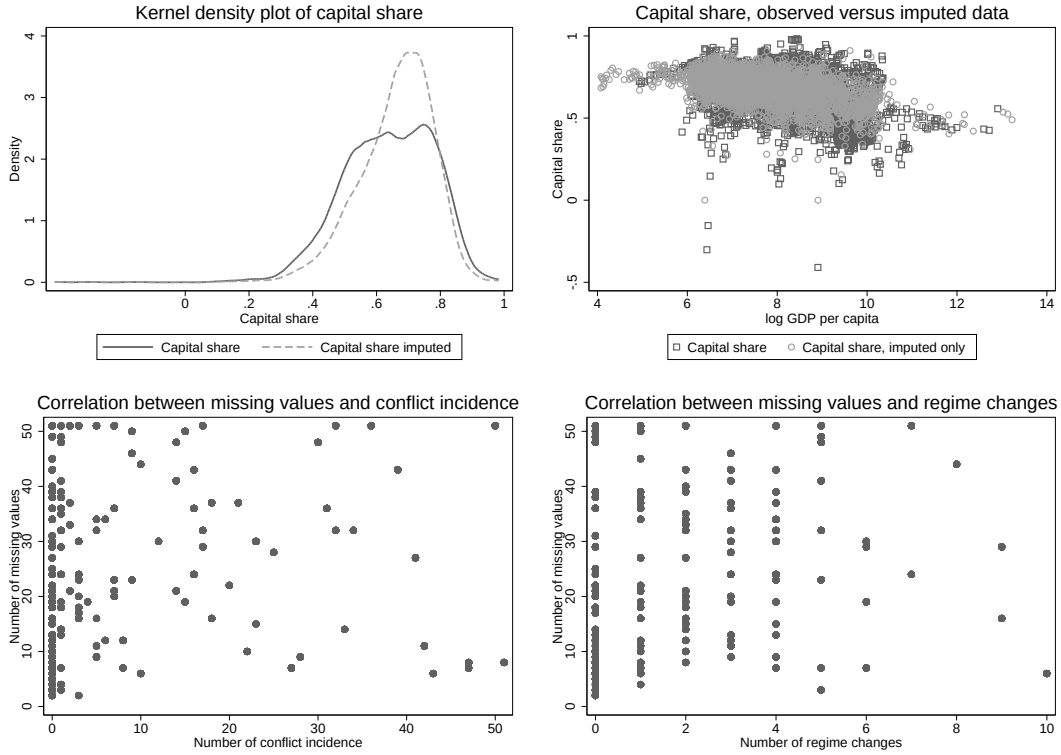
Figure 6.3 reports various diagnostics plots, as suggested by (Abayomi, Gelman and Levy 2008). The panel at the upper–left shows a kernel density plot of the distribution of capital share without any imputed observation, i.e. the solid line, and with imputed data, i.e. the dashed line. The difference between observed and imputed values appears to be driven by overprediction at the mean level, resulting in lower dispersion. This is not unexpected, as I set a high ridge prior because of extensive missingness. The imputed observations are then driven towards the mean, as highlighted in section 5.4.1. It is difficult to state the exact consequences on the results. However, if there is an underlying effect from capital share on regime changes and/or conflict incidence, skewing the sample towards the mean *may* underestimate the effects in the analysis.

In the upper-right panel in Figure 6.3 observed and imputed values on capital share is plotted against ln GDP per capita. The plot supports the previous finding, as the imputed observations, the dots, are located closer together, whereas the observed data shows greater dispersion. There is seemingly not a strong correlation between income level and missingness though. The two panels at the bottom of the figure plot the number of missing values on capital share, against the sum of conflict incidence (left-hand panel), and regime changes (right-hand panel) in each country between 1960–2010. Experiencing more conflict incidence or regime change do not seem to be highly correlated with having a high number of missing values. The MAR assumption therefore seems to hold up quite well, and it should not be problematic to use multiple imputation.

Appendix B reports some scatterplots of observed and imputed values from a select group of countries.[16] The first set of countries, reported in Figure B.1, have some missingness on capital share. The scatterplots show that the imputed values are in line with

[16]The countries are selected more or less randomly, but they have differing levels of development, inequality level, and regime and conflict history. They scatterplots are only meant serve to illustrate how the imputation program have assigned values to the originally missing observations.

Figure 6.3: Imputation diagnostics, capital share



the trend in the observed data. In order to further validate the results from the imputations for those countries which are exclusively assigned imputed, I use an approach that entails comparison of these countries with a select group of 'model countries', which are considered reasonably similar.[17]. Figure B.2 report a few examples of countries I have compared.

Overall, the imputed data is not expected to bias the results, but I still need to be cautious when analyzing the data.

6.3 Regime and conflict history

A regime change is more likely to take place in close proximity to a previous regime change, and evidence suggest that the probability of regime change declines as a regime ages (Clague et al. 2003; Strand et al. 2012). Countries that have experienced conflict in the past, also have a higher risk of returning to conflict (Collier, Hoeffler and Söderbom

[17]By 'reasonably similar' I mean that the model countries have similar values on some of the core variables; GDP per capita, population, regime and conflict history. This is by no means an infallible test of the results, but it is an easy and available technique. The same approach was used by Hegre et al. (forthcoming, 12) to manually impute values for countries without coverage on an explanatory variable.

2008; Fearon and Laitin 2003; Hegre et al. forthcoming), the so-called 'post-conflict risk'.

As explained in chapter 5, the 'dynamic' multinomial logit model is characterized by its inclusion of past values of the dependent variable, on the right-hand side of the equation. They are included as dummy variables for each category of the dependent variable (except for the reference category), at time t-1: 'inconsistency and peace' (IP), 'democracy and peace' (DP), 'autocracy and conflict' (AP), 'inconsistency and conflict' (IC) and 'democracy and conflict' (DP). By including these variables I expect that the 'state' of a country in the previous year, explains a large part of its character this year.

In order to capture the effect of conflict- and regime history further back in time than t-1, I include six variables that record the natural logarithm of the number of years in each of the six states up to t-2. This is in line with Hegre et al. (forthcoming), and Hegre et al. (2011).[18] The variables are jointly referred to as 'history' variables.

6.4 Control variables

In order to reduce the risk of omitted variable bias I include a number of control variables that are widely used the literature on democratization, as well as in the civil conflict literature.[19] The first two, economic development and population, are part of the baseline model. The rest of the variables serve as controls in the more extensive models. In the forecasting period, from 2011–2050, all control variables are assumed exogenous and kept constant at their 2010-level, in line with the argument in section 6.2.

6.4.1 Economic development

The most common proxy of economic development is income level, measured by GDP per capita. GDP per capita is one of the most robust variables within the civil war literature (Hegre and Sambanis 2006). Income can both be related to conflict onset (Collier and Hoeffler 2004; Fearon and Laitin 2003), as well as conflict duration (Collier, Hoeffler and Söderbom 2004). Income has been included in models of civil conflict as a proxy for the economic opportunity cost of war (Collier and Hoeffler 2004), and for some aspect of

[18] I have constructed the variables based on the definition of 'conflict history' in Hegre et al. (forthcoming, 2011). The time in each status starts counting in 1946, which is the first year with available conflict data.

[19] Omitted variable bias occurs when variables not included in the analysis are correlated with both the dependent and the independent variables, and in that way distorting the estimates of the regression analysis. Whether the effects are under- or overestimated depends on the strength and direction of the correlation (Skog 2009, 253).

state capacity (Fearon and Laitin 2003). In contrast to findings by conflict researchers, the relationship between regime changes, and in particular democratization, and the level of income, is highly contested (see e.g. Przeworski and Limongi 1997; Diamond 1992; Przeworski et al. 2000; Boix and Stokes 2003).

I include the variable *Ln of GDP per capita*, which is the natural logarithm of GDP per capita. The variable is originally from Strand et al. (2012), and is constructed by combining data from various sources to maximize available information. They use Maddison's (2006) GDP data, which are measured in 1990 International Geary-Khamis dollars.[20] It is supplemented with data from the World Bank (2010) and Penn World Table (2011). The variable originally has coverage from 1816–2008, and I have extrapolated values for an additional two years.

6.4.2 Population size

Smaller countries are statistically more likely to be democratic than are countries with larger populations (Diamond 2002). Nevertheless, the empirical results on the relationship between population size and democratization/democratic stability is unclear (Hegre, Knutsen and Rød 2012, 5). Population increases has also been found to increase the risk of conflict onset (Hegre and Sambanis 2006; Fearon and Laitin 2003; Collier and Hoeffler 2004), but it has, on the other hand, not been proven to lengthen a conflict in time (Collier, Hoeffler and Söderbom 2004).

In order to account for the possible effect of population size I include the variable *ln of population*, which is the natural logarithm of a country's total population, measured in 1000s. I take the natural logarithm of the observations, both to ensure that extreme outliers do not influence the results disproportionately, but also to account for the declining marginal effect from population on the risk of conflict (Collier and Hoeffler 2004). The population variable is collected from Hegre et al. (forthcoming), and is based on population data from the World Population Prospect (United Nations 2007).

6.4.3 Economic growth

Not only has the risk of civil conflict and democratization been related to the level of economic development, it has also been related to economic growth. Several studies have found that economic growth reduces the risk of civil war (Collier and Hoeffler 2004;

[20] Geary Khamis dollars are also referred to as 'international dollars', or 1990 US$. In the following I use the term US$, but it always refers back to the 1990 US$.

Hegre and Sambanis 2006). Economic crises are also found to increase the risk of regime breakdown (Przeworski and Limongi 1997), and the positive effect of GDP growth on the probability of democratization is documented by (Hegre, Knutsen and Rød 2012). I include a variable *GDP growth*, which measures the annual increase in ln GDP per capita. It is based on the GDP per capita indicator from Strand et al. (2012).

6.4.4 Oil and gas exporter

The relationship between primary commodity dependence and armed conflict is well-examined (Ross 2004, 2006; Fearon and Laitin 2003; Hegre and Sambanis 2006). Fearon and Laitin (2003, 81) argue that states dependent on oil revenues tend to have weaker state capacity and poor bureaucratic quality, given their level of development. Primary commodities is also thought to increase the opportunities for financing rebellion, thus increasing the risk of conflict (Collier and Hoeffler 2004).

Natural resources are also proposed to affect the probability of democratization negatively. The 'taxation effect' implies that autocratic governments who possess a lot of natural resource wealth, impose little taxation, thereby increasing the acceptance of authoritarian rule, and decreasing the demand for democracy (Ross 2001, 332). Additionally, the 'spending effect' from pleasing patrons, reduces the demand for democracy from other elite groups (Ibid.).

The variable *oil and gas* is a dummy variable telling whether a country receives more than one third of its export revenue from oil or gas. The variable is originally from Fearon and Laitin (2003), who apply World Bank data coded in 5 years intervals for the period 1960 to 2005. I use a version of the variable from Hegre et al. (forthcoming) with interpolation between the five year intervals.

6.4.5 Education

Modernization theory accentuates the importance of education for democratization, and for regime stability (Lipset 1959), and it is supported by empirical studies (e.g. Barro 1999). Educational attainment is also thought to increase the opportunity cost of rebellion (Collier 2000). Many studies have found a strong negative effect of education on the probability of civil conflict (Collier 2000; Collier and Hoeffler 2004; Collier, Hoeffler and Söderbom 2004).

I use education data from (Hegre et al. forthcoming), originally collected by Lutz and Sanderson (2007). The data is based on individual-level educational attainment data from

Demographic Health Surveys (DHS), Labour Force Surveys (LFS), and national censuses, and covers the period 1970–2000. I follow Hegre et al. (forthcoming, 11), who apply a measure of male secondary education, defined as the proportion of males aged 20-24 years with secondary or higher education, of all males within the same age group.[21] As the variable have 944 missing observations prior to 1970, I impute data using Amelia II, as described in section 5.4.1.

6.4.6 Cultural determinants: Ethnicity and religion

Ethnicity and religion are emphasized as sources of political instability and violence (Huntington 1996; Horowitz 1985; Easterly and Levine 1997). I test two different indicators for ethnicity, and one for religion: The variable *ethnic dominance* is originally from Collier and Hoeffler (2004), and measures whether there are dominant ethnic groups residing within a country's territory. It takes on the value 1, if one single ethno-linguistic group makes up 45 to 90% of the total population, and 0 otherwise. I also test an alternative measure of ethnicity, *ethnic fractionalization* (ELF). Ethnic fractionlization has been included together with its square term (e.g. Hegre and Sambanis 2006), capturing that in a country with two dominant groups, the risk of conflict is greatest. In the same manner I include *religious fractionalization*. The fractionalization variables report the probability of two randomly drawn individuals in a country come from different ethnolinguistic/religious groups (Fearon and Laitin 2003, 78).[22] The three variables are collected from Hegre et al. (forthcoming). As the fractionalization variables have missing observations, I impute data using Amelia II, as described in section 5.4.1.

6.4.7 Proximity to independence

Fearon and Laitin (2003, 81) hypothesize that "a newly independent state, which suddenly loses the coercive backing of the former imperial power and whose military capabilities are new and untested" will stand a weaker chance of facing up to potential insurgents, thus increasing the probability of internal armed conflict. I include a variable that measures the proximity to independence, adopted from Strand et al. (2012). This variable is a

[21] The original sample only includes 120 countries. In order to expand coverage, Hegre et al. (forthcoming, 12) imputes values based on male secondary enrollment data from Barro (2000). The remaining 27 countries that were still left with missing observations were imputed manually by using model countries they found to be reasonably similar in terms of educational profile and geographic location. For a detailed description of the procedure and the modeling countries, see Hegre et al. (forthcoming, 12).

[22] See Fearon and Laitin (2003, 78) for an extensive description of the two fractionalization measures.

decay function and measures the time since a country gained independence. The half-life parameter is set to 8 years.

6.5 The dataset

The analysis on observed data is conducted for the time-period 1960–2010. Although the dataset from Strand et al. (2012) extends back to the 1800s, and the conflict data covers 1946–2010, the time-span is restricted due to data availability on the explanatory variables. When deciding which time-period to use for the analysis, I have tried to strike a balance between maximizing the total number of observations (regime and conflict transitions), and minimizing the uncertainty and potential bias caused by extensive imputation. The complete data set includes 7363 observations between 1960 and 2010. It covers 164 countries altogether, defined by the the Gleditsch and Ward independent state system (Gleditsch and Ward 1999). For the forecasting period the number of countries is 160. Descriptive statistics of all variables can be found in Appendix A. In the empirical analysis all independent variables are lagged by one year, in order to reduce the threat of reverse causality (Hegre and Sambanis 2006, 514)

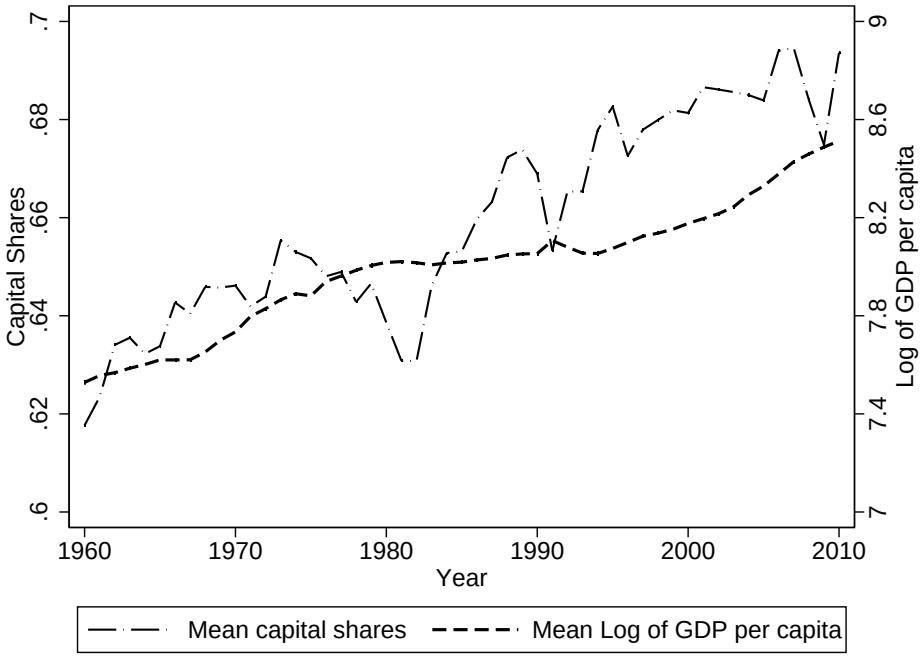
Chapter 7

Empirical analysis

The purpose of the following analysis is to demonstrate that the level of economic inequality is important for how a democratization process develops, its outcome, and its overall amount of civil conflict. What differentiates this study from previous work, is the way in which I measure the *total* effects of democratization on civil conflict. I achieve this by, firstly, analyzing the multinomial logit model where both regime changes and civil conflict are studied in conjunction, thus capturing the reciprocal causality between the two events. Secondly, by measuring the long-run effects over a time period of 40 years. This is done through the use of simulations, as explained in section 5.3. The simulation procedure allows me to aggregate both the direct and indirect effects of inequality, and in addition, taking onto account the influence of previous regime and conflict history by updating the transition probabilities each year over the course of the simulations. I try to identify the "ideal type processes", presented in Table 4.1, by testing the hypotheses from section 5.5.

The chapter proceeds as following: Before undertaking the statistical analysis, I scrutinize the relationship between capital share, the indicator measuring economic inequality, and GDP per capita. There are numerous possible model specifications to base the simulations on. I therefore conduct an out-of-sample evaluation, and choose the model that demonstrates the best predictive ability. Next, I start off with the short-run analysis, and estimate the multinomial logit model of regime and conflict 'states', presented in chapter 5. The main part of the analysis deals with the long-run effects of inequality on democratization and civil conflict. Initially, I examine the long-run transition-matrix based on predictions from the multinomial model. Thereafter, I present the simulation results. The first part summarizes the global simulation results, and in the second part I conduct a number of "experiments", in order to isolate the effect of economic inequality. The last

Figure 7.1: Mean capital share and mean log GDP per capita, 1960–2010



part of the chapter is devoted to a discussion of uncertainty; uncertainty of predictions, uncertainty from imputation, and the validity of capital share, as a measure of economic inequality.

7.1 Economic inequality and economic development

During the work with this thesis I have become aware of a seemingly intimate relationship between GDP per capita and capital share, my proxy for economic inequality. This section therefore provides some descriptive statistics to shed light on the relationship.

As Figure 7.1 reveals, there is a clear historical trend where the rise in mean capital shares and the rise in GDP follow each other.[1] Although two congruent trends do not automatically imply a causal relationship, the fact that capital share is calculated by the value added in the industrial sector does not speak for spuriousness.

The scatterplots in Figures 7.2 and 7.3 strengthen this notion. Figure 7.2 reports the distribution of values on the capital share variable on the y-axis, against the level of GDP per capita on the x-axis, in 2005, and the country names denote each observation. In

[1] A marked difference between the two variables is that the solid line, representing capital share, shows much more volatility than the dashed line, representing GDP per capita, which is a lot smoother. The volatility of capital share was discussed in section 6.2.

Figure 7.2: The relationship between capital share and GDP per capita in 2005

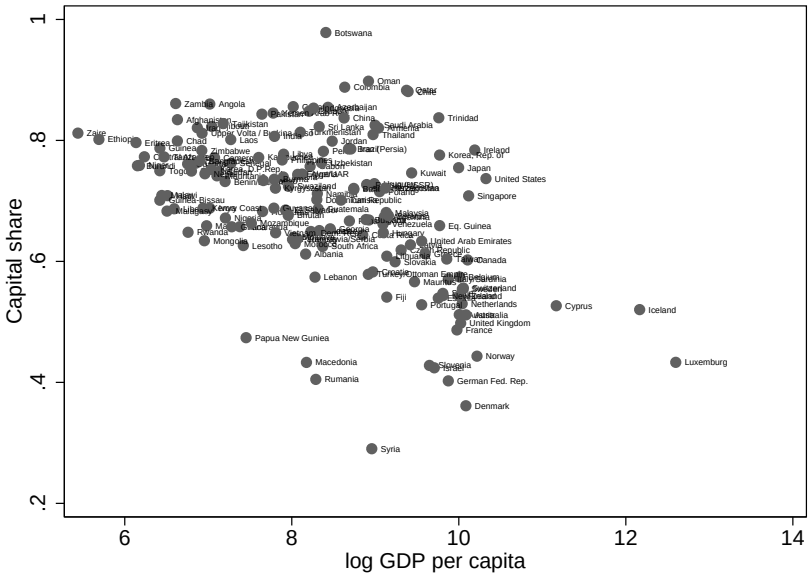
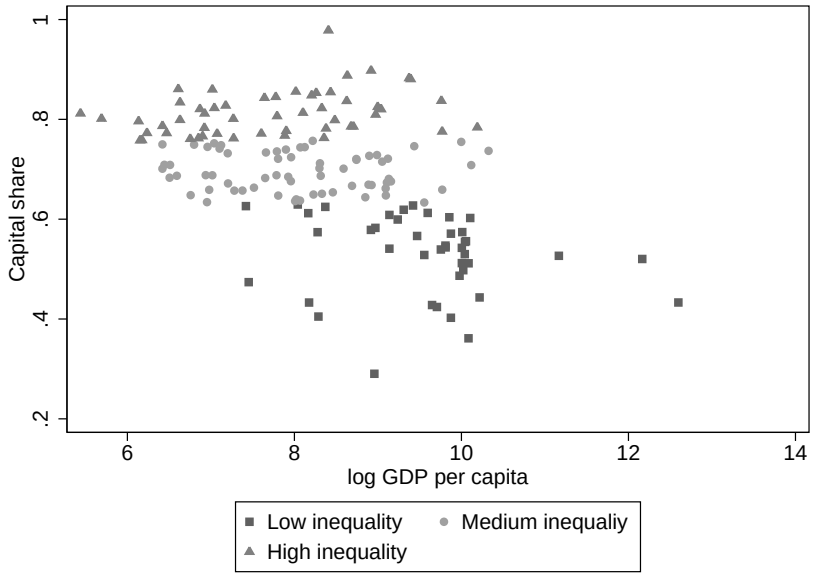


Figure 7.3 the observations are divided according to their level of inequality, based on a threefold categorization scheme. The high inequality group are countries with capital share value above the 75th percentile on the capital share variable, the low inequality group consist of countries with values below the 25th percentile, and a middle category accounts for all countries between the 25th and 75th percentile. The cutoff points are equal to 0.631, for the lower 25th percentile, and 0.758 for the 75th percentile. The squares represent the low inequality countries, the circles denote countries with an intermediate level of inequality, and the triangles denote the high inequality countries. The trend is very clear; there is a strong connection between having high(er) inequality, and having low(er) GDP per capita. The three groups are almost perfectly linearly fitted, where high, medium and low GDP per capita corresponds to high, medium and low levels of inequality. The negative correlation between capital share and log GDP per capita is -0.466, and the correlation was found to be statistically significant by Ortega and Rodriguez (2006). Hence, in this dataset, countries with high inequality tend to be poor countries.[2] However, whether it is low income that causes countries to have an unequal distribution of the wealth, or whether low income is a consequence of economic inequality, is undecided. The above information is a call for caution, and I try to keep in mind this correlation in

[2]The average level of economic development within these groups should also be of interest. The high inequality countries have a mean GDP per capita of 2230 US$ (1990), for the intermediate countries mean income per capita is 4105 US$, and the low inequality countries have a GDP per capita of 16,317 US$. The GDP values are from 2010.

Figure 7.3: The relationship between capital share and GDP per capita in 2005. Countries divided into low-, medium-, and high capital share



the following analysis.

7.2 Choosing models

I operate with a hierarchical system of statistical models, where the baseline model is the building block. The baseline model includes capital share, GDP per capita, population, dummy variables for the regime and conflict categories at t-1, as well as the regime and conflict history variables. The choice of GDP per capita and population for the baseline model is in accordance with Hegre and Sambanis (2006) and Hegre, Knutsen and Rød (2012).

As the dependent variable has six categories, there are five equations in the multinomial logit model. When the control variables are included in the model the number of parameters to be estimated becomes very large. To guard against empty cells in the transitions matrix in Table 5.1, and extreme multicollinearity, I make use of inequality restrictions on the parameter estimates. The restrictions can help simplify the models. Usually, prior information about the regression parameters forms the basis for inequality restrictions (Greene 1997, 411).[3]

[3]I operate with the following guidelines: 1) Are some of the estimated beta-coefficients similar in strength and direction? 2) Are there extreme parameter-estimates that calls for constraining them to a maximum- or minimum-level? 3) Are there weak- or non-existing estimates which justify that the

Table 7.1: Overview of Model Candidates

Model	Variables
0	Baseline model without capital share
1	Baseline model: capital share, GDP per capita, population, regime and conflict history, dummies for state at t-1
2	Baseline model with inequality restrictions on capital share
3	Full model including all control variables
4	Full model; ethnic dominance replaced with ethnic fractionalization
5	Baseline model including education and oil and gas exporter
6	Baseline model including interaction between capital share and oil and gas exporter

Because of time and space constraints I decided to only focus on one model specification in the final analysis. In order to chose the best model I specified model candidates, which are presented in Table 7.1. The table reports the name of the model, and the variables that are included. To select the best-performing model I make use of out-of-sample prediction assessment. With a split-sample design I estimate all of the candidate models on the time period between 1960–2000, and thereafter obtain predictions from the simulation program for the period between 2001 and 2010. Thereafter I compare the predictions with the observed data for the same period period.[4] This is in line with (Hegre et al. forthcoming). I ran 500 simulations of each model in the selection stage.

A technique for quantifying a model's predictive power is performing a 'ROC analysis' (Ward, Greenhill and Bakke 2010). ROC stands for Receiver Operator Curve, and the ROC analysis is used to identify how many observations a model predicts correctly, i.e the 'true positives', and how many times it misses, i.e. the 'false positives'. The trade-off between predicting true positives and avoiding false positives is illustrated by the area under the ROC, called AUC (Area Under Curve). The ROC is often depicted in a diagram, where the true positive rate is plotted on the y-axis, against the false positive rate on the x-axis.[5] The AUC is equal to "the probability that the simulation predicts a randomly chosen positive observed instance as more probable than a randomly chosen negative one" (Hegre et al. forthcoming, 16). Figure 7.4 illustrates the ROC for model m1. A very poor model, not predicting any better than chance, would have an AUC of 0.5, and generate one new false positive for every new state it is able to accurately identify. This is represented by the diagonal line. The perfectly predicting model would have an AUC of 1.0.

parameter be set to zero? I expect the guidelines to be satisfactory for the purpose of this analysis, as they guard against the problems referred to above.

[4]See Appendix D for the multinomial logit estimates that form the basis for the out-of-sample evaluation.

[5]See Hosmer and Lemeshow (2000, 156–164) for an introduction to Receiver Operator Curves.

Table 7.2: AUCs for models m0–m6 estimated on data for 1960–2000. Predictions compared to observed regime and conflict status 2001–2010. 90% Confidence interval in parentheses.

