Income Inequality, Economic Development, and Political Institutions

D I S S E R T A T I O N
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in International Affairs
and Political Economy

submitted by

Florian Jung

from

Germany

Approved on the application of

Prof. Dr. Manfred Gärtner

and

Prof. Dr. Uwe Sunde

Dissertation no. 3961

epubli GmbH, Berlin 2011

The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, October 27, 2011

The President:

Prof. Dr. Thomas Bieger

Acknowledgments

The writing of this little book opened me the door to a fascinating world where I encountered impressive minds and found satisfying intellectual challenges. Therefore I would like to thank all the people who made it possible.

In particular, I am grateful to Manfred Gärtner for being my *Doktorvater* in the true sense of the word. His dedication to macroeconomics, his sense of humor and permanent support are exceptional and turned life at the Institute of Economics (FGN) into an unforgettable experience. What completed my academic environment at the University of St. Gallen was the arrival of Uwe Sunde whose Ph.D. lecture constituted a pivotal moment for my subsequent research. Not only he gave me a new and stimulating perspective on the economic prerequisites of political institutions, but he also supported my very preliminary ideas, and fueled my ambition with his perpetual pursuit of excellence. Though still being the academic role model for me, over the past years Uwe also became a very good friend whose opinion will always be invaluable.

Additionally, I would like to thank Matteo Cervellati, Björn Griesbach, and Thomas Vischer whom I always enjoyed writing papers with, and all my former colleagues at the FGN, especially Henri Buchsteiner, Susanne Burri, and Thomas Seiler. He who deserves a particular mention is my former colleague Andreas Kleiner. Without his exceptional patience and intellectual clarity this book would be a different one. I am also deeply indebted to Mike Golosov, who invited me to stay as a visiting assistant in research at Yale University, as well as to Björn Brügemann, Alex Debs, and all members of the Economics Department for their generous hospitality and the fruitful discussions. Regarding this stay the financial support of the Swiss National Science Foundation must be mentioned gratefully. Furthermore, I thank Michael Pflüger for successfully encouraging me to pursue a Ph.D. in economics.

Outside the academic world, the persons whom I owe the most profound debt of gratitude are of course my closest friends, my girl friend Julia and her family, and my own family, especially my mother and my sister. Thank you so much for always supporting my work but at the same time providing me with many opportunities to forget about it. This book is also yours. The closing words are dedicated to my father in whose memory this book is written.

Florian Jung

November, 2011

Table of Contents

Preface **XVII**

1 Inequality, Development, and the Stability of Democracy **1**

1.1 Introduction 2

1.2 Model 6

1.2.1 Population Structure and Production 6

1.2.2 Political Power and Utility 8

1.2.3 Timing of Events 11

1.3 Political Equilibrium 12

1.4 Politico-Economic Equilibrium 15

1.4.1 Production Environment and Factor Incomes 15

1.4.2 Endogenous Democracy 16

1.5 Stability of Democracy 20

1.6 Empirical Implications and Historical Evidence 24

1.6.1 The Emergence of the German Reich 25

1.6.2 From Weimar Republic to Nazi Germany 27

1.6.3 Towards a Federal Republic of Germany 32

1.6.4 More Historical Examples 36

1.7 Concluding Remarks 37

1.8 Appendix 40

1.8.1 Proof of Lemma 1.1 40

1.8.2 Proof of Proposition 1.2 44

1.8.3 Proof of Proposition 1.3 45

2 Inequality, Redistribution, and Weak Institutions **47**

2.1 Introduction 48

2.2 Model 51

2.2.1 Heterogeneous Population, Production, and Income 51

2.2.2 Redistributive Conflict, Power, and Utility 54

2.2.3 Timing of Events 56

2.3 Political Equilibrium 57

2.4 Politico-Economic Equilibrium 59

2.5 Equilibrium Redistribution 63

2.6 Empirical Implications 67

2.7 Concluding Remarks 69

2.8 Appendix 71

2.8.1 Proof of Lemma 2.1 71

2.8.2 Proof of Lemma 2.2 76

2.8.3 Proof of Proposition 2.3 77

3 Income, Democracy, and Critical Junctures **81**

3.1 Introduction 82

3.2 Data 87

3.3 Econometric Specification 88

3.4 Empirical Results 89

3.4.1 Main Results 89

3.4.2 Robustness 92

3.4.3 Heterogeneity within the Sample of Colonies 93

3.5 Concluding Remarks 96

3.6 Appendix 98

Bibliography **116**

List of Figures

P.1 Global Spread of Democracy Over the Last Century XVIII

1.1 Political Equilibria with Balanced Productivity Levels 18
1.2 Political Equilibria in a Strength-Dominated Society 22
1.3 Political Equilibria in an Ability-Dominated Society 23
1.4 Structural Change in Germany (1874-1954) 27
1.5 Income Distributions in Germany (1871-1949) 31
1.6 Income Development in Germany (1871-1950) 35
1.7 Political Equilibria for Alternative Parametrization 46

2.1 Population structure and Respective Group Sizes 52
2.2 Factor Distributions and Relative Group Powers 62
2.3 Equilibrium Average Tax Rate 65
2.4 Top View of Equilibrium Average Tax Rate 67
2.5 Gini Coefficients and Equilibrium Tax Rates 68
2.6 Illustration of Reasoning for Proof of Proposition 2.3 80

List of Tables

3.1 Classification of Countries (Acemoglu et al. (2008)) 99
3.2 Classification of Countries (CEPII (2011)) 100
3.3 Summary Statistics for 5- and 10-Year Panel (1960-2000) 101
3.4 Replication of Main Results from Acemoglu et al. (2008) 102
3.5 Correlates of Country Fixed Effects 103
3.6 Main Results with 5-Year Panel 104
3.7 Main Results with 10-Year Panel 105
3.8 Robustness Check: CEPII (2011) Colonies with 5-Year Panel 106
3.9 Robustness Check: CEPII (2011) Colonies with 10-Year Panel 107
3.10 Robustness Check: Vanhanen Index 108
3.11 Robustness Check: Binary Democracy Measures 109
3.12 Robustness Check: Tobit Estimates 110
3.13 Partitioned Colony Sample with 5-Year Panel 111
3.14 Partitioned Colony Sample with 10-Year Panel 112
3.15 Year of Independence with 5-Year Panel 113
3.16 Year of Independence with 10-Year Panel 114
3.17 Settler Mortality 115

Summary

This Ph.D. thesis consists of three chapters and analyzes how economic structures affect political outcomes. In the first chapter, Uwe Sunde and I develop a model to study the endogenous emergence of political regimes in societies in which productive resources are distributed unequally and institutions do not ensure political commitments. Our analysis shows that for any level of development there exists a distribution of resources such that democracy emerges in equilibrium, but there are distributions of resources for which democracy is infeasible in equilibrium irrespective of the level of development. The model also delivers results on the stability of democracy with regard to changes in the economic and demographic environment. The results are consistent with the different political regimes that emerged in Germany after 1871.

In the second chapter, I extend the model from chapter one to scrutinize how inequality in terms of factor endowments affects income redistribution when no institutions exist that allow for binding commitments between different groups of society and the political power of social groups depends on their income. In such an environment the relationship between inequality and redistribution turns out to be non-monotonous. A change in inequality might in- or decrease redistribution, or might have no effect at all, depending on the initial level of inequality. But there always exist minimal state regimes in which the equilibrium tax rate is zero. Apart from that, the Gini coefficient turns out to be an inappropriate measure when analyzing the relationship between inequality and redistribution in weakly institutionalized environments.

In the third chapter, Matteo Cervellati, Uwe Sunde, Thomas Vischer, and I revisit the finding that the cross-country correlation between income per capita and democracy disappears once time and country fixed effects are taken into account. This study departs from the premise that the theory of critical junctures suggests income to have different effects on subsequent institutional development depending on the political-economic trajectory of a particular country. A replication of previous findings in the literature provides evidence for heterogeneous income effects in different subsamples, suggesting that sample composition might be relevant in explaining the weak nexus between income and democracy in the full sample.

Zusammenfassung

Die vorliegende Dissertation besteht aus drei Kapiteln, welche der Frage nachgehen, wie ökonomische Strukturen die Ausgestaltung politscher Institutionen in einer Volkswirtschaft beeinflussen. Im ersten Kapitel entwickeln Uwe Sunde und ich ein Modell, das die endogene Entstehung von politischen Regimen in Gesellschaften nachbildet, in welchen Produktionsfaktoren ungleich verteilt und verbindliche politische Übereinkünfte zwischen verschiedenen Gesellschaftsgruppen unmöglich sind. Die Analyse des Modells zeigt, dass für jedes Einkommensniveau bestimmte Einkommensverteilungen existieren, welche die Entstehung von Demokratie im Gleichgewicht zur Folge haben. Demgegenüber gibt es jedoch Einkommensverteilungen, für welche Demokratie unabhängig vom jeweiligen Einkommensniveau im Gleichgewicht unmöglich ist. Die Ergebnisse sind konsistent mit der Entstehung der verschiedenen politischen Regime in Deutschland seit 1871.

Im zweiten Kapitel erweitere ich das zuvor entwickelte Modell, um zu untersuchen, wie die Ungleichheit von Einkommen deren Umverteilung beeinflusst, wenn keine Institutionen zur Durchsetzung politischer Übereinkünfte existieren, und die politsche Macht einzelner Gesellschaftsgruppen durch ihr Einkommen bestimmt ist. In einer derartigen Umgebung erweist sich die Beziehung zwischen Ungleichheit und Umverteilung als nichtmonoton. Abhängig vom ursprünglichen Niveau an Ungleicheit kann eine Veränderung desselbigen Umverteilung verringern, erhöhen oder gänzlich unberührt lassen. Es existieren jedoch immer Minimalstaatsregime, in denen der gleichgewichtige Steuersatz gleich null ist. Zudem erweist sich der Gini-Koeffizient als ungeeignetes Maß zur Analyse des Verhältnisses zwischen Ungleichheit und Umverteilung in derartigen Umgebungen.

Im dritten Kapitel überprüfen Matteo Cervellati, Uwe Sunde, Thomas Vischer und ich ein aktuelles Forschungsergebnis, demzufolge sich die länderübergreifende Korrelation zwischen Pro-Kopf-Einkommen und Demokratie als statistisch insignifikant erweist, sobald zeitkonstante länder- und periodenspezifische Effekte berücksichtigt werden. Die Reproduktion dieses Ergebnisses liefert Evidenz dafür, dass auch bei Berücksichtigung derartiger Effekte qualitativ unterschiedliche Einkommenskorrelationen in verschiedenen Sub-Samples existieren, welche den schwachen Zusammenhang zwischen Pro-Kopf-Einkommen und Demokratie im gesamten Sample erklären. Dies steht im Einklang mit der Theorie der kritischen historischen Zeitpunkte, welche unterschiedliche Einkommenswirkungen entlang unterschiedlicher polit-ökonomischer Entwicklungspfade suggeriert.

Preface

> "The trend towards democracy now widely visible, is a natural trend, due to a general law of social progress."
>
> Bryce (1921), p. 24.

It is hard to imagine that there exists anything like a law of social development. We economists do believe in regularities though. At the macro level, all other things being equal, we would for example expect technological innovations to foster economic development or reductions in the human capital stock to retard it. However, we are less confident about the factors affecting political development. Some of the most influential scholars in the field even resort to coincidence when sorting out the main determinants of political institutions. At certain times in history, they argue, accidental shocks like wars or colonization, so-called critical junctures, occur and change the institutional environment of a country. How this environment changes then depends on the specificity of certain exogenous factors at that decisive point in time. In the case of colonization for example, the density of the indigenous population and the general disease environment in the colonies apparently had a major impact on the nature of the institutions that emerged endogenously, see Acemoglu, Johnson, and Robinson (2002). Once changed, these institutions are supposed to persist and determine the development path of a particular country for a very long time, in many cases for centuries, without being itself affected by changes in the economic environment, see Acemoglu et al. (2008, 2009). This is almost equally hard to imagine.

Another influential theory on democratization, often perceived as opposing to the critical juncture view, is based on the modernization hypothesis put forward by Lipset (1959) who conjectured that economically advanced countries are more likely to develop and sustain democratic institutions. This idea is very often reduced to the hypothesis that

income levels positively affect the democratic quality of political regimes. However, this is not correct. When reading Lipset (1959) thoroughly it is unambiguous that instead of mere income levels he thinks of a broad concept of economic development – involving deep structural changes like industrialization, urbanization and mass education – that affects the political outcome in independent countries. According to this perspective, causality runs in the opposite direction, from economic to political development. Although it is clearly not sufficient to resolve this ongoing debate, let us have a first glimpse at the data.

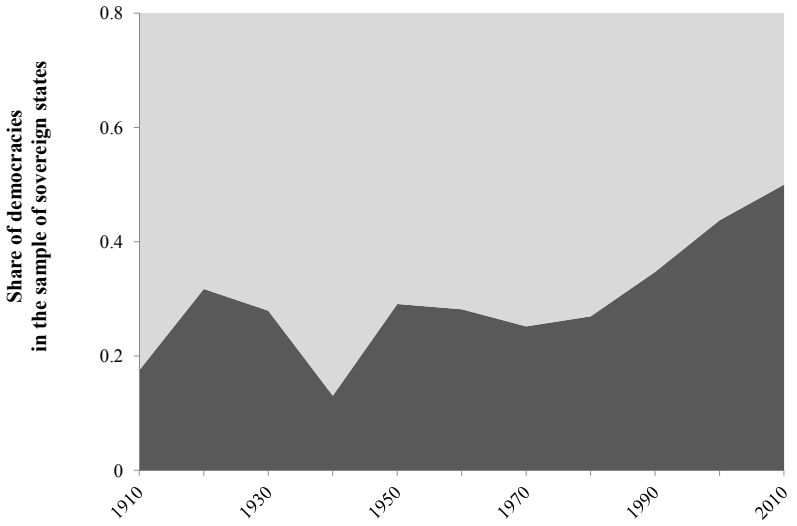

Figure: P.1: Global spread of democracy over the last century.
Data from Polity IV Project, see Marshall and Jaggers (2011).

Figure P.1 shows the spread of democracy over the last one hundred years.[1] Obviously, despite the dubiety of its formalization Bryce's appraisal from 1921 is not contradicted by the data. Starting with a value of below 20% in 1910 the share of democracies in the world more than doubled over the last century. In 2010, exactly half of all sovereign states were democracies. Given the fact that most of these countries developed economically

[1]All countries with a polity score greater or equal to seven in the respective year are counted as democracies. The total number of sovereign states is also taken from the Polity IV dataset. Hence independent countries with a population figure of less than five hundred thousand are not taken into account. To control for an increasing number of sovereign states over the last century Figure P.1 displays shares of democratic countries instead of absolute numbers.

over this period as well, one could be tempted to interpret this correlation in favor of modernization theory. However, Figure P.1 also allows for a different interpretation. Taking a closer look one could argue that obviously all significant increases in the share of democracies occurred in the aftermaths of the three most important historical junctures of the last century: the two World Wars and the fall of the Iron Curtain. In this view, the amplified emergence of democracies results from the specificity of other factors that were crucial at a time when critical junctures placed political institutions at disposal.

In conclusion, it remains unclear how to understand Figure P.1. We cannot disentangle by mere inspection whether it displays a random walk describing a relatively short period in human history, or whether it documents a global convergence of political regimes towards democracy. However, since democracies typically implement policies that are beneficial for the majority of people within society these questions are crucial and need to be addressed. Therefore, this thesis tries to shed some light on the interaction between income inequality, economic development, and political institutions.

In the first chapter, which is joint work with Uwe Sunde, we formalize Lipset's ideas and show theoretically that income inequality seems to be more important for the endogenous emergence of democratic regimes than economic development. And in the second chapter, I extend our model framework to scrutinize the relation between income inequality and redistribution in weakly institutionalized environments. While both chapters analyze the role of income inequality, the first chapter focuses on its implications for the stability of political regimes whereas the second considers the consequences for their size. In the third chapter, which is joint work with Matteo Cervellati, Uwe Sunde and Thomas Vischer, we show empirically that the effect of income on democracy is not spurious as it is claimed by parts of the literature but can be qualitatively affected by institutions that were set up at critical junctures. Our findings suggest that despite the indisputable importance of critical junctures their impact might fade out over time and income might have a positive effect on democracy over the (very) long run. The common denominator of all three chapters is the question how economic structures affect political outcomes.

Chapter 1

Inequality, Development, and the Stability of Democracy

This chapter is joint work with Uwe Sunde from the University of St. Gallen and is published as a discussion paper, see Jung and Sunde (2011).

1.1 Introduction

In the history of modern Germany three critical junctures occurred which required the implementation of a new political regime. They led to the proclamation of the German Reich in 1871, the Weimar Republic in 1919, and the Federal Republic of Germany in 1949. However, despite the strong presence of democratic movements since the first half of the 19th century a stable democracy emerged only at the last juncture. This raises the question why a democracy was not implemented earlier successfully. And what were the reasons for the instability and the eventual breakdown of the democratic Weimar Republic which was overthrown by the Nazi regime? Or, more general: Why do democracies emerge, and what makes some of them last while others vanish?

The importance of political institutions, and in particular that of democracy, for economic development has been one of the most intensely researched areas of the recent years. Democracies typically implement many of the institutions and policies that are thought to be beneficial for economic development, like rule of law, social insurance, or wide-spread education, and thereby allow for a comparably efficient resolution of conflicting interests. Yet, relatively little is known about the determinants of democracy and its stability, even beyond the historical example of Germany. Among the first to address these issues was Seymour Martin Lipset, who conjectured in his famous study that higher levels of economic development and a more equal distribution of resources imply a higher probability for a country to become and to *stay* democratic:

> "Democracy is related to the state of economic development. Concretely, this means that the more well-to-do a nation, the greater the chances that it will sustain democracy. (...) A society divided between a large impoverished mass and a small favored elite would result either in oligarchy (dictatorial rule of the small upper stratum) or in tyranny (popularly based dictatorship)."
>
> Lipset (1959, p. 75)

But irrespective of Lipset's seminal impact on the field of democratization theories, most of the subsequent literature that studies the transitions from oligarchy or autocracy to democracy has concentrated attention exclusively on one of the two factors identified by Lipset, economic development *or* inequality, but not on both. And even more importantly, most of this literature on democratic transitions treats democracy as an absorbing state and thereby assumes that conflicts within such political regimes are solved on the basis of "democratic rules", which obviously implies the existence of some institutionalized environment that ensures these rules to be binding. Assuming democratic rules to be effective seems to be a critical assumption, however, that is unlikely to hold when democracy itself is at stake. Rather, an institutionalized environment cannot be taken for granted when considering the stability of democracy. Or, as Przeworski (2006, p. 312) puts it: "Democracy endures only if it is self-enforcing. It is not a contract because there are no third parties to enforce it." This implies that the stability of democracy needs to be studied in a similar environment as the emergence of democracy from non-democratic rule.

In this paper we consider democracy as an endogenous outcome of a political conflict about the redistribution of incomes within a society in which the income generating factors are distributed unequally. The main novelty of our approach is the consideration of the role of both dimensions, the level of economic development and the distribution of resources, within a heterogeneous society in which no exogenous institutions exist that ensure the possibility to make credible political commitments. Instead, political decisions are made in an environment in which no binding agreements about income redistribution can be made among the different groups of factor owners, and sub-coalitions or single groups can use their de facto power to implement their preferred redistribution scheme against the will of others. In this competition for political power, inequality across several dimensions becomes key for the determination of the politico-economic equilibrium in terms of the political structure and the ex-post allocation of incomes.

The main result of this paper is a novel characterization of the conditions under which democracies emerge or break down in the absence of exogenous institutions that ensure the credibility of political commitments. The equilibrium is characterized by a ruling

coalition that is stable and winning against any other challenging coalition. The equilibrium is a democracy if political decisions are not made by a minority within society but by the overall population. Equilibria where a minority dominates political decisions represent oligarchies.[1] The results provide a characterization of the levels of inequality and development, reflected by the distribution of the different factors in the population and their relative importance in the income generating process, for which democracy or oligarchy emerges in equilibrium. The model also illustrates the consequences of changes in inequality, in terms of population structure and/or factor endowments, or in the economic environment reflected by the economic importance of the different factors, for the stability of democracy. Apart from allowing for a realistic analysis of the stability of political regimes in heterogeneous societies, the approach of considering political regimes as equilibrium in weakly institutionalized environments delivers new insights about the necessary conditions for the emergence and stability of democracy.

The results and implications of the model are consistent with the sequence of political regimes as they emerged in Germany after its unification: the elite-led *German Reich*, the unstable *Weimar Republic* that finally led to the Nazi regime, and the democratic republic after World War II. The three corresponding critical junctures in German history in the years 1871, 1918/19 and 1945 provide an ideal context to illustrate the working of the model. In all three situations, the previous political regime had ceased to exist for exogenous reasons – either due to the unification of previously independent and often competing countries, or due to the loss of one of two immensely costly wars. As a consequence, the shape of the country, the demography and the economic conditions in terms of inequality and economic development had changed dramatically as compared to the respective pre-existing order. This required the emergence of a completely new political regime. The model provides a structural explanation for the very different political regimes that emerged under these conditions: a constitutional monarchy that de facto represented a conservative oligarchy of a landed gentry in the German Reich 1871-1918, a very unstable parliamentary democracy after 1919 that was characterized by several coups and civil conflicts that finally led to the rule of the Nazis 1933-1945, and a stable

[1] The precise definition and classification of equilibria is presented in Section 1.3.

parliamentary democracy after 1945/48.

This paper contributes to a growing literature on endogenous political institutions. Similar to the seminal work of Acemoglu and Robinson (2000, 2001, 2006), it is the redistributive threat by part of the population that brings about a democratic equilibrium. However, in addition to these repercussions of income inequality, the level of economic development is also relevant in the present paper as it affects the economic importance of certain production factors.[2] The model below also differs from most other frameworks that study the endogenous emergence of democracy like e.g. Acemoglu and Robinson (2000, 2001, 2006), Boix (2003), Lizzeri and Persico (2004), Llavador and Oxoby (2005), Gradstein (2007), Cervellati, Fortunato, and Sunde (2008), in that it is not (implicitly or explicitly) assumed that the population consists of different groups among which coalition formation is not a problem or even an issue at all, and that any conflict of interest in democracies can be resolved by credible commitments concerning the policies or the coalitions that are formed. In this respect, our work also differs from Acemoglu and Robinson (2008) who explicitly address the question of regime persistence. The present paper studies the emergence and breakdown of political regimes in an environment in which such credible commitments are not possible, even in democracy. To this end, our analysis builds on the work by Acemoglu, Egorov, and Sonin (2008) who consider the problem of coalition formation in situations where binding agreements among different groups or parties cannot be made, since no party can commit not to eliminate other parties from the ruling coalition in the future. Our model explicitly deals with the concrete problem of coalition formation among distinct groups that represent differently endowed segments of the population and struggle for the redistribution of factor incomes. Finally, since we consider technological progress to be the key driver of income inequality along the lines of Kuznets (1955) or Acemoglu (2002), the determination of political outcomes corresponds with the ideas of Rogowski and MacRae (2008) who deliver various historical examples that are in line with the functioning of our model and thereby complement our case study on Germany.

[2]See Cheibub and Vreeland (2010) for a recent survey on the relationship between economic development and democracy.

The paper is structured as follows. Section 1.2 lays out the model framework, and section 1.3 presents the results concerning the political equilibrium. In section 1.4 the model is nested in a production economy, which allows us to relate the political equilibria to the economic environment in general equilibrium. In section 1.5 we present the main results concerning the emergence and stability of democracy. Section 1.6 illustrates the implications of our model in the context of Germany's history after 1871 and points to various other historical examples. Section 1.7 concludes.

1.2 Model

1.2.1 Population Structure and Production

Consider a static economy that is populated by a unit mass of individuals. These individuals live for one period and leave no bequests. Since consumption is the only component of utility, individuals maximize their disposable incomes. While each individual possesses an identical endowment of labor time, $h > 0$, physical strength and intellectual ability are distributed unevenly in the population.[3] For simplicity, we assume the distribution of both of these characteristics to be dichotomic which means that a share $\gamma > 0$ of individuals possesses one unit of physical strength, denoted by $l = 1$, whereas the complement is left with no physical strength at all, $l = 0$. Likewise, a share $\beta > 0$ of the population possesses intellectual ability, $a = 1$, while all others lack this trait, $a = 0$. We assume physical strength and intellectual ability to be mutually exclusive. Thus the population effectively consists of three distinct groups: the able weaklings, denoted by $\mathcal{A}$, the simple-minded strong, $\mathcal{L}$, and those that possess neither strength nor ability, $\mathcal{P}$.[4] Denote the set

[3] The endowment of labor time h can be normalized to 1 without loss of generality.

[4] In principle, our model society could comprise an arbitrary number of groups, and none of our main results depends on the particular population structure we impose. However, the case with three groups is the least complex to deliver our main results. Increasing the number of groups would complicate the analysis without adding new essential insights.

of groups by $S = \{\mathcal{A}, \mathcal{L}, \mathcal{P}\}$ and the respective size of group $i \in S$ as s_i with

$$s_i = \begin{cases} \beta & \text{if } i = \mathcal{A} \\ \gamma & \text{if } i = \mathcal{L} \\ 1 - \beta - \gamma & \text{if } i = \mathcal{P} \end{cases} \tag{1.1}$$

where $s_i > 0 \ \forall \ i \in S$. Accordingly, the factor endowments of particular group members are given by

$$l_i = \begin{cases} 0 & \text{if } i \in \{\mathcal{A}, \mathcal{P}\} \\ 1 & \text{if } i \in \{\mathcal{L}\} \end{cases} \quad \text{and} \quad a_i = \begin{cases} 0 & \text{if } i \in \{\mathcal{L}, \mathcal{P}\} \\ 1 & \text{if } i \in \{\mathcal{A}\} \ . \end{cases} \tag{1.2}$$

All individuals inelastically supply their endowments on competitive markets to a production sector that uses labor time, strength and ability as separate inputs. Income Y is generated by means of a production function

$$Y \ = \ Y\left(A, H, L, \Lambda\right) \ , \tag{1.3}$$

where $Y(\cdot)$ exhibits constant returns to scale with respect to the input factors H, L and Λ which represent the aggregate levels of working hours, physical strength and ability, respectively. $A > 0$ represents a productivity parameter or vector, reflecting the level of technology. The marginal product of every input factor q is positive but decreasing, i.e. $\partial Y/\partial q > 0$ and $\partial^2 Y/\partial q^2 < 0$. Factor prices are competitive where $\rho = \partial Y/\partial H$ represents the price paid for one unit labor time, $w = \partial Y/\partial L$ gives the remuneration of physical strength, and $\mu = \partial Y/\partial \Lambda$ is the reward for ability. Consequently, the factor income of an individual belonging to group i is given by

$$y_i \ = \ \rho h + w l_i + \mu a_i \quad \text{ with } i \in S \, . \tag{1.4}$$

From the unequal endowment of traits and the remuneration of these traits on competitive markets it follows that factor income is distributed unequally within the population, and individuals with higher endowments earn higher factor incomes. This implies that an

individual in the $\mathcal{P}$-group always receives the lowest factor income $y_{\mathcal{P}}$ in society and

$$y_{\mathcal{P}} < y_{\mathcal{L}}, y_{\mathcal{A}} \,. \tag{1.5}$$

always holds. Note that per-capita income y equals aggregate group income, i.e.,

$$y \;=\; \sum_{i \in S} s_i y_i \;=\; \rho h + w\gamma + \mu\beta \;. \tag{1.6}$$

1.2.2 Political Power and Utility

The given endowment of production factors implies that factor incomes can vary considerably between different groups which gives rise to redistributive conflicts, since we assume the utility of individuals or of members of a certain group not to be affected by the well-being of others. In consequence, a latent conflict between the different groups exists and every group tries to maximize its respective income at the expense of others.[5] All political considerations in the model are therefore reduced to the question of how the income generated by the members of society is redistributed amongst them. We assume that in principle *all* income can be expropriated and redistributed between groups, such that the feasible transfer equals per-capita income $y = \sum_{i \in S} s_i y_i$.[6] In combination with the given production structure the possibility to expropriate all factor income has the important implication that it is always beneficial to employ all available workers in the production process and redistribute their incomes afterwards, as $y_i > 0$ follows from equation (1.4). Since factors are supplied inelastically, there are no hold-up problems or the like through which the political game affects or distorts the production process.

Given the possibility to expropriate factor incomes we need to elaborate on the political dimension of our model and, in particular, consider the question which group or

[5] For simplicity, and contrary to Olson (1965), we assume that no commitment problems exist within groups, i.e., single group members do not free-ride on other members of their group. This implies that our analysis is equivalent to one of a society that consists of three different agents, each representing one income group. Thus, individual members of a group and the group itself can be interchangeably denoted by i. A justification for this assumption is that the collective action required in the case of intra-group conflict is transitory, and hence much easier to sustain, see, e.g., Acemoglu and Robinson (2006).

[6] One could alternatively assume that some subsistence income, for example the factor income from time endowment, can be retained by each individual to ensure that production takes place without changing the main results.

coalition of groups actually makes political decisions and effectively imposes its preferred redistribution scheme on the entire population. As already mentioned before, we consider an environment where no institutions exist that would allow for binding commitments between groups. Thus, no group can make binding offers of how to redistribute income, and no group that is part of the coalition that redistributes income can commit not to exclude other members of that coalition and make political decisions autocratically later on. Given this environment we assume that it is the political power P_i of group i that describes its potential to redistribute factor incomes. To keep the conflict game simple and concentrate on the issue of coalition formation, we model the redistributive conflict as parsimoniously as possible and assume that any group or coalition Q can seize the income of group or coalition $S\backslash Q$ if $P_Q > P_{S\backslash Q}$ holds where $P_Q = \sum_{j\in Q} P_j$ denotes the aggregate power of group or coalition Q. To link the economic and political environment we assume that this political power of a group or coalition is given by its aggregate income, i.e.,

$$P_Q \equiv \sum_{i\in Q} s_i y_i \ . \tag{1.7}$$

This assumption could be motivated by means of a sequential conflict game with perfect information and certain outcome where richer groups can afford more weapons, soldiers, etc., and hence overcome poorer groups in open conflict. Additionally, we assume the power mapping described by equation (1.7) to be bijective such that no two groups can be equal in power, $P_i \neq P_j \ \forall \ i, j \in S$ for $i \neq j$.[7] For notational convenience, we define the most powerful group i_{MAX} to have power P_{MAX} and size s_{MAX}. From this, it follows that the most powerful group is able to make all political decisions alone if, and only if, $2\,P_{MAX} > P_S$ holds where $P_S = \sum_{i\in S} P_i$.

If no group has the power to rule alone, i.e., $2\,P_{MAX} \leq P_S$ the possibility to form a coalition becomes relevant. On the one hand, coalition formation is associated with making concessions to the other members of the coalition with regard to the desired redistribution scheme. Hence, forming part of a coalition is costly in terms of foregone

[7] As will turn out later this assumption is not only convenient but also plausible, since group income and – due to fixed relative group sizes – political power is affected by technological progress and other exogenous factors.

redistribution to the other members of the coalition. On the other hand, being part of a coalition increases political power by pooling resources for a potential conflict with other groups or coalitions. A last aspect of the political environment concerns the question of how the income seized by a particular coalition is redistributed among its members. Since we do not focus on the redistributive implications of our model we assume disposable income $\widetilde{y}_i$ of group i to be determined by

$$\widetilde{y}_i = \widetilde{p}_i y \,, \tag{1.8}$$

with $\widetilde{p}_i$ being the effective relative power of group i given by

$$\widetilde{p}_i = \begin{cases} \frac{P_i}{\sum_{j \in RC} P_j} & \text{if } i \in RC \\ 0 & \text{otherwise} \end{cases} \tag{1.9}$$

where $RC \subseteq S$ denotes the coalition that ultimately redistributes income which we call the ruling coalition.[8] The setting implies that the utility of an individual depends on the disposable income $\widetilde{y}_i$ and therewith on the effective relative power $\widetilde{p}_i$ of the group it belongs to. In its general form the indirect utility function of a member of group $i \in S$ reads

$$u_i = u\left(\frac{\widetilde{y}_i(\widetilde{p}_i)}{s_i}\right) \tag{1.10}$$

with $\dfrac{\partial u_i}{\partial \widetilde{y}_i} > 0$. Since factor income y_i and group size s_i cannot be changed by individuals, the optimization problem amounts to maximizing $\widetilde{p}_i$ in order to maximize lifetime utility, subject to the constraints imposed by the production structure and the political

[8] In this respect, we simply follow Acemoglu, Egorov, and Sonin (2008) by employing a sharing rule that was first used by Gamson (1961) to characterize the sharing of resources amongst coalition members. As several empirical studies suggest, see e.g. Warwick and Druckman (2001) or Ansolabehere et al. (2005), this seems to be a fairly good description of redistribution within coalitions. Given its strong empirical regularity it is often even referred to as Gamson's *law*, see Frechette, Kagel, and Morelli (2005). However, we only adopt this rule for analytic simplicity. Note that for $RC = S$ from equations (1.6) and (1.7) it follows that $\widetilde{y}_i = s_i y_i$. In general, any rule can be applied without qualitatively affecting our results as long as it satisfies that (a) every member of a coalition that seizes income of others gets a positive share of redistributed income; but (b) this share does not grant any member more power than the sum of all others; and (c) it does not perfectly equalize the power of any two members.

environment, i.e.,

$$\max_{\widetilde{p}_i} u_i\left(\widetilde{y}_i(\widetilde{p}_i)\right) \quad \text{subject to (1.4), (1.6), (1.8) and (1.9).} \tag{1.11}$$

Thus every individual always prefers the coalition in which the relative power of the group $i \in S$ he belongs to is greatest.[9]

1.2.3 Timing of Events

The following description of the non-cooperative ruling coalition formation and redistribution game that is played by every generation completes the timing of events. The sequence of events that a particular generation experiences throughout its lifetime is given by

A. Birth, realization of endowments and factor incomes.

B. Ruling coalition formation and redistribution game Γ:

B.1 At the initial stage of the game $k = 0$ an agenda setter is randomly determined from all groups and proposes a sub-coalition (that includes herself).

B.2 The members of this sub-coalition vote sequentially in random order over the proposal (and all non-members automatically vote against it):

B.2.1. If the proposal is not supported by a winning coalition, the game proceeds to step B.3.

B.2.2. If the proposal is supported by a winning coalition and consists of

a) all voting groups, then they all form the *RC* and the game proceeds to step C.

b) a proper subset, then all groups that are not part of this proposal are excluded from participation in the game by redistributing their factor incomes toward the members of the winning subset. Now a new stage $k + 1$ begins and the game proceeds to step B.3.

[9] Since the utility of an individual is determined by the structure of the *RC*, our game is *hedonic* in the sense of Dreze and Greenberg (1980).

B.3 From all (remaining) groups a new agenda setter is randomly determined among all groups that have not yet acted as agenda setter at the current stage of the game k; she proposes a sub-coalition (including herself), and the game proceeds to step B.2. If all remaining groups have been agenda setters at the current stage k, then they all form the *RC* and the game proceeds to step C.

C. Consumption of disposable income and death.

1.3 Political Equilibrium

We start our analysis of political equilibria with a central Lemma on the equilibrium outcome of the game described above.

Lemma 1.1. *In game* Γ *there exist subgame perfect Nash equilibria (SPNEs) in pure strategies which all lead to the same RC.*

Proof. See Appendix. □

The intuition for the equilibrium characterization of the ruling coalition *RC* is as follows. First, a *RC* must – by the nature of the game – be winning in the sense that it is powerful enough to outgun any alternative coalition that may challenge it at the current stage k of the game. And second, every *RC* must be stable such that none of its proper subcoalitions will be winning and become the new *RC* at a subsequent stage of the game $\hat{k} > k$.[10] Apart from that, we can also characterize the *RC* in terms of its size.

Lemma 1.2. *The RC consists of all groups if and only if the most powerful group is dominated by the rest of society, i.e.* $P_S \geq 2\,P_{MAX} \iff s_{RC} = 1$.

Proof. This proof is straightforward since we know from the proof of Lemma 1.1 that the *RC* must be a subset of all winning and stable coalitions. Due to the bijective power

[10] This second equilibrium property goes back to Bernheim, Peleg, and Whinston (1987) and their concept of a Perfectly Coalition-Proof Nash Equilibrium which was already studied in several other contexts, see for example Moreno and Wooders (1996) or Einy and Peleg (1995). See also Acemoglu, Egorov, and Sonin (2008) for a modification in the context of political games. Note that this reasoning corresponds to the conceptualization in terms of the set Ω in the proof of Lemma 1.1, which gives a formal definition of the *RC*.

mapping, a coalition of two groups cannot be stable, since one group always dominates the other, and therefore could always successfully propose an even smaller coalition that only contains itself at a subsequent stage of the game. Hence, $|RC| \neq 2$ always holds where $|RC|$ denotes the cardinality of set RC. Thus, it immediately follows $P_S \geq 2\,P_{MAX} \iff s_{RC} = 1$. □

Before we proceed, it is worth commenting briefly on the underlying concept of society, in particular concerning the ability and the incentive for certain income groups to secede in order to escape taxation. In our model, it is the exploitation of political power rents that constitutes a centripetal force and prevents society from falling apart.[11] Secessions are ruled out endogenously in equilibrium, since the groups who would be better off on their own, the net tax payers, are not powerful enough to split from the RC, whereas the RC, who would be powerful to split from the rest of society has no incentive to do so, because this would make its members worse off.[12]

Note that so far, the political equilibrium was characterized without any reference to political concepts. But the equilibrium itself can be interpreted as reflecting a particular political regime. To simplify the terminology, we first introduce a simple classification of equilibria that follows directly the conceptual distinction of political regimes made by Lipset (1959) in the introductory quote.

Definition 1.1. *In equilibrium the political regime is ...*

1. *... a democracy if* $s_{RC} = 1$ *;*

2. *... a mass dictatorship if* $0.5 \leq s_{RC} < 1$ *;*

3. *... an oligarchy if* $s_{RC} < 0.5$ *.*

In the context of our model we define an oligarchy as a RC that represents the minority of the population and imposes policies on the rest of society.[13] On the opposite, we call

[11] Even though this result might contradict the empirical observation of an increasing number of sovereign states over the last century, it should be kept in mind that this model exclusively focuses on economic mechanisms and thereby ignores other factors like ethnic, religious or cultural identity, which play a prominent role in separation processes of political entities in reality. In our model, we take the size of the polity as exogenously given, for instance due to geographical or historical reasons. For a model where state size is determined endogenously, see, e.g., Alesina and Spolaore (1997, 2003).

[12] In this respect our model very much differs from Boix (2003) whose results depend on the assumption of asset specific factor mobility and the existence of some outside option for the owners of mobile assets.

[13] Naturally, one might give an even more detailed definition of oligarchies, depending on which group

every political system a democracy when the RC embraces the entire population and hence all income groups. In this case, all groups of society are bound together by the fact that no smaller coalition is winning and stable. Then even the small minorities play an active role in policy determination and are actively integrated by all others. From this definition of a democracy, one must distinguish a popularly based or mass dictatorship in which the ruling coalition represents only one single group that constitutes the majority of the population. In such a political regime a minority of the people is expropriated and not involved in political decision-making.[14]

This distinction between a democracy and a mass dictatorship is not obvious from a normative perspective, since in both cases the majority of the population is involved in the redistribution decision.[15] However, in a mass dictatorship, the largest group has the power to dominate all other groups of society that are minorities and extract redistribution from them. It is this monopoly of political power within a mass dictatorship that contradicts the typical connotation of a democracy in which different groups of society can express their will and influence public decisions.[16] With this terminology in mind, we state the following proposition regarding the different types of political regimes.

Proposition 1.1. *In equilibrium, the political regime is ...*

1. *... a democracy if and only if the most powerful group is dominated by the rest of society,* $2\,P_{MAX} \leq P_S \iff s_{RC} = 1$ *;*

2. *... a mass dictatorship if and only if society is strictly dominated by a single group that represents the majority of the population,* $2P_{MAX} > P_S \wedge s_{MAX} \geq 0.5 \iff 0.5 \leq s_{RC} < 1$ *;*

rules. For example, an oligarchy of group $\mathcal{P}$ could be denoted as an ochlocracy (the rule of the mob), whereas an oligarchy of group $\mathcal{A}$ or $\mathcal{L}$ represents a plutocracy (the rule of the rich in the respective situation).

[14]Note that our notion of a mass dictatorship fundamentally differs from the concepts of partial democracies or restricted franchise as in Acemoglu and Robinson (2006) or Lizzeri and Persico (2004) respectively which both rest on the implicit assumption that binding commitments between different groups can be made.

[15]One could argue that it effectively makes no difference for the political outcome whether a homogeneous majority directly dictates the public actions (redistribution in the concrete case), or whether the same majority competes in a democratic ballot with opposing groups who *de jure* have the right to vote, but will *de facto* fail in achieving their political goals. This would be in line with the famous reasoning of Aristotle (1943) who defined democracy as an inferior form of government where the state is ruled by the many who only pursue their private interests.

[16]This distinction between democracies and mass dictatorships not only links to the introductory quote of Lipset (1959) but is also related to de Tocquevilles (1864) famous thoughts on the *tyranny of the majority*.

3. *... an oligarchy if and only if society is strictly dominated by a single group that represents a minority of the population,* $2P_{MAX} > P_S \wedge s_{MAX} < 0.5 \iff s_{RC} < 0.5$.

Proof. This Proposition follows directly from Lemmata 1.1 and 1.2 and the application of Definition 1.1. □

The necessary and sufficient conditions contained in Proposition 1.1 map any distribution of factor endowments to a unique political regime in equilibrium.

1.4 Politico-Economic Equilibrium

1.4.1 Production Environment and Factor Incomes

This section extends the previous analysis by endogenizing factor incomes with respect to the distribution of strength and ability. To illustrate the main points, we adopt a *CRS* specification of the production function

$$Y = (A_a\Lambda + A_l L)^{\sigma} H^{1-\sigma} , \tag{1.12}$$

with $0 < \sigma < 1$ and normalize the individual time endowment h to 1. Without being essential for the results, this specification provides a simple way to model redistributive conflicts along the development path by differentiating between ability-augmenting and strength-augmenting productivity parameters A_a and A_l with $A_a, A_l > 0$.[17] Assuming perfectly competitive markets, the reward for every production factor equals its marginal product. Given expressions (1.4) and (1.12), individual factor income of a member of group i therefore becomes

$$y_i = (A_a\Lambda + A_l L)^{\sigma} H^{1-\sigma} \left[\frac{(1-\sigma)}{H} + \sigma \frac{(A_a a_i + A_l l_i)}{(A_a\Lambda + A_l L)} \right] \quad \text{with } i \in S . \tag{1.13}$$

[17] This specification of the production function is formally equivalent to the production of a homogeneous commodity in two distinct sectors, one employing exclusively ability together with time, and the other exclusively physical strength together with time. Variations in productivity parameters affect income levels as well as the shares of total income generated by an exclusive production factor while the income share devoted to labor time remains constant. Note that our results are qualitatively unaffected when using a *CES* production function instead.

For the following analysis, let us define $\lambda_i = s_i y_i / y$ as the share of total income that is produced by group i. Note that this expression also reflects the relative power of group i, i.e., $\lambda_i = P_i / P_S$. With the distribution of resources as in (1.2) and using the information contained in equation (1.6), equation (1.13) can be rewritten as

$$\lambda_{\mathcal{P}} = (1 - \beta - \gamma)(1 - \sigma) \tag{1.14}$$

for the $\mathcal{P}$-group,

$$\lambda_{\mathcal{L}} = \gamma(1 - \sigma) + \gamma\sigma \frac{A_l}{(A_a\beta + A_l\gamma)} \tag{1.15}$$

for the $\mathcal{L}$-group, and

$$\lambda_{\mathcal{A}} = \beta(1 - \sigma) + \beta\sigma \frac{A_a}{(A_a\beta + A_l\gamma)} \tag{1.16}$$

for the $\mathcal{A}$-group, respectively. As can easily be seen from equation (1.15) given a certain value of β the relative power of the $\mathcal{L}$-group increases in the importance of strength in the production process reflected by A_l or in the size of the group γ, i.e., $\partial\lambda_{\mathcal{L}}/\partial A_l$, $\partial\lambda_{\mathcal{L}}/\partial\gamma > 0$. This reasoning analogously holds for the other groups. And of course, any change that makes one group relatively more powerful makes the others relatively weaker and vice versa.

On the basis of these expressions, we can now characterize a unique politico-economic equilibrium for any given distribution of production factors in the population.

1.4.2 Endogenous Democracy

Every such politico-economic equilibrium reflects the subgame perfect equilibrium of the game described in section 1.2.3 including income production, formation of the RC and redistribution. From Proposition 1.1 it can be seen that the particular political regime emerging in equilibrium depends on the power and on the size of the most powerful group. Since in general any of the three groups can be the most powerful we have to consider both criteria for every group in society. Setting the relative power equations (1.14), (1.15)

and (1.16) equal to one half and solving for β yields

$$\beta_{\lambda_{\mathcal{P}}=0.5} = \frac{0.5-\sigma}{1-\sigma} - \gamma \tag{1.17}$$

$$\beta_{\lambda_{\mathcal{L}}=0.5} = \frac{\sigma - 0.5 + \gamma(1-\sigma)}{0.5 - \gamma(1-\sigma)} \frac{A_l}{A_a} \gamma \tag{1.18}$$

$$\beta_{\lambda_{\mathcal{A}}=0.5} = \frac{0.5-\sigma}{2(1-\sigma)} + \frac{\sqrt{f(\gamma) + A_a^2(\sigma-0.5)^2}}{2A_a(1-\sigma)} + \frac{A_l}{2A_a}\gamma \tag{1.19}$$

with $f(\gamma) = A_l\gamma\left[A_l\gamma(\sigma-1)^2 + A_a(1+\sigma(1-2\sigma))\right] > 0 \,\forall\, \sigma \in (0,1)$. These conditions represent the combinations of parameters for which the relative power of a particular group is just equal to the power of all other groups together. Applying the same reasoning with regard to group size delivers the parametric conditions for the size of a particular group to represent exactly half of total population. The respective loci read

$$\beta_{s_{\mathcal{P}}=0.5} = 0.5 - \gamma, \quad \gamma_{s_{\mathcal{L}}=0.5} = 0.5 \quad \text{and} \quad \beta_{s_{\mathcal{A}}=0.5} = 0.5\ . \tag{1.20}$$

While all equilibria can be solved analytically, and the characterization of equilibria presented in Section 1.3 generally applies, we illustrate the results by ways of parametric examples but to highlight the main results as well as their intuition. To illustrate our analytical results we set $A_l = A_a = 1$ and $\sigma = 0.5$ as a benchmark example. In this case, the income share of mere labor time which is distributed equally across all individuals equals 0.5.

Figure 1.1 presents the corresponding allocation of politico-economic equilibria.[18] The γ-β space is decomposed into different areas of γ-β combinations that imply particular equilibrium constellations. From Lemma 1.1 it follows that there exists a unique equilibrium, in terms of RC and the corresponding redistribution scheme, for each single γ-β combination, i.e., everywhere in the admissible γ-β space. The corresponding characterization of the respective political regime follows from Proposition 1.1.

Given the population structure in our model the admissible γ-β space is restricted by the $\gamma-\beta$ axes for $0 < \gamma, \beta < 1$ and the straight line $\beta = 1-\gamma$ and thus constitutes a triangular

[18]Note that the three-group version of our model is the simplest structure that allows to derive all types of equilibria, including the grand coalition, and to analyze the results in a two-dimensional space.

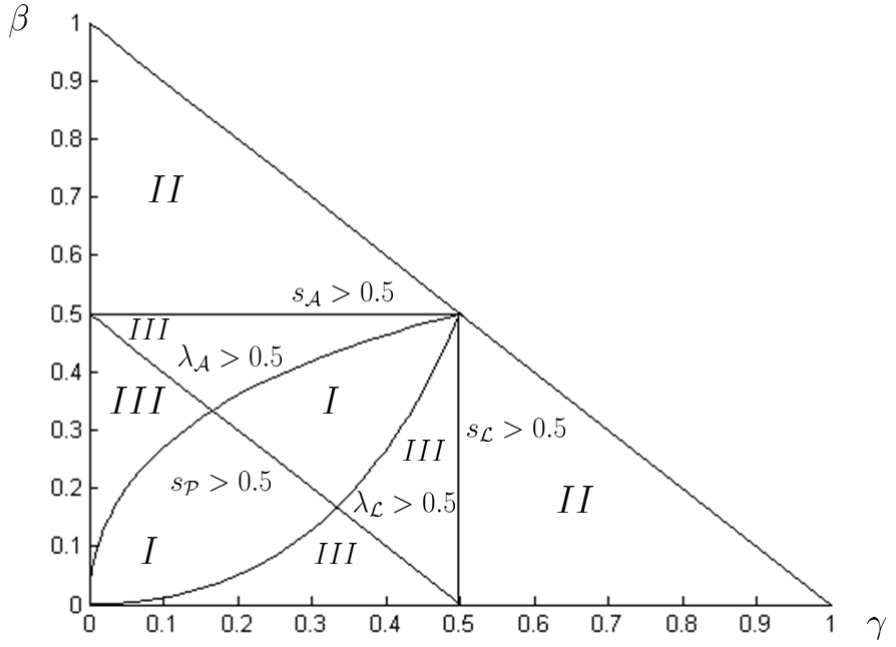

Figure 1.1: Political equilibria with balanced productivity levels ($A_l = A_a = 1, \sigma = 0.5$).

space.[19] Within this admissible γ-β space there is another triangular area of interest. It is defined by the relative size loci given in equations (1.20). All γ-β combinations outside this triangular denote situations where one group represents the majority of the population. For example, South-West of the straight line $\beta = 0.5 - \gamma$ the relative size of the $\mathcal{P}$-group is greater than one half. Similarly, for combinations of γ and β above the the horizontal line at $\beta = 0.5$ group $\mathcal{A}$ constitutes the absolute majority whereas for combinations of γ and β to the right of the vertical line at $\gamma = 0.5$ group $\mathcal{L}$ represents more than half of all people in society.