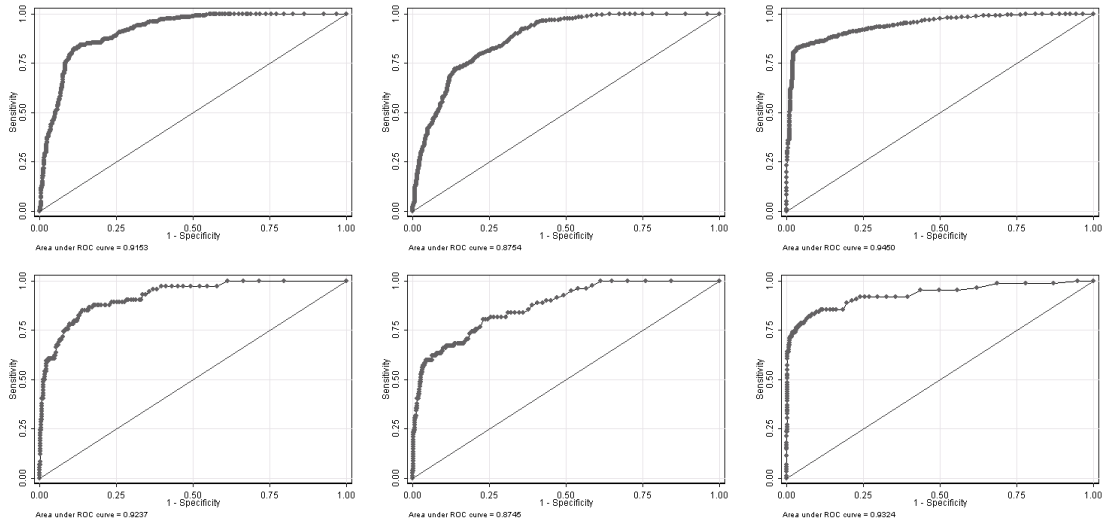
Model	AP	IP	DP	AC	IC	DC
0	0.9170 (0.9010, 0.9331)	0.8733 (0.8557, 0.8910)	0.9441 (0.9327,0.9554)	0.9125 (0.8797, 0.9453)	0.8747 (0.8359, 0.9135)	0.9203 (0.8818, 0.9589)
1	0.9153 (0.8991, 0.9316)	0.8754 (0.8581, 0.8927)	0.9450 (0.9340, 0.956)	0.9237 (0.8934, 0.9541)	0.8745 (0.8368, 0.9132)	0.9323 (0.8976, 0.9671)
2	0.9159 (0.8998, 0.9321)	0.8744 (0.8568, 0.8920)	0.9469 (0.9362, 0.9576)	0.9126 (0.8811, 0.9441)	0.8723 (0.8319, 0.9127)	0.9220 (0.8852, 0.9589)
3	0.6193 (0.5850, 0.6535)	0.5201 (0.4920, 0.5500)	0.7058 (0.6835, 0.7282)	0.6055 (0.5483, 0.6627)	0.6625 (0.6157, 0.7092)	0.4173 (0.3938, 0.4409)
4	0.5886 (0.5595, 0.6176)	0.5236 (0.5010, 0.5462)	0.6677 (0.6476, 0.6879)	0.5823 (0.5336, 0.6310)	0.6426 (0.6014, 0.6838)	0.4732 (0.4467, 0.4998)
5	0.9055 (0.8876, 0.9234)	0.8640 (0.8456, 0.8825)	0.9463 (0.9355, 0.9571)	0.9138 (0.8847, 0.9429)	0.8793 (0.8434, 0.9152)	0.9325 (0.9039, 0.9611)
6	0.7276 (0.6925, 0.7628)	0.7042 (0.6762, 0.7322)	0.7729 (0.7504, 0.7954)	0.7565 (0.7004, 0.8126)	0.8000 (0.7536, 0.8465)	0.8215 (0.7723, 0.8707)

In order to evaluate the predictions from the out-of-sample test I construct dichotomous variables for each category of the dependent variable, as well as for each regime category and for conflict status. Doing this, I am able to test how well the models performs in comparison to each other, by focusing on their ability to identify the correct regime and conflict state. The results from the out-of-sample evaluation are reported in Table 7.2. Table 7.2 reports the AUCs, together with a 90% confidence interval, for the *incidence* of the different combinations of regime and conflict status.[6]

I first focus on model m1, the baseline model including the capital share variable, and contrasts it with the more extensive models, m3–m6. Model m0 without capital share, and model m2, with constraints on capital share, are discussed in sections 7.2.1 and 7.3. As shown in Table 7.2, model m1 has better predictive power than the more extensive models. The baseline model does a fairly good job identifying the combinations of regime and conflict. It is best able to categorize democracies in peace, whereas the poorest job is done categorizing the inconsistent regimes in conflict. Model m1 correctly identifies 94.5% of all democratic and peaceful country-years, but only 87.45% of all country-years

[6]See Appendix D for the do-file used to conduct the ROC analysis.

Figure 7.4: ROC graphs 2001–2010, Model 1. Upper left: AP. Upper middle: IP. Upper right: DP. Lower left: AC. Lower middle: IC. Lower right: DC.



with inconsistent institutions and conflict.[7].

Because of the complexity of the dependent variable, including too many controls drastically reduces the AUC. The most extensive models, m3 and m4, predicts very poorly, and in some instances not any better than chance(!) The only model that is close to the performance of the baseline model is model m5, which includes the two strongest predicting control variables from model m3 and m4, *education* and *oil and gas exporter*. Model m5 performs almost as good as, and for a couple of categories even better, than the baseline model. But, as the confidence intervals are overlapping, and the improvement in log likelihood when performing a log-likelihood test is insignificant, I prefer to continue the analysis with the more parsimonious model, m1.

As seen in Appendix D, some control variables are significantly related to the dependent variable, which is a call for caution. However, including the controls for the most part only cause minor changes in the coefficients for the baseline predictors. For capital share the estimate in the equation for IP changes slightly. The change is equivalent to approximately 1 standard error in the baseline model. As I am interested in how all five estimates for capital share affect regime changes and conflict incidence I do not suspect it to pose a problem, as the overall trend in the models is consistent. Leaving out these controls, therefore, is not expected to alter the simulation results fundamentally.

[7]The percentages imply that if one randomly observed incidence of democracy is drawn from the sample, the probability that this unit will have a higher predicted probability of being a democracy in peace, than a randomly drawn country-year which is not an observed democracy in peace, is 94.5%

7.2.1 The value-added of economic inequality

What is the value added of moving beyond economic development, and focusing on the effect of economic inequality? The out-of-sample evaluation in figure 7.2 reports results from model m0, which is a baseline model with the capital share variable removed. The results are somewhat discouraging, as there is no large improvement in model m1's predictive abilities, and their confidence intervals overlapping. However, log likelihood increases from -2709 to -2704 (see Tables 7.3 and 7.4) between models m0 and m1, and a log likelihood test shows a significant improvement in log likelihood in model m1.

In order to further evaluate the difference in predictive power between m0 and m1, I identify all predicted regime changes and conflict onsets. From the out-of-sample prediction I combine the three categories with conflict – AC, IC, and DC – and the three categories for no conflict – AP, IP and DP. Moreover, I combine all predictions that reported autocracy (AP and AC), inconsistency (IP and IC), and democracy (DP and DC). Interestingly, when comparing observed and predicted conflict onsets and regime changes (results are not reported), model m0 and m1 fail to classify the exact same cases, using a threshold of 0.5 as the classification criterion. Both models predict conflict onset quite poorly, and democratic transitions somewhat better. The AUC for model m0 is 0.735 for democratic regime changes and 0.598 for conflict onset, whereas for model m1 the numbers are 0.708 for democracy and 0.587 for conflict onset.[8]

I also inspect some specific historical cases where both models were unable to identify the onsets and regime changes correctly: Haiti experienced a conflict onset in 2004. Haiti's predicted risk of conflict is 0.254 in model m1, and 0.196 in model m0. Turning to democratic regime changes, model m0 reports a higher risk of democratic regime change in Paraguay in 2003 (0.108 in m0 versus 0.076 in m1), whereas the risk of a democratic transition in Albania in 2005 was predicted to be 0.136 in m1 and only 0.112 in m0. Countries with higher inequality are assigned a somewhat larger probability of conflict, and lower probability of a democratic transition in model m1. However, at the same time it provides more inaccurate predictions in those cases where countries experienced conflicts or democratic transitions contrary to their level of inequality. Thus, if accepting the theoretical argument about the effect of inequality, there is seemingly a value-added of including capital share in the model.

As the aspect of *time* is important in this thesis, the hypotheses from section 5.5 are

[8]The reason why these numbers differ, although I said they classified countries identically, is that the ROC analysis considers all possible thresholds between 0 and 1, whereas I only inspected the p=0.5 threshold.

divided into short-run and long-run effects. The short-run effects of inequality relate to the estimation of and results from the multinomial logit model, which I present in the next section.

7.3 The short-run effects

Table 7.3: Multinomial Logit Analysis 1960–2010: model m0, baseline with capital share removed

	IP		DP		AC		IC		DC	
Inconsistent&Peace t–1	5.894***	(0.216)	4.418***	(0.296)	0.885+	(0.480)	6.164***	(0.492)	3.921***	(1.014)
Democracy&Peace t–1	3.752***	(0.368)	8.068***	(0.381)	1.879**	(0.648)	4.751***	(0.782)	8.149***	(0.784)
Autocracy&Conflict t–1	0.430	(0.526)	-10	(.)	4.001***	(0.203)	4.847***	(0.505)	4.113***	(0.877)
Inconsistent&Conflict t–1	5.627***	(0.447)	4.025***	(0.742)	3.846***	(0.512)	9.245***	(0.629)	7.068***	(0.917)
Democracy&Conflict t–1	5.002***	(1.102)	8.084***	(1.039)	4.830***	(1.103)	6.719***	(1.306)	11.17***	(1.242)
ln(time in AP)	0	(.)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in IP)	0.312***	(0.0917)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.693***	(0.108)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.576***	(0.114)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.452**	(0.166)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	0.938***	(0.161)
ln(GDP per capita)	0.120	(0.0795)	0.591***	(0.110)	-0.377***	(0.0869)	-0.0764	(0.115)	0.397**	(0.145)
ln(Population)	0.0268	(0.0512)	0.178**	(0.0666)	0.216***	(0.0512)	0.367***	(0.0768)	0.498***	(0.0905)
Constant	-4.570***	(0.769)	-10.90***	(1.103)	-2.432**	(0.750)	-9.002***	(1.197)	-14.89***	(1.608)
N	7291									
ll	-2709.3									

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The analysis is conducted on a panel dataset consisting of 164 countries between 1960–2010. All the independent variables are lagged by one year. I compare the estimates from model m1 (Table 7.4), with model m0 without capital share (Table 7.3), and model m2, where capital share is constrained (Table 7.5).

Table 7.4 reports the results from a multinomial logit analysis with 'regime type and conflict status' as the dependent variable. Model m1 includes capital share, log of GDP per Capita and log of Population, in addition to the lagged regime and conflict variables and the history variables. 'Autocracy & Peace' is the base outcome, and the estimates are reported in log odds, with standard errors in parentheses.[9] Equation IP contains the

[9]The coefficient for AC in equation DP is restricted to -10 because it originally had an extreme value. It indicates that the likelihood that what was an autocracy in conflict last year transitions into a democracy in peace this year is close to zero. The history variables are restricted to zero in all equations except the one relevant for the particular regime and conflict state. E.g. equation IP only includes ltsip.

Table 7.4: Multinomial Logit Analysis 1960–2010: model m1, baseline model

	IP		DP		AC		IC		DC	
Capital Share	0.971	(0.726)	-0.956	(0.919)	1.578^{+}	(0.865)	1.554	(0.981)	0.258	(1.206)
Inconsistent&Peace t–1	5.914***	(0.217)	4.413***	(0.297)	0.904^{+}	(0.480)	6.179***	(0.492)	3.928***	(1.015)
Democracy&Peace t–1	3.784***	(0.369)	8.075***	(0.382)	1.913**	(0.648)	4.784***	(0.783)	8.154***	(0.785)
Autocracy&Conflict t–1	0.422	(0.526)	-10	(.)	3.992***	(0.203)	4.833***	(0.505)	4.109***	(0.877)
Inconsistent&Conflict t–1	5.635***	(0.447)	4.000***	(0.742)	3.850***	(0.513)	9.252***	(0.629)	7.072***	(0.918)
Democracy&Conflict t–1	5.015***	(1.102)	8.068***	(1.039)	4.838***	(1.103)	6.736***	(1.306)	11.16***	(1.242)
ln(time in AP)	0	(.)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in IP)	0.308***	(0.0919)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.669***	(0.109)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.567***	(0.115)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.450**	(0.167)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	0.946***	(0.161)
ln(GDP per capita)	0.137^{+}	(0.0805)	0.576***	(0.111)	-0.337***	(0.0894)	-0.0419	(0.117)	0.412**	(0.149)
ln(Population)	0.0159	(0.0518)	0.195**	(0.0680)	0.201***	(0.0516)	0.348***	(0.0780)	0.495***	(0.0924)
Constant	-5.274***	(0.933)	-10.28***	(1.279)	-3.698***	(1.025)	-10.17***	(1.408)	-15.16***	(1.861)
N	7291									
ll	-2704.3									

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7.5: Multinomial Logit Analysis 1960–2010: model m2, restrictions on capital share

	IP		DP		AC		IC		DC	
Capital Share	1.108**	(0.355)	-1.108**	(0.355)	1.108**	(0.355)	1.108**	(0.355)	0	(.)
Inconsistent&Peace t–1	5.919***	(0.217)	4.409***	(0.297)	0.902^{+}	(0.480)	6.179***	(0.492)	3.927***	(1.015)
Democracy&Peace t–1	3.787***	(0.368)	8.072***	(0.382)	1.902**	(0.648)	4.775***	(0.782)	8.149***	(0.784)
Autocracy&Conflict t–1	0.419	(0.526)	-10	(.)	3.992***	(0.203)	4.834***	(0.505)	4.111***	(0.877)
Inconsistent&Conflict t–1	5.631***	(0.447)	4.004***	(0.742)	3.847***	(0.513)	9.247***	(0.629)	7.070***	(0.918)
Democracy&Conflict t–1	5.015***	(1.102)	8.069***	(1.039)	4.834***	(1.103)	6.731***	(1.306)	11.16***	(1.242)
ln(time in AP)	0	(.)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in IP)	0.307***	(0.0920)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.669***	(0.108)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.570***	(0.114)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.451**	(0.167)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	0.945***	(0.160)
ln(GDP per capita)	0.138^{+}	(0.0797)	0.573***	(0.110)	-0.349***	(0.0871)	-0.0534	(0.115)	0.405**	(0.146)
ln(population)	0.0134	(0.0513)	0.197**	(0.0674)	0.206***	(0.0511)	0.355***	(0.0769)	0.499***	(0.0910)
Constant	-5.359***	(0.809)	-10.17***	(1.127)	-3.319***	(0.802)	-9.833***	(1.227)	-14.97***	(1.612)
N	7291									
ll	-2704.7									

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

estimates for the 'Inconsistency & Peace' outcome, equation DP contains the estimates for the 'Democracy & Peace' outcome, and so on.

Most notable in Table 7.4 is the coefficients for the lagged regime and conflict variables, as well as the history variables, which are strongly significant. Previous regime and conflict history thus explain a lot of the variation in regime and conflict status between countries. For example, being an inconsistent regime in peace at time t-1 increases log odds of being an inconsistent regime in peace, relative to an autocracy in peace at time t, with 5.914. This is equivalent to an increase in odds ratio of 370! Moreover, increasing the natural logarithm of time as an inconsistent regime in peace up to time t-2 by 1, increases log odds of being an inconsistent regime in peace, relative to an autocracy in peace, with 0.308. Put differently, log odds of being in IP relative to AP increases with $0.308 * \ln(\text{time in IP})$. Thus, the marginal effect of history is assumed to be more pronounced in the first years in a given 'state', and diminishing as a country remains stable in this same state. That previous regime and conflict history has strong effects underlines the importance of using simulations when estimating long-run effects. Not updating the history variables over time, thus keeping the transition probabilities constant, may alter the long-run outcome considerably.

Capital share, the proxy for economic inequality, does not perform as well as expected. It is only significant at the 10%-level in equation AC, thus weakly supporting hypothesis H_3. Increasing the capital share by one unit increases log odds of being an autocracy in conflict, rather than an autocracy in peace, with 1.56. This finding may indicate that inequality increases the risk of conflict in autocracies *prior* to a regime change. As I have not coded lags between the events in the dependent variable, such delayed effects are not captured in the short-run analysis. Moreover, Figure 6.2 showed that over twice as many regime changes take place in close proximity to a conflict onset, although not in the exact same year.[10]

Although the other coefficients for capital share fail to reach the threshold for significance they point in the expected directions: Increasing capital share increases log odds of being an inconsistent regime in peace with 1, and for being an inconsistent regime in conflict with 1.56, relative to an autocracy in peace. This is in line with hypotheses H_{2a} and H_{2b}. Furthermore, increasing capital share decreases the log odds of being a democracy in peace, relative to an autocracy in peace by 0.940, in line with the expectation from hypothesis H_1.

[10]Nicaragua, for example, experienced a long and bloody transition where a civil conflict broke out in the authoritarian regime, lasted throughout the years with institutional inconsistency during the 1980s, and ended with the introduction of democratic institutions in 1990.

For the two other control variables, GDP per capita and population, the estimates are strongly significant in most equations. Higher GDP per capita increases log odds of being a democracy in peace, it decreases log odds of being both an autocracy and an inconsistent regime in conflict, relative to an autocracy in peace. Interestingly, increasing income per capita increases the probability of being a democracy in conflict relative to autocracy in peace. It shows that higher income may have a strong effect on being democratic, and that this effect trumps the robust finding in the literature that economic development decreases the likelihood of being in a state of war. Moreover, having a larger population increases the probability of being in any one of these regime and conflict states, rather than autocracy in peace.

Comparing the estimates for GDP per capita in model m1 in Table 7.4, with those of model m0 in Table 7.3, shows that the coefficient estimates and standard errors for GDP are fairly the same. A VIF test confirms that there is no problem with multicollinearity between capital share and GDP per capita, as could be expected due to their high correlation.[11]

The results from the baseline model indicate that the direction of the effect from capital share is in line with the short-run hypotheses, although not strongly significant. As I am interested in whether higher or lower levels of economic inequality increases the risk of being in different states, the exact strength of the effect is less important. If I am willing to assume that the effect of capital share on the log odds of being in each of the states are identical, I can constrain model m1 based on the predictions of the short-run hypotheses.

Model m2 is a baseline model where I impose equality restrictions on the parameter estimates of capital share. In line with hypothesis H_2 I equalize the parameters for capital share in the IP and IC equations, as I hypothesize that increasing inequality increases the risk of being in inconsistency. Further, based on hypotheses H_{2b} and H_3 I equalize the parameters in the AC and IC equations, as I also expect increasing inequality to increase the risk of conflict in an autocracy. Hypothesis H_1 makes it reasonable to set capital share in the equation for DP equal to the negative of the parameter in the equation for IP, as increasing capital shares is hypothesized to increase the risk of being in IP, but decrease the risk of being in DP. Lastly, I make no predictions about the state of DC and I therefore I restrict this parameter to 0. With these constraints I am only estimating one parameter for the capital share variable, thus making the model more parsimonious.

[11] VIF stands for Variance Inflator Factor, and it shows how much of the variance of a coefficient estimate that is being inflated by multicollinearity. Multicollinearity increases standard errors and, makes it more difficult to find significant effects.

Table 7.5 reports the results from a multinomial regression of model m2. The parameter estimate for capital share is (+-)1.108, an average of the estimates in the unconstrained model m1. Moreover, the coefficients are now significant at the 5% level in the equation for IP, DP, AC and IC. As expected, a stricter model makes it easier to get significant results. The estimates for GDP per capita and population are almost the same as in model m1. A likelihood ratio test of models m0 and m2 reports a significant improvement in log likelihood at the 1% level with model m2. The out-of-sample evaluation of model m2 is also reported in Table 7.2, and is in line with model m1.

Although the effects of capital share are not as statistically strong as expected, based on the results in Table 7.5 with the constrained capital share variable, I consider it sufficiently justified to continue with the long-run analysis. As model m2 does not predict much better than model m1 (see Table 7.2), and as I am reluctant towards imposing too many restrictions on the parameters for simulation purpose, I continue the long-run analysis with model m1, the baseline model *without* constraints on capital share.

7.4 The long-run effects

In the first part of this section I present long-run effects based on predictions from the estimated multinomial logit model. It serves as a preliminary attempt at corroborating the long-run hypotheses. Moreover, the findings can serve as a comparison to the results from the simulation model. Next, section 7.4.2 reports the global simulation results. Due to a control problem with GDP per capita, I try to isolate the effect of economic inequality by conducting counterfactual 'experiments' at the country-level. Results from the experiments are presented in section, 7.4.3.

7.4.1 The long run transition probability matrix

It is possible to calculate analytically the behavior of the transition probability matrix presented in chapter 5.2.1 in the long run, for a fixed set of values for the explanatory variables and constant transition probabilities (cf. Waner 2004). I previously referred to this as a first order Markov chain. In the example below the 'Markov property', i.e. that the system only depends upon the state at t-1 and not on any previous time steps, does not hold, as I include information on conflict and regime history further back in time. Thus it is in effect no longer a 'true' first order Markov chain. To illustrate the long-run distribution of the various regime and conflict constellations I calculate predicted probabilities, based on the estimates from the multinomial logit model m1 in Table 7.4. A

Table 7.6: The predicted transition probability matrix for low inequality: Regime and Conflict status at t vs. at t-1

	AP	IP	DP	AC	IC	DC
AP t-1	.932	.037	.000	.027	.002	.001
IP t-1	.053	.770	.119	.004	.050	.004
DP t-1	.010	.018	.926	.002	.003	.041
AC t-1	.197	.019	.272	.419	.070	.024
IC t-1	.029	.301	.049	.038	.543	.041
DC t-1	.009	.036	.482	.021	.011	.441

comparison of the long-run distribution when all countries are assumed to have either low inequality or high inequality is of substantial interest.[12] I hold all other variables, except the regime and conflict status at time t-1, at their median level for 2010.[13] Doing this, I am controlling for GDP per capita. The only parameter that differs between Table 7.6 and Table 7.7 is capital share. It is important to note that since the predictions derive from a fixed set of values on the explanatory variables, they only present a partial, and somewhat 'artificial', picture of the world. Still, they can serve as a comparison to the simulation results, and illustrate the difference between holding the transitions probabilities constant and allowing them to vary.

The two transition matrices, when all countries are assumed to have either low or high inequality, are reported in Tables 7.6 and 7.7. The first thing to note is that the states of conflict are more unstable than the states of peace, regardless of inequality level. But, when inequality is low more countries transition away from conflict at time t-1 to time t. Hence, conflicts are presumably longer when inequality is high, than when it is low. Further, when inequality is low, the probability of transition away from inconsistency is somewhat greater than when inequality is high. Moreover, the probability of a transition from any one of the states at time t-1, to a democracy in peace (DP) at time t, is lower when inequality is high. This implies that it is more difficult for a country with high

[12] I define low inequality as the lower 25th percentile on capital share in 2010, which is equal to 0.636. High inequality is the upper 75th percentile, which is equal to 0.754.

[13] I make use of the statistical package Clarify (King, Tomz and Wittenberg 2000) when calculating the predicted probabilities. I base the predictions on the 2010 level of the explanatory variables, as this is the last year with observed data and the starting point for the simulations. The 2010 median value is 8.523 for log GDP per capita and 9.228 for population. The history variables are held at their 2010 *mean* level in order to avoid underprediction of democracies, as the median values are all zero. The do-file used to calculate the transition matrices is found in Appendix E.

Table 7.7: The predicted transition probability matrix for high inequality: Regime and Conflict status at t vs. at t-1

	AP	IP	DP	AC	IC	DC
AP t-1	.923	.041	.000	.032	.002	.001
IP t-1	.049	.789	.099	.004	.055	.004
DP t-1	.011	.022	.913	.003	.004	.047
AC t-1	.185	.020	.231	.462	.078	.023
IC t-1	.025	.293	.039	.039	.567	.037
DC t-1	.009	.041	.445	.025	.013	.467

inequality to introduce democratic institutions, which supports hypothesis H_1.

As this thesis investigates the long-run effects of democratization, the implications of the transition matrices over time is of particular interest. From the transition probabilities in Tables 7.6 and 7.7 I can calculate 'steady state' transition probabilities, i.e. the fixed probabilities the system converges to at some point in the future. I assume that the transition probabilities are constant, and call this transition matrix P. The initial distribution of countries in the various states is called the system's *distribution vector*, v. The 2010 distribution of countries is used as the 'initial' distribution, as it is the last year with observed data. In order to find the distribution of countries in 2011, after 1 time step, the distribution in 2010 is multiplied with the transition matrix P, found in Tables 7.6 and 7.7. For 2012 P has to be multiplied with the distribution in 2011 ($vP * P$, or $v * P^2$), and so on (see Waner 2004). I start off by calculating the distribution of countries in each state in 2050, thus making the following computation: $v * P^{40}$. The equilibrium behavior of the system, also referred to as the steady state matrix, is found by taking the power of matrix P, until P approaches a fixed matrix, P^{∞}. The probabilities in each column of the matrix are then equal.[14]

The long-run distribution of countries, when all countries are assumed to have either high or low inequality and take on the mean values on the other explanatory variables, is summarized in Table 7.8. The first row presents the 2010 distribution, which is equal under both scenarios. The two rows below report the distribution in 2050, under the assumption of an 'exogenous' shock to the level of economic inequality. The long-run hypotheses predict that countries with higher inequality are more likely to be inconsistent regimes

[14]For a description of markov systems and the use of transition matrices to calculate steady-state behavior, see Waner (2004) or Greene (1997). The do-file used for calculation of the steady state distribution can also be found in Appendix E.

Table 7.8: Long run distribution of countries in the regime and conflict states, low vs. high inequality: Estimates based on transition matrices in Tables 7.6 and 7.7

	Year	**AP**	**IP**	**DP**	**AC**	**IC**	**DC**	**Total**
Observed	2010	21	41	73	8	9	8	160
		0.131	0.256	0.456	0.050	0.056	0.050	1.0
Low inequality	2050	38	19	90	3	3	7	160
		0.238	0.119	0.563	0.018	0.018	0.044	1.0
High inequality	2050	38	24	81	4	5	8	160
		0.238	0.150	0.506	0.025	0.03	0.050	1.0

(H_7), and to be in a state of conflict (H_4) in the long run. Comparing the two distributions shows that there is indeed a higher risk of being an inconsistent regime in the long run, moving from low to high inequality. The proportion of countries in inconsistency is found by summarizing the proportion in inconsistency and peace (IP) and the proportion in inconsistency and conflict (IC), which equals (0.150+0.03) 0.18 when inequality is high, and (0.119+0.018) 0.137 when inequality is low. Moreover, summarizing AC, IC and DC when inequality is high and low, respectively, shows that the model predicts 0.025 more countries in conflict status when inequality is high. Hypothesis H_6 states that there is more conflict in inconsistent regimes than democracies in the long run, regardless of inequality level. This proposition is supported. The share of inconsistent regimes in conflict is 0.172 when inequality is high, and 0.143 when inequality is low. The share of democracies in conflict, on the other hand, is only 0.09 when inequality is high, and 0.072 when inequality is low.[15]

Changes to the level of inequality seem to alter the distribution of countries in the various regime and conflict 'states'. Due to the specific values of the exogenous parameters underlying the transition matrices in Tables 7.6 and 7.7 it is difficult to generalize the effects to all countries. Moreover, the matrices are not able to take into account the effects of regime and conflict history, as the calculations are based on the crude assumption of constant transition probabilities. In the following, I discard this assumption as I present the results from the simulations. I start off with the global simulation results.

[15]I also calculated the steady state distribution vector, and it turns out to be identical to the distribution vector for 2050. This indicates that 40 years is a sufficient amount of time to reach a stable matrix, and it provides support for the decision to simulate until 2050.

7.4.2 Global simulation results

With model m1 as a basis I estimate long-run effects by applying the simulation procedure presented in chapter 5.3.2. As explained in chapter 6 the dependent variable 'Regime and Conflict Status' is endogenous to the model, and the same holds for the history variables and dummy variable for the status at t-1. The inequality measure, capital share, GDP per capita and population are assumed exogenous, hence their values are fixed in the forecasting period. Appendix C, Figure C.2 provides an overview of the total number of simulated regime changes and conflict onsets. I run 5000 simulations for the final analysis.[16]

The long-run hypotheses present very clear predictions about the observable effects of economic inequality on democratization and civil conflict in the long run. One way of illustrating the relationship between inequality, conflict incidence and regime changes, is by summarizing the results from the simulations, and extracting the proportion of countries in the various regime and conflict states. Further, I can extract the same proportions for countries with high and low values on capital share. The do-file used to summarize the simulation results can be found in Appendix C.

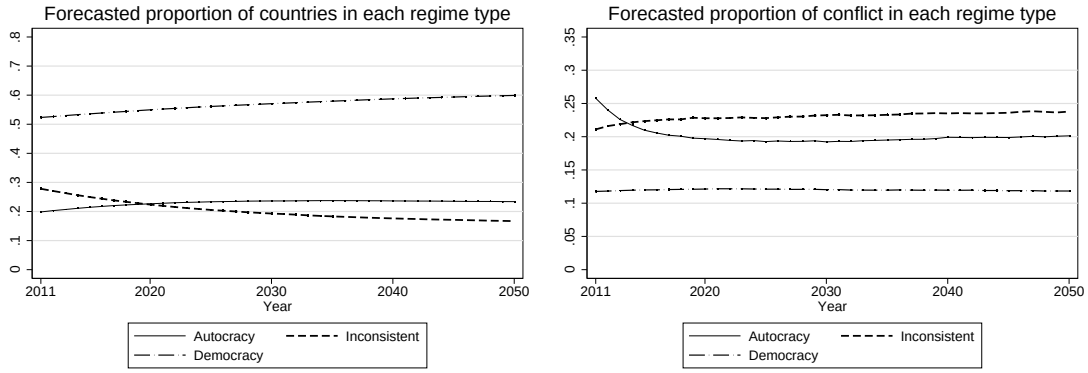
Figure 7.5 reports the long-run distribution of regimes and civil conflict in all countries. The solid line represents the autocratic regimes, the dashed line represents the inconsistent regimes, and the long dashed line represents the democracies. The panel on the left-hand side shows that model m1 predicts an increasing number of democratic and autocratic transitions in the long run, and a corresponding decrease in the number of inconsistent regimes. After 40 years there are 60% democracies, 18% inconsistent regimes, and 22% autocracies.[17] It provides support to the argument that inconsistent regimes are inherently unstable, and with time will transition into the closest equilibrium, either a consistent democracy or autocracy (Gates et al. 2006, 906).