Since the political regime in equilibrium depends not only on the size but also on the relative power of the most powerful group in society we must also consider the latter criterion. Figure 1.1 shows the relative power loci that correspond to equations (1.18) and (1.19). The concave, upward-sloping locus represents all γ-β combinations for which $\lambda_{\mathcal{A}} = 0.5$ holds. Above this line, the members of group $\mathcal{A}$ generate more than half of total income, $\lambda_{\mathcal{A}} > 0.5$, and therefore constitute the single most powerful group that can dominate in open conflict against any other group or coalition of groups. A

[19] Note that all points lying on one of the three boundaries are not considered in the following since they represent societies with less than three groups, i.e., γ-β combinations for which the size of (at least) one group is zero.

larger endowment of ability than given by this condition – in terms of a higher value of β or combinations of γ and β above this threshold – makes the group $\mathcal{A}$ even more dominant. In this case the political equilibrium is either a mass dictatorship (depicted by areas II) or an oligarchy (areas III) depending on the respective $\gamma - \beta$ combination. The corresponding condition for group $\mathcal{L}$ to be more powerful than the sum of all others is represented by the convex, upward-sloping locus. To the right of this line described by equation (1.18), i.e., for higher values of γ, group $\mathcal{L}$ is strictly dominating and constitutes the ruling elite. Finally, note that Figure 1.1 does not contain a graphical representation of equation (1.17). Since the $\mathcal{P}$-group is disadvantaged in both relevant dimensions it can only rule the state on its own if the income share devoted to the common production factor, $1 - \sigma$, become sufficiently large. Then, the size effect can compensate for disadvantages in factor endowments and a dictatorship of the poor mass can be an equilibrium outcome.[20]

The first main result that emerges from this discussion is the characterization of the conditions, in particular of the distribution of resources in the economy, under which democracy can emerge. These conditions are summarized in terms of areas I in Figure 1.1 which represent all combinations of γ and β for which a democracy arises as an equilibrium. As the figure illustrates, democracy is an equilibrium only when inequality is moderate along the two dimensions γ and β, i.e., for intermediate values. The higher the fraction of individuals with strength or ability within society, the more likely becomes a mass dictatorship in which the respective largest group rules the state on its own. For example, in the northern area II in Figure 1.1 the members of the $\mathcal{A}$-group dominate the political decisions and in the East it is the $\mathcal{L}$-group that dominates all others. Note that in principle a democracy could emerge everywhere in the $\gamma - \beta$ space whereas mass dictatorships can by definition only occur outside the inner triangular area. Thus the admissible $\gamma - \beta$ space for democracies is larger than the one for mass dictatorships. All remaining areas III denote oligarchies where the state is ruled by a single group that represents a minority of the population.

[20] More precisely, $s_{\mathcal{P}}(1-\sigma) > 0.5$ must hold for this to be the case which can only be satisfied for $\sigma < 0.5$ and $s_{\mathcal{P}} > 0.5$. Thus the $\lambda_{\mathcal{P}} = 0.5$ locus can only emerge in the south-western corner of Figure 1.1 for $\sigma < 0.5$. A graphical representation of this case is provided in the Appendix.

1.5 Stability of Democracy

Having identified the conditions for the emergence of democracy, the model also delivers results on its stability with respect to two dimensions: first, it allows for an analysis of secular changes in the distribution of production factors via variations of β and γ, and second, it can be used to trace the consequences of economic development in terms of secular changes in the relative importance of production factors in the income generating process, i.e., variations in A_l and A_a.[21]

The effects of changing the distribution of production factors for a given level of economic development, i.e., for a given combination of A_l and A_a, can already be inferred from the previous discussion of Figure 1.1. In particular, one can directly derive the consequences of *ceteris paribus* changes in the population structure for the politico-economic equilibrium. Applications for such an analysis are numerous. With regards to changes in β one could think for example of massive schooling programs that change the distribution of ability whereas epidemics or improvements in health provision can affect the distribution of strength γ within society. There might also be changes in the population structure that affect both dimensions simultaneously, like asymmetric population growth due to war casualties, ethnic cleansing, displacements, group specific birth rates caused by a quality-quantity trade-off or immigration of individuals with particular endowments of ability and strength. It is obvious that the results will depend on the status quo before the change in population structure, as well as on the distribution of the other factor. Massive increases in β will lead to an equalization of power and make democracy more likely if applied to an economy with relatively few able individuals, and hence increase the likelihood of democracy. Whereas in a situation in which only a few individuals do not have ability, i.e., β is high, such a policy might induce a concentration of political power, and make democracy less likely. In the benchmark case given above, it is a fairly balanced distribution of production factors that provides the optimal environment for democracy to emerge in equilibrium.[22]

[21] Note that the model framework does not account for other non-economic factors that have been considered as being important for the stability of democracy by political scientists, like e.g. civic culture or democratic values, see Almond and Verba (1963) or Putnam (1993).

[22] Since the relative importance of both exclusive production factors is equal in this case, i.e., $A_l = A_a$,

A different picture arises when the effects of changes in the relative productivity of the different factors, reflected by A_l and A_a, on the politico-economic equilibrium are taken into account. Such changes might for example be caused by unbalanced technological progress like skill-biased technological change, by natural disasters or by war. Before going to the characterization of the implications for the politico-economic equilibrium, it is worth noting that for any productivity environment there is always a scope for democracy. This is summarized in the following proposition.

Proposition 1.2. *There always exist admissible $\gamma-\beta$ combinations for which a democracy emerges in equilibrium ...*

1. *... irrespective of the productivity environment A_a and A_l for $0 < \sigma < 0.5$.*

2. *... given a particular productivity environment A_a and A_l for $0.5 \leq \sigma < 1$.*

Proof. See Appendix. □

The results of this Proposition are particularly noteworthy from a policy perspective. They essentially state that the structure of the population, in terms of inequality in factor endowments, rather than the level of development, is the central determinant for democracy if a sufficiently large income share goes to the factors that are distributed equally (i.e., for σ being sufficiently small). In this case democracy can be established for any productivity environment by ensuring a suitable distribution of factors or factor incomes. In other words, democracy is feasible regardless the level of economic development. This implication modifies the introductory statement by Lipset, suggesting that the level of development or income is of secondary importance for the emergence and the stability of democracy compared to the distribution of factors. A similar but less pronounced result holds if the income share going to unequally distributed factors (σ) is relatively high. According to the proposition, democracy is also always a possible equilibrium outcome, but the necessary factor distribution depends on the level of development in terms of the particular productivity environment.

The reverse statement is not true, however, as there exist certain factor distributions

the bisectrix constitutes a symmetry axis of the political landscape.

for which the equilibrium outcome is never a democracy, irrespective of how the productivity environment looks like.

Proposition 1.3. *For any* $0 < \sigma < 1$*, and irrespective of the productivity environment* A_a *and* A_l*, there exist admissible* $\gamma - \beta$ *combinations for which ...*

1. *... a mass dictatorship emerges in equilibrium.*

2. *... an oligarchy emerges in equilibrium.*

Proof. See Appendix. □

Hence, the model suggests that there are limits for the possibility to implement democracies by mere technology or income transfers. To illustrate the implications of variations in the relative importance of factors in the income generating process, we change the baseline scenario and consider two stylized cases. The first one refers to a society in which physical strength is much more important than ability in the production process. This we take into account by setting $A_l = 20$ and $A_a = 1$.

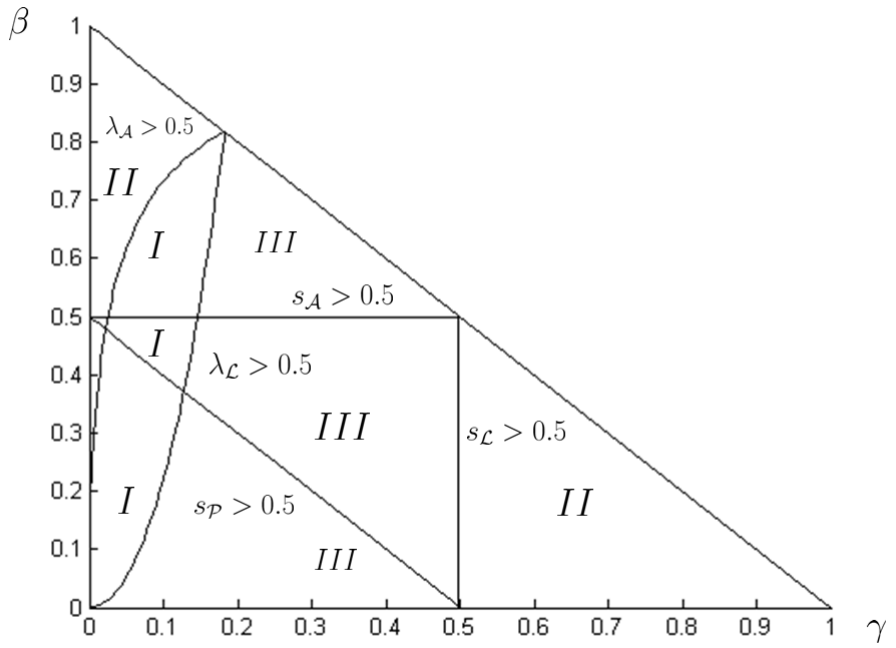

Figure 1.2: Political equilibria in a strength-dominated society ($A_l = 20, A_a = 1, \sigma = 0.5$).

The politico-economic equilibria for a society with strength as the dominant factor of production are depicted in Figure 1.2. Again, as in Figure 1.1, area I represents

democracies whereas in all areas II a mass dictatorship occurs for sure. Finally, all areas III represent oligarchies of the respective minority that is most powerful. The most immediate result of this case is that there is much more scope for oligarchies. Additionally, democracy only emerges as outcome in societies in which strength is a relatively scarce resource, i.e., γ has a low value, whereas it can emerge for a large range of values of β. If γ is too high, a change in β has virtually no effect on the politico-economic equilibrium.

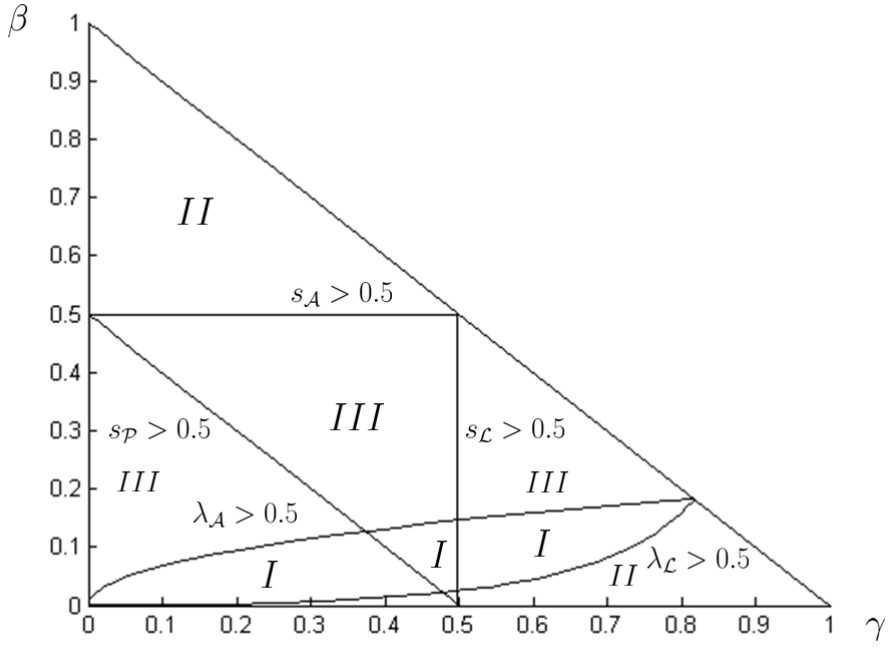

Figure 1.3: Political equilibria in an ability-dominated society ($A_l = 20, A_a = 400, \sigma = 0.5$).

A different, yet somewhat symmetric picture emerges in the second stylized case when considering a developed society. This case represents a society in which physical strength has lost its relative importance and ability has become the predominant factor in the income generating process. In our static model we replicate this kind of skill-biased technological change in a very simplified manner by assuming A_l to stay constant and increasing A_a to 400. This scenario is depicted in Figure 1.3. The figure suggests that changes in γ only affect the equilibrium outcome if the distribution of β is not too high, similar to the previous case. In an economy of this type, in which ability is by far the most important factor for production, even small variations in β, for example due to immigration of high-skilled workers or some other asymmetric change in the demographic

structure, can have far-reaching implications for the politico-economic equilibrium, up to the point that democracy becomes infeasible in equilibrium. In this respect, the model can rationalize to what extent demographic change, in particular with respect to the distribution of low-skilled and high-skilled labor, may provide a challenge for existing democracies. This way, the model can also give some guidance as to what are the likely consequences of drastic demographic changes or policies.[23]

1.6 Empirical Implications and Historical Evidence

To illustrate the model's implications, we begin by discussing evidence from three critical junctures in Germany's recent history, each of which was breaking grounds for the emergence of a completely new political regime: the period following its unification in 1871, the aftermath of World War I, and the period after Germany's capitulation in World War II in 1945. These three dates mark crucial turning points in German history which were preceded by substantial changes in territory, population, and the economic environment in terms of inequality and the corresponding production structure. In all three cases the previous regime had ceased to exist and a completely new political regime had to emerge while the process leading to these breaks was not related to the domestic struggle for political power in any respect. These exogenous breaks from the past were the consequences of a unification process and of the unforeseen outcome of two self-inflicted major wars. In this regard, the three critical junctures in German history provide a perfect setting to illustrate the working of our model. In all three situations different groups of society were confronted with the possibility, or even need, to implement a new political regime that served their purpose. As stipulated by the model, the emerging regime had to be stable and self-enforcing against the background of imminent civil war and open conflict. We conclude our empirical discussion by providing further historical examples from different countries and epochs where wars or exogenous shocks in technology changed the distribution or importance of production factors and subsequently triggered institutional change.

[23] An example would be the one-child policy conducted by the Chinese government which might not be sufficient as a regime-stabilizing measure in the long run since – despite its potentially preserving effects on the population structure – changes in the technological environment are not taken into account.

1.6.1 The Emergence of the German Reich

Prior to 1871, there was a large number of independent German regional kingdoms and principalities. The largest and most dominant of these kingdoms were Prussia and Austria, which formed the so-called *pentarchy* together with England, France and Russia during the 18th and early 19th century. What later was to become Germany therefore essentially consisted of largely independent feudal states, each with a monarch or sovereign that controlled political power. In many of the German states, there were some liberal and democratic movements in the first half of the 19th century that were combined with some tendencies towards a German nation state, and which culminated in the revolution of 1848 and the famous constitution of the Frankfurt assembly. These tendencies were successfully repressed by the leading elites in the years after 1849, however. Eventually, a German nation state emerged under the primacy of Prussia, the so-called "Lesser German solution". At the same time Austria formed a multi-ethnic state with Hungary in 1867. After three unification wars in 1864 against Denmark, in 1866 against Austria and in 1870/71 against France, the German Reich was founded in 1871 in the Hall of Mirrors of the Palace of Versailles near Paris with the proclamation of the Prussian king Wilhelm I. as the first German emperor. As head of the state he appointed Otto von Bismarck first Chancellor (*Reichskanzler*) of the German Reich who at that time also served as head of the executive of the kingdom of Prussia.

The unification of then 17 more or less independent states within one German Reich raised the question about the appropriate political regime.[24] Different interest groups tried to shape the Reich according to their ideas. There was a strong landed gentry and nobility, which had dominated the small states, but there were also tendencies trying to establish a parliamentary democracy following the ideals of the revolution of 1848. However, the democratic movement was split into different factions. Liberal forces were seeking to establish a moderate democracy with monarchistic elements and restricted franchise, while left-wing social democrats wished to establish a democracy with universal franchise and a radical reform of the economic system. Bismarck proposed electoral rules along the lines

[24]For a detailed description of the different political tendencies see Botzenhart (1993), chapters 8 and 9.

of the Frankfurt assembly, but he also deeply despised a truly democratic regime. In fact, to maintain the aristocratic order he repeatedly threatened to prohibit social-democratic and catholic-centrist parties. Bismarck's government was "... in principle hostile towards parties and constitution" (Botzenhart, 1993, p. 119). At several occasions he made clear that he would rather destroy all democratic elements and mount a coup to re-establish a corporatist state under the leadership of the nobility than to concede minimal democratic reforms. And until World War I, the nobility and the conservatives successfully prevented any attempt to implement democratic reforms that would grant the parliament effective influence over the executive or the military whose leaders largely belonged to the nobility.[25]

The success of the conservative nobility can be understood in light of the geographic, economic and demographic structure of the Reich at the time of its foundation. By the 1870s, industrialization had just began in Germany and was mainly concentrated in the Western part of the Reich around Rhine and Ruhr whereas huge areas ranging from the east of the Rhine all the way up to the Baltic Sea beyond *Königsberg* (today's Kaliningrad) and including large states like Bavaria, Wuerttemberg, or Saxonia, were still agrarian. During the 1870s, almost two thirds of the population still lived in the countryside and about half of the active population worked in agriculture while the primary sector contributed a little less than 40% of total net domestic product.[26] These patterns of the economic structure of the German Reich are illustrated in Figure 1.4. At the same time, a huge fraction of Germany was still in the hand of the landed gentry. In 1879 about 40% of all manors in Prussia and Pomerania belonged to them.[27]

In light of Proposition 1.1, this made an oligarchy under the landed gentry the natural candidate for a politico-economic equilibrium in the German Reich. Considering the rural society in Germany around 1871 as one in which physical strength was the dominant factor of production and given the relatively large and strong landed elite, one can infer the high likelihood that the political regime would be characterized by an oligarchy, as illustrated by the areas III in Figure 1.2. Despite the progressing industrialization, the changes in

[25] See also Winkler (1993), p. 610.

[26] Urbanization data is from Bähr (2004). Data on the economic structure is taken from Hoffmann (1965), Tables 1.6 and 2.20, where agriculture includes agriculture, forestry and fishing.

[27] The data for landownership reflect the social structure of ownership of manors, in terms of nobility, middle class and corporate owners; source is Buchsteiner (1993).

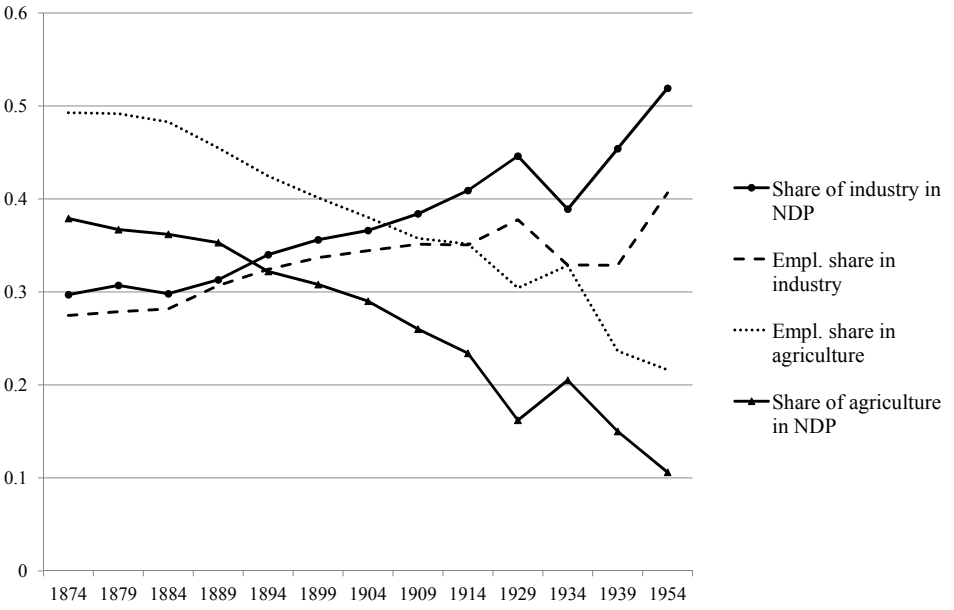

Figure 1.4: Economic Development and Structural Change in Germany.
Five-year averages calculated from Hoffmann (1965), Tables 1.6, 2.20.

the demographic structure due to differential fertility, and the corresponding evolution in inequality, this political regime remained stable until World War I.

1.6.2 From Weimar Republic to Nazi Germany

Germany's entry into World War I was largely independent from domestic politics, and in particular, it was not driven by a democratic movement. In fact, political forces from the entire spectrum agreed in a "class truce" (*Burgfrieden*) to postpone major domestic political reforms to the time after the war. Also the outcome of this war, the defeat of Germany and Austria-Hungary in 1918, came unexpected for the German political leaders and ended with the abdication and exile of the German Kaiser. Hence, the monarchy ceased to exist and an intense political struggle arose about the most appropriate political regime to be implemented. The consequences of the self-inflicted war for Germany were disastrous in terms of national debt, geographic losses and war casualties. More than two million people died on the battlefields, compared to around 14 million soldiers and a total

population of around 67 million in 1914, Eastern Prussia and Alsace-Lorraine were lost and the victorious powers imposed massive reparation obligations on post-war Germany. Moreover, many of the soldiers had been traumatized by their experiences in the trenches and had become callous in the face of the tremendous violence they had seen. In the riots that broke out at the end of the war, many therefore did not shy away from using violence against fellow citizens whom they accused of recklessly forcing them to sacrifice their lives on the battlefield. According to the *Dochstoßlegende* put forward by the military elite the war was not ended on the battle field by foreign military but by domestic socialist forces who were accused of having stabbed in the back of the German army. Despite riots with numerous casualties, an outright civil war could just be avoided. All these factors brought about a fundamental change in the political and economic environment.

Given these circumstances, a coalition of moderate conservatives, liberals and social democrats managed to implement a parliamentary democracy, known as the Weimar Republic.[28] The constitution of Weimar stipulated universal suffrage and control of the parliament over the executive but also contained elements of a strong sovereign in the person of the president who had the power to suspend the parliament and to install a strong executive which could act independently of the legislative in times of crises. In the 1920s, starting with the immense economic burden imposed by the reparation payments and culminating in the Great Depression Germany suffered a permanent instability of its political regime.

In this period, the Weimar Republic saw an ongoing polarization at both ends of the political spectrum given by communists and social democrats on the left, and ultra-conservatives and national socialists on the right. Supported by the ultra-conservative president and former general von Hindenburg, conservative politicians with an inherently anti-democratic attitude implemented deflationary policies and countered the parliament's protests by weakening the parliament more and more.[29] In a series of emergency decrees, the parliament was eventually suspended by Brüning in 1930 despite the obvious strengthening of the Nazi-movement. Brüning was followed by von Papen, formally a conservative

[28]The following description mainly draws on Botzenhart (1993) chapters 13 and 14.
[29]See Schulz (1992), in particular chapters 2 and 10.

from the center, but with an ultra-conservative attitude who intended the establishment of a corporate state under the leadership of the landed aristocracy, and his "cabinet of barons".[30] Despite the widespread belief that the biggest threat to the republic came from the right-wing Nazi-movement, many conservative politicians including von Papen still believed they could contain the Nazis once in power. Von Papen, in a sort of coup, also dismissed the Prussian state government and became acting commissioner of Prussia in addition to being German chancellor. This virtually meant the elimination of the fundamentals of a federal state since police, judiciary and administration of the largest German state were now directly controlled by the chancellor. These newly created structures substantially facilitated the implementation of a dictatorship by the Nazis after the January elections of 1933 that brought Hitler into power.[31] In these elections, the Nazis only collected somewhat less than 44 % of the votes despite the fact that the election campaign and the elections were accompanied by substantial violent repressions of left-wing parties and voters by Nazi paramilitary gangs.[32] The Nazis had to form a coalition with an ultra-conservative party to get to power, but once Hitler was installed as chancellor, it took the Nazis only a bit more than one year to turn Germany into a dictatorship using the loopholes in the Weimar constitution. Power was granted by the fact that Hitler had control over the largest paramilitary army that he used openly to threaten with civil war in the early 1930s.[33] What followed was a political regime that executed unparalleled atrocities in the holocaust and various waves of ethnic cleansing, that started a war which was without comparison in history, and that ultimately led to the destruction of Germany and its political system.