The panel on the right-hand side of Figure 7.5 reports the forecasted proportion of conflict in each regime type. The results show that there is more conflict in inconsistent regimes in the long run. The proportion of inconsistent countries in conflict increases, from approximately 21% to 24%, over 40 years. The share of autocracies in conflict is reduced during the first 15 years, before it stabilizes at a steady-state level of 20%. Only

[16]Hegre et al. (forthcoming, 29) show that both mean and variance of the simulation result converge to a stable estimate, when the number of simulations exceed 500–1000. As my dependent variable has twice as many categories, and 4 times as many transitions, more simulations are needed. I consider 5000 simulations to be sufficient.

[17]Keep in mind that the results are not driven by increases in the general level of development, or changes in the world population, which are kept constant at the 2010 level throughout the simulations.

Figure 7.5: Long-run distribution of regimes and civil conflict



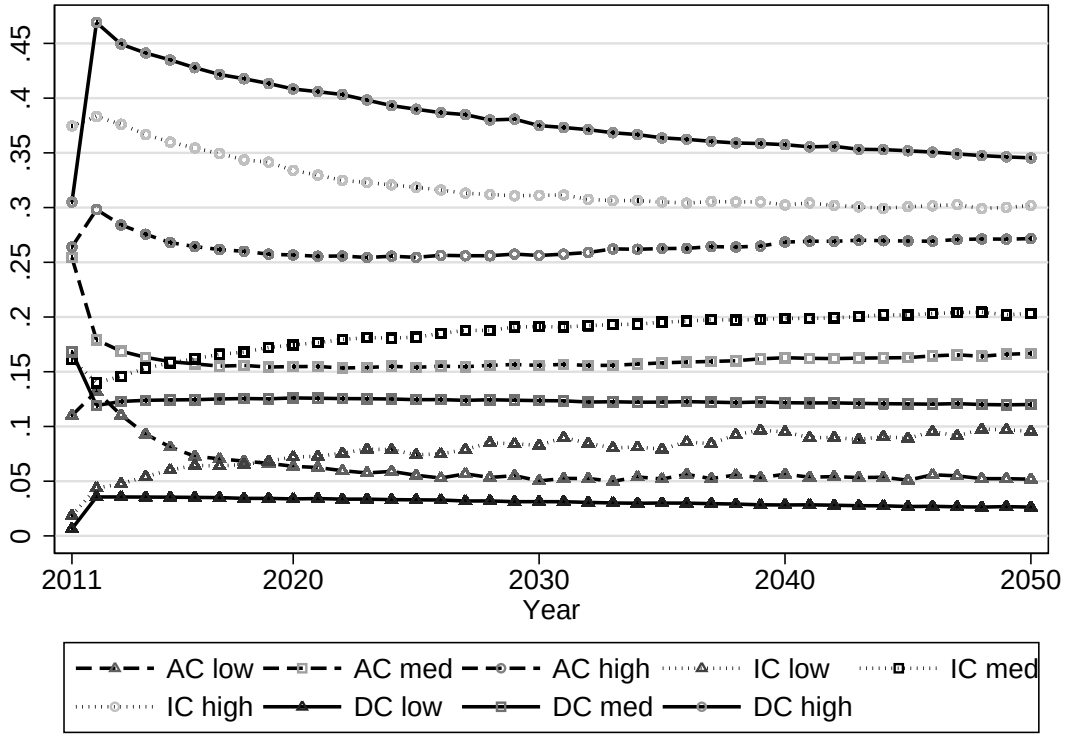
12% of democracies are in conflict after 40 years, and the share is more or less constant during the entire period of measurement. The results give a preliminary verification of hypotheses H_6, which states that there is more conflict incidence in inconsistent regimes in the long run.

The crux of the theoretical argument is that economic inequality matters both for the institutions a country is able to preserve over time, as well as for the observable amount of conflict. Moving to hypotheses H_4 and H_6, Figure 7.6 reports the long-run distribution of civil conflict, by inequality level and regime type. All countries in the dataset are divided into three groups, based on their value on capital share between 2011–2050. The 'intermediary inequality' group is defined as all countries above the 25th- and below the 75th percentile on the capital share variable. In Figure 7.6 this group has squared markers on their lines. The high inequality group is situated above the 75th percentile, and is marked with circles. The low inequality countries are situated below the 25th percentile on capital share and are marked with triangles in Figure 7.6. Within each inequality group, the countries are further subdivided into their respective regime type. Democracies have solid lines, autocracies have dashed lines, and inconsistent regimes have dotted lines.

Hypothesis H_4 states that there is more conflict incidence in high inequality countries in the long run, and Figure 7.6 supports this. The three uppermost lines represent the high inequality countries, and they have clearly the most conflict after 40 years.[18] Further, the intermediary inequality group is situated in the middle, close to, but clearly separated from, the low inequality group are the bottom of the figure. Surprisingly, there is most conflict in high inequality democracies (solid line at the top of Figure 7.6). Although the

[18]The sharp increase in the line for high inequality democracies in 2011 represents the adjustment from observed to simulated data.

Figure 7.6: Long-run distribution of civil conflict by inequality level and regime type. Tripartite inequality classification.



proportions of simulated conflict in all three regime types approach each other in the long run, they never coincide. What is probably driving the result is cases like Colombia and India, as well as some other democratic and conflict prone countries, with high levels of inequality.[19] The finding is contrary to hypothesis H_6, as it predicts more conflict in the inconsistent regimes, regardless of inequality level, in the long run.

The amount of conflict differs much between the three groups. As many as 30% of all high inequality countries are in conflict after 40 years, whereas only 3.5% of the low inequality countries. In the group with medium inequality 15% are in conflict in the long run.[20] There is also divergence internally in a regime type. Whereas only 3.5% of all low inequality democracies are in conflict after 40 years, as much as 35% of the high inequality democracies are in conflict in the long run.

Based on the results in Figure 7.6, economic inequality seems to matter more than political institutions for the proportion of conflict in the long run. Despite differences in the level of conflict between regimes, there is a markedly larger difference between

[19]Colombia, Peru, Zambia, Botswana, India, Bangladesh and Indonesia are the countries which are high inequality and democratic in 2010.

[20]The proportions not reported in Figure 7.6, but are based on an average over the three regime types.

the groups with high and low inequality, supporting hypothesis H_4. The validity of hypothesis H_6, on the other hand, seems to be contingent on the level of inequality. The most 'dangerous' constellation is a high inequality democracy.

Schatzman (2005) finds evidence that the introduction of democracy in Latin America, a region with many high-inequality countries, increased rebellion and conflict in these countries. She highlights the importance of 'democratic quality', in terms of inclusion and legitimacy, to reduce the spells of violence (Schatzman 2005, 306). Thus, the group of high inequality democracies may be 'poor quality' democracies, where opposition groups readily challenge the government. Another explanation is that as high inequality countries often are poor, the observation is purely an income effect. However, looking at the data reveals that median income level is considerably higher in high inequality democracies, than in high inequality inconsistent regimes and autocracies. Income, therefore, does not seem to be the entire explanation of this finding. Still, as I am not controlling for GDP I cannot be completely sure where the effect comes from. It is not straight forward to parse out the effect of GDP in the simulations, but one way of doing it is by conducting counterfactual 'experiments' at the country-level, as I attempt in the following.

7.4.3 'Experiments'

I construct 5 experiments to serve as a tool for testing the long-run hypotheses. In the first four experiments I change the level of economic inequality for the select countries from 2010 an onwards, and investigate the effect on conflict incidence and regime changes, as well as on the probability of being in a given regime state, over time. The 25th and 75th percentile of the capital share variable is used as a benchmark for low and high inequality, respectively. This equals to capital shares of 0.631 and 0.758. By changing the level of inequality, while keeping GDP constant, the effect of inequality is isolated. In the last experiment I alter the regime type in Thailand in 2010, and trace the indirect effects on conflict incidence through regime changes. The results from each experiment are compared with the 'baseline scenario', where regime changes and conflict incidence are determined by the observed history, without any manipulation. To implement the experiments I alter the last historical observation, i.e. the 2010 observation, and the forecasted observations in the select countries before the simulation procedure is conducted.[21] The procedure is

[21]Because of the way in which the simulation procedure is set up, the manipulation in the dataset is done before the multinomial logistic model is calculated. Altering the last historical observation can in theory have an impact on the simulation results, by changing the beta coefficients in the multinomial model. I have checked all the estimated 'scenario' models, and the coefficients and standard errors are

inspired by a similar analysis conducted in Hegre et al. (2011).

I have chosen a bundle of 'experiment countries'. These countries could in theory be replaced by some other countries, as their predicted behavior is a function of the results of the multinomial logit model, not of any idiosyncratic features of the country. This does not mean that they are chosen randomly, though. There are a number of ways in which to could go about choosing countries. The most intuitive is to focus on variation in the key variables, which are the level of inequality, regime and conflict type, regime and conflict history, and economic development.

Economic inequality is the main explanatory variable and to make hypothetical changes in both directions; from a low level of inequality towards a high level, and vice versa, could be meaningful. Yet, I do not consider it as important as to select countries with different income levels. Varying the level of economic development over the scenarios is another way of 'controlling' for GDP. If the results point in the same direction in both high- and low-income countries it increases robustness. Variation in the level of GDP per capita is therefore the first selection criterion.

Regime type and conflict status is the object of inquiry, thus it is reasonable to select countries with different observed status in 2010. On the other hand, as this variable is endogenous to the model, a country's status in a single year should not make much of a difference for the long-run results. What does make a difference, though, is the length of time a country has been in its pre-simulation status. As shown in table 7.4, the variables that capture regime and conflict history have very strong predictive power. Changing the level of inequality does not alter the historical past, and hence, I cannot be 100% sure that the findings are not driven by this 'shadow of the past'. The second criteria for choosing experiment countries has been to secure a certain amount of variation in regime and conflict history. The experiment countries, together with the scenarios, are specified in Table 7.9.

Before I present the results a brief description of the experiment countries are in place: **Thailand** is a country with deep economic disparities, and a turbulent political past. Although civil conflict and political unrest has afflicted this Southeast Asian kingdom, it has managed quite well economically and its GDP per capita was 8900 US$ in 2010, approaching the upper 75th percentile of the GDP per capita indicator. Thailand is coded as an inconsistent regime in conflict in 2010, and in scenario 3 I decrease the level of inequality to the 25th percentile of the capital share variable, whereas in scenario 6 I hypothesize that Thailand became a democracy in conflict in 2010.

close to identical. Hence, I do not expect this approach to impact on the simulation results.

Table 7.9: Overview of Scenarios

Nr. Scenario	Country	Baseline (observed status 2010)	Alteration
1		Baseline	No alterations
2	Thailand	High Inequality	Low inequality
3	Ethiopia	High inequality	Low inequality
4	Laos	High inequality	Low inequality
5	Portugal	Low inequality	High inequality
6	Thailand	Inconsistent and conflict	Democracy and conflict

Ethiopia has been coded an inconsistent regime in conflict since 1995. It has had a very high and stable level of inequality the entire period from 1960. Its growth record is also poor, placing it in the lower 25th percentile of the GDP indicator. In scenario 4 I change the level of inequality to the lower 25th percentile of capital share. As Ethiopia represents a low-income country with an unstable political history, it can be contrasted with Thailand economically.

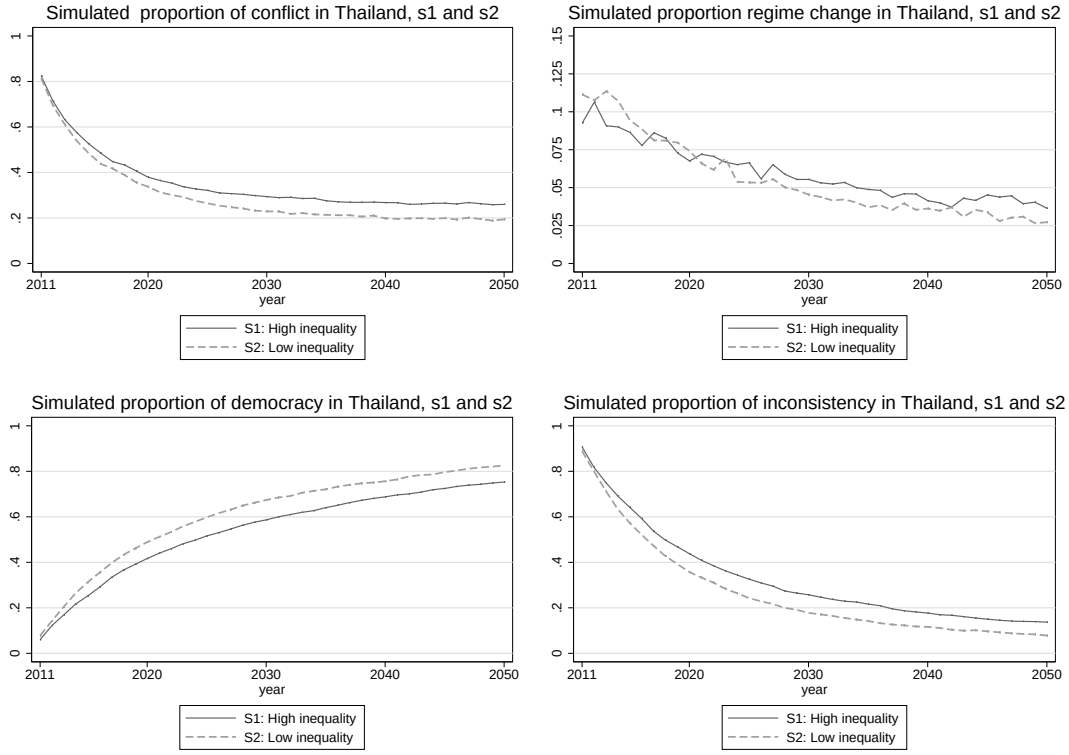
Laos is a poor country, as well as a country with high economic inequality, similar to many of the other autocratic countries in the dataset.[22] There has seemingly not been much pressure for democratization in Laos, as the country has been a peaceful and stable autocracy all the way back to 1960. In scenario 4 I change the level of inequality in Laos from high to low. Laos can be compared with Portugal, as they are both stable, consistent regimes, but with a large disparity in income levels. It can also be compared with Ethiopia, investigating the effect from different regime and conflict history.

The last 'experiment' country is **Portugal**. Portugal is a stable democracy in southern Europe, selected because of its low level of inequality and its relatively high income level. In scenario 5 I change its level of inequality to the upper 75th percentile on capital share. Portugal can be contrasted with Thailand politically, and Laos economically.

With their varying levels of income and diverse institutional and conflict history, the example countries exhaust the most important combinations on the GDP per capita and history variables. The following results serve as a test of long-run hypotheses reported in section 5.5, as well as to identify the 'ideal types' of democratization presented in Table 4.1. For scenarios 2–5 the simulated proportion of conflict, the simulated proportion of country-years with a regime change, simulated share of democracy, and simulated share of inconsistency, under the baseline scenario and with the alteration, are reported. The do-file used to summarize and present the results from the simulations can be found in

[22]Its GDP per capita was 1819 1990 US$ in 2010 and capital share was .80.

Figure 7.7: Simulated effects of introducing low inequality in Thailand in 2010



Appendix C.

Changing the level of economic inequality

The first experiment is conducted on Thailand, where the capital share variable is changed from its initial level of 0.771, to 0.631, the lower 25th percentile on the capital share variable. Figure 7.7 summarizes the long-run effects. The upper-left and upper-right panel report the risk of conflict, and the simulated proportion of country-years with a regime change, under scenarios S1 and S2: The initial risk of conflict is 0.8, or 80%, under both scenarios, as shown in the upper-left panel. There is a strong downward-sloping trend in the simulated proportion of conflict, both in S1, the solid line, and S2, the dashed line. Moreover, the share of conflict under high inequality is continuously higher than under low inequality. After 40 years, 'excess conflict', i.e. the gap between simulated conflict in scenarios S1 and S2, is 6.56%. The finding supports hypothesis H_4. The relative risk of conflict is 1.34 higher in scenario S1, with high inequality, than in S2 with low inequality. This implies relative risk reduction of 25%, if Thailand is to have a decline in its level of inequality.[23]

[23]In order to calculate the *relative risk*, I divide the share of simulated conflict in S1 on the share of conflict in S2: $0.2596/0.194 = 1.34$. If the relative risk is exactly 1, it suggests that the presence of a factor

In order to illustrate the gain from taking into account the reciprocal causality between conflict and regime changes, as well as the effect from previous regime and conflict history, I make two comparisons. The first comparison is between the direct effect of lowering inequality on the risk of conflict from model m1, and the total effects observed in Figure 7.7. In this case I only consider the direct effect on the probability of being in a state of inconsistency and conflict (IC), disregarding the risk of a regime change. To obtain this direct effect I calculate the predicted probability of being in IC, in line with the procedure in section 7.4.1, where I calculated the transition matrices. Instead of using the median values for GDP per capita and population, I use the specific values of the explanatory variables for Thailand in 2010.[24] Doing this, I find that the direct effect of reducing the level of inequality is a 1.62% reduction in the risk of being in inconsistency and conflict.

A second comparison can be made between 'excess conflict' in the simulations and 'excess conflict' in the high inequality scenario in the long run transitions matrices in section 7.4.1. I conduct the same estimations as was used to produce the results in Table 7.8 (See Appendix C for the do-file), but apply Thailand's values, as explained in the above paragraph. 'Excess conflict' is in this case 4.2% in the baseline scenario with high inequality. In comparison, the 'excess conflict' in Figure 7.7 was 6.56% in the long run. The finding shows that when the transition probabilities are not updated each year based on changes to the regime and conflict history variable, the total effect of a reduction in inequality on the long-run level of conflict is underestimated by 2.36%.

Why the difference? In Figure 7.7, a decrease in the level of inequality, firstly, decreases the probability of being in a state of conflict within the current regime type, but it also decreases the probability of regime instability, as seen in the panel at the upper-right. These are the direct effect from inequality. Secondly, a lower probability of conflict further decreases the probability of conflict in the next year, but it also decreases the probability of regime change. This reciprocal causation between institutional changes and conflict is the indirect effect from inequality. If the simulations, in a given year, report Thailand in peace, the history variables are updated, and serve as an additional 'peace-generator', by increasing the probability of being in the respective state next year. Aggregated

---

does not impact on the outcome, whereas a factor above 1 implies that the risk is greater when the factor is present. The *relative risk reduction* from reducing the level of inequality is $(0.2596 - 0.194)/0.2596 = 0.2527$, or 25%.

[24] In line with the estimations in section 7.4.1, I use the Clarify package to generate the predictions (see King, Tomz and Wittenberg 2000). The value for ln GDP is 9.096, the value for population is 11.084, and ln(time in IC) is set to 1.098. The values for capital share is 0.771 and 0.631 under the high and low inequality scenarios, respectively.

over many years, the effect on conflict incidence from decreasing the level of inequality is therefore exacerbated. This is visible in the upper-left panel of Figure 7.7, as an increasing gap between the scenarios S1 and S2. The example above illustrates the importance of simulating in order to capture all of these effects. It provides a more precise picture of the overall impact of an initial increase or decrease in inequality on the level of conflict over time.

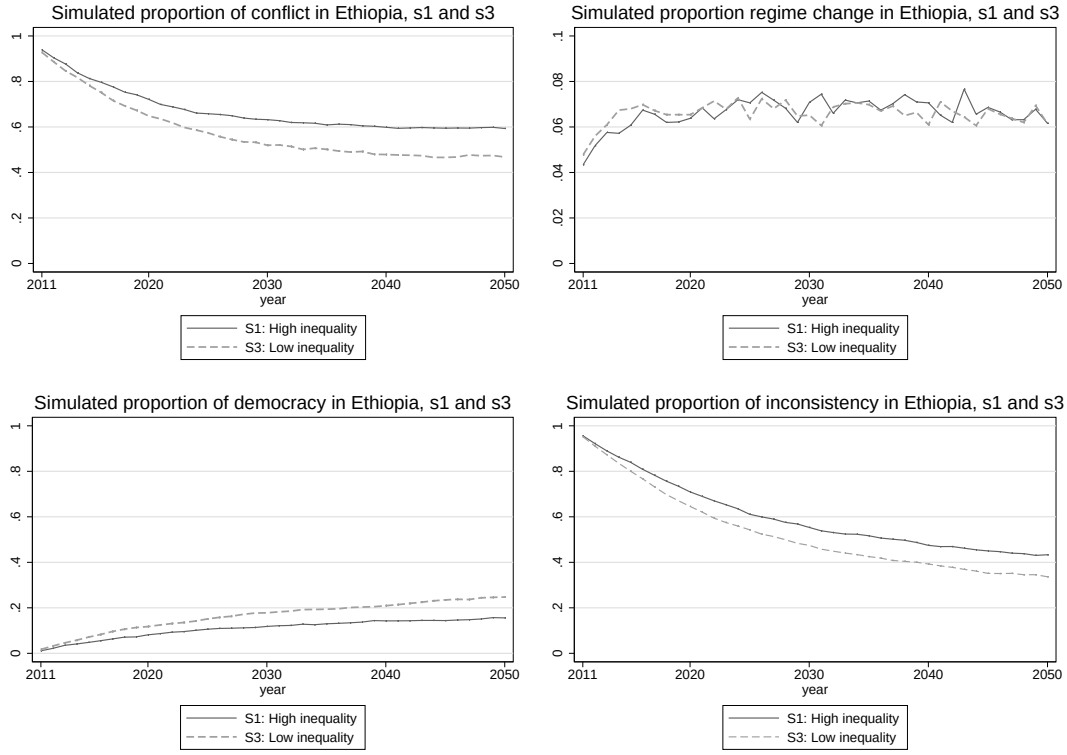
The panel at the upper-right reports simulated country-years with a regime change. During the first 5 years, there are more country-years with regime changes under S2, with low inequality. However, from 2025 an onwards, the simulated proportion of regime change decreases below the baseline scenario, with high inequality. After 40 years scenario S2 reports 0.92% less regime changes than the baseline scenario. This equals a relative risk reduction of 25%. The finding provides support for hypothesis H_5. I expect the initial increase, and then decrease in the share of regime changes in scenario S2 to be caused by an increase in transitions to democracy, followed by a stable proportion of simulated democracy.

The two panels at the bottom report the simulated share of democracy and inconsistency under scenarios S1 and S2. For a country like Thailand, with a relatively high income level, the probability of having democratic institutions is very high in the long run. Around 80% of the simulations report Thailand as a democracy after 40 years. Parallel, the long-run probability of being in inconsistency declines to approximately 10%. The overall trend is independent of inequality, but it is exacerbated when inequality is decreased. 'Excess democracy' is 6.28%, after 40 years. This finding support hypothesis H_7.

Thailand is an example of ideal type 2 in Table 4.1; a country with a long-drawn democratization process, characterized by political instability and subsequent civil conflicts. Reducing the level of inequality in Thailand should, if my assumptions are correct, illustrate the difference between ideal type 1, a swift and peaceful democratization, and ideal type 2. This is partly possible, as the simulated share of democracy increases faster and reported conflict incidence decreases faster in scenario S2 with low inequality. However, it is clear that the underlying trend towards less conflict and more democracy is strong, regardless of inequality level. This is partly an expression of Thailand's level of income, but most importantly it illustrates the convergence to a steady-state level of conflict and democracy. In the next scenario I investigate whether 'excess conflict' or political instability is greater for a country with a lower level of economic development.

Figure 7.8 summarizes the effects of changing the level of inequality in Ethiopia. The

Figure 7.8: Simulated effects of introducing low inequality in Ethiopia in 2010



upper-left panel reports simulated proportion of conflict under the baseline scenario, where Ethiopia has a capital share of 0.822, and in scenario S3, when capital share is reduced to 0.631. There is an underlying decreasing trend in the risk of conflict in Ethiopia, in both scenarios. However, after a couple of years the risk of conflict under low inequality (dashed line), decreases faster than under high inequality (solid line). Simulated conflict stabilizes after approximately 30 years under both scenarios. However, it is considerably lower under scenario S3, and 'excess conflict' is 12.56% after 40 years. Substantively, this is not an insignificant change. The relative risk of conflict Ethiopia, and countries similar to it, is 0.78 lower when inequality is low. This equates to a relative risk reduction of 21%. Ethiopia thus has a lower relative risk reduction than Thailand, where the relative risk reduction was 25%.[25]

Simulated country-years with a regime change under S1 and S3 is reported in the

[25] As I reduce or increase the capital share variable in each country to the 25th or 75th percentile on the capital share variable, the absolute change differs somewhat between the experiments. The reductions are between 1.2 and 1.6 standard deviation on the capital share variable. I take this into consideration as I interpret the results, and I am careful directly comparing the relative risk reduction between countries. However, in Thailand, the capital share variable was reduced somewhat less than in Ethiopia. Thus, the effect in Thailand relative to Ethiopia is even greater than the comparison suggests.

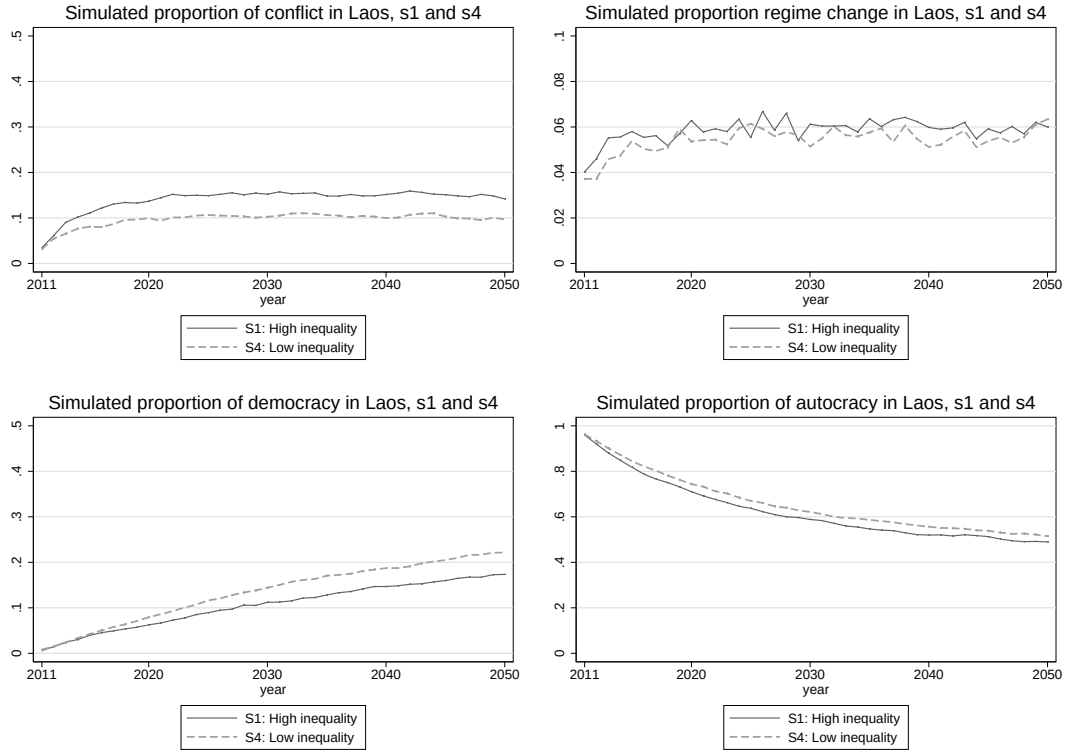
upper-right panel of Figure 7.8. Ethiopia has more reported regime change than Thailand does. Approximately 6% of all simulated country-years report a regime change in Ethiopia. However, after 40 years there is no visible difference between the high and low inequality scenario, contrary to hypothesis H_5. It indicates that most of the reduction in conflict in Ethiopia in the long run is a consequence of the direct effect from inequality on conflict, and not indirect effects through a decrease in the number of regime changes.

The prospects for democracy in a conflict prone and poor country like Ethiopia is much worse than in Thailand. Although there is an upwards sloping trend, indicating more democracy in the long run, only 15.6% of the simulations report democracy in Ethiopia after 40 years under the baseline scenario, whereas 43% report inconsistent institutions (right-hand panel). As Ethiopia's history has never included democratic institutions, the increase in simulated democracy still provides hope for similar countries. Moreover, 'excess democracy' under scenario S3, with low inequality, is 9.2%. This finding indicates a potential positive impact from redistribution on the probability of democracy in Ethiopia.

I expect the difference in simulated proportion of conflict in Ethiopia and Thailand to be ascribed to differences in economic development. Thailand has almost 8000 1990 US$ higher income per capita than Ethiopia, which reflects in Figures 7.7 and 7.8. Both the simulated long-run incidence of conflict and inconsistency is much higher in the low-income country Ethiopia. The persistently higher share of simulated regime change in Ethiopia is also expected to be an artifact of low income. Thailand, on the other hand, will mostly likely manage to consolidate democracy, at least if accepting these results as valid. However, the results may be too optimistic. Despite its growth records, Thailand has experienced political instability and violence, with 8 out of the 20 last years spend in conflict status. Moreover, distribution of newly earned income has been skewed. Thus, as average income has increased, so has inequality (Warr 2007, 152). Warr (2007) ascribes this to failure of the educational system, a persistent rural-urban divide, and weak institutions unwilling to tackle the problems. Thailand is an anomaly in this respect, although it fits well with the theoretical argument outlined in this thesis.