Again, the model presented in this paper can shed light on the mechanisms that led to the emergence of a weak and unstable democracy that eventually gave way to the Nazi regime. Despite the war and the associated losses of territory and population, the population structure in 1919 had remained largely unchanged compared to 1871. Yet, compared to 1871, the economic environment had changed substantially by the end of

[30] See Dederke (1996), p. 247-250.
[31] See Dederke (1996), p. 249.
[32] See Botzenhart (1993), p. 171 and Dederke (1996).
[33] See Winkler (1993), p. 613.

World War I, as is illustrated in Figure 1.4. By the early 1920s, the employment share in agriculture had fallen by roughly 20 percentage points compared to the time after German unification and only about 20% of the net domestic product was produced in the primary sector between 1925 and 1935. At the same time, industrial production had become more important than agriculture and had received a further push by managed efforts to make war production more efficient.[34] Correspondingly, the population in the cities and industrial centers had grown more than twofold leading to a significant increase in the urbanization rate with roughly two thirds of the population now living in urban areas.[35] Nevertheless, there possibly exists no parallel of another industrialized society in which a pre-industrial elite could retain as much political power as the landed gentry in the Weimar Republic, see Winkler (1993). The ongoing economic development caused a structural change in the economic environment. Assuming in the absence of reliable data that working hours and capital intensities per person were roughly equal and constant in both sectors, Figure 1.4 implies that the relative productivity A_a/A_l between the two sectors was much larger in the Weimar Republic than in times of the German Reich.

In terms of the stylized model, this substantially higher relative productivity results in a smaller income gap between the two high-income groups, $\mathcal{L}$ and $\mathcal{A}$ as compared to the agricultural society of 1871.[36] In fact, the available data on pre-tax income inequality as measured by the the Pareto coefficient α seems not to contradict this interpretation.[37] Figure 1.5 below shows Pareto's α over the period 1871 to 1938 for three of the major states of the German Reich, Baden, Prussia, and Saxony, as well as for the Weimar Republic as a whole.[38] The figure suggests that equality of incomes was on average lower

[34] See Schulz (1987), chapters 4 and 5 for details.

[35] See Bähr (2004).

[36] Note that this does not imply a monotonous relation between an increase in the relative productivity and income inequality in our model.

[37] The Pareto-coefficient used to be a common measure of inequality. When observing the distribution of incomes Pareto (1896) assumed the number N of people earning at least income x to be best described by some function $N = Ax^{-\alpha}$ with A and α being constants. According to this reasoning, in a society with a high α a smaller fraction of individuals earns an income equal or above x than compared to a society with a lower α. Despite all obvious shortcomings of this measure, see e.g. Lorenz (1905) or Bresciani-Turroni (1939), in this respect higher values of α are considered to represent more equal income distributions ceteris paribus. To the best of our knowledge, there are no other data on income inequality available for the time period under consideration.

[38] Baden, Prussia and Saxony represent about 70% of total population living in the German Reich, see Hohorst, Kocka, and Ritter (1975), and almost 75% of that in the Weimar Republic, see Petzina, Abelshauser, and Faust (1978). Population data is available only for 1871, 1890, 1910, 1925, 1933, and

in the times of the German Reich than in the times of the Weimar Republic, as reflected by the lower values of Pareto's α in the period before World War I.

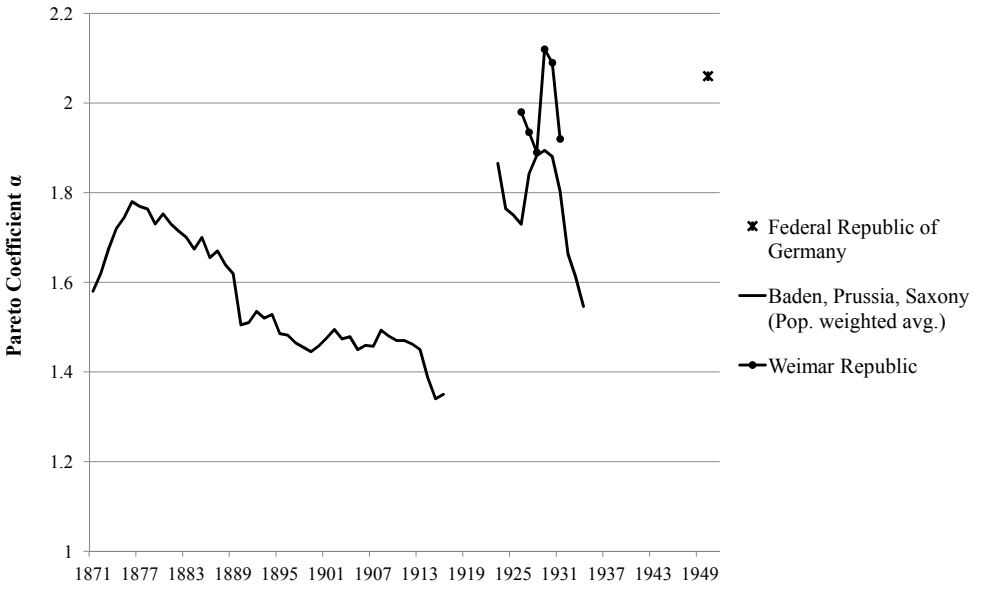

Figure 1.5: Income Distributions in Germany.
Data from Hoffmann (1965), Table 2.123, Hohorst, Kocka, and Ritter (1975), and Petzina, Abelshauser, and Faust (1978).

The data also show that equality increased during the first years after the foundation of the Reich in 1871, but then slowly decreased over time until World War I.[39] Considering this measure of income distribution, it therefore appears as if incomes in the Weimar Republic were more equally distributed than in the German Reich.[40] And as Dumke (1991) points out, this difference mainly stems from changes in the composition of the high-income groups. Given the evidence presented in Figures 1.4 and 1.5, it therefore appears as if in the Weimar Republic the economic environment is best described by the one depicted in Figure 1.1. With moderate levels of β and slightly higher values of γ, the

1939. The remaining data are calculated by assuming exponential population growth between these dates.

[39]In 1891 a major tax reform was implemented in Prussia in order to collect more reliable income data. Thus, the true value of the Pareto coefficient might have already been lower in the years before. With Prussia representing around 90% of the population in our sample this could explain the pronounced drop of the Pareto coefficient in 1891.

[40]This conclusion is also in line with the assessment of Kuznets (1955) regarding the evolution of income inequality in Germany in the 18th and 19th century, as well as with the evidence cited by Acemoglu and Robinson (2000).

equilibrium emerged in an area of Figure 1.1 where all three regimes had some relevance, somewhere in the central lower part. This explains the inherent instability of the Weimar Republic despite its even by modern standards progressive democratic design.

The dynamics during the Weimar Republic that eventually led to the collapse of democracy and the emergence of the Nazi regime can also be rationalized in light of the model. Despite their political influence at the beginning of the Weimar Republic, the landed aristocracy lost influence slowly but significantly during the 1920s as a consequence of the ongoing structural changes shown in Figure 1.4, see also Zollitsch (1999).[41] This creeping descent of the political power of the nobility together with the presence of a structural indeterminacy in the Weimar Republic, in terms of different feasible equilibria, might have provided the scope for a single, unscrupulous man with extraordinary political talents to ultimately implement and stabilize an oligarchy of the "strong", the Nazis, as reflected by area III.[42] By repression, expropriation and terror, the Nazis formed the economy in terms of population composition and inequality in a way to stabilize their regime and rule out any political change until their ultimate defeat in 1945. In fact, Figure 1.5 suggests that inequality increased substantially during the first years of the Nazi regime, reflected in a pronounced drop of Pareto's α.[43]

1.6.3 Towards a Federal Republic of Germany

The war and the terror caused by the Nazis devastated large parts of the world, and also Germany. Until 1945, acts of war and inhumanity like the holocaust and different measures of ethnic cleansing had cost the lives of more than 10 million Germans. Huge territorial losses like almost the entire Prussia and a virtually complete destruction of infrastructure and production facilities marked a deep caesura (the so-called "hour zero",

[41] This is reflected in political reforms at the expense of the aristocracy that were implemented in this period. For example, from 1928 on, manor districts were legally no longer considered as political entities on a communal level in Prussia which effectively ended the political dominance of land owners in such communities.

[42] Such an interpretation taking into account the interplay between the structural setting and individual talent would be in line with the famous study of Bullock (1964) on the reasons for the rise of Hitler. One could also argue that given the directed changes in the economic structure and its high popularity, Nazism eventually turned out being a mass dictatorship during the late 1930s and is therefore better represented by area II.

[43] See also Jeck (1968) and Morrisson (2000).

Stunde Null) in German history. The landed elite had lost most of their estates, and had been decimated substantially by war and repression.[44] As a result, Germany 1945 looked entirely different from what it used to be in the early days of the Weimar Republic, or even at the beginning of the Nazi dictatorship. The political regime had to be completely re-established under the occupation by the four allied forces, the USA, England, France and the Soviet Republic. In the so-called "Bizone", the territory occupied by the USA and England, the Germans were given the freedom to develop a federal structure and a constitution that would enable the formation of a democratic political system. The drafts of the new constitutions for federal states (*Länder*) and for communities had to be authorized by the respective occupying powers. This happened fairly quickly in the years after the war without a significant exertion of influence on the constitutional details by the allies. The construction of a national state was impeded by the diverging interests of the allied forces with the Soviets trying to expand their direct influence in Germany. Struggles between the Western allies and the Soviets culminated in different policies of containment that led to the Cold War, and eventually to the establishment of two German states in 1948.[45] The Federal Republic of Germany (FRG) as a parliamentary democracy was established in the Western part of Germany comprising the American, English and French occupation zones. In their zone, the Soviets installed the so-called German Democratic Republic (GDR) as a socialist state which was essentially undemocratic and heavily controlled by the Soviets. Since the possibility of external influence is beyond the consideration of our model we focus on the FRG in the following.[46]

The constitution of the federal republic stipulated a modern parliamentary democracy with universal and equal franchise. After the demise of the old elite, the defeat of the Nazis, and under the impression of the Soviet influence in the GDR, the political climate in Germany was very moderate and pragmatic. There was a consensus among the three

[44] As in World War I, large parts of the officer corps consisted of men of the nobility but the Nazis became increasingly distrustful about the support of the nobility during the war, which is why the nobles lost influence in the military. As a reaction to several attempts to assassinate Hitler, culminating in the attempt of July 20, 1944, by von Stauffenberg and his group, the Nazi regime killed many officers they accused of conspiracy, a large fraction being of noble decent, see e.g. Reif (1999).

[45] See Eschenburg (1983), p. 375.

[46] Historically, this procedure appears to be justified also on grounds of the fact that the GDR collapsed in 1989 and joined the FRG in the process of German re-unification in 1990. The political system of the FRG was sustained in re-unified Germany.

leading political movements, the social democrats, the conservatives and the liberals, to establish a political regime that would avoid the loopholes and problems of the Weimar Republic while keeping some of its progressive features.[47] Social democrats and conservatives had their own proposals, but in a famous "constitutional convent", the draft of a new constitution (*Grundgesetz*) was finished in just two weeks. The constitution emerged from this draft after it was ratified by the constitutional assembly and authorized by the allied forces without serious objections, is – with minor modifications – still in effect today.

In the light of the model presented in this paper, the emergence of a stable democracy that encompasses all parts of society and that in many ways still represents a model democracy for large parts of the world today can be rationalized by the very equilibrated economic and political interests after the devastation and the defeat in the self-inflicted Second World War. Abstracting from the immense war-related destruction of production facilities, the industrial structure was very similar to that in the Weimar times, as exemplified by the virtually identical employment shares in agriculture over the period from the 1920s to the late 1940s.[48] From Figure 1.4 one can also see that the primary sector had lost further economic importance during the Weimar Republic and the Nazi period. However, when again assuming constant working hours and capital intensities per person the relative productivity between both sectors seems to not have changed dramatically indicating a fairly similar economic environment to that in 1919. Thus, the German economy at that time might still be best described by a level of development as depicted in Figure 1.1 – with the most salient difference being the much lower inequality after World War II due to a more equal demographic structure.

According to the study of Morrisson (2000) who analyzes the income distribution within Germany from 1870 to 1990 inequality had decreased after World War II compared

[47]Scholars like Dahrendorff (1965) went as far as to argue that the destruction of the traditional social structures by the Nazis, and the ultimate defeat of Nazi Germany constituted a prerequisite for the emergence of the liberal and modern democracy after 1945. According to Dahrendorff (1965, p.155), it was the social inequality and the persistence of traditional corporative structures that had impeded any progress towards a modern democracy in the *German Reich* and in the *Weimar Republic*. In his view, this was mainly because it was in the interest of the politically dominant groups of society, consistent with the discussion above.

[48]Mitchell (2003, Table B1) provides data for the first year directly after the war. According to his numbers, the shares of economically active men working in agriculture evolved from 23.3 % in 1925 to 22.5 % in 1933 when Hitler came to power, to 22.6 % in 1946. For women, the respective numbers are 43.3 %, 40.7 % and 40.6 %.

to the pre-war era, and then remained almost stable at historically low levels. With all groups of society being of similar importance, and the need to focus attention on reconstruction and development, the politico-economic equilibrium in this situation is likely to be one of democracy, as represented by areas I. To complement this picture, Figure 1.6 displays the development of income levels in Germany using the data from Maddison (2003). With income levels after World War II being even lower than after World War I or at the end of the Weimar Republic, the data seems to lend further support to the model prediction that inequality and the distribution of resources, rather than the level of income, crucially affect the institutional equilibrium.

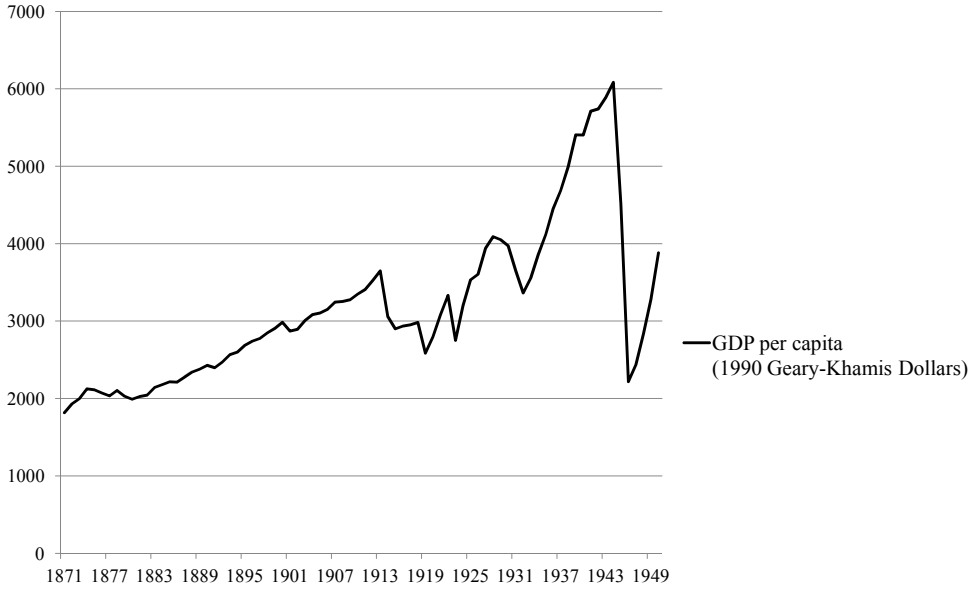

Figure 1.6: Income Development in Germany.
Data from Maddison (2003), Table 1c.

In the aftermath of World War II, the reconstruction and development in Germany was fairly balanced, the recovery was fast and accompanied by policies intended to stabilize the political and social situation. Led by later chancellor Ludwig Erhard, the currency and the welfare system were reformed in order to flatten out wealth inequality and to level economic conditions, at least to a certain extent. The goal was to implement a so-called "social market economy" which would mitigate social conflicts, ensure a moderate political

climate and thereby stabilize the economy by maintaining the conditions necessary for a democratic politico-economic equilibrium. The recovery was complemented by international support through the Marshall plan, which was intended as a measure of policy of containment against the Soviets.[49]

1.6.4 More Historical Examples

The theoretical result that economic inequality, reflected by the distribution of resources, might be the crucial factor for institutional stability as well as for institutional change appears to be consistent with many other observations in history. In fact, Rogowski and MacRae (2008) argue that the majority of historians, including economic historians, agree that exogenous changes in inequality due to warfare or shocks in technology or demographics have been the major determinant for institutional change in human history.[50]

The different historical events consistent with this line of causality include the emergence and collapse of the democracy in Ancient Greece, the collapse of the Roman Republic into a de-facto dictatorship, the rise of feudalism under the Merovingians, the tendencies towards liberation of tenants in response to the Black Death in the 14th century, the reformation movement in the 15th and 16th century, the rise of absolutism in the 17th century, as well as the first and second waves of democratization in the early 19th and in the 20th century, respectively. Rogowski and MacRae (2008) make the convincing case that all these events were triggered by an initial exogenous shock that changed inequality and subsequently led to a change in political institutions. In line with the theoretical predictions of the model above, changes in inequality and institutions appear to occur in both directions, indicating that the stability of institutions in general, and of democracy in particular, might crucially depend on inequality. For instance, democratic reforms in ancient Greece were induced by a drop in inequality which was caused by changes in military technology that made a lightly armored infantry, the *hoplites*, favorable over the

[49] See also Eschenburg (1983), p. 432-445.

[50] Rogowski and MacRae suggest that these exogenous shocks are likely to have changed both inequality and institutions, and propose an explanation that is based on standard political economy arguments regarding the link between inequality and public goods provision. Although this explanation is consistent with different degrees of franchise that are optimal from the perspective of the rich political entrepreneurs, it also rests on the implicit assumption of an institutionalized environment unlike the theory presented above.

former dominant knights. Conversely, the increasing professionalisation of the army in the Roman Empire, or the adoption of the stirrup in the middle ages, gave more economic and political power to smaller elites. This led to transitions towards more oligarchic structures, reflected in the breakdown of the Roman Republic and the rise of Caesarism, or the rise of feudalism in medieval Europe. Likewise did demographic or technological shocks, such as the Black Death or the movable type, which led to a reduction in inequality, eventually give rise to a more equal distribution of political power. The wars of the Napoleonic times as well as the World Wars required huge armies and workforces, thereby also causing a reduction in political and economic inequality.

In summary, the historical evidence collected by Rogowski and MacRae (2008) suggests that exogenous changes in inequality were the major driver behind major institutional changes in history. One aspect that appears noteworthy in this context is the fact that most of the historical examples of changes in inequality were also associated with an increase in economic living standards. In some cases, however, inequality increased, whereas in others it decreased, pointing towards the importance of inequality, rather than income levels, for the emergence and stability of political institutions.

1.7 Concluding Remarks

This paper has developed a model of political institutions in which democratic or oligarchic rule emerges as equilibrium outcome of a political struggle for redistribution in a weakly institutionalized environment where no binding agreements between different groups of society can be made. The results show that factual inequality along several dimensions, in terms of the distribution of factors in the economy as well as of their importance in the income generating process, is key for the resulting political institutions. Democracies can emerge only in fairly balanced economic environments whereas alternative conditions give rise to various forms of oligarchies or mass dictatorships. The main results are robust to several extensions of the model, including the consideration of more income groups or the implementation of other production technologies, and are in line with historical evidence on changes in political regimes. In particular, the model can rationalize the emergence

of different political regimes in Germany after its unification in 1871. Apart from that, our results suggest that the advent of democracy neither is an indispensable event in the process of development nor necessarily marks the beginning of an era of eternal stability of democracy.

The results have several relevant implications. First, democracy might not be the automatic outcome of economic development, consistent with the evidence produced by Acemoglu et al. (2008, 2009). The fact that different political regimes can emerge for comparable levels of income and economic development is also suggested by the historical data for Germany as provided by Figure 1.6. Second, the model shows that the distribution of factors or incomes, rather than the level of economic development, very often is key for democracy to emerge. Again, this result is supported by the observations from Germany's history. Facing fairly equal levels of income at three critical junctures, a stable democracy in Germany emerged only after the third juncture in 1945, when the distribution of factors and incomes was more balanced than at the previous junctures, see also Figure 1.5. Our results suggest that democratization is possible at every level of economic development if the distribution of production factors lies within a certain range. Third, the model also implies a note of caution in that the situation of a stable democracy does not necessarily constitute an absorbing state. Even if democracy eventually emerges, it might not be stable in the long run but only be a temporary phenomenon. Hence, democracy might fail if income discrepancies and redistributive tensions between the different social groups become too large. The model characterizes the conditions under which this is the case. In particular, ongoing technical change and economic development that affects different groups of society in different ways, as well as increases inequality and polarization may potentially lead to a breakdown of the democratic equilibrium and to the (re)emergence of an oligarchy or a mass dictatorship.

The model presented in this paper suggests various directions for future research. Several implications of the model can be tested empirically, including the prediction that democracies are more likely to emerge in balanced economic environments with fairly equal factor incomes. In this context, it would also be interesting to investigate the structure of democracy, in terms of the emerging ruling coalitions, under different scenarios of

economic development and resource distribution, as well as the corresponding patterns of income redistribution. Another interesting avenue for future research would be to link the model closer to the empirical and theoretical concepts of polarization and fractionalization, as developed by Esteban and Ray (1994, 2008) and Alesina et al. (2003). Finally, a dynamic version of the model could be used to investigate the interdependencies of the political regime and the corresponding policies on the one hand, and endogenous factor accumulation and technological change, as well as the associated changes in the income distribution on the other.

1.8 Appendix

1.8.1 Proof of Lemma 1.1

We first show that for any group there exists a pure strategy profile σ^* that is a $SPNE$ and leads to a unique RC.

Part I. Existence. This part of the proof follows the structure of the proof of *Lemma 1* in Acemoglu, Egorov, and Sonin (2008). First consider the preferred coalition of agenda setter $i \in S_k$ at stage $k \in \{k \in \mathbb{N}_0 : k \leq 2\}$ of game Γ where S_k is the set of all (remaining) groups whose income has not been redistributed away up to the current stage of the game, i.e., $S_0 = S$, $S_k \subset S \, \forall \, k > 0$ and $S_k \neq \varnothing \, \forall \, k$. Let $\mathscr{P}(S_k)$ denote the power set of S_k and let $\mathcal{I}_i = \{\mathcal{I} \in \mathscr{P}(S_k) : i \in \mathcal{I}\}$ be the set of all coalitions that include group i whereas $F_i = \{F \in \mathcal{I}_i \backslash S_k : 2P_i > P_F\}$ represents the set of all coalitions in which group i is more powerful than the other coalition members at the current stage of the game excluding the set of all (remaining) groups. For notational convenience we set $P(\varnothing) = \infty$. Let $W_k = \{W \in \mathscr{P}(S_k) : P_W > 0.5 P_{S_k}\}$ be the generic set of winning coalitions and denote the set of stable coalitions as $E_k = \{E \in \mathscr{P}(S_k) : [\nexists \, Q \subset E : 2P_Q > P_E \wedge |Q| = 1]\}$. Additionally, let us define the union of the set of coalitions that are both winning and stable and the set of all (remaining) groups at the current stage of the game which is given by $R_k = [W_k \cap E_k] \cup S_k$ where the coalition that exhibits the lowest aggregate power in this set is given by

$$\Omega_k = \underset{X \in R_k}{\operatorname{argmin}} P(X) \ .$$

Then, the preferred proposal of an agenda setting group i at stage k of the game is given by

$$\Pi_{i,k} = \underset{X \in \mathcal{I}_i \cap R_k}{\operatorname{argmin}} P(X)$$

and the pure strategy profile for group i reads

$$\sigma^*_{i,k} = \begin{cases} \text{agenda-setting stage: } i \text{ proposes } \Pi_{i,k} \\ \text{voting stage: } i \text{ votes} \begin{cases} yes & \text{if } \Pi_{j,k} = \Pi_{i,k} \vee \Pi_{j,k} \in F_i \\ no & \text{otherwise .} \end{cases} \end{cases}$$

where $\Pi_{j,k}$ denotes the proposal made by group $j \in S_k$ on which groups currently vote. Now we need to prove that the pure strategy profile σ^* which is a vector of $\sigma^*_{i,k} \forall\, i,k$ constitutes a $SPNE$. Since we consider a finite game it is sufficient to show that there exists no one-shot deviation from $\sigma^*_{i,k}$ which is profitable for group i at any given history $\hbar$ of the game. In order to do this we need to distinguish two cases each one itself containing two subcases, since in this sequential game any group i is either an agenda setter (*case A*) or a voter (*case B*) at a given history of the game, and any proposed redistribution policy can either be rejected (*subcase 1*) or accepted (*subcase 2*).

Case A

In this case we show that group i cannot benefit from making a proposal $\pi_{i,k} \in \mathcal{I}_i$ that differs from that stipulated by $\sigma^*_{i,k}$. We need to distinguish two subcases.

Subcase A.1. Let us assume that there exists such an alternative proposal $\pi_{i,k} \neq \Pi_{i,k}$ and that $\Pi_{i,k}$ is rejected if proposed. Then, obviously $\pi_{i,k}$ must be accepted if proposed as otherwise group i would not benefit from making this proposal. By definition we know that $\Pi_{i,k} \in \mathcal{I}_i \cap R_k$ holds. Suppose first that $\Pi_{i,k} = \Omega_k$. In this case, $\Pi_{i,k}$ is only rejected by others if $\{i\} = \Omega_k$ holds. But then making a proposal $\pi_{i,k} \neq \Pi_{i,k}$ cannot be beneficial for group i. Now assume $\Pi_{i,k} \neq \Omega_k$. This implies $\{j\} = \Omega_k$ with $j \neq i$. Obviously, since $\Pi_{i,k}$ is not accepted in this case there can exist no $\pi_{i,k} \in \mathcal{I}_i$ which would not also be rejected. For this reason no deviation from $\sigma^*_{i,k}$ can be beneficial in the given subcase.

Subcase A.2. Let us now suppose that there exists an alternative proposal $\pi_{i,k} \neq \Pi_{i,k}$ and that $\Pi_{i,k}$ is accepted if proposed. Note that by the nature of the game $\pi_{i,k} \in R_k$ holds as no proposal $\pi_{i,k} \notin W_k$ can be accepted and no proposal $\pi_{i,k} \notin E_k$ will be accepted. Furthermore, no proposal $\pi_{i,k} \notin \mathcal{I}_i$ can be made by group i. Hence $\pi_{i,k} \in \mathcal{I}_i \cap R_k$ needs to hold.

Given our assumption of a bijective power mapping $\pi_{i,k} \neq \Pi_{i,k}$ then implies $P_{\Pi_{i,k}} < P_{\pi_{i,k}}$ since $\Pi_{i,k} = \text{argmin}_{X \in \mathcal{I}_i \cap R_k} P(X)$. With regards to the optimization problem (1.11) we can therefore conclude that it is not beneficial for group i to propose $\pi_{i,k}$ instead of $\Pi_{i,k}$ in the given subcase.

Case B

Subcase B.1. Suppose that instead of voting *yes* according to $\sigma^*_{i,k}$ voter i would be better off if he voted *no*. Since the votes of the other groups do not depend on the decision of group i such a behavior could only cause a rejection of a proposal that would have been accepted otherwise if group i is pivotal for the decision outcome. In every other case such a deviation has no effect on equilibrium outcome and therefore cannot be beneficial. For this reason, let us assume that group i is pivotal for the decision outcome and that it votes *no* contrary to $\sigma^*_{i,k}$.[51]

To understand why no such deviation can be beneficial for $\Pi_{j,k} = \Pi_{i,k}$ is trivial since from the perspective of group i there exists no better proposal on which will be voted on at the given stage of the game according to σ^* (see *Case A*). Note that if $\{i\} \neq \Omega_k$ then $\Pi_{i,k} = S_k$. In this case voting *yes* of the pivotal group i would preserve the status quo and prevent the possibility of $\Pi_{j,k} = \Omega_k$ being made afterward. Voting *no* and thereby rejecting a proposal $\Pi_{j,k} = \Pi_{i,k}$ can therefore not be beneficial. Now suppose that $\Pi_{j,k} \in F_i$ holds. Since F_i consists of all coalitions in which group i is more powerful than all other coalition members, it is clear that whenever one of those coalitions is proposed and accepted given $\{i\} \neq \Omega$, then $\widetilde{p}_i$ becomes maximal at the subsequent stage of the game. Thus, in the absence of redistribution costs it cannot be beneficial to vote *no* if $\Pi_{j,k} \in F_i$. Note that $S_k \notin F_i$ such that voting *yes* for $\Pi_{j,k} \in F_i$ always implies group i to become an agenda setter again at the subsequent stage of the game. We can therefore conclude that it is not beneficial to vote *no* contrary to $\sigma^*_{i,k}$ for any group $i \in S$ at any stage of the game.

Subcase B.2. Now suppose that instead of voting *no* according to $\sigma^*_{i,k}$ group i would be better off if it voted *yes*. Again, this could only affect equilibrium outcome if group i's

[51] For consistency, and without loss of generality, the strategy of non-pivotal or indifferent voters is also assumed to be characterized by σ^* in the following.

decision is pivotal and leads to the acceptance of a proposal that would have been rejected otherwise. Let us assume it does. Here, we only need to consider the case where $\Pi_{j,k} \neq \Pi_{i,k}$ and $\Pi_{j,k} \notin F_i$ holds. In this case, it is obvious that a deviation from $\sigma^*_{i,k}$ cannot be beneficial for group i since such a decision would lead to an unstable coalition in which group i is not the most powerful group. Given this, income of group i would be redistributed away at the subsequent stage of the game if such a proposal was accepted. We can therefore conclude that it is not beneficial to vote *yes* contrary to $\sigma^*_{i,k}$ for any group $i \in S$ at any stage of the game. □

Part II. Uniqueness. Finally, we need to show that all $SPNEs$ lead to the same RC. Note that the assumption of a bijective power mapping implies that in equilibrium different RCs cannot be equal in aggregate power. To see this suppose, by contradiction, that $P_M = P_Q$ holds for the two equilibrium coalitions $M, Q \in \mathscr{P}(S) \setminus \{\varnothing\}$ which are not identical, $M \neq Q$. Obviously, the bijective power mapping directly rules out $|M| = |Q| = 1$ in the given case. Additionally, a coalition of two groups can never be an equilibrium outcome, because, due to the bijective power mapping, it would not be stable as the stronger group could always propose a winning subcoalition only containing itself at a later stage of the game. And uniqueness in the case of the grand coalition comprising all three groups is trivial. Therefore we can conclude that in equilibrium any two coalitions M and Q can only be equal in power, $P_M = P_Q$, when they are identical, $M = Q$.

Under strategy profile σ^* the resulting RC does not depend on the moves of nature. Therefore the $SPNEs$ in our finite coalition formation and redistribution game with perfect information can only lead to different RCs if a pivotal group i is indifferent about her action at a certain decision node. Suppose first that group j is not part of the equilibrium coalition and is indifferent at a given history of the game $\hbar$. Then $|\Omega| = 1$ immediately follows and group j cannot be pivotal. Now suppose that the pivotal group i is part of different equilibrium coalitions and is indifferent at a given history of the game $\hbar$. This can only be the case if (at least) two actions lead to the same equilibrium payoff which requires – given the optimization problem (1.11) and the political power of group i – the aggregate power of (at least) two different RCs to be the same. With regards to

our former reasoning this is impossible. Thus there cannot exist two different equilibrium coalitions between which any pivotal group $i \in S_k$ is indifferent at a given history of the game $\hbar$. This establishes the proof of Lemma 1.1. □

1.8.2 Proof of Proposition 1.2

Proof of Proposition 1.2.1. The proof shows that for $0 < \sigma < 0.5$ there exist $\gamma - \beta$ combinations for which a democracy emerges in equilibrium irrespective of the productivity environment A_a and A_l.

Note first that from $0 < \sigma < 0.5$ it follows directly that $0 < 1 - \frac{1}{2(1-\sigma)} < 0.5$ which implies with regard to equation (1.17) that the $\beta_{\lambda_{\mathcal{P}}=0.5}$-locus intersects the β-axis at positive values. Thus, we must consider the relative power loci of all groups for the following analysis. Suppose that $\gamma^* = \beta^* = \frac{0.5-\sigma}{2(1-\sigma)} + \epsilon$ where ϵ is some positive parameter which is infinitely small such that $0 < \gamma^*, \beta^* < 1$. Using equation (1.14) we find that

$$\lambda_{\mathcal{P}}(\gamma^*, \beta^*) = 0.5 - 2\epsilon(1-\sigma) \leq 0.5$$

always holds. Applying the same reasoning to equation (1.15) yields

$$\lambda_{\mathcal{L}}(\gamma^*, \beta^*) = 0.25 + \epsilon(1-\sigma) + \sigma\left(\frac{A_l}{A_a + A_l} - 0.5\right)$$

where obviously $\lim_{\epsilon \to 0} \lambda_{\mathcal{L}}(\gamma^*, \beta^*) \leq 0.5$ for $0 < \sigma < 0.5$. It is straightforward to obtain an analogous result for $\lambda_{\mathcal{A}}(\gamma^*, \beta^*)$ when using equation (1.16). We can therefore conclude that for $\gamma^* = \beta^* = \frac{0.5-\sigma}{2(1-\sigma)} + \epsilon$ with $0 < \sigma < 0.5$ and small enough values of ϵ no group is powerful enough to rule the state on its own irrespective of the productivity environment A_a and A_l. From Proposition 1.1 it then follows that a democracy emerges as equilibrium outcome. □

Proof of Proposition 1.2.2. The proof shows that for $0.5 \leq \sigma < 1$ and a given productivity environment A_a and A_l there exist $\gamma - \beta$ combinations for which a democracy emerges in equilibrium.

Since $0.5 \leq \sigma < 1$ holds we do not need to consider the $\beta_{\lambda_{\mathcal{P}}=0.5}$ locus in the following.

From equation (1.19) it can easily be seen that within the admissible $\gamma - \beta$ space the $\beta_{\lambda_{\mathcal{A}}=0.5}$ locus is a continuously differentiable and monotonically increasing function in γ which passes through the origin. Additionally, we know that for any given value of γ an increase in β increases the relative power of the $\mathcal{A}$-group but decreases the relative power of all others. Thus, for any given γ the value of β which is necessary to reach a certain relative power (like 0.5) must be higher for the $\mathcal{A}$-group than for the $\mathcal{L}$-group. Therefore the $\beta_{\lambda_{\mathcal{A}}=0.5}$ locus always lies above the $\beta_{\lambda_{\mathcal{L}}=0.5}$ locus within the admissible $\gamma - \beta$ space. And since both loci do not intersect and therefore cannot be identical due to $\lambda_{\mathcal{P}} > 0$ we can conclude that, for a given productivity environment A_a and A_l, there must always exist a set of $\gamma - \beta$ combinations such that $\lambda_{\mathcal{L}} \leq 0.5 \wedge \lambda_{\mathcal{A}} \leq 0.5$ holds. This establishes the proof. □

1.8.3 Proof of Proposition 1.3

Proof of Proposition 1.3.1. The proof shows that for $0 < \sigma < 1$ there exist $\gamma - \beta$ combinations for which a mass dictatorship emerges in equilibrium irrespective of the productivity environment A_a and A_l.

With regard to equations (1.17) and (1.20) we see that the $\beta_{\lambda_{\mathcal{P}}=0.5}$ locus is a parallel of the $s_{\mathcal{P}}$ locus where the former always lies below the latter. Thus for $\sigma \in (0, 0.5)$ there always exist admissible $\gamma - \beta$ combinations for which $\lambda_{\mathcal{P}} > 0.5 \wedge s_{\mathcal{P}} > 0.5$ holds independent of the productivity environment A_a and A_l. This also is true for $\sigma \in [0.5, 1)$ since we already know from the proof Proposition 1.2.2 that at least for small enough values of γ there must exist $\gamma - \beta$ combinations directly above the $\beta_{\lambda_{\mathcal{A}}=0.5}$ locus which satisfy these two conditions with regard to group $\mathcal{A}$. □

Proof of Proposition 1.3.2. The proof shows that for $0 < \sigma < 1$ there exist $\gamma - \beta$ combinations for which an oligarchy emerges in equilibrium irrespective of the productivity environment A_a and A_l.

Let us focus on the $\beta_{\lambda_{\mathcal{L}}=0.5}$ locus and highlight some of its properties. As can easily be seen from equation (1.18) the $\beta_{\lambda_{\mathcal{L}}=0.5}$ locus has a pole at $\gamma = 0.5/(1-\sigma)$ and intersects the abscissa at most twice for $\gamma_1 = 0$ and $\gamma_2 = (0.5-\sigma)/(1-\sigma)$ with $0 < \gamma_2 < 0.5 \; \forall \, \sigma \in (0, 0.5)$

and $\gamma_2 \leq 0 \,\forall\, \sigma \in [0.5, 1)$. Also $\partial\lambda_{\mathcal{L}}/\partial\gamma > 0 \,\forall\, \gamma \geq \max[\gamma_1, \gamma_2]$ holds. Thus for $\sigma \in [0.5, 1)$ the $\beta_{\lambda_{\mathcal{L}}=0.5}$ locus is a continuously differentiable and monotonously increasing function in the admissible γ set which passes through the origin. This implies that there always exists a set of $\gamma - \beta$ combinations below the $\beta_{\lambda_{\mathcal{L}}=0.5}$ locus for which $\lambda_{\mathcal{L}} > 0.5 \wedge s_{\mathcal{L}} < 0.5$ holds irrespective of the productivity environment A_a and A_l. The same is true for $\sigma \in (0, 0.5)$ since $\gamma_2 = \max[\gamma_1, \gamma_2] < 0.5$ implies that a set of such $\gamma - \beta$ combinations exists to the left of the $s_{\mathcal{L}}$ locus. This establishes the proof. □

Exemplary Figure for $\sigma < 0.5$

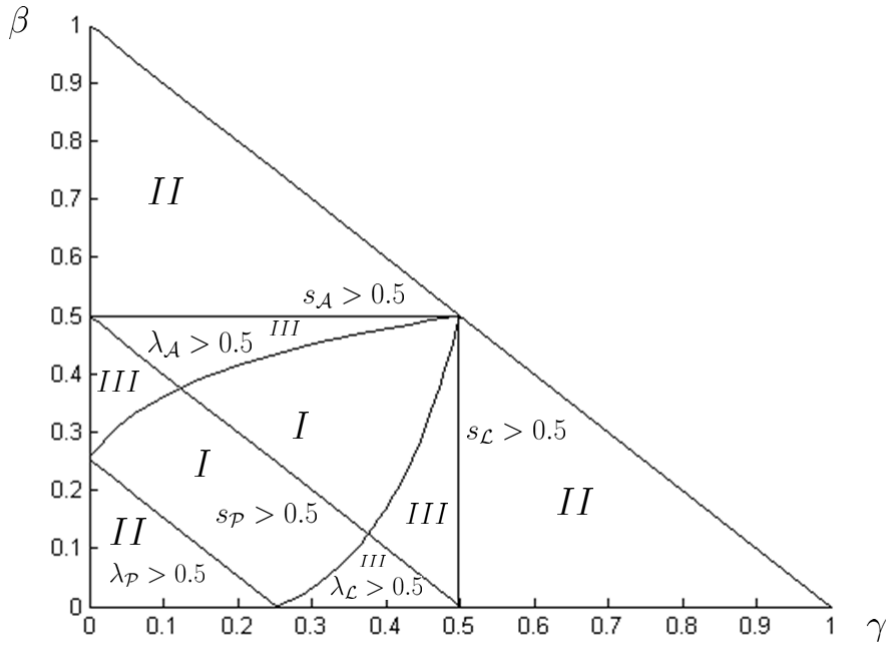

Figure 1.7: Political equilibria for $A_l = A_a = 1, \sigma = 0.33$.

Chapter 2

Inequality, Redistribution, and Weak Institutions

2.1 Introduction

One of the main policy issues in every society concerns the redistribution of income. In reality, the fraction of income that is redistributed by the policy maker varies significantly across countries. Standard politico-economic theory tries to explain this by means of the median voter theorem according to which higher pretax income inequality leads to higher redistribution. However, this logically appealing approach, first formulated by Meltzer and Richard (1981), fails in most cases when confronted with the data. This well-known phenomenon is called the 'paradox of redistribution'.

Various attempts have been made to resolve this paradox. While some studies have questioned the empirical results on conceptional grounds others have put forward alternative theoretical explanations. However, most of these seemingly different theories share a key feature with the median voter framework which is the (implicit) assumption that political conflicts about redistribution are solved within an institutionalized environment. This institutional environment is assumed to be exogenous and strong enough to commit individuals to some given political modus operandi; typically a democratic one. In this respect, most of these alternative models are essentially modified versions of the median voter framework that attach different political weights to individuals. Thus, despite its steady emphasis in empirical studies, see for example Persson and Tabellini (2003), Alesina and Glaeser (2004), Ramcharan (2010), or Karabarbounis (2011), the effect of institutions on the relation between inequality and redistribution has rarely been studied in the theoretical literature. But taking a strong institutional environment as given can be problematic for two reasons. First of all, in a large number of countries strong institutions do not exist or are very weak at least. For these kind of countries the existing literature offers almost no explanation on the relationship between income inequality and redistribution. And second, even if only countries with strong institutions are considered it is important to understand which share of total redistribution results from the mere existence of such institutions.

Therefore, we relax the assumption that an environment with strong institutions exists. It is the main novelty of our approach that we present a model of factual political power

where different groups within society cannot credibly commit themselves to some political constraints. Contrary to the median voter framework and all related modifications the *de facto* power of individuals in our model cannot be tamed by some *de jure* restrictions which are assumed exogenously. By doing this, we provide a theory that is independent of the particular political regime and can therefore be applied to non-democracies also. As it turns out, the relation between income inequality and redistribution in a weakly institutionalized environment becomes more complex than suggested by the previous literature. In our model, the relationship between inequality and redistribution is non-monotonous. A change in inequality might in- or decrease redistribution, or might have no effect at all, depending on the initial income distribution. Additionally, we find that there always exist minimal state regimes in which the equilibrium tax rate is zero. We also show that the Gini coefficient is not an appropriate measure when analyzing the relationship between inequality and redistribution in weakly institutionalized environments.

This paper contributes to the large literature on inequality and redistribution following the seminal work of Meltzer and Richard (1981) who first formulated the median voter theorem. By demonstrating the impact of institutional strength on the link between inequality and redistribution on purely theoretical grounds our model serves at least as a note of caution for all empirical studies which use cross-sectional data without controlling for this factor, see for example Perotti (1996), Bassett, Burkett, and Putterman (1999), Milanovic (2000), or Moene and Wallerstein (2001). At the same time it is also conducive to all other, in this respect, conceptually more elaborated studies – which consider within country variation only, see Husted and Kenny (1997), Gouveia and Masia (1998), Rodriguez (1999), Georgiadis and Manning (2007), Lind (2007), or Ramcharan (2010), control directly for measures of institutions, see Persson and Tabellini (2003), Alesina and Glaeser (2004), Iversen and Soskice (2006) or Shelton (2007), or account for unobserved country heterogeneity by estimating fixed effects, see Lindert (1996), Bellettini and Berti Ceroni (2007), or Karabarbounis (2011) – since our model delivers theoretical arguments for their appropriate but mostly intuitive approach.

Apart from this, our work contributes to the theoretical literature on the relation between inequality and redistribution. Numerous models have been developed to explain

deviations from the equilibrium tax rate in the median voter framework due to geographical mobility of the rich voters, Epple and Romer (1991), a two-dimensional policy space, Roemer (1998), differences in the social origin of voters, Piketty (1995), or in the weight of votes, Saint Paul and Verdier (1996), due to vote-buying by rich voters, Breyer and Ursprung (1998), the simultaneous consideration of wealth inequality, Krusell and Rios-Rull (1999), capital market imperfections and heterogeneous endowments, Benabou (2000) and Harms and Zink (2003), the prospect of social upward mobility, the so-called POUM hypothesis, Benabou and Ok (2001), or differences in culture or the historical experience, Corneo and Gruener (2002) and Luttmer and Singhal (2011). Others emphasize the role of lobbying activities of rich voters, Rodriguez (2004), Esteban and Ray (2006), and Campante and Ferreira (2007), differences in beliefs about the causes or justness of income inequality, Alesina and Angeletos (2005), and Benabou and Tirole (2006), the role of national identity that might overcome group thinking, Shayo (2009), of tax evasion, Traxler (2009), of altruism and the moral evaluation of the behavior of others, Cervellati, Esteban, and Kranich (2010), of uncertainty about the adoption of redistribution policies, Mattozzi (2010), or the corruption of politicians, Balafoutas (2011).[1] Our framework which is an extension of the model by Jung and Sunde (2011) differs from all of the above as it is not nested in an institutional environment that guarantees the possibility of binding commitments between different social groups of society. For this reason, our analysis of the relation between inequality and redistribution is independent of the particular political regime whereas all of the above models are applied within democracies only.

Though the latter property is also a common feature of most models on endogenous political regimes which are usually centered around the question of how income is redistributed within society and therefore related to our work, see for example Acemoglu and Robinson (2001, 2006), Lizzeri and Persico (2004), Llavador and Oxoby (2005), or Cervellati, Fortunato, and Sunde (2008), our framework allows for an analysis of more complex income patterns by considering the possibility of coalition formation explicitly. In this respect, our work is also related to the literature on game theoretic equilibrium concepts

[1] Given the large amount of literature on the relation between inequality and redistribution we do not claim our list of empirical and theoretical studies to be exhaustive. The most recent review of the pertinent literature is provided by Borck (2007).

of coalition formation; especially to the studies of Bernheim, Peleg, and Whinston (1987) and Acemoglu, Egorov, and Sonin (2008).

Lastly, by demonstrating that minimal state regimes with zero redistribution can emerge in equilibrium when weakly institutionalized environments are considered our findings also add to the literature on the income effects of inequality via redistribution, see for example Bertola (1993), Alesina and Rodrik (1994), or Persson and Tabellini (1994), and of institutions respectively, see Acemoglu, Johnson, and Robinson (2002), Glaeser et al. (2004), or Acemoglu and Robinson (2010) for example.

The paper is structured as follows. Section 2.2 lays out the model framework, and sections 2.3 and 2.4 present the results concerning the political and the politico-economic equilibrium. In section 2.5 we analyze the relationship between inequality and redistribution in equilibrium. And in Section 2.6 we work out the empirical implications of our model. Section 2.7 concludes.

2.2 Model

2.2.1 Heterogeneous Population, Production, and Income

The basic structure of our model follows Jung and Sunde (2011). Accordingly, we consider a static model society that is populated by a unit mass of individuals. These individuals live for one period and leave no bequests. By birth, all individuals are endowed with labor time, and some of them additionally possess physical strength and/or intellectual ability. While each individual has an identical endowment of labor time, $h > 0$, at his disposal, physical strength and intellectual ability are distributed unevenly in the population.[2] For simplicity, we assume that the distribution of both of these characteristics is dichotomic. This means, a share $0 < \gamma < 1$ of individuals possesses physical strength, denoted by $l = 1$, whereas the complement $1 - \gamma$ possesses no physical strength, $l = 0$. Likewise, a share $0 < \beta < 1$ of the population is endowed with intellectual ability, $a = 1$, while a share $1 - \beta$ possesses no intellectual ability at all, $a = 0$. Contrary to Jung and Sunde (2011), we

[2]In the following, the endowment of labor time h will be normalized to one without loss of generality.

do not assume physical strength and intellectual ability to be mutually exclusive traits. Thus, the population in our model effectively consists of four distinct groups: the twofold-privileged, strong and intelligent elite, denoted by $\mathcal{E}$, the able weaklings, $\mathcal{A}$, the simple-minded strong, $\mathcal{L}$, and those that possess neither strength nor ability, $\mathcal{P}$.[3] Denote the set of groups by $S = \{\mathcal{P}, \mathcal{L}, \mathcal{A}, \mathcal{E}\}$, and the size of group $i \in S$ as s_i.[4] Figure 2.1 below shows an exemplary population structure and the respective group sizes for some $0 < \gamma^*, \beta^* < 1$ which is represented by point $D(\gamma^*, \beta^*)$ in the admissible γ-β space.[5]

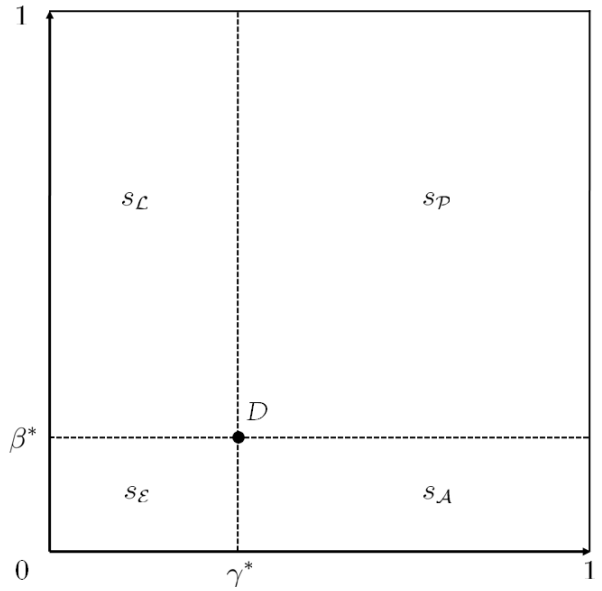

Figure 2.1: Population structure and respective group sizes.