In scenario 4 I investigate whether the long-run effects are similar for a country that is autocratic, rather than inconsistent, at the outset. Laos is a representative of low-income autocracies that have yet to start on a democratization process. I conduct this experiment to be able to investigate into ideal type processes 1 and 3. Although Laos is poor and has a long autocratic history it has not experienced civil conflict since 1960. Findings corroborating with the long-run hypotheses would strengthen the argument that inequality matters for democratization. In scenario 4 the level of inequality in Laos is

Figure 7.9: Simulated effects of introducing low inequality in Laos in 2010

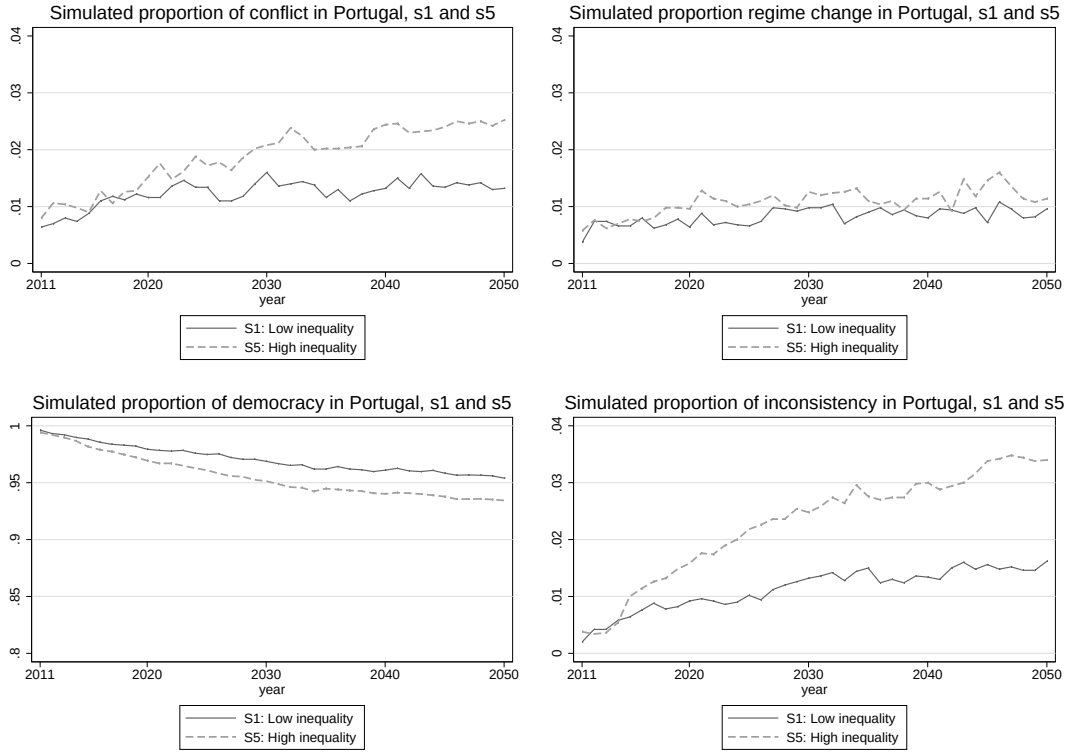


moved from 0.80 to 0.631. The results are reported in Figure 7.9.

The upper-left panel of Figure 7.9 reports the proportion of conflict in Laos under scenarios S1 and S4. In line with the two previous experiments, simulated conflict incidence is higher when inequality is high. Although Laos already has an initially low conflict level, reducing inequality reduces the level of conflict considerably. After 40 years, 'excess conflict' is 4.5% under the baseline scenario, S1. Hence, reducing the level of inequality leads to a reduction in the relative risk of conflict of 31.7% in the long run. Scenario 4 thus corroborates with hypotheses H_4. The upper-right panel of Figure 7.9 reports the simulated proportion of country-years with a regime change. There is little effect from changing the level of inequality on political stability in Laos, although there are somewhat more simulated regime change in the long run in scenario S1 with high inequality, supporting hypothesis H_5.

As reported in the lower-left panel, the simulated share of autocracy declines over time in Laos, but remains around 50% in both scenarios. It indicates that autocracy is persistent, regardless of inequality level. That the simulated proportion of autocracy is somewhat higher when inequality is low is somewhat surprising; 'excess autocracy' is 2% in scenario S4, with low inequality. However, this can be interpreted as a confirmation that low inequality creates a more stable environment where consistent democracies and

Figure 7.10: Simulated effects of introducing high inequality in Portugal in 2010



autocracies institutions prevail at the expense of the inconsistent. I inspected the results for the inconsistent regimes as well, and found a decrease in the relative risk of being in inconsistency of 22% when inequality is low. This further strengthens hypothesis H_7. It is also in line with this thesis' main argument, namely t hat high inequality creates more demand for democratization in authoritarian regimes.

The prospects for democracy in Laos are in line with those in Ethiopia. After 40 years approximately 20% of the simulations report democracy. Moreover, 'excess democracy' is 4.8% in the long run. Thus, the chances that Laos will become a democracy within the next 40 years are relatively slim, although redistribution may potentially speed up democratization.

With a persistently high share of autocracy, a higher simulated proportion of conflict incidence, and a higher share of institutional inconsistency, in the long run in scenario S1, ideal type 3, the failed democratization, is ambiguously identified in the case of Laos. There are evidently several dynamics at work, and thus somewhat difficult to get a consistent picture of the situation.

In scenario 5 I make the counterfactual assumption that a low inequality country transform into a high inequality country in 2010. The example country is the high-income, stable democracy Portugal. The most notable feature in the panels of Figure 7.10 is the

apparent volatility in the simulated risk of conflict and inconsistency, the upper-left and lower-right panels, respectively. This is only an artifact of the scaling on the axis.[26]

Looking at the risk of conflict first; in the upper-left panel of Figure 7.10 the solid line (S1), representing low inequality, lies below the dashed line (S5), representing high inequality. Although slow at first, the effect from inequality increases gradually, and after 40 years 'excess conflict' is 12%, supporting hypothesis H_4. This may seem like an insignificant effect, but it does in fact represent a doubling in the number of simulations reporting conflict.

The upper-right panel reports the simulated country-years with a regime change. The risk of regime changes is overall very low in Portugal – approximately 1%, under both scenarios. Further, it is difficult to separate the simulated regime change under the two scenarios. However, there seems to be a slight increase in the number of changes in the long run when inequality is high.

If economic inequality in Portugal should increase, there is no immediate danger to democracy, as shown in the lower-left panel of Figure 7.10. A long democratic history and a high income level supports continued democracy under both scenarios, although there is 2% less simulated democracy after 40 years under scenario S5. Furthermore, 'excess inconsistency' is 1.8% under scenario S5, supporting the proposition that there is more inconsistency when inequality is high (H_7).

The above experiments focus on the direct effects of changing the underlying economic structure. Another way in which inequality affect conflict incidence, is indirectly through regime changes. I can study these indirect effects explicitly by hypothetical altering the regime type in a given country. I have selected Thailand for this experiment.

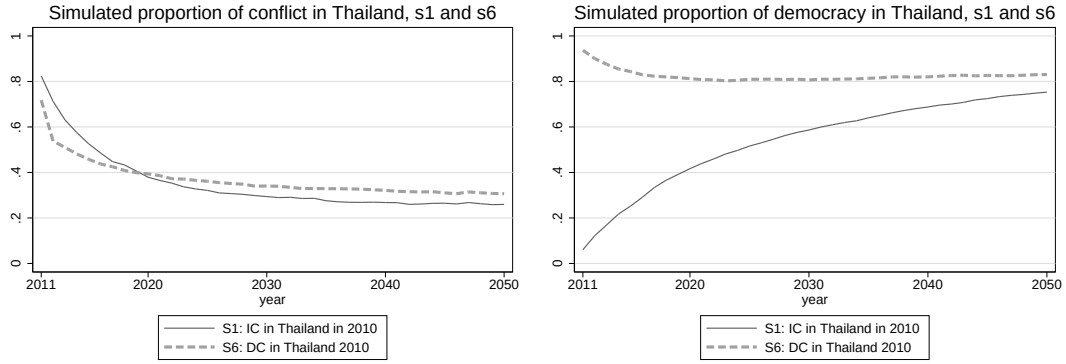
Indirect effects from regime changes

In scenario 6 I alter the status of Thailand in 2010 from an inconsistent regime in conflict to a democracy in conflict, while keeping its level of inequality constant, at a high level. The results are reported in Figure 7.11.

The panel on the left-hand side of Figure 7.11 reports the simulated proportion of conflict in Thailand, whereas the panel on the right-hand side reports simulated proportion of democracy. Moving from an inconsistent regime in conflict (S1) to a democracy in conflict (S6), causes an immediate, and steeper, decrease in simulated conflict under scenario S6,

[26]The axis in Figure 7.10 goes from 0 to 0.04, due to the overall low risk of conflict and of having inconsistent institutions. Scaling up to the same level as the other figures would result in smoother curves.

Figure 7.11: Simulated effects of introducing democracy in Thailand in 2010



as shown by the dashed line. The effect lasts about 10 years. The finding is contrary to previous studies, which show that the risk of conflict increases in the immediate aftermath of institutional changes (see e.g. Cederman, Hug and Krebs 2010). Moreover, after 40 years the excess share of conflict is 5% under scenario S6, which is surprising. However, it corroborates with the global simulation results, where high inequality democracies had the most conflict incidence in the long run. Thus, I do not find support for hypothesis H_6.

The experiment indicates that the effect from inequality on the long-run incidence of conflict found in scenarios 2–5 partly can be attributed to the indirect effects through regime changes. This is in line with my theoretical argument.

The right-hand panel of 7.11 reports simulated proportion of democracy under scenarios S1 and S6. The effect of introducing democracy in 2010 on the simulated proportion of democracy is strong and persistent. After 40 years the simulated proportion of democracy is higher in scenario S6, where I introduced 'democracy and conflict' in 2010. Further, Thailand's high underlying probability of democracy is striking.

Summary of 'experiments'

Conducting the counterfactual 'experiments' have provided support for most of the long-run hypotheses. Firstly, high inequality causes more conflict incidence in the long run than do low inequality, keeping the level of economic development constant. Hypothesis H_4 is therefore supported. Secondly, hypothesis H_5 is provided some supported by the above analysis. There are generally a higher proportion of country-years reporting a regime change under high inequality, although the degree of this effect differs between the countries. Thirdly, in line with hypothesis H_7 there is a higher share of simulated institutional inconsistency in the long run when economic inequality is high, and the result

seems to be robust to differing levels of income and regime and conflict history. Lastly, comparing scenarios S1 to S6 does not corroborate with hypothesis H_6, as the proportion of conflict incidence is in fact higher after 40 years when democracy was introduced in Thailand in 2010.

Based on the 'experiments' I was also able to identify some of the characteristic of the three 'ideal type' processes outlined in section 4.5. The expected manifestations of ideal type 2 is both increasing conflict incidence, more inconsistency *and* more regime instability in the long run. I was able to identify all of these effects in Thailand, and Portugal, although Ethiopia, did not see any decrease in simulated country-years with a regime change when inequality was reduced. I was also able to detect 'ideal type' 1, as it predicts a swift democratization when inequality is low. 'Ideal type' 3, where an autocratic regime persists, but where there are subsequent episodes of violent conflict is ambiguously identified in Laos, although there was a slight increase in the simulated share of autocracy when inequality was low.

The importance of economic development is also salient. Countries with lower income have a much higher baseline risk of conflict, and are less likely to become democratic in the long run. Needless to say, there are some important issues of uncertainty that have to be considered. I now turn to the uncertainty and robustness of the results.

7.5 Uncertainty of the results

All scientific research must consider the issue of uncertainty (King, Keohane and Verba 1994). Potential sources of errors in the results should be discussed, and their sensitivity to alterations must be assessed. There are three issues which I find particularly pressing in this thesis; the uncertainty in my simulations, the potential bias induced through the extensive use of imputations, and the objections against capital share as a valid measure of economic inequality.[27]

7.5.1 Uncertainty in simulations

With regards to the simulations there are in particular three sources of uncertainty that may impact on the results:

Uncertainty about the **correct model specification** is highlighted by Hegre et al. (forthcoming, 29) as one source of uncertainty. As I conducted an out-of-sample evalua-

[27]The second and third issues are somewhat related, as the capital shares variable includes a considerable amount of imputed data.

tion and chosen the model with the best predictive ability, I tried to take this uncertainty into consideration. One issue worthy of critique is that I disregarded the more extensive models, even though some of the control variables were statistically related to the dependent variable. I chose the more parsimonious model, as it performed as good as the more extensive models in the out-of-sample evaluation. Nevertheless, given more time in the finishing phase of this project, initial tests should have been conducted to guard against potential erroneous effects from leaving out these controls. Preferably, I should have run the entire simulation analysis on a model including these controls as well. There may also exist other model specifications that represent the observed data in a better way than the models presented in Table 7.1. However, it would be beyond the scope of this thesis to carry out more extensive testing.

Another source of uncertainty relates to the **parameter estimates** in the multinomial logit model. The standard errors reflect the uncertainty of the parameter estimates, and by relying on 'realizations' of the parameter estimates from the variance-covariance matrix (see section 5.3.2) some of the uncertainty of the underlying model is captured in the simulation results (Hegre et al. forthcoming, 29). It also, to a certain degree, remedies the uncertainty deriving from the limited **number of simulations**. As referred to earlier on, Hegre et al. (forthcoming, 29) show that both mean and variance converge to stable estimates when the number of simulations exceed 500–1000 simulations. My dependent variable has twice as many categories and more simulations are therefore needed. I still consider 5000 simulations to be sufficient to obtain correct estimates.

The inequality measure, capital share, is a potential source of bias in several ways. The first concern is the **forecast-scenarios** used in the simulations, the second concern is its amount of missingness, and a third concern is related to measurement validity. The first issue is discussed in this section, as it directly relates to the simulations, and the other two are discussed in the following sections.

In model m1 the average capital share value between 2006–2010 is used to proxy the underlying level of economic inequality in 2010, and thus in the forecasting period. I previously argued, in section 6.2, why I consider this to be the best solution. In order to test how sensitive the simulations are to minor changes in the capital share scenarios I construct two other forecast-scenarios; one where the capital share value in each country is averaged over the last ten years, i.e. 2001–2010 (scenario 2), and another which simply extends the observed value in 2010 (scenario 3).

The results are presented in Appendix F. Figure F.1 shows that the amount of simulated global conflict in the left-hand panels is almost identical in the baseline scenario

and the two other scenarios. The same holds for the share of countries in each of the 6 regime and conflict states, as reported in the panels on the right-hand side. This is not unexpected, as the global, median capital share value is almost the same in all three forecast-scenarios.[28] Moreover, it shows that my decision to use the average capital share value between 2005-2010, instead of the 2010 value, is not expected to have been decisive for the results.

7.5.2 Are the results driven by the imputed data?

I have imputed around 50% of the observations in the capital share variable. This sounds very extensive, and indeed it is. Nevertheless, the variable has considerably better coverage, both over time and cross-sectionally, than the the Gini coefficient, the widely used proxy for economic inequality. Alternatively I could have restricted the time-frame, or removed countries with very poor coverage. However, as listwise deletion may also bias the results, and because I would lose about half of my dataset if I removed the missing observations, I considered imputation methods the best available option. Instead, in section 6.2.1, I tried to diagnose the imputed data. I did not discover anything alarmingly, although the extensive use of imputations is in itself a reason to be careful.

The imputed capital share variable has a slightly higher mean value than the capital share variable without imputations, as seen in Appendix A. This implies that the missing observations are assigned somewhat higher values, on average, than the non-missing observations. The reason why is that countries with much or all missing values on average have a lower mean value of GDP, are more often inconsistent or autocratic, and have somewhat more reported conflict. The imputed observations also have lower dispersion, as demonstrated in section 6.2.1. It is not clear-cut how exactly these relations influence the results. Given a non-causal relationship between capital share and conflict and/or regime change, the results may be slightly over confident. However, if there is an underlying effect from inequality it may also be somewhat underestimated as many of the observations are clustered around the mean. Both of these possibilities should be kept in mind. Moreover, it illustrates how difficult it is to operate with indicators that have extensive missingness. The fact that some of the countries without coverage can be

[28]The median capital share value is 0.703 in the baseline scenario, 0.710 in scenario 2, and 0.707 in scenario 3. The experiment countries in Table 7.9 are also assigned more or less the same values in the two other forecast-scenarios (the value in scenario 2 and 3 are reported after the country name): Portugal 0.55 and 0.60; Ethiopia 0.81 and 0.82; Thailand 0.78 and 0.72; and Laos 0.79 and 0.79. Thus, the scenario results are expected to be almost identical independent of which of the three forecast-scenarios is applied.

described as more 'high-risk' countries also shows that it is not entirely satisfactory to remove them from the sample either.

In order to find out whether the results are driven by the imputed data I follow Regan and Bell (2010, 7) and construct a dummy variable where the missing observations are coded 1 and the non-missing observations are coded 0, and include it in the regression. The dummy variable, firstly, tells whether the results are robust when the effects of the imputed data are controlled for. Secondly, it shows whether the missing observations are statistically more prone to being in the various states, even after controlling for the other variables (Ibid.). If the 'missing variable' significantly predicts the dependent variable there is potentially a problem with non-random missingness. The results (see Appendix G) show that for the baseline model the imputation dummy, denoted 'missing', only significantly predicts a reduction in the probability of being a 'democracy in conflict' (DC), relative to an 'autocracy in peace' (AP). To note, this is the one category I did not make predictions about in the hypotheses, and which also had a very weak effect in the multinomial logit model. The parameter estimates are not close to significant in any of the other equations. There are only minor changes in the estimates for capital share compared to model m1. Thus, there are no clear signs that the imputed data significantly alter, or bias, the results.

7.5.3 Capital share - a valid measure of inequality?

Issues of validity within the social sciences deal with whether or not the indicators used in the empirical assessment of theory meaningfully capture the ideas contained in the concepts (Adcock and Collier 2001, 531). One pressing issue in this thesis is whether capital share is a valid measure of inequality.

As discussed in section 6.2, inequality has many manifestations; social, political, economic etc. In the context of democratization and civil conflict I decided to focus on economic inequality, because it is the type of inequality most directly linked to the theoretical literature I base the analytical framework on (see Boix 2003; Acemoglu and Robinson 2006). I consider economic inequality to capture important underlying driving force in the tug-of-war over the political institutions. That being said, what are the main caveats of capital share?

Capital share is, first of all, a very crude measure. It only captures a certain type of economic inequality, namely the relative income between registered laborers and capital owners. When capital shares increase less of the surplus goes to the workers, and hence, economic inequality is presumed to increases. An important consideration is the relation

between the industrial sector and other productive sectors of the economy. Capital share does, by definition, not capture economic inequality in rural communities or in countries where the economy is predominantly agriculturally based. Case studies from Latin America (see e.g. Booth 1991) show how mobilization both within agrarian communities and urban settlements comprised of wage-laborers contributed to the strength and support of the popular upheavals. Incorporating both sectors also allows better capture the second element of Boix' theory, namely the degree of asset specificity, i.e. whether a lot of wealth is held in land rather than in capital. I chose to look more broadly at economic inequality, and disregarded degree of asset specificity in the theoretical argument, as well as in the empirical investigation. This simplification is in line with Houle (2009). Further, capital share does also not capture redistribution of income from the government. The effect from government redistribution is expected to increase the inequality gap between rich and poor nations further, as it is more often affluent, Western countries that have strong welfare states and a well-developed system of government transfers.

As capital share derives from the value added in the industrial sector, the measure is sensitive to fluctuations in the economy and to specific economic activity. Countries that receive large revenues from natural resource wealth, like oil or gas production, may on average have more fluctuating and higher capital shares. I controlled for countries that were oil and gas exporters in the more extensive models, but chose not to rely on them in the simulations, as argued above. Moreover, I also tested a model which included an interaction term between capital share and the natural resource indicator. The interaction term was not significant, and the model's overall performance was far poorer than the baseline model (see Appendix D and Table 7.2).

To sum up, capital share capture economic inequality in capital intensive industries, but is expected to less well capture economic inequality in low-capital, tertiary sector or agricultural societies where inequalities are found along other dimensions. It is therefore expected to predict democratization in middle-income countries relatively well, but civil conflict in low-income agriculturally based societies less well. Clearly not flawless, there are not very many good alternative indicators of economic inequality. Moreover, as the world has become increasingly industrialized over the last decades, and the trend is expected to continue, capital share may prove to become a more relevant and precise measure of economic inequality in the years to come.

Chapter 8

Conclusion

The dynamic relationship between violence and institutional changes was the point of departure of my thesis. In line with two prominent contributions within the literature on democratization, Boix (2003) and Acemoglu and Robinson (2006) (A&R), I have argued that high inequality creates more demand for democratization and more willingness to use violence, either to introduce democracy or to preserve the autocratic institutions.

The core of A&R's (2006) and Boix' (2003) argument is that economic inequality proxy the level of social conflict over the political institutions between the citizens and the elites. The dispute concerns the redistribution of wealth and resources in society – the citizens prefer extensive redistribution whereas the elites prefer zero redistribution. The main prediction of Boix is that democratization is impossible in unequal societies, because the costs of redistribution that accrue to the elites in a democracy are high. A&R, on the contrary, predict that both high and low levels of inequality is a hindrance for democratization; at high levels the democratic equilibrium is unattainable, whereas at low levels the demand for democracy is low or non-existing.

I developed an analytical framework drawing predominantly on the arguments of Boix (2003) and Acemoglu and Robinson (2006). However, inspired by the vast array of quantitative studies that investigate the link between institutional changes and civil conflict (see e.g Hegre et al. 2001; Gleditsch, Hegre and Strand 2009; Cederman, Hug and Krebs 2010), I placed more emphasis on democratization's effect on civil conflict. As it is found that democratization increases the risk of conflict initially, I wanted to investigate how the two phenomena related to each other over time, taking the reciprocal causality between regime changes and conflict into account.

The framework highlighted three 'ideal types' of democratization: a swift, peaceful, and complete democratization; a long-drawn, conflict prone, and incomplete democra-

tization; and a failed attempt at democratization, where the authoritarian institutions endure and there is more or less violence. With these 'ideal types' as a point of departure I focused on how economic inequality shapes the course and outcome of democratization, as well as its level of violence. I proposed that inconsistent regimes may constitute a 'compromise' solution between the elites and the citizens when an autocratic or democratic institutional equilibrium is impossible to achieve. Thus, I argued that high inequality is not necessarily an impediment for democratization, but it makes it difficult to complete.

Two questions guided the inquiry. I tried to understand whether *economic inequality increases the risk of regime changes and of being in a state of civil conflict*, and to find out *how economic inequality relates to democratization and civil conflict in the long run.*

8.1 Summary of main findings - and their implications

My findings indicate that economic inequality may have an impact on the course and outcome of democratization and for the amount of civil conflict observed over time. In the analysis, higher inequality was related to more conflict incidence in the long run, as well as a more country-years with inconsistent institutions. The results seem to imply that high inequality makes democratization more violent and long-drawn. In the following I briefly summarize some of the main findings of the empirical analysis.

I started the analysis investigating the short-run effects of economic inequality, using multinomial logistic regression analysis of quantitative data from 164 countries observed between 1960 to 2010. To measure economic inequality I used the variable *capital share*, defined as the the share of value added in the industrial sector that accrues to the owners of capital, as opposed to the share that goes directly to the laborers. When capital share is high, economic inequality is expected to be high.

The short-run analysis provided some support for the argument that higher inequality increases the risk of being in a state of civil conflict, and the risk of regime changes. For example, increasing capital share by 1 standard deviation increased log odds of being an autocracy in conflict, relative to an autocracy in peace, by 1.36. The estimate was significant at the 10% level. The regression coefficients all pointed in the directions expected from the analytical framework, although they were not as statistically strong as anticipated: Increasing capital share by 1 standard deviation increased log odds of being an inconsistent regime in peace with 0.7, and for being an inconsistent regime in conflict with 1.52, relative to an autocracy in peace. Moreover, increasing capital share decreased

log odds of being a democracy in peace, relative to an autocracy in peace.

For the purpose of investigating how economic inequality related to democratization and civil conflict in the long run, I extended a simulation procedure developed in Hegre et al. (forthcoming). Based on the parameter estimates from the multinomial logit model the simulation program calculates probabilities for being in each regime and conflict state, draws realized outcomes from a probability distribution, and updates the transition probabilities based on the the estimates in the variance-covariance matrix and values assigned to the exogenous and endogenous explanatory factors. The procedure allowed me to capture the overall impact of inequality on the endogenous relationship between regime changes and civil conflict.

In order to isolate the effect of economic inequality I conducted 5 counterfactual 'experiments' at the country-level. By comparing the simulation results when conflict status and political regimes were determined by the observed history without any intervention (the 'baseline scenario'), with the results after a hypothetical change in a country's capital share value from 2010 onwards, I was able to investigate the effects of inequality on on the long-run incidence of conflict and on the prospects of democratization.

In line with the 'ideal type' 2, the long-drawn and conflict prone democratization process, I found that high inequality was associated with more civil conflict in the long run, regardless of the initial regime type of a country. After 40 years, 'excess conflict', i.e. the difference in simulated proportion of conflict when economic inequality is reduced by 1.6 standard deviation on the capital share variable, was 12.56% in Ethiopia. Hence, moving from an initially high to a low level of inequality leads to a relative risk reduction of 21% of conflict for the poor and conflict prone country Ethiopia, and other countries similar to it.

High inequality was also associated with more institutional inconsistency in the long run, the second feature of ideal type 2. Further, I found a slightly higher share of regime changes under high inequality in Thailand, Portugal and Laos. The findings corroborate with the two components of my analytical framework, namely that high(er) inequality both increases the risk of conflict and regime changes.

In line with 'ideal type' 1, the swift and peaceful democratization, I found an increase in the reported share of democracy in all scenarios where capital share was reduced to a low level. An example is Ethiopia, where 'excess democracy' was 9.25% after 40 years. However, I also found a higher share of autocracy in Laos in the long run. This also corroborate with the theoretical propositions, as consistent institutions are expected to be more compatible with low levels of inequality; in an autocratic regime the demand for

democracy is less pronounced, whereas in a democracy low inequality is compatible with a stable democratic equilibrium.

The reason why the long-run effects are quite substantial relates to two mechanisms at work in the statistical model. Economic inequality has, firstly, a direct effect on the risk of conflict and on regime changes, and secondly, an indirect effect through the reciprocal causality between regime changes and conflict. I calculated the direct effect of inequality on the predicted probability of Thailand being in a state of inconsistency and conflict. I found that by reducing capital share in Thailand with the same amount as was done in the simulation 'experiment', i.e. approximately 1.5 standard deviation on the capital share variable, only led to a a 1.62% reduction in the risk of being an inconsistent regime in conflict. Moreover, when I estimated the long-run effect of a reduction in capital share based on a transition matrix that assumed constant transition probabilities, I found that 'excess conflict' was 4.2% after 40 years. This was somewhat less than the effects observed in the simulation results, where 'excess conflict' was 6.56% after 40 years. The findings show that without making regime changes and conflict endogenous, and applying the simulation procedure, the total effect of a reduction in inequality on the long-run level of conflict is underestimated. No previous study has attempted to capture all of these mechanisms simultaneously. This analysis shows that it is important, in order to get a more precise impression of the aggregated effects of socio-economic changes on civil conflict and political institutions.

Substantively the results from the analysis are not trivial and they have some implications. One implication is that countries with high inequality that have started to democratize, like Thailand and Ethiopia, are likely to see positive effects from redistribution. For Thailand, redistribution seems to be particularly pertinent. The country has experienced considerable increases in income per capita the last 20 years, but inequalities have increased simultaneously. Moreover, violent conflict and political instability proves persistent, and is to some extent blamed on increasing inequality and social tensions (Warr 2007). If the results are taken at face value, reducing economic inequality may both increase Thailand's, and other similar countries', chances of civil peace and a swifter democratization. This is in line with, amongst others, Acemoglu and Robinson (2001), who highlight the difficulty of consolidating democracy in unequal societies.

8.2 Discussion

In order to further evaluate the value-added of this study it is necessary to compare and contrast my findings with those from other relevant studies. It is also important to identify the main caveats, and the areas of improvement.

This thesis speaks to several fields of studies. It first of all relates to the democratization literature. Second, it speaks to studies which have not found an effect from income inequality on the risk of civil conflict. Third, it is relevant for studies that deal with the conflict-inducing effects from democratization and regime changes.

My analytical framework relies heavily on the works of Boix (2003) and Acemoglu and Robinson (2006). Boix and A&R did not conduct extensive statistical testing, but there are others who have scrutinized their theories. One of them is Houle (2009). He does not find statistical effects of inequality, measured by capital shares, on the probability of democratization. Although at first glance my findings seem to contradict with Houle's conclusion, a further inspection reveals that the discrepancy may not be as great.

First, Houle only includes transitions directly from autocracy to democracy in his analysis, and disregards the inconsistent regimes. As he is interested in testing the theories of Boix and A&R, his choice of a dichotomous regime classification is reasonable. However, in line with Epstein et al. (2006), I have argued that failing to distinguish between autocracies and democracies and the inconsistent regimes, may disguise important dynamics of democratization.

Second, although Houle (2009) does not find any evidence that economic inequality matters for democratization, he does find evidence that high economic inequality destabilizes democracy, and hence, that inequality is bad for democratic consolidation. My findings also seem to point in this direction, as high inequality is related with a lower simulated share of democracy in the long run.

Third and, in my opinion, most importantly, Houle does not study long-run effects. As argued in the above section, this seems to matter; when I capture the reciprocal causality between regime changes and civil conflict, the effects of lower economic inequality on democratization are exacerbated.

One notable study within the civil war literature that reject grievances as a source of rebellion is Collier and Hoeffler (2004). The main difference between this study and theirs, is that Collier and Hoeffler include income inequality measured as a Gini coefficient. To use the Gini coefficient to scrutinize the effect of inequality on conflict onset has been common practice (see e.g. Fearon and Laitin 2003; Hegre, Gissinger and Gleditsch 2003). I have discussed the limitations of the Gini coefficient on several occasions, and to sum

up; due to extensive missingness and a small sample size, it is not unexpected that there is lack of evidence of a statistical relationship between inequality and civil conflict. This problem is noted by Fearon and Laitin (2003, 85) who say that "the poor quality of the inequality data, [...], does not allow us to go beyond the claim that there appears to be no powerful cross-national relationship between inequality and onset [...]." That I find a stronger connection between inequality and conflict may be attributed to the use of the capital share variable, and the fact that I applied multiple imputation technique to reduce extent of missingness. This has not been done in any of the previous studies.

This thesis can hopefully supplement previous research that has focused on the short-run effects of institutional changes in general, and democratization in particular, on the risk of civil conflict (e.g. Cederman, Hug and Krebs 2010; Gleditsch, Hegre and Strand 2009; Hegre et al. 2001). My preliminary conclusion is that democratization *may* reduce the level of conflict in the long run, but that democratization is not necessarily a peace-promoter. How unstable the transition period is, i.e. the amount of regime changes observed during its course, seems to impact on the amount of civil conflict. The findings are in line with the arguments of Hegre et al. (2001, 44). Hegre and his colleagues also stated that a country is more likely to end up in the democratic equilibrium in the long run. My results suggest that this conclusion may have to be modified. How unequal a society is could have an impact on the possibility of reaching a stable equilibrium, as higher inequality in most of my country-cases was connected with a somewhat higher share of regime changes, and more institutional inconsistency.

The relationship between economic inequality and GDP per capita, i.e. between relative income and absolute income, has lurked in the background throughout the analysis. Previous studies have, as described above, dismissed the significance of inequality. In my opinion, it has been prematurely. The long-run analysis showed an effect from economic inequality, controlling for GDP per capita. I expect that if I had removed economic inequality from the equation, the already strong income effect would have been exacerbated. As there is a high correlation between capital share and GDP per capita, the independent effect from inequality would then be captured by GDP per capita instead.

8.2.1 Some caveats

As highlighted above, I believe this thesis makes some valuable contributions, both to the literature on democratization, as well as to the literature on institutional changes and civil conflict. However, I embarked on an ambitious project and acknowledge that there are short-comings to this study. There are arguments that could have been more carefully

developed, and improvements in the analysis I should have made. Particularly time-, but also space constraints, unfortunately did not make it possible to deal with all of these issues. I will briefly mention some caveats, the areas of improvement I find most pressing, and suggest some remedies for future studies.

Although my findings indicate an effect of economic inequality in the long-run, there is uncertainty connected to the results. First, I have not been able to estimate the exact uncertainty related to for example 'excess conflict'. During the work with this thesis I made attempts at constructing such estimates based on the information from the simulation results. It proved difficult, and the lack of time in the finishing process unfortunately did not allow me to pursue this further.

Second, it is difficult to separate the uncertainty of the transition matrix from the statistical uncertainty. There are events outside of my model that over time may impact on for example Thailand's transition probabilities. This will also reflect in the long-run effects. However, if accepting the underlying statistical model as valid, economic inequality does matter; the long-run effects from changes to the level of inequality are relatively large, and of substantial interest.

Third, the analytical framework focuses on civil conflict that emerge because of an underlying struggle over the political institutions. In the Uppsala/PRIO armed conflict dataset they are so-called 'conflicts over government'. I argued, in section 6.1.2, why I did not prefer to exclude conflicts which are coded with a different incompatibility. It is not always easy to code a conflict as one or the other, and sometimes the incompatibility may change during the course of the conflict. The multinomial logit analysis therefore included all observed civil conflicts. Moreover, in the results I assume that all conflicts are conflicts over government, although I realize that it is a crude assumption.

Fourth, I have relied on extensive use of imputed data in this analysis. This is by no means an ideal solution, but I have argued why I considered it to be the best available option. I tried to test the imputed data in line with suggestions made in previous studies (see e.g. Abayomi, Gelman and Levy 2008), in order to reveal potential problems that could introduce bias in my results. I did not find this to be the case. Still. there is still good reason to try to avoid such an extensive use of imputed data in the future, to reduce uncertainty. However, listwise deletion is, in my opinion, not a better option, and indicators with better coverage should rather be sought.

The model of regime and conflict transitions should also be improved. For once, it should be extended with interaction terms between the main explanatory variables and the regime and conflict status at time t-1, in order to estimate explicitly institutional

transitions in the short-run analysis. Second, simulation with more complex models that include additional control variables could brush any doubts concerning omitted variable bias. For example, both the variable for oil and gas exporter and education proved to have statistical effects in the multinomial logit models, and should be further tested.

One of the main challenges of this study has been the search of appropriate measures of economic inequality. *Capital share* was my attempt at introducing a more valid measure, but during the work with this thesis it became clear that this variable also has some drawbacks. As previously highlighted, capital share only captures inequality between capital owners and laborers in the industrial sector. This is perhaps satisfactory when estimating the effects on democratization in middle-income countries, but it is less valid for in low-income, agrarian societies. Constructing new indicators with the ability to capture aspects of economic inequality relevant to both the study of civil conflict *and* democratization is therefore a next important step. A second problem with capital share is its correlation with per capita income. Future studies should preferably construct indicators which are not as closely related to GDP, to assure valid measurement.

8.2.2 To conclude

This thesis suggests that complete democratization may be more difficult when economic inequality is high. Moreover, high inequality is expected to make the democratization process more violent. The findings highlight the importance of not only focusing on the level of income, but also on the distribution of income, in studies of democratization and civil conflict. It is difficult to disentangle the causality between inequality and income. However, the case of Thailand and some Latin American countries, as highlighted in this thesis, show that for low-income countries, economic growth may exacerbate an already skewed income distribution. The crux seems to be how to induce sustainable economic growth and at the same time reduce persistent economic inequalities, so that it is possible reduce the level of conflict, and introduce durable democratic regimes.

Chapter 9

Bibliography

Abayomi, Kobi, Andrew Gelman and Marc Levy. 2008. "Diagnostics for multivariate imputations." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 57(3):273–291.

Acemoglu, Daron and James A. Robinson. 2001. "A Theory of Political Transitions." *American Economic Review* 91(4):938–963.

Acemoglu, Daron and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy.* New York: Cambridge University Press.

Adcock, Robert and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95(3):529–546.

Alan Heston, Robert Summers and Bettina Aten. 2011. "Penn World Table Version 7.0." Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, May 2011.

Barro, Robert J. 1999. "Determinants of Democracy." *Jounal of Political Economy* 107(6):158–183.

Barro, Robert J. 2000. "Inequality and Growth in a Panel of Countries." *Journal of Economic Growth* 5(1):5–32.

Beetham, David. 1999. *Democracy and Human Rights.* Cambridge: Polity Press.

Boix, Carles. 2003. *Democracy and Redistribution.* Cambridge: Cambridge University Press.

Boix, Carles. 2008. "Economic Roots of Civil Wars and Revolutions in the Contemporary World." *World Politics* 60(3):390–437.

Boix, Carles and Susan C. Stokes. 2003. "Endogenous Democratization." *World Politics* 55(4):517–549.

Bollen, Kenneth A. 1979. "Political Democracy and the Timing of Development." *American Sociological Review* 44(4):572–587.

Bollen, Kenneth A. and Robert W. Jackman. 1995. "Income Inequality and Democratization Revisited: Comment on Muller." *American Sociological Review* Vol. 60, No. 6:pp. 983–989.

Booth, John A. 1991. "Socioeconomic and Political Roots of National Revolts in Central America." *Latin American Research Review* 26(1):33–73.

Bryman, Alan. 2004. *Social Research Methods*. Oxford: Oxford University Press.

Burkhart, Ross E. and Michael S. Lewis-Beck. 1994. "Comparative Democracy: The Economic Development Thesis." *American Political Science Review* 88(4):903–910.

Cederman, Lars-Erik, Nils B. Weidmann and Kristian S. Gleditsch. 2011. "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison." *American Political Science Review* 105(3):478–495.

Cederman, Lars-Erik, Simon Hug and Lutz F. Krebs. 2010. "Democratization and civil war: Empirical evidence." *Journal of Peace Research* 47(4):377–394.

Clague, Christopher, Philip Keefer, Stephen Knack and Mancur Olson. 2003. Property and Contract Rights in Autocracies and Democracies. In *Democracy, Governance and Growth*, ed. Stephen Knack. Ann Arbor: The University of Michigan Press pp. 78–109.

Collier, Paul. 2000. "Rebellion as a Quasi-Criminal Activity." *Journal of Conflict Resolution* 44(6):839–853.

Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(4):563–595.

Collier, Paul, Anke Hoeffler and Måns Söderbom. 2004. "On the Duration of Civil War." *Journal of Peace Research* 41(3):253–273.

Collier, Paul, Anke Hoeffler and Måns Söderbom. 2008. "Post-conflict risks." *Journal of Peace Research* 45(4):461–478.

Collier, Ruth Berins. 1999. *Paths Towards Democracy: The Working Class and Elites in Western Europe and South America*. Cambridge: Cambridge University Press.

Dahl, Robert A. 1971. *Polyarchy: Political Participation and Opposition*. New Haven: Yale University Press.

Deininger, Klaus and Lyn Squire. 1996. "A New Data Set Measuring Income Inequality." *The World Bank Economic Review* 10(1):565–591.

Diamond, Larry. 1992. "Economic Development and Democracy Reconsidered." *American Behavioral Scientist* 35(4):450–499.

Diamond, Larry. 2002. "Thinking About Hybrid Regimes." *Journal of Democracy* 13(2):21 – 35.

Dunning, Thad. 2008. *Crude Democracy: Natural Resource Wealth and Political Regimes*. Cambridge: Cambridge University Press.

Easterly, William and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *The Quarterly Journal of Economics* 112(4):1203–1250.

Eckstein, Harry. 1973. "Authority Patterns: A Structural Pattern for Inquiry." *American Political Science Review* 67(4):1142–1161.

Epstein, David L., Robert Bates, Jack Goldstone, Ida Kristensen and Sharyn O'Halloran. 2006. "Democratic Transitions." *American Journal of Political Science* 50(3):551–569.

Fearon, James D. 1995. "Rationalist explanations for war." *International Organization* 49(3):379–414.

Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.

Filiu, Jean Pierre. 2011. *Yhe Arab Revolustion. Ten Lessons form the Democratic Uprising*. London: Hurst & Company.

Gandhi, Jennifer and Adam Przeworski. 2007. "Authoritarian Institutions and the Survival of Autocrats." *Comparative Political Studies* 40(11):1279–1301.

Gates, Scott, Håvard Hegre, Mark P. Jones and Håvard Strand. 2006. "Institutional Inconsistency and Political Instability: Polity Duration, 1800–2000." *American Journal of Political Science* 50(4):893–908.

Gleditsch, Kristian S. and Michael D. Ward. 1999. "A Revised List of Independent States since the Congress of Vienna." *International Interactions* 25(4):393–413.

Gleditsch, Nils Petter, Håvard Hegre and Håvard Strand. 2009. Democracy and Civil War. In *Handbook of War Studies III*, ed. Manus Midlarsky. Ann Arbor, MI: University of Michigan Press pp. 155–192.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg and Håvard Strand. 2002. "Armed Conflict 1946–2001: A New Dataset." *Journal of Peace Research* 39(5):615–637.

Goldstone, Jack A., Robert H. Bates, David L. Epstein, Ted Robert Gurr, Michael B. Lustik, Monty G. Marshall, Jay Ulfelder and Mark Woodward. 2010. "A Global Model for Forecasting Political Instability." *American Journal of Political Science* 54(1):190–208.

Greene, William H. 1997. *Econometric Analysis*. 3rd. ed. Upper Saddle River, NJ: Prentice-Hall.

Grossman, Herschell I. 1991. "A General Equilibrium Model of Insurrections." *The American Economic Review* 81(4):912–921.

Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.

Gurr, Ted Robert. 1974. "Persistence and Change in Political Systems, 1800–1971." *American Political Science Review* 68(4):1482–1504.

Gurr, Ted Robert. 2000. *People Versus States, Minorities at Risk in the New Century*. Washington D.C.: United States Institute of Peace Press.

Gurr, Ted Robert and Will H. Moore. 1997. "Ethnopolitical Rebellion: A Cross-Section Analysis of the 1980s with Risk Assessments for the 1990s." *American Journal of Political Science* 41(4):1079–1103.

Hadenius, Axel and Jan Teorell. 2005*a*. "Assessing Alternative Indicies of Democracy." Quality of Government Institute, University of Gothenburg.

Hadenius, Axel and Jan Teorell. 2005*b*. "Cultural and Economic Prerequisites of Democracy: Reassessing Recent Evidence." *Studies in Comparative International Development* 39(4):87–106.

Hegre, Håvard, Carl Henrik Knutsen and Espen Geelmuyden Rød. 2012. "The Determinants of Democracy: A Sensitivity Analysis." Unpublished.
URL: *http://folk.uio.no/hahegre/Papers/Determinantsdemocracy.pdf*

Hegre, Håvard, Håvard Mokleiv Nygård, Håvard Strand, Scott Gates and Ranveig Drolsum Flaten. 2011. "The Conflict Trap." Presented to American Political Science Association Annual Meeting 2011, Seattle, WA.
URL: *http://folk.uio.no/haavarmn/ConflictTrapPaper.pdf*

Hegre, Håvard, Joakim Karlsen, Håvard Mokleiv Nygård, Håvard Strand and Henrik Urdal. forthcoming. "Predicting Armed Conflict 2010–2050." *International Studies Quarterly* 2013.
URL: *http : //folk.uio.no/hahegre/Papers/PredictionISQ_Final.pdf*

Hegre, Håvard and Nicholas Sambanis. 2006. "Sensitivity Analysis of Empirical Results on Civil War Onset." *Journal of Conflict Resolution* 50(4):508–535.

Hegre, Håvard, Ranveig Gissinger and Nils Petter Gleditsch. 2003. Globalization and Internal Conflict. In *Globalization and Armed Conflict*, ed. Gerald Schneider, Katherine Barbieri and Nils Petter Gleditsch. Lanham, MD: Rowman and Littlefield pp. 251–276.

Hegre, Håvard, Tanja Ellingsen, Scott Gates and Nils Petter Gleditsch. 2001. "Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816–1992." *American Political Science Review* 95(1):33–48.

Hewitt, J. Joseph. 2008. The Peace and Conflict Instability Ledger: Ranking States on Future Risks. In *Peace and Conflict 2008*, ed. J. Joseph Hewitt Hewitt, Jonathan Wilkenfeld and Ted Robert Gurr. Boulder, CO: Paradigm.

Honaker, James and Gary King. 2010. "What to Do about Missing Values in Time-Series Cross-Section Data." *American Journal of Political Science* 54(2):561–581.

Honaker, James, Gary King and Michael Blackwell. 2011. "Amelia II: A Program for Missing Data." *Journal of Statistical Software* 45:1–47.

Horowitz, Donald L. 1985. *Ethnic groups in conflict*. Berkeley, CA: University of California Press.

Hosmer, David W and Stanley Lemeshow. 2000. *Applied Logistic Regression. Second Edititon.* New York, NY: Wiley.

Houle, Christian. 2009. "Inequality and Democracy: Why Inequality Harms Consolidation but Does Not Affect Democratization." *World Politics* 61(4):589–622.

Høyland, Bjørn and Håvard Nygård. 2012. "Missing Uncertainty and Uncertain Missingness in the Study of Civil War Onset.".

Huntington, Samuel P. 1968. *Political Order in Changing Societies.* New Haven, CT and London: Yale University Press.

Huntington, Samuel P. 1996. *The Clash of Civilizations and the Remaking of World Order.* New York: Simon & Schuster.

Hurwitz, Leon. 1973. "Contemporary Approaches to Political Stability." *Comparative Politics* 5:449–463.

Inglehart, Ronald and Christian Welzel. 2006. *Modernization, Cultural Change and Democracy –The Human Development Sequence.* Cambridge: Cambridge University Press.

Karl, Terry Lynn. 1990. "Dilemmas of Democratization in Latin America." *Comparative Politics* 23(1):1–21.

King, Gary and Langche Zeng. 2001. "Logistic Regression in Rare Events Data." *Political Analysis* 9(2):137–163.

King, Gary, Michael Tomz and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2):347–361.

King, Gary, Robert O. Keohane and Sidney Verba. 1994. *Designing Social Inquiry. Scientific Inference in Qualitative Research.* Princeton, NJ: Princeton University Press.

Kirkpatrick, David D. 2011. "Military Flexes its Muscles as Islamists Gain in Egypt." Online.
URL: *http : //www.nytimes.com/2011/12/08/world/middleeast/egyptian-general-mukhtar-al-mulla-asserts-continuing-control-despite-elections.html?_r = 2*

Knutsen, Carl Henrik. 2011. The Economic Effects of Democracy and Dictatorship PhD thesis Department of Political Science, University of Oslo Oslo: .
URL: *http://folk.uio.no/carlhk/publications/CHKdissertationForPrint.pdf*

Lambert, Peter J. 2001. *The Distribution and Redistribution of Income.* New York: Palgrave.

Lichbach, Mark I. 1989. "An Evaluation of Does Economic-Inequality breed Political-Conflict studies." *World Politics* 41(4):431–470.

Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53(1):69–105.

Little, Roderick, J. A. and Donald B. Rubin. 2002. *Statistical Analysis With Missing Data.* 2nd ed. Hoboken, NJ: Wiley.

Little, Walter. 1997. *Democratization.* Cambridge: Polity Press chapter Democratization in Latin America, 1980-1995, pp. 174–194.

Lutz, Wolfgang; Anne Goujon, Samir K.C. and Warren Sanderson. 2007. Reconstruction of Population by Age, Sex and Level of Educational Attainment for 120 Countries 1970–2000. Technical report IIASA Laxenburg, Austria: .
URL: *http://www.iiasa.ac.at/Admin/PUB/Documents/IR-07-002.pdf*

Maddison, Angus. 2006. *The World Economy.* Paris: OECD Publishing.

Mansfield, E. D. and G. H. Snyder. 2002. "Democratic transitions, institutional strength, and war." *Internatiional Organization* 56(2):297–337.

Marshall, Monthy G. and Keith Jaggers. 2002. "Polity IV Project – Dataset Users' Manual." College Park: Program Center for International Development and Conflict Management, University of Maryland. Manual.

McFaul, Michael. 2004. "Democracy promotion as a world value." *The Washington Quarterly* 28(1):147–163.

Meltzer, Allan and Scott Richards. 1981. "A rational theory of the size of government." *Journal of Political Economy* 89(5):914–927.

Muller, Edward N. 1988. "Democracy, Economic Development, and Income Inequality." *American Sociological Review* 53(1):50–68.

Muller, Edward N. 1995. "Income Inequality and Democratization: Reply to Bollen and Jackman." *American Sociological Review* 60(6):990–996.
URL: *http://www.jstor.org/stable/2096437*

Muller, Edward N. and Erich Weede. 1990. "Cross-National Variations in Political Violence: A Rational Action Approach." *Journal of Conflict Resolution* 34(4):624–651.

Muller, Edward N. and Erich Weede. 1998. "Rebellion, Violence and Revolution: A Rational Choice Perspective." *Journal of Peace Research* 35(1):43–59.

News, BBC. 2012. "Guide: Syria Crisis.".
URL: *http* : *//www.bbc.co.uk/news/world-middle-east-13855203*

Olson, Mancur. 1965. *The Logic of Collective Action.* MA: Harvard University Press.

Ortega, Daniel and Francisco Rodriguez. 2006. "Are Capital Shares Higher in Poor Countries? Evidence from Industrial Surveys." Unpublished manuscript.
URL: *http*: *//frrodriguez.web.wesleyan.edu/docs/working_papers/Are_capital_shares.pdf*

Østby, Gudrun. 2008. "Polarization, Horizontal Inequalities and Violent Civil Conflict." *Journal of Peace Research* 45(2):143–162.

Przeworski, Adam. 1991. *Democracy and the Market: Political and Economic Reforms in Eastern Europe and Latin America.* Cambridge: Cambridge University Press.

Przeworski, Adam and Fernando Limongi. 1997. "Modernization: Theory and Facts." *World Politics* 49(2):155–183s.

Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub and Fernando Limongi. 2000. *Democracy and Development. Political Institutions and Well-Being in the World, 1950–1990.* Cambridge: Cambridge University Press.

Raknerud, Arvid and Håvard Hegre. 1997. "The Hazard of War: Reassessing the Evidence for the Democratic Peace'." *Journal of Peace Research* 33(4):385–404.

Regan, Patrick M. and Sam R. Bell. 2010. "Changing Lanes or Stuck in the Middle: Why Are Anocracies More Prone to Civil Wars?" *Political Research Quarterly* 63(4):747–759.

Reynal-Querol, Marta. 2002. "Ethnicity, Political Systems, and Civil Wars." *Journal of Conflict Resolution* 46(1):29–54.

Ross, Michael. 2001. "Does Oil Hinder Democracy." *World Politics* 53(3):325 –61.

Ross, Michael. 2004. "What Do We Know About Natural Resources and Civil War." *Journal of Peace Research* 41(3):337–356.

Ross, Michael. 2006. "A Closer Look at Oil, Diamonds and Civil War." *Annual Review of Political Science* 9(1):265–300.

Rost, Nicolas, Gerald Schneider and Johannes Kleibl. 2009. "A Global Risk Assessment Model for Civil Wars." *Social Science Research* 38(4):921–933.

Sambanis, Nicholas. 2004*a*. "Using Case Studies to Expand Economic Models of Civil War." *Perspectives on Politics* 2(2):259–279.

Sambanis, Nicholas. 2004*b*. "What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* 48(6):814–858.

Schatzman, Christina. 2005. "Political Challenge in Latin America: Rebellion and Collective Protest in an Era of Democratization." *Journal of Peace Research* 42(3):291–310.

Schumpeter, Joseph A. 1976. *Capitalism, Socialism and Democracy.* New York: Harper-Perennial.

Skog, Ole-Jørgen. 2009. *Å Forklare Sosiale Fenomener. En regresjonsbasert Tilnrming.* Gyldendal Norsk Forlag.

Small, Melvin and David J. Singer. 1992. *Resort to Arms: International and Civil War 1816-1980.* Beverly Hills: Sage.

Stewart, Frances. 2002. Horizontal Inequalities: A Neglected Dimension of Development. QEH Working Paper No. 1 Queen Elisabeth House, University of Oxford.
URL: *http://www.crise.ox.ac.uk/pubs/workingpaper1.pdf*

Stewart, Frances, Graham Brown and Luca Mancini. 2005. Why Horizontal Inequalities Matter: Some Implications for Measurement. CRISE (Centre for Research on Inequality, Human Security and Ethnicity) Working Paper No. 19 Queen Elizabeth House, University of Oxford.
URL: *http://www.crise.ox.ac.uk/pubs/workingpaper19.pdf*

Strand, Håvard. 2007. Reassessing the Civil Democratic Peace PhD thesis Department of Political Science, University of Oslo.

Strand, Håvard, Håvard Hegre, Scott Gates and Marianne Dahl. 2012. "Democratic Waves? Global Patterns of Democratization, 1816-2008." Unpublished.
URL: *http://folk.uio.no/hahegre/Papers/WhyWaves_2012.pdf*

Taylor, Howard M. and Samuel Karlin. 1998. *An Introduction to Stochastic Modeling, 3rd ed.* San Diego, CA and London: Academic Press.

Themnér, Lotta and Peter Wallensteen. 2011. "Armed conflict, 1946–2010." *Journal of Peace Research* 48(4):525–536.

UCDP/PRIO. 2010. *Armed Conflict Dataset Codebook Version 4-2009*. Uppsala Conflict Data Program (UCDP) Centre for the Study of Civil Wars, The Peace Research Institute Oslo (PRIO).
URL: *http://www.prio.no/CSCW/Datasets/Armed-Conflict/UCDP-PRIO/Old-Versions/929/*

United Nations. 2007. *World Population Prospects. The 2006 Revision*. New York: UN.

Vanhanen, Tatu. 2000. "A New Dataset for Measuring Democracy, 1810-1998." *Journal of Peace Research* 37(2):251–265.

Vreeland, James Raymond. 2008. "The Effect of Political Regime on Civil War: Unpacking Anocracy." *Journal of Conflict Resolution* 52(3):401–425.

Waner, Stefan. 2004. "Summary of Chapter 9 in Finite Mathematics/Finite Mathematics & Applied Calculus, Topic: Markov Systems." Online.
URL: *http://people.hofstra.edu/Stefan_Waner/RealWorld/Summary8.html*

Ward, Michael D., Brian D. Greenhill and Kristin M. Bakke. 2010. "The Perils of Policy by p-value: Predicting civil conflicts." *Journal of Peace Research* 47(4):363–375.

Warr, Peter. 2007. "Long-term Economic Performance in Thailand." *ASEAN Economic Bulletin* 24(1):138–163.

Weingast, Barry R. 1997. "The Political Foundations of Democracy and the Rule of Law." *American Political Science Review* 91(2):245–263.

World Bank. 2010. "World Development Indicators." Washington, DC: The World Bank.
URL: *http://data.worldbank.org/data-catalog/world-development-indicators/wdi-2010*

Zanger, Sabine C. 2000. "A Global Analysis of the Effect of Political Regime Changes on Life Integrity Violations, 1977–93." *Journal of Peace Research* 37(2):213–233.