Regarding point D it is obvious that any increase in at least one of the two shares of exclusive production factors, $d\gamma^* \wedge d\beta^* \geq 0$, increases the relative size of the $\mathcal{E}$-group, $ds_{\mathcal{E}} \geq 0$. Hence all points lying in the north and/or east of point D represent larger relative sizes of group $\mathcal{E}$ whereas points in the south and/or west denote smaller relative

[3]The model of Jung and Sunde (2011) considers only three distinct groups and is therefore not appropriate for studying the relation between inequality and redistribution as will become clear below. In general, our model society could comprise up to n groups. However, we restrict our analysis to four groups for the sake of simplicity.

[4]For simplicity, we assume that no commitment problems exist within groups such that single group members do not free ride on the expense of others. Thus, we can analyze the society as consisting of four different agents, each representing one income group, i.e., individual group members and groups can interchangeably be denoted by i. This is contrary to Olson (1965) but may be justified in the case of transitory inter-group conflicts where collective action within groups is much easier to sustain, see Acemoglu and Robinson (2006).

[5]All points lying on one of both axes or on the vertical or horizontal line, $\gamma = 1$ or $\beta = 1$, are not considered in the following since they represent cases where the number of income groups within society is smaller than four.

sizes. When considering γ-β combinations lying in the northwest or the southeast of point D we know that the overall change of the relative size of group $\mathcal{E}$ is given by $ds_{\mathcal{E}} = \gamma^* d\beta^* + \beta^* d\gamma^*$. From this we can see that all points representing larger relative sizes of group $\mathcal{E}$ must lie above the line passing through D with slope $-\beta^*/\gamma^*$. For any point below this line the relative size of the $\mathcal{E}$-group is smaller than in point D whereas for all points lying on this line it is the same. Of course, analogous reasoning applies for all other groups which implies that changes in γ^* and β^* that leave the size of the $\mathcal{E}$-group unaffected do alter the relative size of all others – except for $\gamma^* = \beta^* = 0.5$. In this special case, $-\beta^*/\gamma^*$ equals $-(1-\beta^*)/(1-\gamma^*)$ and movements along the line with slope $-\beta^*/\gamma^*$ have no impact on the relative size of two groups, the $\mathcal{E}$- and the $\mathcal{P}$-group.[6]

Apart from that, our assumptions concerning the factor endowments of an individual belonging to group i can be summarized by

$$l_i = \begin{cases} 0 & \text{if } i \in \{\mathcal{P}, \mathcal{A}\} \\ 1 & \text{if } i \in \{\mathcal{L}, \mathcal{E}\} \end{cases} \quad \text{and} \quad a_i = \begin{cases} 0 & \text{if } i \in \{\mathcal{P}, \mathcal{L}\} \\ 1 & \text{if } i \in \{\mathcal{A}, \mathcal{E}\} \; . \end{cases} \tag{2.1}$$

Let us further assume that all individuals supply their endowments inelastically on competitive markets and that labor, strength and ability are employed as separate inputs in the production sector. Aggregate income Y is generated by means of a CRS production function

$$Y \; = \; Y(H, L, A) \tag{2.2}$$

where H, L and A represent the aggregate levels of working hours, physical strength and ability, respectively. The marginal product of every input factor is positive but decreasing. Factor prices are competitive such that the price paid for the provision of one unit of every input factor equals its marginal product. The respective factor prices are denoted by $\rho = \partial Y/\partial H$, $w = \partial Y/\partial L$ and $\mu = \partial Y/\partial A$. Hence, any individual belonging to

[6] Similarly, movements along the line with slope $(1-\beta^*)/\gamma^*$ passing through $\gamma^* = \beta^* = 0.5$ leave the relative sizes of the two other groups, $\mathcal{L}$ and $\mathcal{A}$, unchanged.

group i earns the factor income

$$y_i \;=\; \rho + wl_i + \mu a_i \qquad \forall\; i \in S\,. \tag{2.3}$$

The unequal endowment of input factors and their remuneration on competitive markets implies an unequal income distribution within society where individuals with richer endowments earn higher factor incomes. From the assumptions about population structure, it follows directly that members of the $\mathcal{E}$-group always earn the highest factor income per capita, whereas the factor income of individuals in the $\mathcal{P}$-group is always the lowest, i.e., $y_{\mathcal{P}} < y_{\mathcal{L}}, y_{\mathcal{A}} < y_{\mathcal{E}}$. In the given setting, per-capita income equals aggregate group income, thus

$$y \;=\; \sum_{i \in S} s_i y_i \;=\; \rho + w\gamma + \mu\beta \tag{2.4}$$

always holds.

2.2.2 Redistributive Conflict, Power, and Utility

We assume individual utility to be not affected by the well-being of others. This gives rise to redistributive conflicts. In the model, a latent conflict between different groups of society exists and every group tries to maximize its own income at the expense of others. The ability to do this depends on the political power of a particular group. As in Jung and Sunde (2011), this political power P_i of group i is determined by its aggregate income such that $P_i \equiv s_i y_i$.[7] Additionally, we assume this power mapping to be bijective. Hence no two groups or coalitions can be equal in power, $P_Q \;\neq\; P_M \;\forall\; Q, M \in \mathscr{P}(S) \setminus \{\varnothing\}$ for $Q \neq M$, and in every period the ranking of power of all social groups is well-defined. Given the fact that a realistic description of a conflict game is not the purpose of this paper, we model the redistributive conflict as parsimonious as possible and simply assume that any group or coalition $Q \in \mathscr{P}(S) \setminus \{\varnothing\}$ can seize the income of group or coalition

[7] In this respect, our model is related to the *one dollar, one vote* principle and stands in the tradition of earlier studies like Becker (1983), Grossman and Helpman (1996), or Benabou (2000) who also emphasize the positive effect of income on political influence, opposite to the *one man, one vote* approach underlying the median voter theorem. However, in contrast to these models ours is not restricted to a democratic policy space only.

$S\backslash Q$ if $P_Q > P_{S\backslash Q}$ holds where $P_Q \equiv \sum_{j\in Q} P_j$ denotes the aggregate power of group or coalition Q. Thus, the most powerful group whose power we denote as P_{MAX} is able to make all political decisions autonomously if, and only if $2\,P_{MAX} > \sum_{i\in S} P_i$ holds.

If no group has the power to rule alone, i.e., $2\,P_{MAX} \leq \sum_{i\in S} P_i$, forming a coalition becomes a relevant option.[8] However, since we consider an environment where no institutions exist that would allow for binding commitments between groups and every group tries to maximize its own disposable income only, no group can rely on promises of others. Before we can turn to a complete characterization of political equilibria in such a situation we need to specify a last aspect of the political environment. It concerns the question how revenues from redistribution are shared within the ruling coalition. As in Jung and Sunde (2011), we adopt the approach of Acemoglu, Egorov, and Sonin (2008) and assume that the share of transferable income seized by group $i \in S$ is determined by its effective relative power $\widetilde{p}_i$ within the coalition that redistributes income. It is defined by

$$\widetilde{p}_i = \begin{cases} \frac{P_i}{P_{RC}} & \text{if } i \in RC \\ 0 & \text{otherwise} \end{cases} \tag{2.5}$$

where $P_{RC} \equiv \sum_{j\in RC} P_j$ denotes the aggregate power of the coalition that ultimately redistributes income which we from now on call the ruling coalition $RC \subseteq S$.[9] Note that effective relative power is always defined conditional on the respective RC. In principle, we assume that the aggregate income of every group can be completely expropriated and redistributed among others such that the feasible transfer equals per-capita income.[10]

[8] In general, forming a coalition means the pooling of power resources at the cost of lower redistributed income per capita.

[9] As Acemoglu, Egorov, and Sonin (2008) or Jung and Sunde (2011), we abstract from bargaining processes within ruling coalitions and impose this sharing rule exogenously. This rule was first used by Gamson (1961) to characterize the sharing of resources amongst coalition members and seems to be a fairly good description as several empirical studies suggest, see, e.g., Warwick and Druckman (2001) or Ansolabehere et al. (2005). It might be worthwhile to analyze the implications of other empirically less relevant sharing rules in the context of our model.

[10] We could also assume that some subsistence income is obtained by each individual to ensure that production takes place without qualitatively affecting our main results; or alternatively, as in Acemoglu and Robinson (2000), consider an untaxable and technologically inferior informal sector where the final good is produced with using labor time only.

Disposable income $\widetilde{y}_i$ of group i is then described by

$$\widetilde{y}_i = \widetilde{p}_i y \,. \tag{2.6}$$

Finally, we can write out the indirect utility function of members of group $i \in S$ which reads in its general form

$$u_i = u\left(\widetilde{y}_i(\widetilde{p}_i)\right) \tag{2.7}$$

with $\partial u_i / \partial \widetilde{y}_i > 0$. To be precise we would need to divide disposable income by the respective group size additionally. But with regards to the individual optimization problem we can neglect this constant denominator since group sizes cannot be changed by individuals. The same holds true for factor endowments and factor incomes. Thus, the optimization problem amounts to maximizing $\widetilde{p}_i$ in order to maximize lifetime utility subject to the constraints imposed by the production structure and the political environment:

$$\max_{\widetilde{p}_i} u_i\left(\widetilde{y}_i(\widetilde{p}_i)\right) \quad \text{subject to (2.3), (2.4), (2.5) and (2.6).} \tag{2.8}$$

This implies that every group always prefers the coalition in which its relative power is greatest. This must not be confused with a positive effect of P_i on u_i, as the latter does not monotonically increase in the former, which will become clear below.

2.2.3 Timing of Events

To complete our model framework we need to elaborate on the non-cooperative ruling coalition formation and redistribution game that is played by every generation. We assume that every generation experiences the following sequence of events throughout its lifetime

1. Birth, realization of endowments and factor incomes.
2. Ruling coalition formation and redistribution game Γ:
   - 2.1 An agenda setter is randomly determined from all (remaining) groups.
   - 2.2 The agenda setter proposes a subcoalition that includes himself to all (remaining) groups.

2.3 The members of this subcoalition vote sequentially in random order over the proposal (and all non-members automatically vote against the proposal); if all groups that support the proposal form a winning coalition, the game proceeds to step 2.4, otherwise to step 2.5.

2.4 If the proposal includes all groups of the current stage k of the game, then they all form the *RC* and the game proceeds to step 3. If the proposal consists of a proper subset, then all groups that are not part of this proposal are excluded by redistribution of their factor incomes to the members of the subset which causes some (arbitrarily small) costs ϵ;[11] in this case, a new stage $k+1$ begins with step 2.1.

2.5 A new agenda setter is determined randomly among all (remaining) groups that have not yet acted as agenda setter at the current stage of the game k, and the game proceeds to step 2.2; if all (remaining) groups have been agenda setters at the current stage k, then they all form the *RC* and the game proceeds to step 3.

3. Consumption of disposable income and death.

2.3 Political Equilibrium

Regarding the political equilibrium of our model we state the following Lemma.

Lemma 2.1. *In game* Γ *there exist subgame perfect Nash equilibria (SPNEs) in pure strategies which all imply the same redistribution scheme.*

Proof. See Appendix □

Intuitively, it is easy to understand how the *RC* that determines the redistribution scheme looks like. First, by the nature of the game it must be winning such that it is

[11] Different to Jung and Sunde (2011) we have to assume some arbitrary small costs of redistribution to exclude path dependency and thereby ensure uniqueness of the equilibrium. In principle, this would call for a reduction of individual utility by a per-capita share of total costs incurred. But since in equilibrium these redistribution costs incur only once at most and are assumed to be arbitrarily small, we neglect this reduction of individual utility for the sake of simplicity.

powerful enough to outgun any alternative coalition at the current stage of the game k. Second, it must be stable in the sense that it contains no proper subcoalition which will be winning at a subsequent stage of the game $\hat{k} > k$.[12] And third, if there exist more than one coalition which satisfy both properties the RC will be the coalition with least aggregate power, since the optimization problem $\max \widetilde{p}_i$ is solved by minimizing the denominator in condition (2.5).[13]

We can now characterize the RC in terms of its structure. Let us define $\lambda_i \equiv P_i / \sum_{i \in S} P_i$ as the relative power of group i, then we already know from our former reasoning that the RC will consist of one group if and only if $\lambda_{MAX} > 0.5$ holds where λ_{MAX} represents the relative power of the most powerful group within society.[14] When λ_{MAX} exactly equals one half it is easy to see that the grand coalition $RC = S$ constitutes the equilibrium outcome, since no coalition excluding the most powerful group can be winning and no coalition including it can be stable in this case. Additionally, we can state the following Lemma that summarizes which groups are not part of the RC for a given factor distribution when $\lambda_{MAX} < 0.5$ holds.

Lemma 2.2. *Given* $\lambda_{MAX} < 0.5$ *then ...*

1. *...* $\lambda_{MIN} \geq \overline{\lambda} - \underline{\lambda} \iff i_{MAX} \notin RC.$

2. *...* $\lambda_{MIN} < \overline{\lambda} - \underline{\lambda} \wedge \lambda_{MIN} \geq \lambda_{MAX} - \overline{\lambda} \iff \underline{i} \notin RC.$

3. *...* $\lambda_{MIN} < \overline{\lambda} - \underline{\lambda} \wedge \lambda_{MIN} < \lambda_{MAX} - \overline{\lambda} \wedge \underline{\lambda} \geq \lambda_{MAX} - \overline{\lambda} \iff i_{MIN} \notin RC.$

4. *...* $\lambda_{MIN} < \overline{\lambda} - \underline{\lambda} \wedge \underline{\lambda} < \lambda_{MAX} - \overline{\lambda} \iff RC = S.$

Proof. See Appendix.[15] □

With regard to Lemma 2.2 three notions about the political equilibrium in our model seem to be noteworthy. First, $\lambda_{MAX} \leq 0.5 \Leftrightarrow |RC| \geq 3$ holds where $|RC|$ denotes the

[12] This second equilibrium property resembles the notion of a Perfectly Coalition-Proof Nash Equilibrium which was first introduced by Bernheim, Peleg, and Whinston (1987). Acemoglu, Egorov, and Sonin (2008) apply it to a class of political games.

[13] This reasoning corresponds to set Ω in the proof of Lemma 2.1 which gives a formal definition of the RC.

[14] The denotation of relative powers is used analogously to that of absolute powers, see section 2.2.2.

[15] Note that $\lambda_{MAX} < 0.5$ is redundant for conditions 2 and 3 of Lemma 2.2 since the former condition is always satisfied if one of the latter holds. However, we chose this explicit formulation as it might be easier to understand intuitively. For the same reason, we refrain from a further simplification of the given inequalities which will be done below.

cardinality of set RC. In the given setting, a coalition of two groups never constitutes an equilibrium outcome, since the more powerful group can always propose a coalition that only contains itself at the subsequent stage of the game. Second, $\lambda_{MAX} \leq 0.5 \Rightarrow \bar{i} \in RC$ is true, i.e., the second most powerful group $\bar{i}$ is always part of the RC and hence receives a positive fraction of redistributed income in any case given $\lambda_{MAX} \leq 0.5$. And third, though an increase in income always increases the power of a group this does not necessarily translate into an increase in its effective political power. As Lemma 2.2.1 shows, under certain circumstances the most powerful group can be very well excluded from the RC by a coalition of the poorer groups. This result distinguishes our model from most others and explains our former notion that the utility of a group does not monotonically increase in its political power, see page 56.

2.4 Politico-Economic Equilibrium

Before we can turn to the relationship between inequality and redistribution we must complement the equilibrium analysis by considering the economic equilibrium in our model. We specify equation (2.2) in the following way

$$Y = A^{\sigma} L^{\theta} H^{1-\sigma-\theta} \tag{2.9}$$

with $\sigma, \theta > 0 \wedge \sigma + \theta < 1$ and assume perfectly competitive markets such that the reward for every production factor equals its marginal product.[16] Given expressions (2.3) and (2.9), and the normalized individual time endowment factor income of a member of group i is described by

$$y_i \;=\; A^{\sigma} L^{\theta} H^{1-\sigma-\theta} \left[\frac{(1-\sigma-\theta)}{H} + \frac{\sigma a_i}{A} + \frac{\theta l_i}{L} \right] \quad \text{with } i \in S \; . \tag{2.10}$$

Using the information contained in Figure 2.1 and in equations (2.1) and (2.4), we can

[16] Since our focus here is on income inequality only we take the current technology level as given. For a version of the model that takes technology levels into account see Jung and Sunde (2011).

write out the relative power for every group of society which reads

$$\lambda_{\mathcal{P}} = (1-\beta)(1-\gamma)(1-\sigma-\theta) \tag{2.11}$$

for the $\mathcal{P}$-group,

$$\lambda_{\mathcal{L}} = \gamma(1-\beta)(1-\sigma-\theta) + (1-\beta)\theta \tag{2.12}$$

for the $\mathcal{L}$-group,

$$\lambda_{\mathcal{A}} = \beta(1-\gamma)(1-\sigma-\theta) + (1-\gamma)\sigma \tag{2.13}$$

for the $\mathcal{A}$-group, and

$$\lambda_{\mathcal{E}} = \gamma\beta(1-\sigma-\theta) + \gamma\sigma + \beta\theta \tag{2.14}$$

for the $\mathcal{E}$-group, respectively. In our framework the relative power of group $i \in S$ by construction increases in the size of a group and/or in the income share devoted to this group which can easily be seen from equations (2.11) - (2.14). The total effect accounting for possibly opposing changes in γ and β is a bit more complex though. But similar to our former reasoning concerning relative group sizes we can easily determine the group specific iso-power curve along which the relative power of a group is constant by setting the respective total derivative to zero. For example, given a particular γ-β combination (small) changes in the endowment shares leave the relative power of the $\mathcal{E}$-group unaffected if and only if

$$\frac{d\beta}{d\gamma} = -\frac{\beta(1-\sigma-\theta)+\sigma}{\gamma(1-\sigma-\theta)+\theta} \tag{2.15}$$

holds, i.e., if the ratio of (small) changes in the endowment shares equals the slope of the iso-power curve.[17] Whenever $d\beta/d\gamma$ deviates from this value the relative power of group $\mathcal{E}$ changes. Like before, let us consider some arbitrary point $D(\gamma^*, \beta^*)$ in the admissible γ-β space for any given $0 < \sigma, \theta < 1$. Then all points lying below (above) the line running through D with a slope value according to equation (2.15) represent γ-β combinations for which the relative power of the $\mathcal{E}$-group is smaller (larger) than in point D. Note that there exist infinitely many iso-power curves for each group which cannot intersect each

[17] Suppose $d\sigma = d\theta = 0$ for a moment.

other and that relative power becomes maximal when the size of a group approaches one. Thus, given two different iso-power curves of the $\mathcal{E}$-group for example we know that the one lying closer to the northeastern corner of the admissible γ-β space, where $\gamma = \beta = 1$, represents a higher level of relative power of group $\mathcal{E}$. And since analogous reasoning applies for all other groups we know exactly for any given value of γ and β how changes in one or both variables affect the power distribution within our model society.

Before we turn to our main point of interest which is the share of income redistributed in equilibrium let us now focus on a special set of iso-power curves. We already know from our former reasoning that for the composition of the RC it is crucial whether there exists one group which is more powerful than all others together or not, i.e., whether $\lambda_{MAX} > 0.5$ holds or not. It hence seems worthwhile to scrutinize the four respective iso-power curves a little closer. Therefore we set equations (2.11) - (2.14) equal to one half and solve for β which gives

$$\beta_{\lambda_{\mathcal{P}}=0.5} = 1 - \frac{1}{2(1-\gamma)(1-\sigma-\theta)} \tag{2.16}$$

for the $\mathcal{P}$-group,

$$\beta_{\lambda_{\mathcal{L}}=0.5} = 1 - \frac{1}{2[\gamma(1-\sigma-\theta)+\theta]} \tag{2.17}$$

for the $\mathcal{L}$-group,

$$\beta_{\lambda_{\mathcal{A}}=0.5} = \frac{0.5-(1-\gamma)\sigma}{(1-\gamma)(1-\sigma-\theta)} \tag{2.18}$$

for the $\mathcal{A}$-group, and

$$\beta_{\lambda_{\mathcal{E}}=0.5} = \frac{0.5-\gamma\sigma}{\gamma(1-\sigma-\theta)+\theta} \quad . \tag{2.19}$$

for the $\mathcal{E}$-group, respectively. Equations (2.16) - (2.19) allow us to identify all areas in the admissible γ-β space where the most powerful group is powerful enough to rule alone. For a graphical analysis we need to assume certain parameter values for σ and θ. Figure 2.2 below shows the four group-specific iso-power curves for $\sigma = \theta = 0.16$, i.e., for a situation in which the share of income devoted to labor time $(1-\sigma-\theta)$ equals roughly two-thirds.

First, note that all iso-power curves have the expected slope. According to equation (2.15) the $\lambda_{\mathcal{E}} = 0.5$ locus is a convex function in the admissible γ-β space since any increase in γ must be accompanied by a decrease in β to keep the relative power level

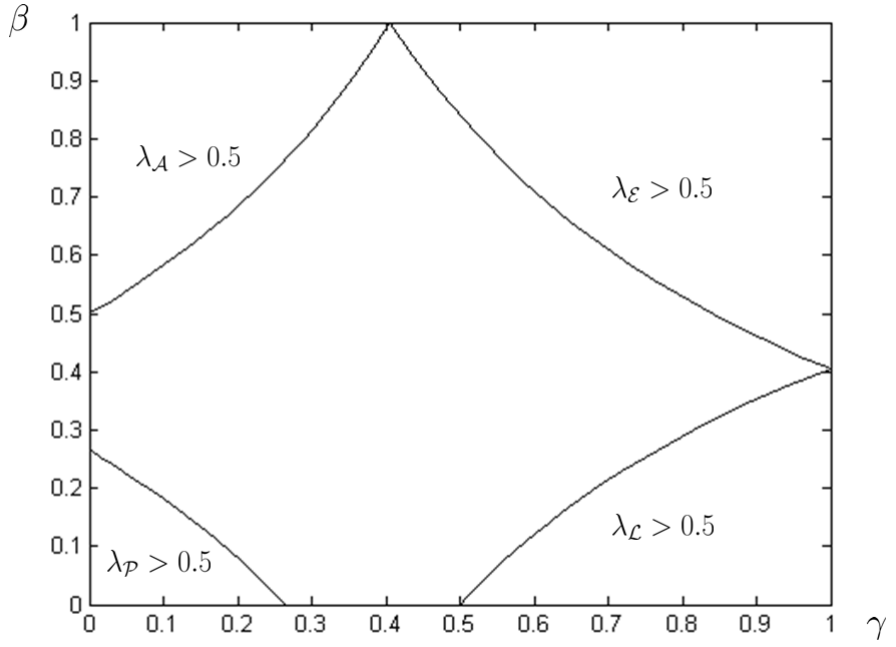

Figure 2.2: Factor distributions and relative group powers for $\sigma = \theta = 0.16$.

constant whereas the $\lambda_{\mathcal{L}} = 0.5$ locus for example is supposed to be concave as in case of group $\mathcal{L}$ a higher β must compensate for higher values of γ. For the given parametrization depicted in Figure 2.2 all four iso-power curves pass through the admissible γ-β space.