Appendices

Appendix A

Descriptive statistics

A.1 Coding of transition periods on the regime type variable

Table A.4: Coding of transition periods on *regime type* variable: Country, time-period, regime type, and conflict status

Country name	**Time period**	**Regime type coded**	**Observed conflict (1 or more years)**
Haiti	1986–1987	Autocracy	No
Haiti	1999	inconsistent	No
Haiti	2004–2005	Autocracy	Yes
Dominican Republic	1961	Autocracy	No
Dominican Republic	1963–1965	Democracy	Yes
Guatemala	1985	Autocracy	Yes
Honduras	1980–1981	Autocracy	No
El Salvador	1979–1983	Inconsistent	Yes
Nicaragua	1979–1980	Autocracy	Yes
Ecuador	2005	Inconsistent	No
Peru	1978–1979	Autocracy	No
Peru	2000	Inconsistent	No
Brazil	1964	Inconsistent	No
Uruguay	1964–1965	Democracy	Yes
Spain	1975–1977	Autocracy	No
Portugal	1974–1975	Autocracy	No
German Democratic Republic	1989-1990	Autocracy	No
Hungary	1989	Inconsistent	No
Czechoslovakia	1968	Autocracy	No
Croatia	1999	Inconsistent	No
Serbia	2006–2008	Inconsistent	No
Greece	1974	Autocracy	No
Cyprus	1963–1967	Democracy	No
Cyprus	1974	Democracy	No
Rumania	1989	Autocracy	Yes
Guinea-Bissau	1998–1999	Inconsistent	Yes
Guinea-Bissau	1998–1999	Inconsistent	Yes

Mali	1991	Autocracy	No
Senegal	1962	Inconsistent	No
Benin	1963–1964	Autocracy	No
Benin	1990	Autocracy	No
Niger	1991	Autocracy	Yes
Cote DIvore	1999	Autocracy	No
Cote DIvore	2002–2008	Inconsistent	Yes
Burkina Faso	1977	Autocracy	No
Liberia	1991–1996	Inconsistent	Yes
Liberia	2003–2005	Inconsistent	No
Sierra Leone	1997–2001	Inconsistent	Yes
Ghana	1969	Autocracy	No
Ghana	1978	Autocracy	No
Ghana	1991	Autocracy	No
Togo	1991–1992	Autocracy	No
Nigeria	1978	Autocracy	No
Nigeria	1998	Autocracy	No
Gabon	1990	Autocracy	No
Chad	1978–1984	Autocracy	Yes
Congo	1991	Autocracy	No
DR Congo	1960–1964	Autocracy	Yes
DR Congo	1992–5005	Autocracy	Yes
Uganda	1966	Democracy	No
Uganda	1979	Democracy	Yes
Uganda	1985	Inconsistent	Yes
Burundi	1965	Inconsistent	Yes
Burundi	1992–1995	Autocracy	Yes
Burundi	2001–2004	Autocracy	Yes
Somalia	1991–2010	Autocracy	Yes
Ethiopia	1974	Autocracy	Yes
Ethiopia	1991–1994	Autocracy	Yes
Angola	1991–1996	Autocracy	Yes
Zimbabwe	1979	Democracy	Yes
South Africa	1992–1993	Inconsistent	No
Lesotho	1998–2001	Democracy	Yes
Madagascar	1991	Inconsistent	No
Morocco	1961–1962	Autocracy	No
Sudan	1964	Autocracy	Yes
Sudan	1969–1970	Inconsistent	Yes
Sudan	1985	Autocracy	Yes
Iran	1979–1981	Autocracy	Yes
Turkey	1960	Inconsistent	No
Iraq	2003–2008	Autocracy	Yes
Syria	1960	Democracy	No
Lebanon	1975–2004	Inconsistent	Yes
Yemen	1989–2008	Autocracy	Yes
Yemen, People's Republic	1990	Autocracy	No
Kuwait	1990	Autocracy	No
Afghanistan	1978–1988	Autocracy	Yes
Afghanistan	1992–1995	Autocracy	Yes
Afghanistan	2001–2008	Autocracy	Yes
Republic of Korea	1978	Inconsistent	No
Pakistan	1969–1972	Inconsistent	Yes
Thailand	1968	Autocracy	No

Thailand	1973	Autocracy	No
Thailand	1977	Autocracy	Yes
Cambodia	1970–1971	Autocracy	Yes
Cambodia	1975	Autocracy	Yes
Cambodia	1979–1992	Autocracy	Yes
Cambodia	2000–2002	Inconsistent	No
Laos	1961–1974	Autocracy	Yes
Republic of Vietnam	1965–1972	Autocracy	Yes
Republic of Vietnam	1975	Inconsistent	No
Philippines	1986	Autocracy	Yes
Fiji	2000	Inconsistent	No

Table A.1: N for independent variables

Variable	N	Mean	Std. Dev.	Min.	Max
Capital share	3373	0.637	0.142	-0.409	0.983
Capital share, imputed	7363	0.661	0.118	-0.409	0.983
ln GDP per Capita	7363	8.020	1.179	4.080	13.226
ln Population	7363	9.046	1.549	4.808	14.117
Education	6398	.580	.280	.022	1.315
Education, imputed	7363	.560	.276	.022	1.315
GDP Growth	7291	.021	.069	-.954	1.273
Oil and gas exporter	7363	.512	.500	0	1
Ethnic dominance	7363	.498	.500	0	1
Ethnic fractionalization	7178	.475	.268	.002	1
Ethnic fractionalization, imputed	7363	.474	.266	.002	1
Religious fractionalization	7178	.378	.218	0	.783
Religious fractionalization, imputed	7363	.377	.216	0	.783
Proximity to independence	7363	.125	.223	0	1

Table A.2: Left-hand table: Regime type. Right-hand table: Conflict incidence

Category	Frequency	Percent
Autocracy	3,106	42,18
Inconsistent	1,594	21,65
Democracy	2,663	36,17
Total	7,363	100.00

Category	Frequency	Percent
No conflict	6,144	83,44
Conflict	1,219	16,56
Total	7,363	100.00

Table A.3: Dependent Variable: 'Regime type and conflict status'

Category	Frequency	Percent
AP	2,546	34.58
IP	1,284	17.44
DP	2,314	31.43
AC	560	7.61
IC	310	4.21
DC	349	4.74
Total	7,363	100.00

Appendix B

Economic Inequality

Table B.1: Missing countries on the capital share variable: List of countries included in the dataset with 100% missingness, and their average imputed capital share value

Country name	
Haiti	0.718
Guyana	0.649
Switzerland	0.553
Bosnia	0.679
Belarus	0.679
Guinea-Bissau	0.697
Equatorial-Guniea	0.691
Mali	0.733
Mauritania	0.691
Niger	0.743
Guinea	0.756
Liberia	0.669
Togo	0.721
Chad	0.756
Zaire	0.711
Djibouti	0.733
Angola	0.766
Mozambique	0.708
Zimbabwe	0.733
Namibia	0.638
Yemen	0.710
Bahrein	0.689
Afghanistan	0.757
Turkmenistan	0.767
Tajikistan	0.772
Uzbekistan	0.786
Kazakhstan	0.708
Democrati Rebublic of Korea	0.710
Bhutan	0.721
Burma	0.742
Republic of Vietnam	0.761
Average	0.710

Figure B.1: Observed and imputed values on capital share from a select group of countries, 1960–2010

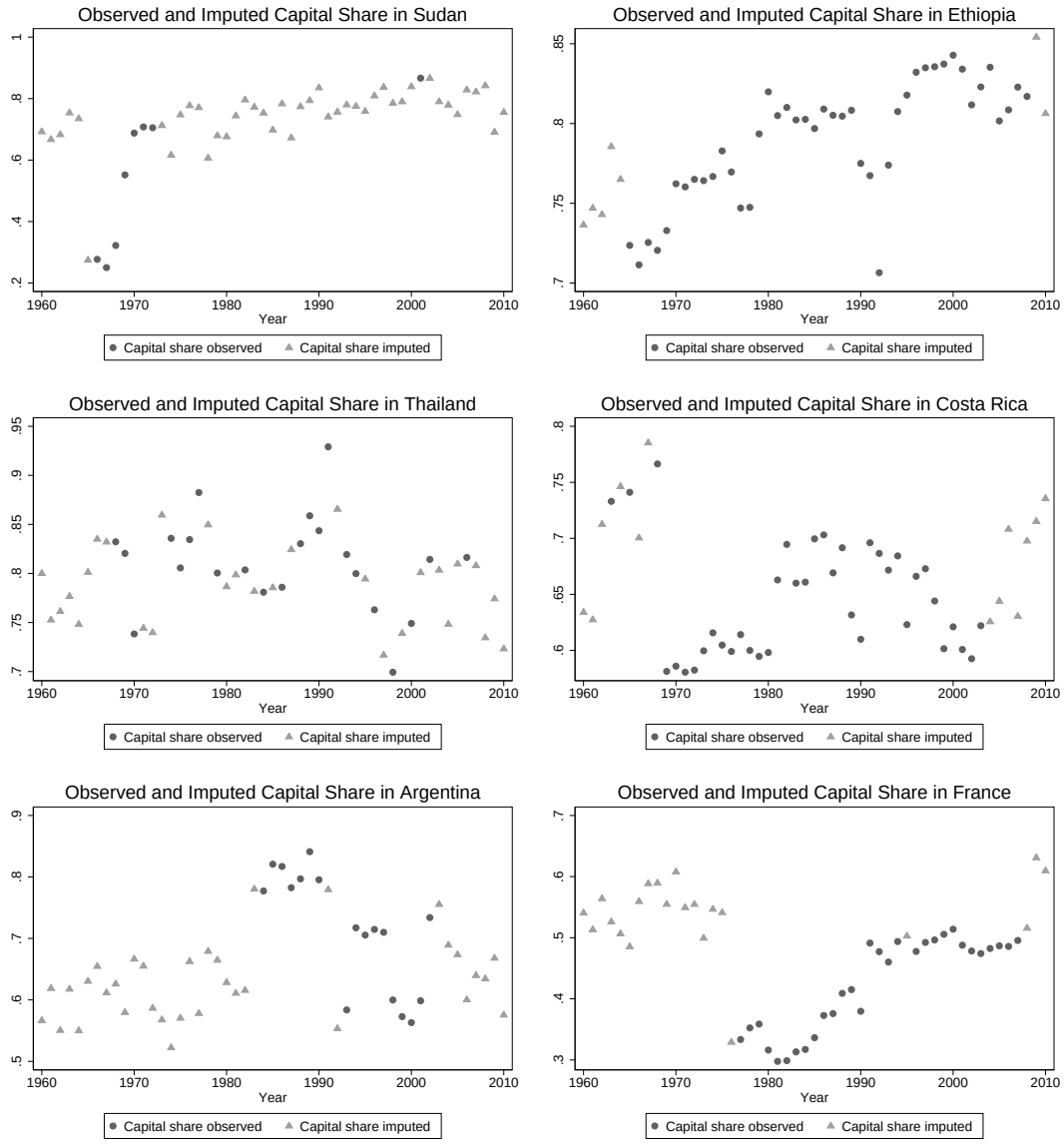


Figure B.2: Comparison of countries without coverage on capital share with select model countries, 1960–2010

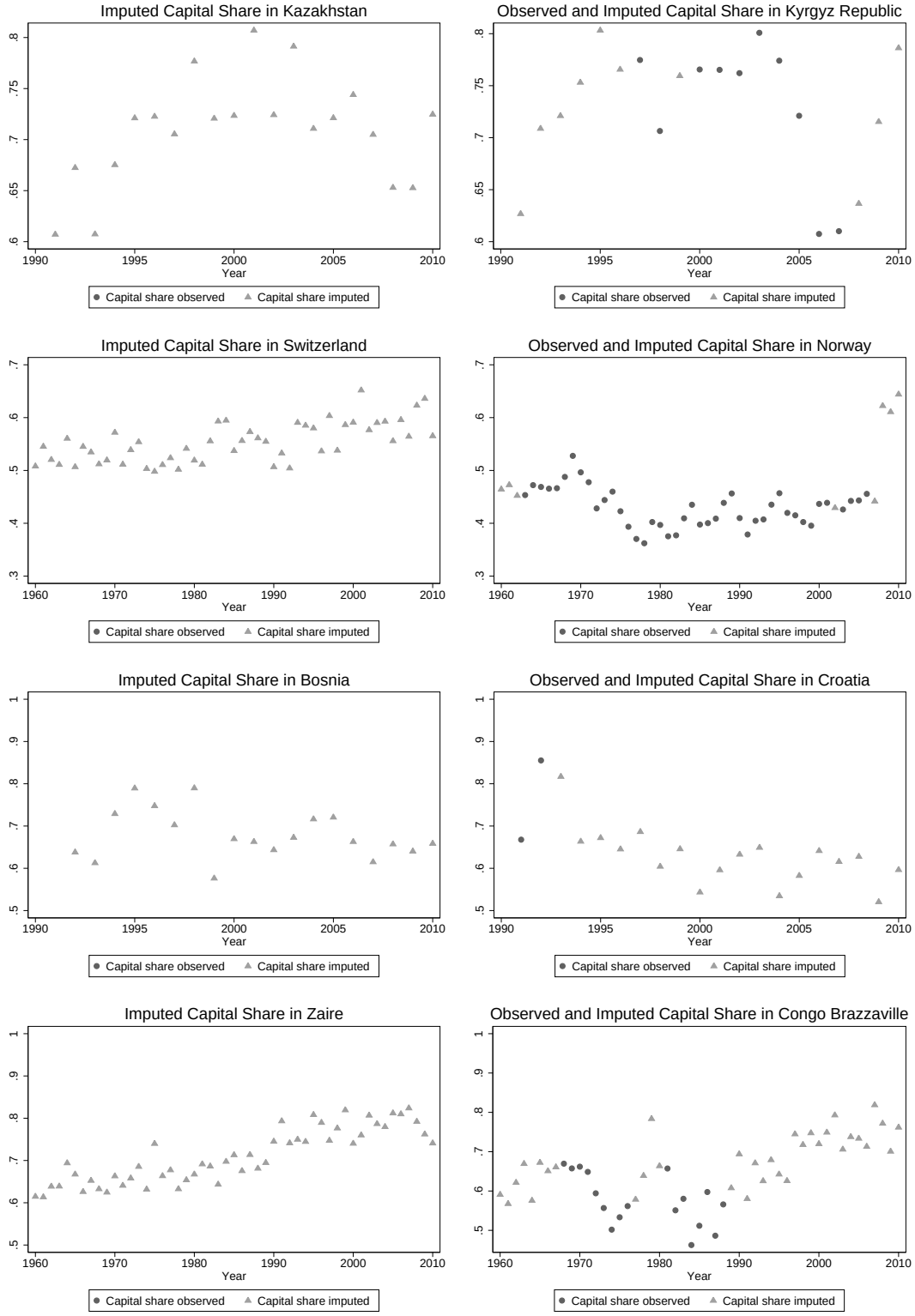


Table B.2: Ranking of 20 most equal and unequal countries, 2011–2050.
Capital share, conflict status, regime type and regime instability

	Country	**Capital Share**	**Conflict status 2010**	**Regime type 2010**	**Nr. regime changes 1960–2010**
High inequality	Botswana	0.920	0	3	1
	Zimbabwe	0.826	0	2	3
	Afghanistan	0.826	1	1	2
	Ethiopia	0.823	1	2	1
	Pakistan	0.819	1	2	6
	Indonesia	0.814	0	3	2
	India	0.810	1	3	0
	Azerbajan	0.809	0	2	3
	Oman	0.805	0	1	0
	China	0.803	0	1	0
	Laos	0.801	0	1	0
	Sri Lanka	0.799	0	2	3
	Kazakhstan	0.794	0	1	2
	Togo	0.794	0	2	2
	Yemen	0.791	1	1	2
	Bangladesh	0.790	0	3	9
	Chad	0.788	1	2	2
	Sudan	0.788	1	1	4
	Guinea	0.787	0	2	5
	D.R Congo	0.786	0	2	1
Low inequality	Sierra Leone	0.350	3	0	8
	Iceland	0.439	0	3	0
	Denmark	0.450	0	3	0
	Slovenia	0.481	0	3	0
	Cyprus	0.496	0	3	0
	Rumania	0.504	0	3	2
	Luxembourg	0.506	0	3	0
	Germany	0.509	0	3	0
	Lithuania	0.523	0	3	0
	Israel	0.536	1	3	0
	Estonia	0.546	0	2	0
	France	0.546	0	3	0
	United Kingdom	0.550	0	3	3
	Australia	0.555	0	3	0
	Norway	0.555	0	3	0
	Austria	0.562	0	3	0
	New Zealand	0.564	0	3	0
	Portugal	0.569	0	3	0
	Netherlands	0.570	0	3	0
	Belgium	0.581	0	3	0

Appendix C

Simulations

C.1 Parameterfile

Figure C.1: Parameterfile for model m1

```
                                        Parameters.txt
//Parameters for estimation command and simulation
// MODEL 1: Baseline model
UNITID; gwno
TIMEID; year
DEP; type
//When using lag(var,0) function it has to be calculated before use
SUPPVAR; timeinstatus; (lag(type,1.0)==lag(type,0.0))?lag(timeinstatus,1.0)+1:1
SUPPVAR; regime;
((lag(type,1.0)==0.0||lag(type,1.0)==3.0)?1.0:((lag(type,1.0)==1.0||lag(type,1.0)
==4.0)?2.0:((lag(type,1.0)==2.0||lag(type,1.0)==5.0)?3.0:0.0)))
SUPPVAR; lts; ln(lag(timeinstatus,1.0))
SUPPVAR; ap; lag(type,0.0)==0.0?1:0
IDEP; ip; lag(type,0.0)==1.0?1:0
IDEP; dp; lag(type,0.0)==2.0?1:0
IDEP; ac; lag(type,0.0)==3.0?1:0
IDEP; ic; lag(type,0.0)==4.0?1:0
IDEP; dc; lag(type,0.0)==5.0?1:0
IDEP; ltsap; lag(lts,0.0)*lag(ap,1.0)
IDEP; ltsip; lag(lts,0.0)*lag(ip,1.0)
IDEP; ltsdp; lag(lts,0.0)*lag(dp,1.0)
IDEP; ltsac; lag(lts,0.0)*lag(ac,1.0)
IDEP; ltsic; lag(lts,0.0)*lag(ic,1.0)
IDEP; ltsdc; lag(lts,0.0)*lag(dc,1.0)
IDEP; lcs3; lagh(lcs3,0.0)
IDEP; gdp; lagh(gdp,0.0)
IDEP; lpop2; lagh(lpop2,0.0)
LINKS; neighbors
IFS; if year >= 1960 & year <= 2010
BETADRAWS; 1
SPLIT; none
SIMULATIONS; 5000
ITERATIONS; 40
STARTTIME; 2010
ENDPARAMS; baseoutcome(0) cons(1/20 100/104 600)
RESULTFILE; E:\Work\stvhavar\Ranveig\Application\m1\results.txt
LOGFILE; E:\Work\stvhavar\Ranveig\Application\m1\log.txt
```

C.2 Do-file for summary of global simulation results

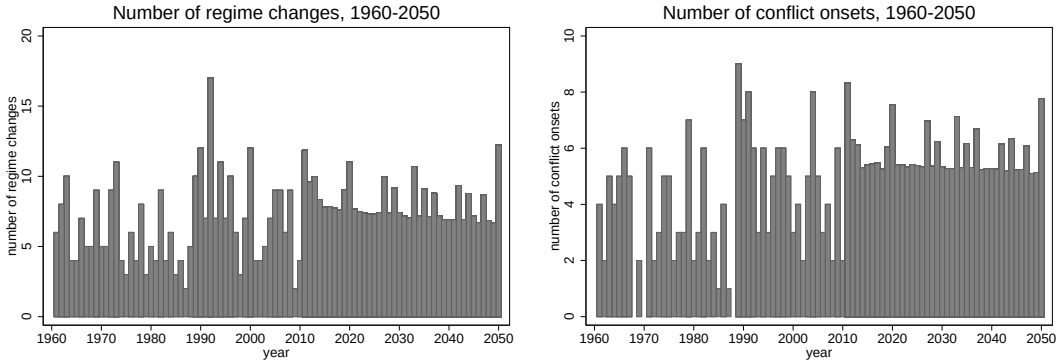
```
*********************************DESCRIPTION OF DO-FILE***************
/*
* This do-file calculates mean, variance, and percentiles of shares
*   of countries in conflict,
globally and regionally.
* This do-file must be run from stat-hiperf
* It also requires that the following local macros are set:
*lastdata
*lastyear
*local dropbox directory
*local model
*first scenario
*last scenario
*numlist (i.e the local models specified)
* Input to the first part is resultfiles from the simulations,
*    Results\_‘scenario’.txt
* Results are placed in "WorldYearResult_‘model’.txt",
in directory E:\work\stvhavar\Masteroppgave\Results    */
*********************************************************************

*********************************************************************
*                                INPUT SECTION                      *
*********************************************************************
clear
clear matrix
set mem 100g
set matsize 11000
set maxvar 10000

/* NAME OF SCENARIO TO SUMMARIZE */
/* The ‘lastdata parameter’ sets the range for which observed
data will be plotted.
Simulated data should start at lastdata + 1 */
local lastdata = 2010
local lastyear = 2050
*********************************************************************
local dropbox_directory = "E:\Work\stvhavar\Ranveig\Application"

*************************************************
************Analyse Results Section**************
*************************************************
cd ‘"‘dropbox_directory’"’
capture mkdir Figures
cd ‘work_directory’
capture mkdir Figures

local firstscenario = 1
local lastscenario = 1

forvalues scenario = ‘firstscenario’(1)‘lastscenario’ {
foreach mline of numlist 1  {
clear
cd ‘"‘dropbox_directory’"’
local infile = "m‘mline’\s‘scenario’\results.txt"
insheet using ‘infile’, names comma
set more off
*drop if year == ‘lastdata’

sort simno gwno year
drop if simno == simno[_n-1] & gwno == gwno[_n-1] & year == year[_n-1]
compress
save "m‘mline’\Results_‘scenario’.dta", replace
summ
local intcut = 50

/* Code region variable */
cd ‘"‘dropbox_directory’"’
capture drop region
capture drop regmarker

/*Gen individual type variables for each regime and conflict state*/
gen type0 = 0
replace type0 = 1 if type == 0
gen type1 = 0
replace type1 = 1 if type == 1
gen type2 = 0
replace type2 = 1 if type == 2
gen type3 = 0
replace type3 = 1 if type == 3
gen type4 = 0
replace type4 = 1 if type == 4
gen type5 = 0
replace type5 = 1 if type == 5
summ type*

/* ******************** */
/* GLOBAL DISTRIBUTION */
/* ******************** */
sort year
capture drop sh_t*
by year: egen sh_w_t0 = mean(type0)
by year: egen sh_w_t1 = mean(type1)
by year: egen sh_w_t2 = mean(type2)
by year: egen sh_w_t3 = mean(type3)
by year: egen sh_w_t4 = mean(type4)
by year: egen sh_w_t5 = mean(type5)

gen sh_w_A = sh_w_t0 + sh_w_t3
gen sh_w_I = sh_w_t1 + sh_w_t4
gen sh_w_D = sh_w_t2 + sh_w_t5
gen sh_w_c = sh_w_t3 + sh_w_t4 + sh_w_t5
gen sh_w_p = sh_w_t0 + sh_w_t1 + sh_w_t2

/*by inequality level*/
sort gwno year
*Three groups: high,low and medium inequality countries*
gen ineq2=.
replace ineq2=1 if lcs3<0.631 & lcs3!=.  /*Baseline m1*/
replace ineq2=2 if lcs3>0.631 & lcs3<0.758
replace ineq2=3 if lcs3>0.758
sort year

*low inequality - threefold classification*
by year: egen sh_w_t0_low2 = mean(type0) if ineq2==1
by year: egen sh_w_t1_low2 = mean(type1) if ineq2==1
by year: egen sh_w_t2_low2 = mean(type2) if ineq2==1
by year: egen sh_w_t3_low2 = mean(type3) if ineq2==1
by year: egen sh_w_t4_low2 = mean(type4) if ineq2==1
by year: egen sh_w_t5_low2 = mean(type5) if ineq2==1

gen sh_w_A_low2 = sh_w_t0_low2 + sh_w_t3_low2
gen sh_w_I_low2 = sh_w_t1_low2 + sh_w_t4_low2
gen sh_w_D_low2 = sh_w_t2_low2 + sh_w_t5_low2
gen sh_w_c_low2 = sh_w_t3_low2 + sh_w_t4_low2 + sh_w_t5_low2
gen sh_w_p_low2 = sh_w_t0_low2 + sh_w_t1_low2 + sh_w_t2_low2

*medium inequality - threefold classification*
by year: egen sh_w_t0_med2 = mean(type0) if ineq2==2
by year: egen sh_w_t1_med2= mean(type1) if ineq2==2
by year: egen sh_w_t2_med2 = mean(type2) if ineq2==2
by year: egen sh_w_t3_med2 = mean(type3) if ineq2==2
by year: egen sh_w_t4_med2 = mean(type4) if ineq2==2
by year: egen sh_w_t5_med2 = mean(type5) if ineq2==2

gen sh_w_A_med2 = sh_w_t0_med2 + sh_w_t3_med2
gen sh_w_I_med2 = sh_w_t1_med2 + sh_w_t4_med2
gen sh_w_D_med2= sh_w_t2_med2 + sh_w_t5_med2
gen sh_w_c_med2 = sh_w_t3_med2 + sh_w_t4_med2 + sh_w_t5_med2
gen sh_w_p_med2 = sh_w_t0_med2 + sh_w_t1_med2 + sh_w_t2_med2
```

```
*high inequality - threefold classification*
by year: egen sh_w_t0_high2 = mean(type0) if ineq2==3
by year: egen sh_w_t1_high2 = mean(type1) if ineq2==3
by year: egen sh_w_t2_high2 = mean(type2) if ineq2==3
by year: egen sh_w_t3_high2 = mean(type3) if ineq2==3
by year: egen sh_w_t4_high2 = mean(type4) if ineq2==3
by year: egen sh_w_t5_high2 = mean(type5) if ineq2==3

gen sh_w_A_high2 = sh_w_t0_high2 + sh_w_t3_high2
gen sh_w_I_high2 = sh_w_t1_high2 + sh_w_t4_high2
gen sh_w_D_high2 = sh_w_t2_high2 + sh_w_t5_high2
gen sh_w_c_high2 = sh_w_t3_high2 + sh_w_t4_high2 + sh_w_t5_high2
gen sh_w_p_high2 = sh_w_t0_high2 + sh_w_t1_high2 + sh_w_t2_high2

outsheet year gwno sh_w_t0 sh_w_t1 sh_w_t2 sh_w_t3 sh_w_t4 sh_w_t5 ///
sh_w_c sh_w_p  sh_w_A sh_w_I sh_w_D ///
sh_w_t0_low2 sh_w_t1_low2 sh_w_t2_low2 ///
sh_w_t3_low2 sh_w_t4_low2 sh_w_t5_low2 sh_w_c_low2 ///
sh_w_p_low2 sh_w_A_low2 sh_w_I_low2 sh_w_D_low2 ///
sh_w_t0_med2 sh_w_t1_med2 sh_w_t2_med2 sh_w_t3_med2 ///
sh_w_t4_med2 sh_w_t5_med2 sh_w_c_med2 sh_w_p_med2 ///
sh_w_A_med2 sh_w_I_med2 sh_w_D_med2 sh_w_t0_high2 sh_w_t1_high2 ///
sh_w_t2_high2 sh_w_t3_high2 sh_w_t4_high2 sh_w_t5_high2 sh_w_c_high2 ///
sh_w_p_high2 sh_w_A_high2 sh_w_I_high2 sh_w_D_high2 sd_sh_w_t0 ///
sd_sh_w_t1 sd_sh_w_t2 sd_sh_w_t3 sd_sh_w_t4 sd_sh_w_t5 sd_sh_w_c ///
sd_sh_w_A  sd_sh_w_I  sd_sh_w_D ///
using "m`mline'\WorldYearResult_`scenario'.txt" ///
if simno == 1, replace
save WorldYearResults_`scenario', replace
}
}
```

C.3 Do-file for summary of country-level simulation results

```
*****************************************************************
*  This do-file summarizes simulation results at the country-level*
*****************************************************************
clear
clear matrix
set mem 80g
set matsize 11000
set maxvar 10000

/* NAME OF SCENARIO TO SUMMARIZE */
/* The `lastdata parameter' sets the range for which observed data will
be plotted. Simulated data should start at lastdata + 1 */
local lastdata = 2010
local lastyear = 2050
local dropbox_directory="E:\Work\stvhavar\Ranveig\Application"
display `""`dropbox_directory'""'

cd `""`dropbox_directory'""'
capture mkdir Figures
cd `work_directory'
capture mkdir Figures

local firstscenario = 1
local lastscenario =  6
forvalues scenario = `firstscenario'(1)`lastscenario' {
foreach mline of numlist 1 {
clear
cd `""`dropbox_directory'""'
local infile = "m`mline'\s`scenario'\results.txt"
insheet using `infile', names comma
set more off

sort simno gwno year
drop if simno == simno[_n-1] & gwno == gwno[_n-1] & year==year[_n-1]
save "m`mline'\Results_`scenario'.dta", replace
local intcut = 50

/*Generate `type' variables for each regime and conflict state*/
gen type0 = 0
replace type0 = 1 if type == 0
gen type1 = 0
replace type1 = 1 if type == 1
gen type2 = 0
replace type2 = 1 if type == 2
gen type3 = 0
replace type3 = 1 if type == 3
gen type4 = 0
replace type4 = 1 if type == 4
gen type5 = 0
replace type5 = 1 if type == 5
summ type*

/*generating regime variable*/
sort simno gwno year
capture drop regime
gen regime =.
replace regime =1 if type==0 | type==3
replace regime =2 if type==1 | type==4
replace regime =3 if type==2 | type==5
inspect regime

/*generating regimechange variable*/
capture drop regchange
gen regchange=0
replace regchange=1 if regime!=regime[_n-1] & ///
gwno==gwno[_n-1] & simno==simno[_n-1]

/* ************************* */
/* COUNTRY-LEVEL DISTRIBUTION */
/* ************************* */

/*Share in each status, by country and year*/
sort year gwno
capture drop sh_cnt*
by year gwno: egen sh_cnt_t0 = mean(type0)
by year gwno: egen sh_cnt_t1 = mean(type1)
by year gwno: egen sh_cnt_t2 = mean(type2)
by year gwno: egen sh_cnt_t3 = mean(type3)
by year gwno: egen sh_cnt_t4 = mean(type4)
by year gwno: egen sh_cnt_t5 = mean(type5)
gen sh_cnt_c = sh_cnt_t3 + sh_cnt_t4 + sh_cnt_t5
gen sh_cnt_A = sh_cnt_t0 + sh_cnt_t3
gen sh_cnt_I = sh_cnt_t1 + sh_cnt_t4
gen sh_cnt_D = sh_cnt_t2 + sh_cnt_t5

/*Share of simulations with a regime change, by country-year*/
by year gwno: egen sh_cnt_regchange = mean(regchange)

outsheet year gwno sh_cnt_t0 sh_cnt_t1 sh_cnt_t2 sh_cnt_t3 sh_cnt_t4 ///
sh_cnt_t5 sh_cnt_c sh_cnt_A sh_cnt_I sh_cnt_D sh_cnt_regchange ///
using "m`mline'\CountryYearResult_`scenario'.txt" if simno == 1, replace
}
}
```

Figure C.2: Observed and simulated regime changes and conflict onsets, 1960–2050



C.4 Do-file (excerpt) for presentation of 'experiments'

```
**************************DESCRIPTION OF DO-FILE*************
* Summarizes and generates figures for several scenarios
* Requires that the following local macros are set:
* local model(s), i.e "mline of numlist.."
* local dropbox directory
* local firstscenario
* local lastscenario
* Also requires that the relevant countries are defined
* Input to first part are files CountryYearResult.txt
for each scenario
*/
************************************************************
*local model = "m1"
*local mname = string(‘model’)
clear
clear mata
set more off
set mem 10g
set maxvar 5000
local dropbox_directory =
 "C:\users\ranveidf\Dropbox\Masteroppgave\Analyse\Simulation"
display ‘"‘dropbox_directory’"’

local firstscenario = 1
local lastscenario = 6

capture log close
set scheme s1mono
capture adopath + m:\ado

* Retrieving simulation data */
cd ‘"‘dropbox_directory’"’
clear
use "Results\ResultsSeveralModels.dta"
drop if _n>=1
save, replace

foreach mline of numlist 1 {
display "model"
forvalues sname = ‘firstscenario’(1)‘lastscenario’ {
display "newscenario "
sleep 3000
local scenario = "s" + string(‘sname’)
clear
insheet using "m‘mline’\CountryYearResult_‘sname’.txt"
gen scenario = .
replace scenario = ‘sname’
gen model = .
replace model = ‘mline’

append using "Results\ResultsSeveralModels.dta"
save "Results\ResultsSeveralModels.dta", replace

} /*scenarios end*/
} /*end forvalues model*/
cd ‘"‘dropbox_directory’"’
*****************************************************
*clear
*cd "Results"
*use "ResultsSeveralModels.dta"
summ

/* Naming relevant countries for graphing purposes */
capture drop country
gen country = ""
replace country = "Egypt" if gwno == 651
replace country = "Thailand" if gwno == 800
replace country = "Portugal" if gwno == 235
replace country = "Ethiopia" if gwno ==530
replace country = "Laos" if gwno==812

capture mkdir Figures
cd "Figures"
sort year gwno

/* Generate scenario-specific variables for graphing purposes */
capture drop sh_cnt_c_m*_s*
foreach mline of numlist 1 {
forvalues s = ‘firstscenario’(1)‘lastscenario’ {
display "figure"
capture drop t_*_‘model’_‘scenario’
local scenario = "s" + string(‘s’)
local model = "m" + string(‘mline’)

*simulated proportion in various states*
gen t_c_‘model’_‘scenario’ = sh_cnt_c if scenario == ‘s’ & model == ‘mline’
gen t_ap_‘model’_‘scenario’ = sh_cnt_t0 if scenario == ‘s’ & model == ‘mline’
gen t_ip_‘model’_‘scenario’ = sh_cnt_t1 if scenario == ‘s’ & model == ‘mline’
gen t_dp_‘model’_‘scenario’ = sh_cnt_t2 if scenario == ‘s’ & model == ‘mline’
gen t_ac_‘model’_‘scenario’ = sh_cnt_t3 if scenario == ‘s’ & model == ‘mline’
gen t_ic_‘model’_‘scenario’ = sh_cnt_t4 if scenario == ‘s’ & model == ‘mline’
gen t_dc_‘model’_‘scenario’ = sh_cnt_t5 if scenario == ‘s’ & model == ‘mline’
gen t_A_‘model’_‘scenario’ = sh_cnt_a if scenario == ‘s’ & model == ‘mline’
gen t_I_‘model’_‘scenario’ = sh_cnt_i if scenario == ‘s’ & model == ‘mline’
gen t_D_‘model’_‘scenario’ = sh_cnt_d if scenario == ‘s’ & model == ‘mline’

by year gwno: egen sh_cnt_c_‘model’_‘scenario’ =
max(t_c_‘model’_‘scenario’)
```

```
by year gwno: egen sh_cnt_t0_‘model’_‘scenario’ =
max(t_ap_‘model’_‘scenario’)
by year gwno: egen sh_cnt_t1_‘model’_‘scenario’ =
 max(t_ip_‘model’_‘scenario’)
by year gwno: egen sh_cnt_t2_‘model’_‘scenario’ =
 max(t_dp_‘model’_‘scenario’)
by year gwno: egen sh_cnt_t3_‘model’_‘scenario’ =
 max(t_ac_‘model’_‘scenario’)
by year gwno: egen sh_cnt_t4_‘model’_‘scenario’ =
max(t_ic_‘model’_‘scenario’)
by year gwno: egen sh_cnt_t5_‘model’_‘scenario’ =
 max(t_dc_‘model’_‘scenario’)
by year gwno: egen sh_cnt_A_‘model’_‘scenario’ =
max(t_A_‘model’_‘scenario’)
by year gwno: egen sh_cnt_I_‘model’_‘scenario’ =
max(t_I_‘model’_‘scenario’)
by year gwno: egen sh_cnt_D_‘model’_‘scenario’ =
max(t_D_‘model’_‘scenario’)

*simulated proportion of regime change*
gen t_rc_‘model’_‘scenario’ = sh_cnt_regchange if scenario == ‘s’
& model == ‘mline’
by year gwno: egen sh_regchange_‘model’_‘scenario’ =
max(t_rc_‘model’_‘scenario’)
}
}
save Results_CountryYears, replace
cd ..
cd "Results/Figures"

foreach mline of numlist 1 {
display "model"
local model = "m" + string(‘mline’)
*cd ‘model’

/*Figures for scenario 2: Low inequality in Thailand from 2010*/
*Simulated share in conflict*
line sh_cnt_c_‘model’_s1 sh_cnt_c_‘model’_s2  year ///
if country == "Thailand" & year>=2011 & year <= 2050, ///
ylabel(0 (0.2)1, grid) xlabel(2011 2020(10) 2050)  ///
lpattern(solid dash) lwidth(medium medthick) ///
legend(label(1 "S1: High inequality") ///
 label(2 "S2: Low inequality") rows(2)) ///
title("Simulated  proportion of conflict in Thailand, s1 and s2")
graph export "‘model’_16_C_Thailand.pdf", replace

*Simulated share  with regime change*
line sh_regchange_‘model’_s1 sh_regchange_‘model’_s2  year ///
if country == "Thailand" & year>=2011 & year <= 2050, ///
ylabel(0 (0.025)0.15, grid) xlabel(2011 2020(10) 2050)  ///
lpattern(solid dash) lwidth(medium medthick) ///
legend(label(1 "S1: High inequality")
label(2 "S2: Low inequality") rows(2)) ///
title("Simulated regime change in Thailand, s1 and s2")
graph export "‘model’_16_regchange_Thailand.pdf", replace

*Probability of being a democracy under scenario 1 and 2*
line sh_cnt_D_‘model’_s1 sh_cnt_D_‘model’_s2  year ///
if country == "Thailand" & year>=2011 & year <= 2050, ///
ylabel(0 (0.2)1, grid) xlabel(2011 2020(10) 2050)  ///
lpattern(solid dash) lwidth(medium medthick) ///
legend(label(1 "S1: High inequality") ///
label(2 "S2: Low inequality") rows(2)) ///
title("Simulated proportion of democracy in Thailand, s1 and s2")
graph export "‘model’_16_D_Thailand.pdf", replace

*Probability of being in inconsistency under scenario 1 and 2*
line sh_cnt_I_‘model’_s1 sh_cnt_I_‘model’_s2  year ///
if country == "Thailand" & year>=2011 & year <= 2050, ///
ylabel(0 (0.2)1, grid) xlabel(2011 2020(10) 2050)  ///
lpattern(solid dash) lwidth(medium medthick) ///
legend(label(1 "S1: High inequality") ///
label(2 "S2: Low inequality") rows(2)) ///
title("Simulated proportion of inconsistency in Thailand, s1 and s2")
graph export "‘model’_16_I_Thailand.pdf", replace

/*Scenario 6: DC instead of IC in Thailand 2010*/

* Figures for scenarios 1 and 6: Simulated incidence of conflict*
line sh_cnt_c_‘model’_s1 sh_cnt_c_‘model’_s6  year ///
if country == "Thailand" & scenario == 1 & year>=2011 & year <= 2050, ///
ylabel(0 (0.2)1, grid) xlabel(2011 2020(10) 2050)  ///
lpattern(solid dash) lwidth(medium thick) ///
legend(label(1 "S1: IC in Thailand in 2010") ///
label(2 "S6: DC in Thailand 2010") rows(2)) ///
title("Simulated proportion of conflict in Thailand, s1 and s6")
graph export "‘model’_13_C_Thailand.pdf", replace

* Figures for scenarios 1 and 6: Simulated share democracy*
line sh_cnt_D_‘model’_s1 sh_cnt_D_‘model’_s6  year ///
if country == "Thailand" & scenario == 1 & year>=2011 & year <= 2050, ///
ylabel(0 (0.2)1, grid) xlabel(2011 2020(10) 2050)  ///
lpattern(solid dash) lwidth(medium thick) ///
legend(label(1 "S1: IC in Thailand in 2010") ///
label(2 "S6: DC in Thailand 2010") rows(2)) ///
title("Simulated proportion of democracy in Thailand, s1 and s6")
graph export "‘model’_13_D_Thailand.pdf", replace
}
```

Appendix D

Out-of-sample evaluation

D.1 Multinomial logit regressions for out-of-sample evaluation

Table D.1: Multinomial Logit Analysis for Out-of-Sample Evaluation, 1960–2000: m0

	IP		DP		AC		IC		DC	
Inconsistent&Peace t–1	5.933***	(0.252)	4.015***	(0.329)	1.266**	(0.488)	6.160***	(0.552)	3.263**	(1.237)
Democracy&Peace t–1	3.667***	(0.396)	7.782***	(0.407)	2.068**	(0.655)	4.926***	(0.818)	7.827***	(0.791)
Autocracy&Conflict t–1	0.457	(0.605)	-10	(.)	4.117***	(0.224)	4.904***	(0.566)	4.192***	(0.878)
Inconsistent&Conflict t–1	5.577***	(0.493)	3.666***	(0.867)	4.130***	(0.554)	9.330***	(0.698)	6.915***	(0.952)
Democracy&Conflict t–1	4.595***	(1.162)	7.773***	(1.045)	4.715***	(1.127)	6.830***	(1.326)	10.81***	(1.246)
ln(time in AP)	0	(.)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in IP)	0.210+	(0.118)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.606***	(0.125)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.633***	(0.129)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.360+	(0.197)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	0.991***	(0.191)
ln(GDP per capita)	0.205*	(0.0915)	0.757***	(0.126)	-0.399***	(0.0966)	-0.0372	(0.136)	0.416*	(0.167)
ln(Population)	-0.00138	(0.0576)	0.174*	(0.0716)	0.193***	(0.0565)	0.346***	(0.0870)	0.429***	(0.0984)
Constant	-5.059***	(0.868)	-12.05***	(1.239)	-2.180**	(0.834)	-9.200***	(1.368)	-14.19***	(1.796)
N	5619									
ll	-2140.2									

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table D.2: Multinomial Logit Analysis for Out-of-Sample Evaluation, 1960–2000. m1

	IP		DP		AC		IC		DC	
Inconsistent&Peace t–1	5.690***	(0.380)	3.192***	(0.526)	0.195	(0.524)	5.310***	(1.029)	1.860	(1.610)
Democracy&Peace t–1	3.434***	(0.490)	6.987***	(0.572)	1.008	(0.683)	4.108***	(1.194)	6.446***	(1.300)
Autocracy&Conflict t–1	0.205	(0.668)	-10	(.)	3.031***	(0.294)	4.036***	(1.036)	2.793*	(1.353)
Inconsistent&Conflict t–1	5.327***	(0.570)	2.829**	(0.961)	3.043***	(0.586)	8.468***	(1.114)	5.509***	(1.404)
Democracy&Conflict t–1	4.347***	(1.197)	6.938***	(1.124)	3.634**	(1.143)	5.991***	(1.586)	9.412***	(1.615)
ln(time in AP)	-0.110	(0.125)	-0.363+	(0.193)	-0.535***	(0.109)	-0.398	(0.444)	-0.661	(0.595)
ln(time in IP)	0.208+	(0.118)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.585***	(0.127)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.627***	(0.130)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.365+	(0.197)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	0.991***	(0.190)
Capital Share	0.482	(0.871)	-0.243	(1.125)	1.324	(0.960)	1.440	(1.110)	0.624	(1.390)
ln(GDP per capita)	0.240*	(0.0939)	0.790***	(0.129)	-0.307**	(0.103)	0.0291	(0.139)	0.483**	(0.175)
ln(Population)	0.00560	(0.0588)	0.198**	(0.0737)	0.197***	(0.0587)	0.346***	(0.0883)	0.444***	(0.100)
Constant	-5.457***	(1.106)	-11.53***	(1.537)	-2.728*	(1.198)	-9.821***	(1.809)	-13.87***	(2.371)
N	5619									
ll	-2123.9									

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table D.3: Multinomial Logit Analysis for Out-of-Sample Evaluation, 1960–2000: m2

	IP		DP		AC		IC		DC	
Capital Share	0.666	(0.440)	-0.666	(0.440)	0.666	(0.440)	0.666	(0.440)	0	(.)
Inconsistent&Peace t–1	5.949***	(0.253)	4.018***	(0.329)	1.277**	(0.488)	6.170***	(0.553)	3.269**	(1.237)
Democracy&Peace t–1	3.690***	(0.396)	7.797***	(0.407)	2.080**	(0.655)	4.942***	(0.818)	7.827***	(0.791)
Autocracy&Conflict t–1	0.450	(0.605)	-10	(.)	4.112***	(0.224)	4.896***	(0.566)	4.189***	(0.878)
Inconsistent&Conflict t–1	5.577***	(0.493)	3.659***	(0.867)	4.128***	(0.554)	9.329***	(0.698)	6.913***	(0.952)
Democracy&Conflict t–1	4.605***	(1.162)	7.761***	(1.045)	4.720***	(1.127)	6.841***	(1.326)	10.81***	(1.246)
ln(time in AP)	0	(.)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in IP)	0.206+	(0.118)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.583***	(0.126)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.629***	(0.130)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.361+	(0.197)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	0.992***	(0.190)
ln(GDP per capita)	0.218*	(0.0920)	0.743***	(0.127)	-0.379***	(0.0974)	-0.0214	(0.137)	0.423*	(0.168)
ln(Population)	-0.00933	(0.0579)	0.182*	(0.0722)	0.188***	(0.0565)	0.341***	(0.0871)	0.430***	(0.0988)
Constant	-5.540***	(0.925)	-11.57***	(1.280)	-2.747**	(0.914)	-9.725***	(1.412)	-14.25***	(1.798)
N	5619									
ll	-2139.1									

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table D.4: Multinomial Logit Analysis for Out-of-Sample Evaluation, 1960–2000: m3

	IP		DP		AC		IC		DC	
Inconsistent&Peace t–1	6.035***	(0.263)	4.099***	(0.336)	1.329**	(0.491)	6.239***	(0.556)	3.286**	(1.238)
Democracy&Peace t–1	3.888***	(0.414)	8.063***	(0.437)	2.123**	(0.664)	5.238***	(0.829)	8.002***	(0.802)
Autocracy&Conflict t–1	0.499	(0.611)	-10	(.)	4.069***	(0.229)	4.788***	(0.572)	4.122***	(0.885)
Inconsistent&Conflict t–1	5.529***	(0.499)	3.541***	(0.874)	4.131***	(0.556)	9.302***	(0.703)	6.737***	(0.958)
Democracy&Conflict t–1	4.682***	(1.167)	7.874***	(1.055)	4.746***	(1.130)	6.905***	(1.333)	10.76***	(1.255)
ln(time in AP)	0	(.)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in IP)	0.194	(0.120)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.552***	(0.131)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.613***	(0.133)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.256	(0.209)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	1.063***	(0.195)
Capital Share	-0.447	(0.885)	-1.325	(1.155)	1.825+	(1.000)	0.878	(1.125)	-0.483	(1.443)
ln(GDP per capita)	0.0256	(0.125)	0.491**	(0.172)	-0.0960	(0.132)	-0.00858	(0.186)	0.414+	(0.242)
ln(Population)	-0.0571	(0.0623)	0.0890	(0.0788)	0.228***	(0.0626)	0.264**	(0.0960)	0.339**	(0.106)
GDP Growth	-1.186	(1.128)	-2.024	(1.550)	0.00659	(1.103)	-3.324*	(1.403)	0.524	(2.396)
Education	1.381**	(0.457)	1.612**	(0.622)	-1.788***	(0.524)	-0.0906	(0.708)	-0.117	(0.945)
Oil and Gas Exporter	-0.850***	(0.183)	-1.069***	(0.253)	-0.207	(0.175)	-0.850**	(0.274)	-1.300***	(0.372)
Ethnic Dominance	0.328+	(0.177)	0.110	(0.235)	-0.0557	(0.174)	0.287	(0.242)	0.430	(0.323)
Religion	0.428	(0.417)	-0.538	(0.576)	-0.0322	(0.415)	0.274	(0.579)	0.0738	(0.812)
Time Since Independence	-0.597+	(0.352)	-0.632	(0.484)	0.0878	(0.400)	-1.707**	(0.575)	-0.867	(0.791)
Constant	-3.474**	(1.311)	-8.495***	(1.791)	-5.184***	(1.450)	-8.915***	(2.025)	-12.63***	(2.650)
N	5619									
ll	-2084.8									

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table D.5: Multinomial Logit Analysis for Out-of-Sample Evaluation, 1960–2000: m4

	IP		DP		AC		IC		DC	
Inconsistent&Peace t–1	6.028***	(0.264)	4.102***	(0.336)	1.300**	(0.491)	6.214***	(0.556)	3.271**	(1.239)
Democracy&Peace t–1	3.891***	(0.414)	8.125***	(0.440)	2.116**	(0.664)	5.235***	(0.829)	8.028***	(0.803)
Autocracy&Conflict t–1	0.460	(0.611)	-10	(.)	4.064***	(0.229)	4.742***	(0.573)	4.081***	(0.885)
Inconsistent&Conflict t–1	5.510***	(0.500)	3.555***	(0.877)	4.091***	(0.556)	9.274***	(0.704)	6.707***	(0.959)
Democracy&Conflict t–1	4.641***	(1.168)	7.869***	(1.055)	4.663***	(1.131)	6.846***	(1.332)	10.76***	(1.254)
ln(time in AP)	0	(.)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in IP)	0.190	(0.120)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.508***	(0.134)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.592***	(0.134)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.248	(0.209)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	1.003***	(0.196)
Capital Share	-0.363	(0.881)	-1.300	(1.142)	1.520	(0.990)	0.892	(1.135)	-0.516	(1.445)
ln(GDP per capita)	-0.0217	(0.125)	0.474**	(0.174)	-0.135	(0.134)	-0.0703	(0.191)	0.358	(0.243)
ln(Population)	-0.0465	(0.0621)	0.0934	(0.0791)	0.240***	(0.0627)	0.270**	(0.0957)	0.362***	(0.105)
GDP Growth	-0.996	(1.131)	-2.210	(1.541)	0.0555	(1.103)	-3.093*	(1.409)	0.404	(2.347)
Education	1.685***	(0.474)	1.517*	(0.646)	-1.607**	(0.538)	0.238	(0.724)	-0.0377	(0.975)
Oil and Gas Exporter	-0.830***	(0.183)	-1.061***	(0.253)	-0.181	(0.176)	-0.819**	(0.275)	-1.287***	(0.373)
Ethnic fractionalization	2.680+	(1.443)	1.029	(1.780)	2.724+	(1.525)	4.315+	(2.271)	5.234*	(2.667)
Ethnic fractionalization sq.	-1.821	(1.493)	-1.337	(1.842)	-2.444	(1.543)	-3.344	(2.283)	-4.638+	(2.603)
Religion	0.289	(0.478)	-0.300	(0.623)	0.0647	(0.458)	0.194	(0.673)	0.204	(0.873)
Time Since Independence	-0.666+	(0.356)	-0.621	(0.489)	0.00108	(0.400)	-1.812**	(0.581)	-0.908	(0.793)
Constant	-3.937**	(1.311)	-8.497***	(1.807)	-5.500***	(1.448)	-9.573***	(2.041)	-13.39***	(2.681)
N	5619									
ll	-2078.2									

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table D.6: Multinomial Logit Analysis for Out-of-Sample Evaluation, 1960–2000: m5

	IP		DP		AC		IC		DC	
Inconsistent&Peace t–1	5.989***	(0.259)	4.043***	(0.334)	1.305**	(0.489)	6.192***	(0.554)	3.279**	(1.238)
Democracy&Peace t–1	3.754***	(0.404)	7.913***	(0.420)	2.135**	(0.657)	5.024***	(0.821)	7.896***	(0.796)
Autocracy&Conflict t–1	0.507	(0.608)	-10	(.)	4.091***	(0.227)	4.883***	(0.568)	4.103***	(0.880)
Inconsistent&Conflict t–1	5.534***	(0.497)	3.570***	(0.872)	4.133***	(0.555)	9.315***	(0.701)	6.760***	(0.955)
Democracy&Conflict t–1	4.678***	(1.166)	7.833***	(1.052)	4.727***	(1.129)	6.874***	(1.329)	10.75***	(1.250)
ln(time in AP)	0	(.)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in IP)	0.203+	(0.119)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.555***	(0.130)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.591***	(0.132)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.278	(0.206)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	1.055***	(0.192)
Capital Share	-0.427	(0.873)	-1.157	(1.129)	1.734+	(0.972)	0.853	(1.117)	-0.335	(1.402)
ln(GDP per capita)	0.00948	(0.118)	0.555***	(0.162)	-0.104	(0.124)	0.0413	(0.169)	0.485*	(0.222)
ln(Population)	-0.0491	(0.0586)	0.124+	(0.0738)	0.229***	(0.0602)	0.301***	(0.0890)	0.389***	(0.101)
Education	1.456***	(0.439)	1.464*	(0.595)	-1.780***	(0.517)	-0.200	(0.644)	-0.366	(0.903)
Oil and Gas Exporter	-0.920***	(0.180)	-1.174***	(0.248)	-0.197	(0.171)	-0.992***	(0.270)	-1.355***	(0.358)
Constant	-3.198**	(1.192)	-9.548***	(1.623)	-5.111***	(1.323)	-9.500***	(1.759)	-13.45***	(2.365)
N	5619									
ll	-2098.2									

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

D.2 Do-file for ROC

```
cd "E:/Work/stvhavar/Ranveig/Application"
insheet using "m‘mline’/Assess/Results_assess.txt", names clear
erase "m‘mline’/Assess/Results_assess.txt"
gen primkey=(gwno*10000)+year
sort simno primkey

/*generating variables for simulated conflict and regime type*/

*Type 0: Autocracy and peace (ap)*
gen sch_ap = 0
replace sch_ap = 2*(ap-ap[_n-1]) if simno == simno[_n-1] ///
& gwno==gwno[_n-1]
gen simons_ap = 0
replace simons_ap = 1 if ap[_n-1]==0 & sch_ap > 0
gen simtrm_ap = 0
replace simtrm_ap = 1 if (ap[_n-1]==1) & sch_ap < 0 & ap!=1

*Type 1: Inconsistent and peace (ip)*
gen sch_ip = 0
replace sch_ip = 2*(ip-ip[_n-1]) if simno == simno[_n-1] & ///
gwno==gwno[_n-1]
gen simons_ip = 0
replace simons_ip = 1 if ip[_n-1]==0 & sch_ip > 0
gen simtrm_ip = 0
replace simtrm_ip = 1 if (ip[_n-1]==1) & sch_ip < 0 & ip!=1

*Type 2: Democracy and peace (dp)*
gen sch_dp = 0
replace sch_dp = 2*(dp-dp[_n-1]) if simno == simno[_n-1] ///
& gwno==gwno[_n-1]
gen simons_dp = 0
replace simons_dp = 1 if dp[_n-1]==0 & sch_dp > 0
gen simtrm_dp = 0
replace simtrm_dp = 1 if (dp[_n-1]==1) & sch_dp < 0 & dp!=1

*Type 3: Autocracy and conflict (ac)*
gen sch_ac = 0
replace sch_ac = 2*(ac-ac[_n-1]) if simno == simno[_n-1] ///
& gwno==gwno[_n-1]
gen simons_ac = 0
replace simons_ac = 1 if ac[_n-1]==0 & sch_ac > 0
gen simtrm_ac = 0
replace simtrm_ac = 1 if (ac[_n-1]==1) & sch_ac < 0 & ac!=1

*Type 4: Inconsistent and conflict (ic)*
gen sch_ic = 0
replace sch_ic = 2*(ic-ic[_n-1]) if simno == simno[_n-1] ///
& gwno==gwno[_n-1]
gen simons_ic = 0
replace simons_ic = 1 if ic[_n-1]==0 & sch_ic > 0
gen simtrm_ic = 0
replace simtrm_ic = 1 if (ic[_n-1]==1) & sch_ic < 0 & ic!=1

*Type 5: Democracy and conflict (dc)*
gen sch_dc = 0
replace sch_dc = 2*(dc-dc[_n-1]) if simno == simno[_n-1] ///
& gwno==gwno[_n-1]
gen simons_dc = 0
replace simons_dc = 1 if dc[_n-1]==0 & sch_dc > 0
gen simtrm_dc = 0
replace simtrm_dc = 1 if (dc[_n-1]==1) & sch_dc < 0 & dc!=1

*Conflict: combining type 3 (ac), 4 (ic) and 5 (dc)*
gen sch_c = 0
replace sch_c = 2*(ac-ac[_n-1]) + (ic-ic[_n-1]) + (dc-dc[_n-1]) ///
if simno == simno[_n-1] & gwno==gwno[_n-1]
```

Table D.7: Multinomial Logit Analysis for Out-of-Sample Evaluation, 1960–2000: m6

	IP		DP		AC		IC		DC	
Inconsistent&Peace t–1	5.971***	(0.258)	4.009***	(0.333)	1.280**	(0.489)	6.192***	(0.555)	3.254**	(1.239)
Democracy&Peace t–1	3.750***	(0.403)	7.845***	(0.416)	2.092**	(0.656)	5.022***	(0.821)	7.899***	(0.797)
Autocracy&Conflict t–1	0.478	(0.606)	-10	(.)	4.120***	(0.226)	4.902***	(0.568)	4.118***	(0.881)
Inconsistent&Conflict t–1	5.482***	(0.496)	3.485***	(0.870)	4.126***	(0.555)	9.216***	(0.700)	6.716***	(0.956)
Democracy&Conflict t–1	4.638***	(1.166)	7.762***	(1.052)	4.727***	(1.128)	6.888***	(1.329)	10.83***	(1.252)
ln(time in AP)	0	(.)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in IP)	0.204+	(0.119)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.585***	(0.128)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.618***	(0.131)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.368+	(0.199)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	0.990***	(0.189)
Capital Share	0.631	(1.055)	0.185	(1.319)	1.279	(1.171)	2.007	(1.348)	1.266	(1.590)
ln(GDP per capita)	0.245*	(0.0953)	0.784***	(0.133)	-0.342***	(0.102)	0.0113	(0.141)	0.452*	(0.179)
ln(Population)	-0.0194	(0.0585)	0.153*	(0.0732)	0.181**	(0.0569)	0.314***	(0.0886)	0.396***	(0.0995)
Oil and Gas Exporter	0.442	(1.213)	1.033	(1.591)	-0.436	(1.357)	1.343	(1.593)	2.461	(2.079)
Capital share*Oil&Gas	-2.108	(1.805)	-3.426	(2.434)	0.326	(1.929)	-3.420	(2.348)	-5.892+	(3.249)
Constant	-5.272***	(1.172)	-11.74***	(1.601)	-3.277**	(1.261)	-10.31***	(1.743)	-14.55***	(2.298)
N	5619									
ll	-2113.6									