However, this is not always the case. Regarding the $\mathcal{P}$-group we can infer from equation (2.11) that its relative power can only exceed one half if the income share devoted to labor time is sufficiently large, i.e., $(1-\sigma-\theta) > 0.5$ is a necessary condition for $\lambda_{\mathcal{P}} > 0.5$ to hold. Otherwise the $\lambda_{\mathcal{P}} = 0.5$ locus in the south-western corner of Figure 2.2 vanishes. Given that $(1-\sigma-\theta) > 0.5$ any additional increase of $(1-\sigma-\theta)$ ceteris paribus shifts the $\lambda_{\mathcal{P}} = 0.5$ locus further north-east in the admissible γ-β space which corresponds to $\partial\beta_{\lambda_{\mathcal{P}}=0.5}/\partial(1-\sigma-\theta) > 0$, see equation (2.16). As $(1-\sigma-\theta)$ approaches one, implying $\sigma, \theta \approx 0$, a completely symmetric picture emerges with the bisectrix, the horizontal line $\beta = 0.5$ and the vertical line $\gamma = 0.5$ being axes of reflection. In this extreme case, every group generates an (almost) identical share of income and (almost) any difference in relative power stems from size effects. In such a society, pretax incomes per capita are distributed (almost) equally among differently endowed individuals.

In the opposite extreme, as $(1-\sigma-\theta)$ approaches zero the run of the power loci

depends on the relation between σ and θ. In such a situation, either the $\lambda_{\mathcal{L}} = 0.5$ or the $\lambda_{\mathcal{A}} = 0.5$ locus might disappear from our picture if the relation between both variables is highly unbalanced. For example, if σ approaches one which implies θ to approach zero due to $\sigma + \theta < 1$, then there exists no γ-β combination anymore for which the $\mathcal{L}$-group can dominate all other groups, since $\lim_{\sigma \to 1} \lambda_{\mathcal{L}} = 0$ follows from equation (2.12). Analogous reasoning applies for the $\mathcal{A}$-group. Only concerning the relative power locus of the $\mathcal{E}$-group we can make a general statement.

Lemma 2.3. *For any admissible combination of σ and θ there always exist γ-β combinations such that the relative power of the $\mathcal{E}$-group is larger than one half, $\lambda_{\mathcal{E}} > 0.5$.*

Proof. Note that $\beta_{\lambda_{\mathcal{E}}=0.5}(\gamma = 1) = (0.5 - \sigma)/(1 - \sigma)$ and $\gamma(\beta_{\lambda_{\mathcal{E}}=0.5} = 1) = (0.5 - \theta)/(1 - \theta)$. Thus, due to $\sigma + \theta < 1$, there always exists (at least) one point of intersection of the $\lambda_{\mathcal{E}} = 0.5$ locus with the right (and) or the upper bound of the admissible γ-β space for any $0 < \sigma, \theta < 1$. And since the $\lambda_{\mathcal{E}} = 0.5$ locus is a continuous function in the admissible γ-β space we can conclude that the above assertion is true. □

According to Lemma 2.3 the $\lambda_{\mathcal{E}} = 0.5$ locus always passes through the admissible γ-β space. This is easy to understand intuitively. In our model, relative group powers are affected by group size and the relative importance of the exclusive production factors in the income generating process. Given that members of group $\mathcal{E}$ possess both exclusive production factors and therefore earn the highest (pretax) per-capita income, increases in size are always sufficient for making them more powerful than all other groups.

2.5 Equilibrium Redistribution

Finally turning to the analysis of redistribution in equilibrium we first need to define a measure for it. As such we will use the share of income that is redistributed in equilibrium

$$t = \frac{1}{y} \sum_{i \in S \setminus RC} s_i y_i = \sum_{i \in S \setminus RC} \lambda_i = 1 - \sum_{i \in RC} \lambda_i \tag{2.20}$$

which we denote by t since it represents the average tax rate in the economy.[18] It is straightforward to calculate this average tax rate for all situations where the most powerful group rules the state on its own. Regarding again Figure 2.2, it is clear that for all γ-β combinations south-east of the $\lambda_{\mathcal{L}} = 0.5$ locus for example the ruling coalition consists of the $\mathcal{L}$-group only and hence the tax rate is given by $t = 1 - \lambda_{\mathcal{L}}$. Analogous reasoning applies to the regions north-east of the $\lambda_{\mathcal{E}} = 0.5$ locus, north-west of the $\lambda_{\mathcal{A}} = 0.5$ locus and south-west of the $\lambda_{\mathcal{P}} = 0.5$ locus. We also know that, by construction of our model, for all γ-β combinations lying directly on one of the iso-power curves where the relative power of a group equals one half the *RC* comprises all groups of society as no other coalition can be stable and winning in this situation. In such knife-edge cases no income will be redistributed in equilibrium and the corresponding tax rate is zero. Let us summarize these properties in the following proposition.

Proposition 2.1. *Given* $\lambda_{MAX} > 0.5 \implies t = 1 - \lambda_{MAX}$. *For* $\lambda_{MAX} = 0.5 \implies t = 0$ *holds.*

Regarding Proposition 2.1 two aspects seem to be noteworthy. First, for all cases where $\lambda_{MAX} > 0.5$ holds, any change in γ and/or β that further increases the relative power of the most powerful group always decreases the equilibrium average tax rate which equals the aggregate relative power of the other groups, i.e., $dt|_{\lambda_{MAX}>0.5} = -d\lambda_{MAX}$. Second, since our former reasoning led us to conclude that the $\lambda_{\mathcal{E}} = 0.5$ locus always passes through the γ-β space for any admissible combination of σ and θ, see Lemma 2.3, we know that the knife-edge case $\lambda_{MAX} = 0.5$ with a minimal tax rate $t = 0$ always exists. However, when $\lambda_{MAX} < 0.5$ holds instead, the equilibrium analysis becomes more complex. In this case, the following proposition which follows directly from Lemma 2.2 will prove to be useful.

Proposition 2.2. *Given* $\lambda_{MAX} < 0.5$ *then ...*

1. *...* $\lambda_{MIN} \geq \overline{\lambda} - \underline{\lambda} \iff t = \lambda_{MAX}$.

2. *...* $\lambda_{MAX} - \overline{\lambda} \leq \lambda_{MIN} < \overline{\lambda} - \underline{\lambda} \iff t = \underline{\lambda}$.

3. *...* $\lambda_{MIN} < \lambda_{MAX} - \overline{\lambda} \leq \underline{\lambda} < \overline{\lambda} - \lambda_{MIN} \iff t = \lambda_{MIN}$.

4. *...* $\lambda_{MIN} + \underline{\lambda} < \overline{\lambda} < \lambda_{MAX} - \underline{\lambda} \iff t = 0$.

[18] Again, we would need to take into account redistribution costs here but as they are assumed to be infinitely small and the ruling coalition forms after (at most) one round of redistribution, we neglect them in the following for the sake of simplicity.

On the basis of Proposition 2.2 we can now calculate the equilibrium average tax rate for every admissible γ-β combination. Figure 2.3 below shows the results of this exercise for the same parametrization we used before. In this sense, Figure 2.2 represents the corresponding ground plot to Figure 2.3. The most obvious conclusion that can be drawn from Figure 2.3 is that in a weakly institutionalized environment the relation between inequality and redistribution becomes quite complex. But there also exist some general patterns which need to be pointed out.

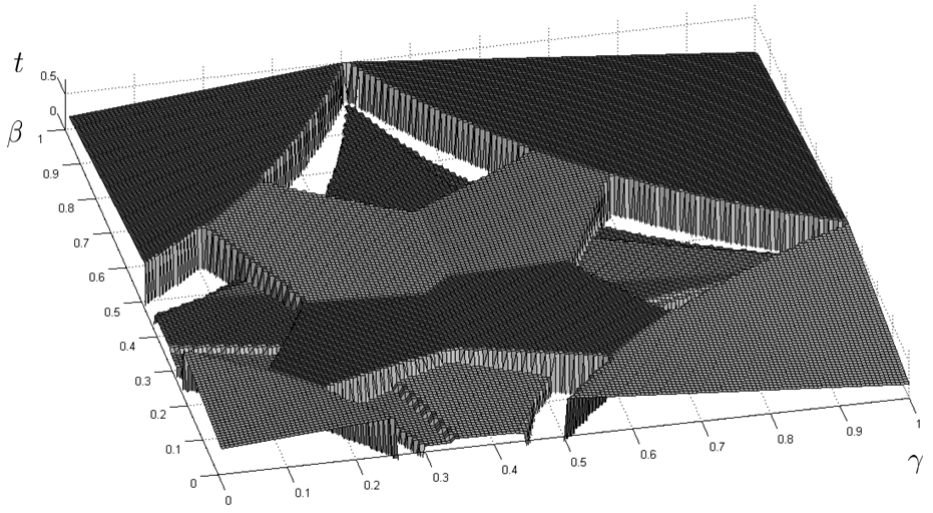

Figure 2.3: Equilibrium average tax rate t for $\sigma = \theta = 0.16$.

Since any group or coalition that redistributes income from others has to have more power than them, the equilibrium tax rate in our model is by construction always (strictly) smaller than fifty percent of income. From this it also seems to be clear that it must be maximal for γ-β combinations that lie close to the $\lambda_{MAX} = 0.5$ loci. However, this notion requires some further specification. It is always true for γ-β combinations for which the relative power of the most powerful group approaches one half from above, since $\lim_{\lambda_{MAX} \to 0.5^+} t = 0.5$. But concerning the left-hand side of the limit, we can already see from Figure 2.3 that this is not always the case. Obviously, there also exist γ-β combinations close to the $\lambda_{MAX} = 0.5$ loci for which $\lambda_{MAX} < 0.5$ holds and the equilibrium tax

rate is zero. These γ-β combinations are represented by the white areas in Figure 2.3. Hence the tax rate does not necessarily become maximal if the relative power of the most powerful group approaches one half from below. To conclude our analysis of equilibrium redistribution, let us state a last proposition concerning the average tax rate which is less intuitive.

Proposition 2.3. *Given $\lambda_{MAX} < 0.5$ then for any admissible combination of σ and θ there always exist γ-β combinations such that the equilibrium tax rate is zero, $t = 0$.*

Proof. See Appendix □

Proposition 2.3 represents one of the central results of our study and therefore deserves some further elaboration. To get a better understanding of the minimal state regimes with an equilibrium tax rate of zero in our model let us look at Figure 2.3 from a bird's eye perspective. This is done in the left panel of Figure 2.4 below. When comparing the white areas in the left panel with the illustration in the right panel of Figure 2.4, which is a small-scale version of Figure 2.2, we can see that the degrees of inequality for which the equilibrium tax rate is zero are described by γ-β combinations that lie close to the $\lambda_{MAX} = 0.5$ loci. With regard to the proof of Proposition 2.2 we can conjecture that this might be a rather general pattern which is independent of the particular parametrization of σ and θ.[19]

Apart from this formal aspect, the finding of a zero equilibrium tax rate within a society where heterogeneous agents compete for political power to enforce their preferred redistribution scheme is in stark contrast to the median voter framework. In Meltzer and Richard (1981), redistribution in equilibrium can only be zero in the trivial case where society is homogeneous and consists of individuals earning identical incomes. This also is true for numerous other models, see for example Krusell and Rios-Rull (1999), Benabou (2000), and Saint Paul and Verdier (1996). However, the finding of an equilibrium tax rate of zero is not novel. In the models of Roemer (1998) or Balafoutas (2011), for example,

[19]The proof of Proposition 2.3 focuses on the $\lambda_{\mathcal{E}} = 0.5$ locus only since it always passes through the admissible γ-β space. It shows that there always exist equilibrium tax rates of zero close to the $\lambda_{\mathcal{E}} = 0.5$ locus. However, the underlying reasoning can be applied analogously to all other loci as well – if they pass through the admissible γ-β space.

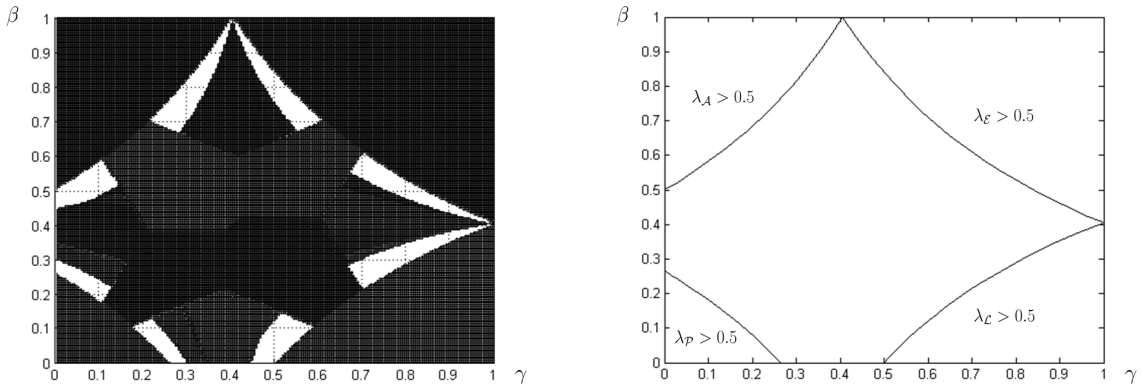

Figure 2.4: Top view of equilibrium average tax rates for $\sigma = \theta = 0.16$ (left) and corresponding iso-power curves (right).

it also represents a possible equilibrium outcome even if individuals earn different pretax incomes. What is special though, is that our result emerges in a model framework with weak institutions. Thus, if redistribution is harmful for the development of income as it is claimed by parts of the pertinent literature, see Alesina and Rodrik (1994) or Persson and Tabellini (1994) for example, then our model implies that for certain degrees of income inequality the weakening of institutions could have a positive effect on income by generating a minimal state regime. Note that this rather counterintuitive finding is contrary to the prevailing view in the literature which posits a positive causal effect of strong institutions on income levels, see Acemoglu (2009) or Acemoglu and Robinson (2010).[20] In closing, let us now focus on the empirical implications of our model.

2.6 Empirical Implications

So far, we talked about income inequality in a rather abstract sense by referring to the endowment shares γ and β. A common measure which is often used in the literature to estimate the effect of income inequality on redistribution is the Gini coefficient, see for example Milanovic (2000), Persson and Tabellini (2003), Iversen and Soskice (2006), Shelton (2007), or Ramcharan (2010). For this reason, it is worthwhile to work out the implications of our model for the relation of both variables. To calculate Gini coefficients, we employ a method that is commonly used in the literature when heterogenous income

[20] This opinion is not unchallenged of course, see for example Glaeser et al. (2004).

groups within society are considered, see Esteban and Ray (1994) or Montalvo and Reynal-Querol (2005) for example. Accordingly, the Gini coefficient in our model reads

$$G = \sum_{i\in S}\sum_{j\in S} s_i s_j \left|y_i - y_j\right| = y\sum_{i\in S}\sum_{j\in S}\left|s_j\lambda_i - s_i\lambda_j\right| \ . \tag{2.21}$$

Using equation (2.21), we calculate ten thousand Gini coefficients and the corresponding tax rates based on one hundred combinations of σ/θ and γ/β respectively. The results are shown in Figure 2.5 below where each of the gray plus signs denotes a combination of a Gini coefficient and the corresponding equilibrium tax rate which was calculated with the same γ-β combination.

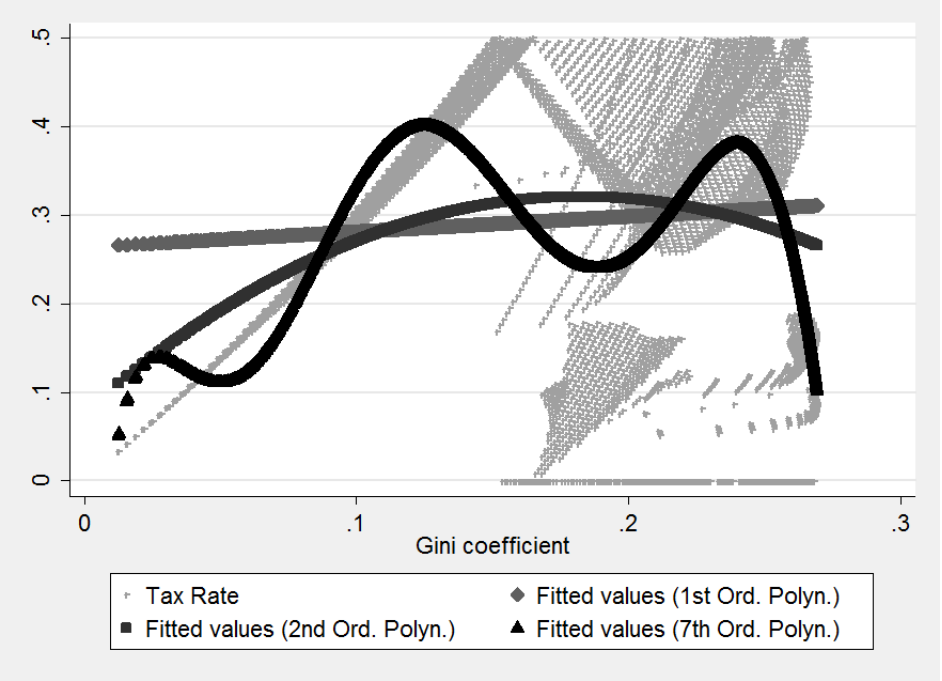

Figure 2.5: Ten thousand Gini coefficients and corresponding equilibrium tax rates.

As it turns out, in our model the equilibrium tax rate is a non-injective function of the Gini coefficient as most Gini values are mapped into several tax rates. This finding does not come as a surprise however. From Figure 2.3 we can easily see that particular tax rates, for example a tax rate of zero, can be described by different combinations of γ and β. This means, when translating the distribution of production factors into a single number, the tax rate, we lose information. And from equation (2.21) we can infer that this also happens when the endowment shares are mapped into Gini coefficients. For this

reason, we would not expect a well-defined function to exist that maps the Gini coefficients into equilibrium tax rates. Thus, in our model multiple equilibrium tax rates are often attached to the same Gini coefficient.

Delusively, naive regressions of the tax rate on the Gini coefficient would yield statistically significant results. If real observations were uniformly distributed over the space of possible equilibrium combinations such a regression would disclose 'true' relationships between both variables as it is represented by the solid lines in Figure 2.5. Conducting first, second or seventh order polynomial regressions of the tax rate on the Gini coefficient would suggest positive and statistically significant relationships which are linear, hump-shaped or M-shaped respectively depending on the particular estimation specification.[21] However, if in reality certain equilibria were more likely to occur than others, we would estimate different regression parameters which would also be statistically significant. For this reason, we can conclude that the Gini coefficient is not a sufficient measure for analyzing the effects of inequality on redistribution; at least when weakly institutionalized environments are considered. Regressions trying to predict a causal effect of the Gini coefficient on redistribution in such environments seem to be conceptually flawed. In general, we conjecture that this might always be the case when distributions of income are reduced to a single measure.[22]

2.7 Concluding Remarks

We presented a model to analyze the relationship between income inequality and redistribution in an environment where no institutions exist that allow for credible binding commitments between distinct social groups. In such an environment this relationship becomes more complex than suggested by the previous literature which does not take the role of institutions into account. In our model, a change in inequality might in- or decrease redistribution, depending on the initial level of inequality, or might even have

[21] We repeated this exercise for three hundred combinations of σ and θ and calculated 2.500 pairs of the gini coefficient and the tax rate for every such combination. Based on a total sample size of 750.000 variable values we again conducted first, second and seventh order polynomial regressions and found the same qualitative patterns.

[22] A similar claim was already made by Karabarbounis (2011) who based it on a simple numerical example instead of giving some theoretical explanation.

no effect at all. We also find that there always exist minimal state regimes in which the equilibrium tax rate is zero. Apart from this, we demonstrate that the Gini coefficient is not an appropriate measure when analyzing the relationship between inequality and redistribution in weakly institutionalized environments.

Therefore, our findings have important implications for the empirical literature on inequality and redistribution. First of all, since the strength of institutions has a huge impact on the relation between both variables it is necessary to control for this factor when conducting cross-sectional estimations. And second, our work indicates that the Gini coefficient is not an appropriate measure for regressions of state size on income inequality at least when considering weakly institutionalized environments. Apart from this, our model suggests several avenues for future theoretical work. First, it might be worthwhile to check the robustness of our results in a setting with n groups. And second, endogenizing the factor supply in our model to allow for distortionary effects of the tax rate would clearly be an appealing venture. Finally, the relation between inequality and redistribution in different political regimes has rarely been studied. For this purpose, linking the study of Jung and Sunde (2011) to our work appears to be a promising point of departure.

2.8 Appendix

2.8.1 Proof of Lemma 2.1

We first show that for any group there exists a pure strategy profile σ^* that is a $SPNE$ and leads to a unique RC. This proof is an extension of the proof of Lemma 1 in Jung and Sunde (2011) for the case of four groups.

Part I. Existence. First consider the preferred coalition of agenda setter $i \in S_k$ at stage $k \in \{k \in \mathbb{N}_0 : k \leq 3\}$ of game Γ where S_k is the set of all (remaining) groups whose income has not been redistributed away up to the current stage of the game, i.e., $S_0 = S$, $S_k \subset S \; \forall \; k > 0$ and $S_k \neq \varnothing \; \forall \; k$. Let $\mathscr{P}(S_k)$ denote the power set of S_k and let $\mathcal{I}_i = \{\mathcal{I} \in \mathscr{P}(S_k) : i \in \mathcal{I}\}$ be the set of all coalitions that include group i whereas $F_i = \{F \in \mathcal{I}_i : 2P_i > P_F\}$ represents the set of all coalitions in which group i is more powerful than the other coalition members at the current stage of the game. Define the generic set of winning coalitions as $W_k = \{W \in \mathscr{P}(S_k) : P_W > 0.5P_{S_k}\}$ and denote the set of stable coalitions as $E_k = \{E \in \mathscr{P}(S_k) : [\nexists\, Q \subset E : 2P_Q > P_E \wedge [P_Q \geq 2\max_{j \in Q} P_j \vee |Q| = 1]]\}$. Additionally, we define the union of the set of coalitions that are both winning and stable and the set of all (remaining) groups at the current stage of the game which is given by $R_k = [W_k \cap E_k] \cup S_k$ where the coalition that exhibits the lowest aggregate power in the set is given by

$$\Omega = \underset{X \in R_k}{\operatorname{argmin}} P(X) \; .$$

Then, the preferred proposal of an agenda setting group i at stage k of the game is given by

$$\Pi_{i,k} = \underset{X \in \mathcal{I}_i \cap R_k}{\operatorname{argmin}} P(X) \; .$$

This does not mean that there exist no other proposals which group i would support on the voting stage at a given history $\hbar$ of the game.Let $A_k \subseteq S_k$ be the set of all groups that have not been acting as an agenda-setter at the current stage of the game yet and let the subset $A_k^+ \subseteq A_k$ be defined as $A_k^+ = \left\{A_k^+ \in A_k : A_k^+ \in \Omega\right\}$. Now, define $\Pi_k = \bigcup\limits_{i \in A_k} \Pi_{i,k}$ as the set of preferred proposals of all groups that have not been acting as an agenda-setter at

the current stage of the game yet. Consequently, the most preferred proposal in view of group i among all the proposals of groups that have not acted as agenda setter yet, can be written as

$$\Psi_{i,k} = \begin{cases} \underset{X \in \Pi_k \cap \mathcal{I}_i \cap R_k}{\operatorname{argmin}} \; P(X) & \text{if } \Pi_k \cap \mathcal{I}_i \cap R_k \neq \varnothing \\ \varnothing & \text{otherwise .} \end{cases}$$

For notational convenience, we define the power of this coalition $\Psi_{i,k}$ to be infinite if it equals the empty set, i.e., $P_{\Psi_{i,k}} = \infty$ for $\Psi_{i,k} = \varnothing$. Then, the pure strategy profile for group i reads

$$\sigma^*_{i,k} = \begin{cases} \text{agenda-setting stage: } i \text{ proposes } \Pi_{i,k} \\ \text{voting stage: } i \text{ votes} \begin{cases} yes & \text{if } \; \Pi_{j,k} \in \mathcal{I}_i \cap R_k \wedge \left[P_{\Pi_{j,k}} \leq P_{\Psi_{i,k}} \vee \left[i \notin \Omega \wedge A^+_k \neq \varnothing \right] \right] \\ & \text{or } \Pi_{j,k} \in F_i \wedge \left[\{i\} \neq \Omega \vee A^+_k = \varnothing \right] \\ no & \text{otherwise .} \end{cases} \end{cases}$$

where $\Pi_{j,k}$ denotes the proposal made by group $j \in S_k$ on which groups currently vote. Now we need to prove that the pure strategy profile σ^* which is a vector of $\sigma^*_{i,k} \forall\, i,k$ constitutes a $SPNE$. Since we consider a finite game it is sufficient to show that there exists no one-shot deviation from $\sigma^*_{i,k}$ which is profitable for group i at any given history $\hbar$ of the game. In order to do this we need to distinguish two cases each one itself containing two sub-cases, since in this sequential game any group i is either a voter (*case A*) or an agenda setter (*case B*) at a given history of the game, and any proposed redistribution policy can either be rejected (*subcase 1*) or accepted (*subcase 2*).