Standard errors in parentheses

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

```
gen simons_c = 0
replace simons_c = 1 if (ac[_n-1]==0 & ic[_n-1]==0 & ///
 dc[_n-1]==0) & sch_c > 0
gen simtrm_c = 0
replace simtrm_c = 1 if (ac[_n-1]==1 | ic[_n-1]==1 | ///
dc[_n-1]==1) & sch_c<0 & (ac!=1 & ic!=1 & dc!=1)

*Autocracy: combining type 0 (ap) and 3 (ac)*
gen sch_a = 0
replace sch_a = 2*(ap-ap[_n-1]) + (ac-ac[_n-1]) ///
if simno==simno[_n-1] & gwno==gwno[_n-1]
gen simons_a = 0
replace simons_a = 1 if (ap[_n-1]==0 & ac[_n-1]==0) & sch_a>0
gen simtrm_a = 0
replace simtrm_a = 1 if (ap[_n-1]==1 | ac[_n-1]==1) ///
& sch_a<0 & (ap!=1 & ac!=1)

*Inconsistent: combining type 1 (ip) and 4 (ic)*
gen sch_i = 0
replace sch_i = 2*(ip-ip[_n-1]) + (ic-ic[_n-1]) ///
if simno==simno[_n-1] & gwno==gwno[_n-1]
gen simons_i = 0
replace simons_i = 1 if (ip[_n-1]==0 & ic[_n-1]==0) & sch_i > 0
gen simtrm_i = 0
replace simtrm_i = 1 if (ip[_n-1]==1 | ic[_n-1]==1) & sch_i < 0 ///
& (ip!=1 & ic!=1)

*Democracy: combining type 2 (dp) and 5 (dc)*
gen sch_d = 0
replace sch_d = 2*(dp-dp[_n-1]) + (dc-dc[_n-1]) ///
if simno==simno[_n-1] & gwno==gwno[_n-1]
gen simons_d = 0
replace simons_d = 1 if (dp[_n-1]==0 & dc[_n-1]==0) & sch_d > 0
gen simtrm_d = 0
replace simtrm_d = 1 if (dp[_n-1]==1 | dc[_n-1]==1) & sch_d < 0 ///
& (dp!=1 & dc!=1)

sort primkey
collapse (mean) ap ip dp ac ic dc simtrm_ap simons_ap simtrm_ip ///
simons_ip simtrm_dp simons_dp simtrm_ac simons_ac simtrm_ic ///
simons_ic simtrm_dc simons_dc simtrm_c simons_c simtrm_a ///
simons_a simtrm_i simons_i simtrm_d simons_d gwno year, by(primkey)

/*Merging in historical data*/
merge primkey using
"E:\Work\stvhavar\Ranveig\Application\Data\type.dta", ///
keep(type) nokeep
sort primkey
capture drop _merge
merge primkey using
"E:\Work\stvhavar\Ranveig\Application\Data\dataframe.dta", ///
keep(country)
drop _merge

/*Generating simulated incidence for type variables*/
gen sim_ap = ap
gen sim_ip = ip
gen sim_dp = dp
gen sim_ac = ac
gen sim_ic = ic
gen sim_dc = dc
gen sim_a = ap + ac
gen sim_i = ip + ic
gen sim_d = dp + dc
gen sim_c = ac + ic + dc
gen sim_p = ap + ip + dp

/*Generating observed conflict and regime type variables*/

*observed: Autocracy and peace (ap)*
gen obs_ap = 0
```

```
replace obs_ap = 1 if type==0

*observed: inconsistent and peace (ip)*
gen obs_ip = 0
replace obs_ip = 1 if type==1

*observed: democracy and peace (dp)*
gen obs_dp = 0
replace obs_dp = 1 if type==2

*observed: autocracy and conflict (ac)*
gen obs_ac = 0
replace obs_ac = 1 if type==3

*observed: inconsistent and conflict (ic)*
gen obs_ic = 0
replace obs_ic = 1 if type==4

*observed: democracy and conflict (dc)*
gen obs_dc = 0
replace obs_dc = 1 if type==5

*observed: autocracy variable*
gen obs_a =1 if type==0 | type==3
replace obs_a=0 if obs_a==.

*gen observed inconsistent variable*
gen obs_i =1 if type==1 | type==4
replace obs_i=0 if obs_i==.

*gen observed democracy variable*
gen obs_d =1 if type==2 | type==5
replace obs_d=0 if obs_d==.

*observed: conflict variable*
gen obs_c =1 if type==3 | type==4 | type==5
replace obs_c=0 if obs_c==.


/* ons: If a regime and/or conflict state starts */
/* trm: If a regime and/or conflict state ends */
gen ons_ap = 0
gen trm_ap = 0
replace ons_ap = 1 if (obs_ap-obs_ap[_n-1]==1) & gwno==gwno[_n-1]
replace trm_ap = 1 if (obs_ap-obs_ap[_n-1]==-1) & gwno==gwno[_n-1]

gen ons_ip = 0
gen trm_ip = 0
replace ons_ip = 1 if (obs_ip-obs_ip[_n-1]==1) & gwno==gwno[_n-1]
replace trm_ip = 1 if (obs_ip-obs_ip[_n-1]==-1) & gwno==gwno[_n-1]

gen ons_dp = 0
gen trm_dp = 0
replace ons_dp = 1 if (obs_dp-obs_dp[_n-1]==1) & gwno==gwno[_n-1]
replace trm_dp = 1 if (obs_dp-obs_dp[_n-1]==-1) & gwno==gwno[_n-1]

gen ons_ac = 0
gen trm_ac = 0
replace ons_ac = 1 if (obs_ac-obs_ac[_n-1]==1) & gwno==gwno[_n-1]
replace trm_ac = 1 if (obs_ac-obs_ac[_n-1]==-1) & gwno==gwno[_n-1]

gen ons_ic = 0
gen trm_ic = 0
replace ons_ic = 1 if (obs_ic-obs_ic[_n-1]==1) & gwno==gwno[_n-1]
replace trm_ic = 1 if (obs_ic-obs_ic[_n-1]==-1) & gwno==gwno[_n-1]

gen ons_dc = 0
gen trm_dc = 0
replace ons_dc = 1 if (obs_dc-obs_dc[_n-1]==1) & gwno==gwno[_n-1]
replace trm_dc = 1 if (obs_dc-obs_dc[_n-1]==-1) & gwno==gwno[_n-1]

gen ons_a = 0
gen trm_a = 0
replace ons_a = 1 if (obs_a-obs_a[_n-1]==1) & gwno==gwno[_n-1]
replace trm_a = 1 if (obs_a-obs_a[_n-1]==-1) & gwno==gwno[_n-1]

gen ons_i = 0
gen trm_i = 0
replace ons_i = 1 if (obs_i-obs_i[_n-1]==1) & gwno==gwno[_n-1]
replace trm_i = 1 if (obs_i-obs_i[_n-1]==-1) & gwno==gwno[_n-1]

gen ons_d = 0
gen trm_d = 0
replace ons_d = 1 if (obs_d-obs_d[_n-1]==1) & gwno==gwno[_n-1]
replace trm_d = 1 if (obs_d-obs_d[_n-1]==-1) & gwno==gwno[_n-1]

gen ons_c = 0
gen trm_c = 0
replace ons_c = 1 if (obs_c-obs_c[_n-1]==1) & gwno==gwno[_n-1]
replace trm_c = 1 if (obs_c-obs_c[_n-1]==-1) & gwno==gwno[_n-1]


**********************ROC analysis***************************
/*ROC for all regime and conflict types*/

/*Incidence*/

*AP*
roctab obs_ap sim_ap
matrix ROC[‘mline’,2] = r(area)
matrix ROC[‘mline’,3] = r(lb)
matrix ROC[‘mline’,4] = r(ub)
capture drop disim_ap
gen disim_ap = 0
replace disim_ap = 1 if sim_ap > 0.5
tab disim_ap obs_ap, row col
/* Sensitivity */
summarize disim_ap if obs_ap == 1
matrix ROC[‘mline’,5] = r(mean)
/* Specificity */
summarize disim_ap if obs_ap == 0
matrix ROC[‘mline’,6] = r(mean)

*IP*
roctab obs_ip sim_ip
matrix ROC[‘mline’,7] = r(area)
matrix ROC[‘mline’,8] = r(lb)
matrix ROC[‘mline’,9] = r(ub)
capture drop disim_ip
gen disim_ip = 0
replace disim_ip = 1 if sim_ip > 0.5
tab disim_ip obs_ip, row col
/* Sensitivity */
summarize disim_ip if obs_ip == 1
matrix ROC[‘mline’,10] = r(mean)
/* Specificity */
summarize disim_ip if obs_ip == 0
matrix ROC[‘mline’,11] = r(mean)

*DP*
roctab obs_dp sim_dp
matrix ROC[‘mline’,12] = r(area)
matrix ROC[‘mline’,13] = r(lb)
matrix ROC[‘mline’,14] = r(ub)
capture drop disim_dp
gen disim_dp = 0
replace disim_dp = 1 if sim_dp > 0.5
tab disim_ip obs_dp, row col
/* Sensitivity */
summarize disim_dp if obs_dp == 1
matrix ROC[‘mline’,15] = r(mean)
/* Specificity */
summarize disim_dp if obs_dp == 0
matrix ROC[‘mline’,16] = r(mean)

*AC*
```

```
roctab obs_ac sim_ac
matrix ROC[‘mline’,17] = r(area)
matrix ROC[‘mline’,18] = r(lb)
matrix ROC[‘mline’,19] = r(ub)
capture drop disim_ac
gen disim_ac = 0
replace disim_ac = 1 if sim_ac > 0.5
tab disim_ac obs_ac, row col
/* Sensitivity */
summarize disim_ac if obs_ac == 1
matrix ROC[‘mline’,20] = r(mean)
/* Specificity */
summarize disim_ac if obs_ac == 0
matrix ROC[‘mline’,21] = r(mean)

*IC*
roctab obs_ic sim_ic
matrix ROC[‘mline’,22] = r(area)
matrix ROC[‘mline’,23] = r(lb)
matrix ROC[‘mline’,24] = r(ub)
capture drop disim_ic
gen disim_ic = 0
replace disim_ic = 1 if sim_ic > 0.5
tab disim_ic obs_ic, row col
/* Sensitivity */
summarize disim_ic if obs_ic == 1
matrix ROC[‘mline’,25] = r(mean)
/* Specificity */
summarize disim_ic if obs_ic == 0
matrix ROC[‘mline’,26] = r(mean)

*DC*
roctab obs_dc sim_dc
matrix ROC[‘mline’,27] = r(area)
matrix ROC[‘mline’,28] = r(lb)
matrix ROC[‘mline’,29] = r(ub)
capture drop disim_dc
gen disim_dc = 0
replace disim_dc = 1 if sim_dc > 0.5
tab disim_dc obs_dc, row col
/* Sensitivity */
summarize disim_dc if obs_dc == 1
matrix ROC[‘mline’,30] = r(mean)
/* Specificity */
summarize disim_dc if obs_dc == 0
matrix ROC[‘mline’,31] = r(mean)

*Conflict*
roctab obs_c sim_c
matrix ROC[‘mline’,32] = r(area)
matrix ROC[‘mline’,33] = r(lb)
matrix ROC[‘mline’,34] = r(ub)
capture drop disim_c
gen disim_c = 0
replace disim_c = 1 if sim_c > 0.5
tab disim_c obs_c, row col
/* Sensitivity */
summarize disim_c if obs_c == 1
matrix ROC[‘mline’,35] = r(mean)
/* Specificity */
summarize disim_c if obs_c == 0
matrix ROC[‘mline’,36] = r(mean)

*Autocracy*
roctab obs_a sim_a
matrix ROC[‘mline’,42] = r(area)
matrix ROC[‘mline’,43] = r(lb)
matrix ROC[‘mline’,44] = r(ub)
capture drop disim_a
gen disim_a = 0
replace disim_a = 1 if sim_a > 0.5
tab disim_a obs_a, row col
/* Sensitivity */
summarize disim_a if obs_a == 1
matrix ROC[‘mline’,45] = r(mean)
/* Specificity */
summarize disim_a if obs_a == 0
matrix ROC[‘mline’,46] = r(mean)

*Inconsistent*
roctab obs_i sim_i
matrix ROC[‘mline’,47] = r(area)
matrix ROC[‘mline’,48] = r(lb)
matrix ROC[‘mline’,49] = r(ub)
capture drop disim_i
gen disim_i = 0
replace disim_i = 1 if sim_i > 0.5
tab disim_i obs_i, row col
/* Sensitivity */
summarize disim_i if obs_i == 1
matrix ROC[‘mline’,50] = r(mean)
/* Specificity */
summarize disim_i if obs_i == 0
matrix ROC[‘mline’,51] = r(mean)

*Democracy*
roctab obs_d sim_d
matrix ROC[‘mline’,52] = r(area)
matrix ROC[‘mline’,53] = r(lb)
matrix ROC[‘mline’,54] = r(ub)
capture drop disim_d
gen disim_d = 0
replace disim_d = 1 if sim_d > 0.5
tab disim_d obs_d, row col
/* Sensitivity */
summarize disim_d if obs_d == 1
matrix ROC[‘mline’,55] = r(mean)
/* Specificity */
summarize disim_d if obs_d == 0
matrix ROC[‘mline’,56] = r(mean)

save "m‘mline’/Assess/ROCResults.dta", replace
```

Appendix E

Transition matrix and steady state distribution

E.1 Do-file for transition matrix

```
*********************************************************************
**********************TRANSITION MATRIX**************************
*********************************************************************
cd "C:\users\ranveidf\Dropbox\Masteroppgave\Analyse\Data"
use DemConf2012, clear
run "../Dofiles/Constraints"
drop if year>2010
label define type 0 "AP" 1 "IP" 2 "DP" 3 "AC" 4 "IC" 5 "DC"

****************************************
*  This dofile creates matrices which
*  are used to find the steady state
*  for the transition probabilities
****************************************

/* Observed distribution between 1960 and 2010: transition matrix ch.5*/

*AP t-1*
tab gwno if type==0 & type[_n+1]==0 & gwno==gwno[_n+1]
tab gwno if type==0 & type[_n+1]==1 & gwno==gwno[_n+1]
tab gwno if type==0 & type[_n+1]==2 & gwno==gwno[_n+1]
tab gwno if type==0 & type[_n+1]==3 & gwno==gwno[_n+1]
tab gwno if type==0 & type[_n+1]==4 & gwno==gwno[_n+1]
tab gwno if type==0 & type[_n+1]==5 & gwno==gwno[_n+1]
*IP t-1*
tab gwno if type==1 & type[_n+1]==0 & gwno==gwno[_n+1]
tab gwno if type==1 & type[_n+1]==1 & gwno==gwno[_n+1]
tab gwno if type==1 & type[_n+1]==2 & gwno==gwno[_n+1]
tab gwno if type==1 & type[_n+1]==3 & gwno==gwno[_n+1]
tab gwno if type==1 & type[_n+1]==4 & gwno==gwno[_n+1]
tab gwno if type==1 & type[_n+1]==5 & gwno==gwno[_n+1]
*DP t-1*
tab gwno if type==2 & type[_n+1]==0 & gwno==gwno[_n+1]
tab gwno if type==2 & type[_n+1]==1 & gwno==gwno[_n+1]
tab gwno if type==2 & type[_n+1]==2 & gwno==gwno[_n+1]
tab gwno if type==2 & type[_n+1]==3 & gwno==gwno[_n+1]
tab gwno if type==2 & type[_n+1]==4 & gwno==gwno[_n+1]
tab gwno if type==2 & type[_n+1]==5 & gwno==gwno[_n+1]
*AC t-1*
tab gwno if type==3 & type[_n+1]==0 & gwno==gwno[_n+1]
tab gwno if type==3 & type[_n+1]==1 & gwno==gwno[_n+1]
tab gwno if type==3 & type[_n+1]==2 & gwno==gwno[_n+1]
tab gwno if type==3 & type[_n+1]==3 & gwno==gwno[_n+1]
tab gwno if type==3 & type[_n+1]==4 & gwno==gwno[_n+1]
tab gwno if type==3 & type[_n+1]==5 & gwno==gwno[_n+1]
*IC t-1*
tab gwno if type==4 & type[_n+1]==0 & gwno==gwno[_n+1]
tab gwno if type==4 & type[_n+1]==1 & gwno==gwno[_n+1]
tab gwno if type==4 & type[_n+1]==2 & gwno==gwno[_n+1]
tab gwno if type==4 & type[_n+1]==3 & gwno==gwno[_n+1]
tab gwno if type==4 & type[_n+1]==4 & gwno==gwno[_n+1]
tab gwno if type==4 & type[_n+1]==5 & gwno==gwno[_n+1]
*DC t-1*
tab gwno if type==5 & type[_n+1]==0 & gwno==gwno[_n+1]
tab gwno if type==5 & type[_n+1]==1 & gwno==gwno[_n+1]
tab gwno if type==5 & type[_n+1]==2 & gwno==gwno[_n+1]
tab gwno if type==5 & type[_n+1]==3 & gwno==gwno[_n+1]
tab gwno if type==5 & type[_n+1]==4 & gwno==gwno[_n+1]
tab gwno if type==5 & type[_n+1]==5 & gwno==gwno[_n+1]


/*The syntax below is used to calculate predicted probabilities,
transition matrices and the long-run distribution
in section 7.4.1 The long run transition probability matrix*/

*************************Low Inequality**************************
*Low inequality is defined as below 25 percentile capital shares in 2010*
sum lcs3 if year==2010, detail

/*using clarify to generate predictions*/
capture drop b*

*estimating mlogit model
estsimp mlogit type ip dp ac ic dc ltsap ltsip ltsdp ltsac ///
ltsic ltsdc lgdpcap llnpop lcs3, ///
constraint(1/20 100/104 600) baseoutcome(0)

*setting predictors at mean and all dummys to 0 - ap=1*
setx ip 0 dp 0 ac 0 ic 0 dc 0
setx lgdpcap 8.523 lcs3 0.6357 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*setting predictors at median and ip to 1*
```

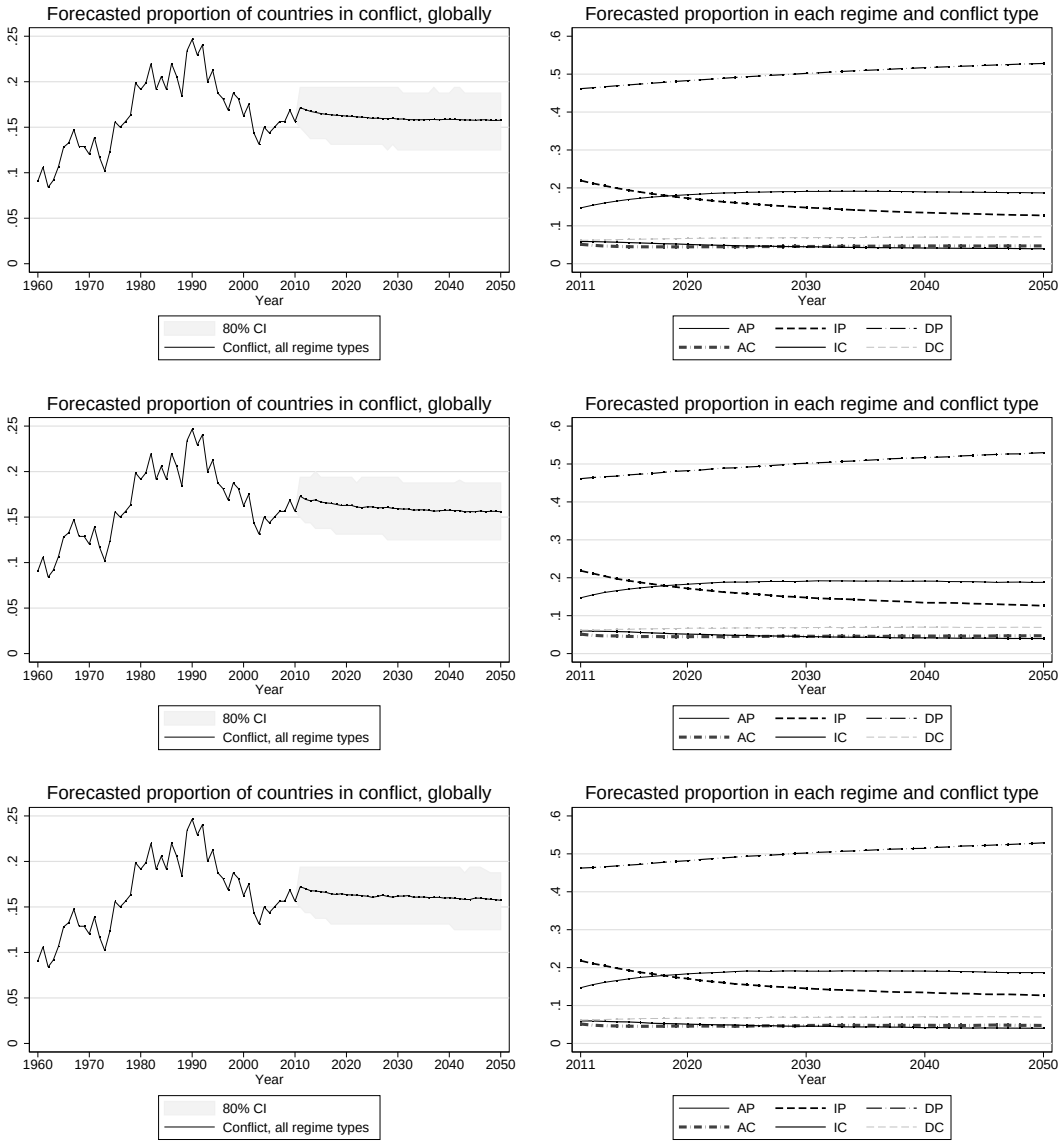
```
setx ip 1 dp 0 ac 0 ic 0 dc 0
setx lgdpcap 8.523 lcs3 0.6357 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*setting predictors at median and dp to 1*
setx ip 0 dp 1 ac 0 ic 0 dc 0
setx lgdpcap 8.523 lcs3 0.6357 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*setting predictors at mean and ac to 1*
setx ip 0 dp 0 ac 1 ic 0 dc 0
setx lgdpcap 8.523 lcs3 0.6357 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*setting predictors at median and ic to 1*
setx ip 0 dp 0 ac 0 ic 1 dc 0
setx lgdpcap 8.523 lcs3 0.6357 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*setting predictors at median and dc to 1*
setx ip 0 dp 0 ac 0 ic 0 dc 1
setx lgdpcap 8.523 lcs3 0.6357 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

matrix input TM_l = (0.9328, 0.0370, 0.0000, 0.0267, 0.0020, 0.0014 \ ///
0.0532, 0.7683, 0.1204, 0.0041, 0.0499, 0.0041 \ ///
0.0104, 0.0182, 0.9258, 0.0022, 0.0028, 0.0407 \ ///
0.1975, 0.0195, 0.2756, 0.4161, 0.0682, 0.0231 \ ///
0.0290, 0.3017, 0.0496, 0.0372, 0.5403, 0.0421 \ ///
0.0085, 0.0366, 0.4834, 0.0198, 0.0110, 0.4408)
matrix list TM_l

*distribution vector in 2010
tab type if year==2010 & cs3<=0.636
matrix input v_l = (1, 4, 35, 0, 0, 0) /*40 countries*/
matrix input v = (21, 41, 73, 8, 9, 8)

*finding the distribution of countries in each state in 2050
matrix vTM_l_50 = v * (TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l* ///
TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l* ///
TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*  ///
TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l*TM_l ///
*TM_l*TM_l*TM_l*TM_l)

matrix list vTM_l_50

*finding steady state for low inequality*
matrix TM2_l = TM_l * TM_l
matrix list TM2_l
matrix TM4_l = TM2_l*TM2_l
matrix TM16_l = TM4_l*TM4_l
matrix list TM16_l
matrix TM256_l = TM16_l*TM16_l
matrix list TM256_l
matrix TM65536_l = TM256_l*TM256_l
matrix list TM65536_l

**********High Inequality**********
*Capital share above 75th percentile equals high inequality*

/*using clarify to generate predictions*/
capture drop b*

*estimating mlogit model
estsimp mlogit type ip dp ac ic dc ltsap ltsip ltsdp ltsac ltsic ltsdc ///
lgdpcap llnpop lcs3, constraint(1/20 100/104 600) baseoutcome(0)

*setting predictors at mean and all dummys to 0 - ap=1*

setx ip 0 dp 0 ac 0 ic 0 dc 0
setx lgdpcap 8.523 lcs3 0.754 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*setting predictors at mean and ip to 1*
setx lgdpcap 8.523 lcs3 0.754 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
setx ip 1 dp 0 ac 0 ic 0 dc 0
simqi, pr

*setting predictors at mean and dp to 1*
setx ip 0 dp 1 ac 0 ic 0 dc 0
setx lgdpcap 8.523 lcs3 0.754 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*setting predictors at mean and ac to 1*
setx ip 0 dp 0 ac 1 ic 0 dc 0
setx lgdpcap 8.523 lcs3 0.754 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*setting predictors at mean and ic to 1*
setx ip 0 dp 0 ac 0 ic 1 dc 0
setx lgdpcap 8.523 lcs3 0.754 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*setting predictors at mean and dc to 1*
setx ip 0 dp 0 ac 0 ic 0 dc 1
setx lgdpcap 8.523 lcs3 0.754 llnpop 9.228
setx ltsap mean ltsip mean ltsdp mean ltsac mean ltsic mean ltsdc mean
simqi, pr

*Transition probabilities between 1960 and 2010*
matrix input TM_h = (0.9237, 0.0411, 0.0000, 0.0312, 0.0024, 0.0014 \ 0.0488, 0.7882, //
0.1000, 0.0044, 0.0547, 0.0039 \ 0.0114, 0.0223, 0.9134, 0.0028, 0.0036, 0.0465 \ ///
0.1854, 0.0205, 0.2350, 0.4598, 0.0767, 0.0226 \ 0.0253, 0.2951, 0.0395, 0.0383, ///
0.5635, 0.0383 \ 0.0087, 0.0417, 0.4453, 0.0237, 0.0133, 0.4673)

*distribution vector in 2010*
sum cs3 if year==2010, detail
tab type if year==2010 & cs3>0.754
matrix input v_h = (10, 11, 7, 5, 6, 2) /*41 countries*/

*finding the distribution of countries in each state in 2050*
matrix vTM_h50 = v * (TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h* ///
TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h* //
TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h*TM_h)
matrix list vTM_h50

*finding steady state for high inequality*
matrix TM2_h = TM_h * TM_h
matrix list TM2_h
matrix TM4_h = TM2_h*TM2_h
matrix TM16_h = TM4_h*TM4_h
matrix list TM16_h
matrix TM256_h = TM16_h*TM16_h
matrix list TM256_h
matrix TM65536_h = TM256_h*TM256_h
matrix list TM65536_h
```

Appendix F

Simulation sensitivity

Figure F.1: Comparison of three forecast-scenarios for capital share

Upper panels: Baseline scenario, model m1. Intermediate panels: Scenario 2 (average capital share 2001–2010). Bottom panels: Scenario 3 (2010 value of capital share)



Appendix G

Imputation sensitivity

Table G.1: Mutlinomial Logit Analysis, 1960–2010, including 'missing dummy'

	IP		DP		AC		IC		DC	
Missing	-0.189	(0.161)	-0.254	(0.221)	-0.124	(0.171)	-0.323	(0.223)	-0.512^{+}	(0.294)
Capital Share	1.015	(0.743)	-0.825	(0.960)	1.373	(0.886)	1.556	(1.002)	0.300	(1.235)
Inconsistent&Peace t–1	5.913***	(0.217)	10.48***	(0.336)	0.890^{+}	(0.480)	6.154***	(0.492)	3.864***	(1.014)
Democracy&Peace t–1	3.779***	(0.369)	14.18***	(0.406)	1.850**	(0.648)	4.715***	(0.783)	8.067***	(0.784)
Autocracy&Conflict t–1	0.781	(0.531)	10	(.)	4.323***	(0.215)	5.175***	(0.510)	4.472***	(0.881)
Inconsistent&Conflict t–1	5.664***	(0.447)	10.15***	(0.754)	3.855***	(0.513)	9.255***	(0.629)	7.062***	(0.918)
Democracy&Conflict t–1	5.044***	(1.102)	14.21***	(1.049)	4.809***	(1.102)	6.730***	(1.307)	11.17***	(1.242)
ln(time in IP)	0.315***	(0.0918)	0	(.)	0	(.)	0	(.)	0	(.)
ln(time in DP)	0	(.)	0.649***	(0.109)	0	(.)	0	(.)	0	(.)
ln(time in AC)	0	(.)	0	(.)	0.564***	(0.115)	0	(.)	0	(.)
ln(time in IC)	0	(.)	0	(.)	0	(.)	0.450**	(0.167)	0	(.)
ln(time in DC)	0	(.)	0	(.)	0	(.)	0	(.)	0.925***	(0.162)
ln(GDP per capita)	0.164^{+}	(0.0844)	0.675***	(0.121)	-0.288**	(0.0934)	-0.0238	(0.122)	0.455**	(0.157)
ln(population)	-0.00191	(0.0536)	0.149*	(0.0756)	0.205***	(0.0530)	0.320***	(0.0813)	0.443***	(0.0972)
Constant	-5.241***	(1.002)	-16.68***	(1.414)	-3.873***	(1.081)	-9.844***	(1.517)	-14.71***	(2.004)
N	7291									
ll	-2841.1									

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$