Case A

Subcase A.1. Suppose that instead of voting *yes* according to $\sigma^*_{i,k}$ voter i would be better off if he voted *no*. Since the votes of the other groups do not depend on the decision of group i such a behavior could only cause a rejection of a proposal that would have been accepted otherwise if group i is pivotal for the decision outcome. In every other case such a deviation has no effect on equilibrium outcome and therefore cannot be beneficial. For this reason, let us assume that group i is pivotal for the decision outcome and that it

votes *no* contrary to $\sigma^*_{i,k}$.

To understand why no such deviation can be beneficial if $\Pi_{j,k} \in \mathcal{I}_i \cap R_k$ holds with $P_{\Pi_{j,k}} \leq P_{\Psi_{i,k}}$ is almost trivial since the latter condition implies that $\Pi_{j,k}$ either equals the previous or the current $\Psi_{i,k}$. Thus from the perspective of group i there exists no better proposal on which will be voted on at the given stage of the game according to σ^*. Voting *no* and thereby rejecting a proposal $\Pi_{j,k} \in \mathcal{I}_i \cap R_k$ can therefore not be beneficial for $P_{\Pi_{j,k}} \leq P_{\Psi_{i,k}}$.

Now suppose that $\Pi_{j,k} \in \mathcal{I}_i \cap R_k$ holds with $i \notin \Omega \wedge A^+_k \neq \varnothing$. In this case, with regards to σ^* rejecting the current proposal will result in the proposal and acceptance of coalition Ω at a subsequent history of the game. Since group i is not part of this coalition it cannot benefit from voting *no* instead of *yes* in such a situation.

Next consider the case where $\Pi_{j,k} \in F_i$ holds and group i is not more powerful than all other groups, $\{i\} \neq \Omega$. Since F_i consists of all coalitions in which group i is more powerful than all other coalition members, it is clear that whenever one of those coalitions is proposed and accepted given $\{i\} \neq \Omega$, group i strictly prefers such a proposal to $\Pi_{i,k}$ as it implies $\widetilde{p}_i$ to become maximal at the subsequent stage of the game. On the opposite, consider a history of the game where $\Pi_{j,k} \in F_i$ and group i is more powerful than all other groups but will not act as an agenda-setter anymore, $A^+_k = \varnothing$. Then, the best possible proposal after a rejection of the current is $\Psi_{i,k}$. Even though voting for $\Pi_{j,k} \in F_i$ causes some additional redistribution cost ϵ for group i, these are outweighed when becoming the sole ruler at the subsequent stage of the game as we assumed ϵ to be arbitrarily small. Thus, group i strictly prefers to vote *yes* for any $\Pi_{j,k} \in F_i$ if $A^+_k = \varnothing$. We can therefore conclude that it is not beneficial to vote *no* contrary to $\sigma^*_{i,k}$ for any group $i \in S$ at any stage of the game.

Subcase A.2. Now suppose that instead of voting *no* according to $\sigma^*_{i,k}$ group i would be better off if it voted *yes*. Again, this could only affect equilibrium outcome if group i's decision is pivotal and leads to the acceptance of a proposal that would have been rejected otherwise. Let us assume it does.

Let us first consider all cases where $\Pi_{j,k} \notin \mathcal{I}_i \cap R_k$ holds. Suppose additionally $\Pi_{j,k} \notin F_i$. In this case, it is obvious that a deviation from $\sigma^*_{i,k}$ cannot be beneficial for group i since

such a decision would lead to an unstable coalition in which group i is not the most powerful group. Given this, income of group i would be redistributed away at the subsequent stage of the game if such a proposal was accepted. Now suppose that $\Pi_{j,k} \notin \mathcal{I}_i \cap R_k$ holds with $\Pi_{j,k} \in F_i \wedge \{i\} = \Omega \wedge A_k^+ \neq \varnothing$. Also in this case voting *yes* instead of *no* is not beneficial for group i since it has not been acting as an agenda-setter yet and strictly prefers to propose and enforce the coalition $\Omega = \{i\}$ at a subsequent history of the game.

We next focus on all cases where $\Pi_{j,k} \notin F_i$ holds with $\Pi_{j,k} \in \mathcal{I}_i \cap R_k$ and $P_{\Pi_{j,k}} > P_{\Psi_{i,k}}$. Note that from $P_{\Pi_{j,k}} > P_{\Psi_{i,k}}$ it follows directly that $\Pi_{j,k} \neq \Psi_{i,k}$ and $\Psi_{i,k} \neq \varnothing$ must hold which rules out that $i \in \Omega \wedge A_k^+ = \varnothing$ can be true. Therefore we only need to distinguish two different cases. First consider that additionally $i \in \Omega \wedge A_k^+ \neq \varnothing$ holds true which implies $\Pi_{j,k} \neq \Omega$.[23] In this case, accepting the current proposal is not beneficial as the better proposal Ω will be made and accepted at a subsequent history of the game according to σ. Next suppose that $i \notin \Omega \wedge A_k^+ = \varnothing$ holds true instead. This implies that $\Psi_{i,k} \neq \varnothing$ will be proposed and accepted at a subsequent history of the game which generates a higher payoff for group i than the current proposal $\Pi_{j,k} \neq \Psi_{i,k}$.

Finally, consider the case where $\Pi_{j,k} \in F_i$ and $\Pi_{j,k} \in \mathcal{I}_i \cap R_k$ hold with $P_{\Pi_{j,k}} > P_{\Psi_{i,k}}$ and $i = \Omega \wedge A_k^+ \neq \varnothing$. Also in this case, group i strictly prefers to refrain from voting *yes* in order to propose and enforce Ω at a subsequent history of the game We can therefore conclude that it is not beneficial to vote *yes* contrary to $\sigma_{i,k}^*$ for any group $i \in S$ at any stage of the game.

Case B

In this case we show that group i cannot benefit from making a proposal $\pi_{i,k} \in \mathcal{I}_i$ that differs from that stipulated by $\sigma_{i,k}^*$. Again, we need to distinguish two different subcases.

Subcase B.1. Let us first assume that there exists such an alternative proposal $\pi_{i,k} \neq \Pi_{i,k}$ and that $\Pi_{i,k}$ is rejected if proposed. Then, obviously $\pi_{i,k}$ must be accepted if proposed as otherwise group i would not benefit from making this proposal.

By definition we know that $\Pi_{i,k} \in \mathcal{I}_i \cap R_k$ holds. Suppose first that $A_k^+ \neq \varnothing$ holds in addition. This implies that $j \in \Omega$ must also be true as otherwise $\Pi_{i,k}$ would not be rejected. Thus, in the given situation a rejection of $\Pi_{i,k}$ can only occur if $P_{\Pi_{i,k}} > P_{\Psi_{j,k}}$,

[23] Note that for $i \in \Omega \wedge A_k^+ \neq \varnothing \Leftrightarrow \Psi_{i,k} = \Omega$.

i.e., $P_{\Pi_{i,k}} > P_{\Omega}$ was to hold which according to σ^* is only possible for $i \notin \Omega$. But then, there can exist no $\pi_{i,k} \in \mathcal{I}_i$ which would not also be rejected.

Now assume that $A_k^+ = \varnothing$ holds instead. In that case again, $\Pi_{i,k}$ would only be rejected if $P_{\Pi_{i,k}} > P_{\Psi_{j,k}}$ was to hold which directly rules out $\Psi_{i,k} = \varnothing$. From this it follows that either $\Pi_{i,k} = \Psi_{i,k}$ or $\Pi_{i,k} = \Omega$ must be true according to σ^*. Thus, either the former inequality does not hold or there can exist no $\pi_{i,k} \in \mathcal{I}_i$ which would not also be rejected. For this reason no deviation from σ can be beneficial in the given subcase.

Subcase B.2. Let us now suppose that there exists an alternative proposal $\pi_{i,k} \neq \Pi_{i,k}$ and that $\Pi_{i,k}$ is accepted if proposed. Note that by the nature of the game $\pi_{i,k} \in R_k$ holds as no proposal $\pi_{i,k} \notin W_k$ can be and no proposal $\pi_{i,k} \notin E_k$ will be accepted. Furthermore, no proposal $\pi_{i,k} \notin \mathcal{I}_i$ can be made by group i. Hence $\pi_{i,k} \in \mathcal{I}_i \cap R_k$ needs to hold.

Given our assumption of a bijective power mapping $\pi_{i,k} \neq \Pi_{i,k}$ then implies $P_{\Pi_{i,k}} < P_{\pi_{i,k}}$ since $\Pi_{i,k} = \operatorname{argmin}_{X \in \mathcal{I}_i \cap R_k} P(X)$. With regards to the optimization problem (2.8) we can therefore conclude that it is not beneficial for group i to propose $\pi_{i,k}$ instead of $\Pi_{i,k}$ in the given subcase. □

Part II. Uniqueness. Finally, we need to show that all $SPNEs$ lead to the same RC. We do this by first emphasizing that the assumption of a bijective power mapping implies that in equilibrium different RCs cannot be equal in aggregate power. To see this suppose to the contrary that $P_M = P_Q$ holds for the two equilibrium coalitions $M, Q \in \mathscr{P}(S) \setminus \{\varnothing\}$ which are not identical, $M \neq Q$. Obviously, the bijective power mapping directly rules out $|M| = |Q| = 1$ in the given case. Additionally, a coalition of two groups can never be an equilibrium outcome, because, due to the bijective power mapping, it would not be stable as the stronger group could always propose a winning subcoalition only containing itself at a later stage of the game. Uniqueness in the case of the grand coalition comprising all four groups is trivial. Hence, we need to distinguish two cases, a case with two coalitions comprising three groups each, and a case with one coalition of three groups and another with one group only. First, suppose that each of the two coalitions comprises three groups, i.e., $|M| = |Q| = 3$. In this case, two groups $i, j \in S$ must be part of both coalitions, $i, j \in M \cap Q$. Given this, it requires the third group l also to be in both coalitions,

$l \in M \cap Q$, for $P_M = P_Q$ to hold which implies $M = Q$ and thereby contradicts our former supposition. Second consider the case where a coalition M with $|M| = 3$ has the same power as some coalition Q with $|Q| = 1$, i.e., $P_M = P_Q$. Obviously, M can only be winning if it incorporates the fourth group. But then $|M| \neq 3$ holds in equilibrium which contradicts our assumption. Therefore we can conclude that in equilibrium any two coalitions M and Q can only be equal in power, $P_M = P_Q$, when they are identical, $M = Q$.

Under strategy profile σ^* the resulting RC does not depend on the moves of nature. Therefore the $SPNEs$ in our finite coalition formation and redistribution game with perfect information can only lead to different RCs if a pivotal group i is indifferent about her action at a certain decision node. Suppose first that group j is not part of the equilibrium coalition and is indifferent at a given history of the game $\hbar$. In this case, it can only be pivotal if it supports a coalition $M \neq \Omega$ with $j \in M$ that is not stable. Note that this creates nothing but some redistribution costs ϵ for group j as its income will be redistributed away in a following stage of the game. Therefore group j will always strictly prefer not to be part of any transitory coalition(s). Now suppose that the pivotal group i is part of different equilibrium coalitions and is indifferent at a given history of the game $\hbar$. This can only be the case if (at least) two actions lead to the same equilibrium payoff which requires – given the optimization problem (2.8) and the political power of group i – the aggregate power of (at least) two different RCs to be the same. With regards to our former reasoning this is impossible. Thus there cannot exist two different equilibrium coalitions between which any pivotal group $i \in S_k$ is indifferent at a given history of the game $\hbar$. This establishes the proof of Lemma 2.1. □

2.8.2 Proof of Lemma 2.2

Proof. Note first that due to our assumption of a bijective power mapping, a two-group coalition cannot be stable, since one group always dominates the other, and therefore could always successfully propose a coalition that only contains itself at the subsequent stage of the game. Hence, $|RC| \neq 2$ always holds where $|RC|$ denotes the cardinality of set RC. Thus, we can immediately conclude that $\lambda_{MAX} \leq 0.5 \iff |RC| \geq 3$ holds. Given this

situation, a coalition of the three least powerful groups is always the first-best solution in order to maximize $\widetilde{p}_i$. Since it only is stable, i.e., it contains no winning subcoalition itself, if $\lambda_{MIN} \geq \overline{\lambda} - \underline{\lambda}$ holds, we can conclude that $0.5 > \lambda_{MAX} \wedge \lambda_{MIN} \geq \overline{\lambda} - \underline{\lambda} \Longleftrightarrow i_{MAX} \notin RC$.

If this condition fails the next best alternative is the exclusion of the second most powerful group. But this is not feasible under the given conditions as this requires $\lambda_{MIN} \geq \lambda_{MAX} - \underline{\lambda}$ to hold. Since then $\lambda_{MIN} \geq \overline{\lambda} - \underline{\lambda}$ would also be true, the exclusion of the most powerful group would also be feasible in that case. Thus, the second most powerful group is *always* part of the RC under the given conditions, i.e., $0.5 > \lambda_{MAX} \iff \bar{i} \in RC$.

For this reason it is in fact the exclusion of the third most powerful group which represents the second-best solution. It will be realized if in the given situation the first best solution is not feasible and $\lambda_{MIN} \geq \lambda_{MAX} - \overline{\lambda}$ holds, i.e., $\lambda_{MIN} < \overline{\lambda} - \underline{\lambda} \wedge \lambda_{MIN} \geq \lambda_{MAX} - \overline{\lambda} \Longrightarrow \underline{i} \notin RC$. Only if this condition also fails a coalition of the three most powerful groups becomes the preferred choice which requires $\underline{\lambda} \geq \lambda_{MAX} - \overline{\lambda}$ to be stable. Thus, $\lambda_{MIN} < \overline{\lambda} - \underline{\lambda} \wedge \lambda_{MIN} < \lambda_{MAX} - \overline{\lambda} \wedge \underline{\lambda} \geq \lambda_{MAX} - \overline{\lambda} \Longrightarrow i_{MIN} \notin RC$.

If in the given situation even the coalition of the three most powerful groups is unstable, the grand coalition $RC = S$ occurs as the equilibrium outcome. Note that $\underline{\lambda} < \lambda_{MAX} - \overline{\lambda}$ implies $\lambda_{MIN} < \lambda_{MAX} - \overline{\lambda}$ to hold. □

2.8.3 Proof of Proposition 2.3

Proof. This proof shows that there always exist γ-β combinations in the admissible γ-β space for which $t = 0$ given $\lambda_{MAX} < 0.5$ with $0 < \sigma, \theta < 1$ and $\sigma + \theta < 1$. From Proposition 2.2 we know that this is the case if and only if the two well-defined conditions (a) $\overline{\lambda} > \lambda_{MIN} + \underline{\lambda}$ and (b) $\lambda_{MAX} > \underline{\lambda} + \overline{\lambda}$ are satisfied.

Let us focus on the $\lambda_{\mathcal{E}} = 0.5$ locus in the following since we already know that it always passes through the admissible γ-β space, see Lemma 2.3. Concerning this locus, the following Lemma will prove to be useful.

Lemma 2.4. $\lambda_{\mathcal{E}} = 0.5 \Longrightarrow \lambda_{MIN} = \lambda_{\mathcal{P}}$

Proof. Note that in general for any $\gamma, \beta > 0.5$ the relative power of the $\mathcal{P}$-group is always the smallest, $\lambda_{MIN} = \lambda_{\mathcal{P}}$, which can easily be seen from equations (2.11) - (2.14). Addi-

tionally, for $\gamma = \beta = 0.5$ we can infer from equation (2.14) that $\lambda_{\mathcal{E}} = 0.25\,(1+\sigma+\theta) < 0.5$ holds for all $\sigma, \theta > 0 \wedge \sigma + \theta < 1$. And since the $\lambda_{\mathcal{E}} = 0.5$ locus is a convex function, see equation (2.15), we can conclude that it never passes through the southwestern area of the admissible γ-β space where $\gamma, \beta \leq 0.5$. Thus, all points lying on the $\lambda_{\mathcal{E}} = 0.5$ locus represent γ-β combinations for which $\gamma, \beta > 0.5$ is true. Therefore, $\lambda_{\mathcal{E}} = 0.5 \Rightarrow \lambda_{MIN} = \lambda_{\mathcal{P}}$. □

It is almost trivial to state that for all γ-β combinations describing the $\lambda_{\mathcal{E}} = 0.5$ locus $\lambda_{MAX} = \lambda_{\mathcal{E}}$ holds and $\lambda_{MAX} > \underline{\lambda} + \overline{\lambda}$, i.e., condition (b) is always satisfied. Furthermore, from Lemma 2.4 it follows that for all γ-β combinations describing the $\lambda_{\mathcal{E}} = 0.5$ locus $\underline{\lambda}$ and $\overline{\lambda}$ equal $\lambda_{\mathcal{L}}$ and $\lambda_{\mathcal{A}}$ respectively. Thus, we can infer that condition (b) also holds for all γ-β combinations lying in-between the $\lambda_{\mathcal{E}} = 0.5$ and the $\lambda_{\mathcal{E}} + \lambda_{\mathcal{P}} = 0.5$ locus which encircles the former in the southwest but never intersects it due to $\lambda_{\mathcal{P}} > 0$.

The next Lemma shows that additionally there always exist γ-β combinations lying on the $\lambda_{\mathcal{E}} = 0.5$ locus which also satisfy condition (a) such that $\overline{\lambda} > \lambda_{MIN} + \underline{\lambda}$ is true which is equivalent to $\overline{\lambda} > 0.25$ since $\lambda_{MIN} + \underline{\lambda} + \overline{\lambda} + \lambda_{MAX} = 1$ and $\lambda_{MAX} = 0.5$.

Lemma 2.5. *There always exist admissible γ-β combinations such that $\lambda_{\mathcal{E}} = 0.5$ holds with $\overline{\lambda} > 0.25$.*

Proof. Note that $\overline{\lambda} > 0.25 \iff \lambda_{MIN} + \underline{\lambda} < 0.25$ given $\lambda_{MAX} = 0.5$. From Lemma 2.4 it follows that for all γ-β combinations describing the $\lambda_{\mathcal{E}} = 0.5$ locus $\overline{\lambda}$ equals either $\lambda_{\mathcal{L}}$ or $\lambda_{\mathcal{A}}$.

Suppose first that $\overline{\lambda} = \lambda_{\mathcal{L}}$. In this case, $\lambda_{MIN} + \underline{\lambda} = \lambda_{\mathcal{P}} + \lambda_{\mathcal{A}} = (1-\gamma)(1-\theta) = 0.25$ for $\gamma^* = (0.75-\theta)/(1-\theta)$. Hence all γ-β combinations which satisfy $\lambda_{\mathcal{P}} + \lambda_{\mathcal{A}} = 0.25$ lie on the vertical line $\gamma^* = (0.75-\theta)/(1-\theta)$. Obviously, $0 < \gamma^* < 1 \;\forall\, 0 < \theta < 0.75$, i.e., this line always passes through the admissible γ-β space for $0 < \theta < 0.75$ which it divides in two regions. If $\gamma < \gamma^*$ then $\lambda_{\mathcal{P}} + \lambda_{\mathcal{A}} > 0.25$ whereas for $\gamma > \gamma^*$ it follows $\lambda_{\mathcal{P}} + \lambda_{\mathcal{A}} < 0.25$. Thus, for all γ-β combinations on the $\lambda_{\mathcal{E}} = 0.5$ locus which lie to the right of $\gamma^* = (0.75-\theta)/(1-\theta)$ we can conclude that they satisfy $\overline{\lambda} > 0.25$. In the special situation, where $0.75 < \theta < 1$ the income share devoted to physical strength becomes so large that $\overline{\lambda} > 0.25$ effectively holds for *all* γ-β combinations lying on the $\lambda_{\mathcal{E}} = 0.5$ locus.

Now consider the case $\overline{\lambda} = \lambda_{\mathcal{A}}$. Repeating the above exercise we find that $\lambda_{\mathcal{P}} + \lambda_{\mathcal{L}} = 0.25$ for $\beta^* = (0.75 - \sigma)/(1 - \sigma)$. Analogously, this horizontal line always passes through the admissible γ-β space for $0 < \sigma < 0.75$. Here, $\overline{\lambda} > 0.25$ holds for all γ-β combinations which lie on the $\lambda_{\mathcal{E}} = 0.5$ locus or above this line. For $0.75 < \sigma < 1$ this is the case for *all* γ-β combinations describing the $\lambda_{\mathcal{E}} = 0.5$ locus. Since, due to our assumption $\sigma + \theta < 1$, at most one of both variables, σ or θ, can take on values larger than 0.75 we summarize the following.

First, given $\sigma \vee \theta > 0.75$ we already know that in either case $\overline{\lambda} > 0.25$ holds for *all* γ - β combinations lying on the $\lambda_{\mathcal{E}} = 0.5$ locus. Second, if $\sigma, \theta < 0.75$, both, the vertical and the horizontal line pass through the admissible γ-β space. And as the $\lambda_{\mathcal{E}} = 0.5$ locus is a convex function that never passes the southwestern area, such that $\gamma, \beta \leq 0.5$ holds, it cannot run through the admissible γ-β space without intersecting (at least) one of the two lines. Thus, for $\sigma, \theta < 0.75$ there always exist admissible γ-β combinations for which the $\lambda_{\mathcal{E}} = 0.5$ locus (partially) lies to the right of $\gamma^* = (0.75 - \theta)/(1 - \theta)$ and/or above $\beta^* = (0.75 - \sigma)/(1 - \sigma)$ implying $\overline{\lambda} > 0.25$ to hold in this case also. This establishes the proof. □

From Lemma 2.5 we know that for any given $0 < \sigma, \theta < 1$ there always exist admissible γ-β combinations along the $\lambda_{\mathcal{E}} = 0.5$ locus for which $\overline{\lambda} > \lambda_{MIN} + \underline{\lambda}$. And according to our former reasoning the complete $\lambda_{\mathcal{E}} = 0.5$ locus is encircled by a southwestern belt, the $\lambda_{\mathcal{E}} + \lambda_{\mathcal{P}} = 0.5$ locus, and all γ-β combinations lying in-between both loci denote situations where $\lambda_{MAX} > \underline{\lambda} + \overline{\lambda}$ holds what was required to be proved. □

For a better understanding of our general reasoning the proof of Proposition 2.3 is built upon, Figure 2.6 below shows all relevant loci for the parametrization given in the main text.

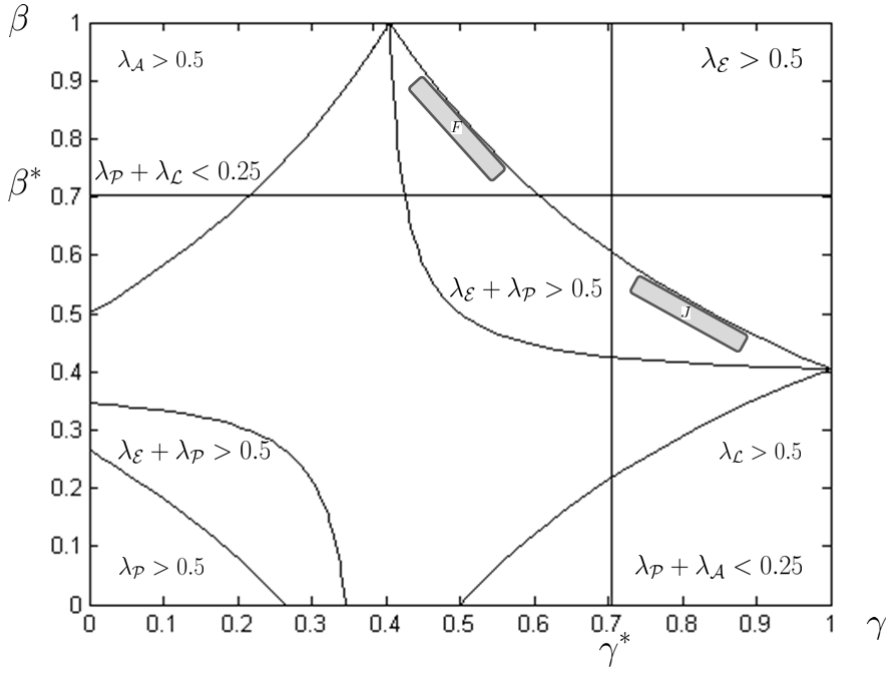

Figure 2.6: Illustration of reasoning for proof of Proposition 2.3 ($\sigma = \theta = 0.16$).

In this specific case, both, the vertical and the horizontal line, pass through the admissible γ - β space with $\beta^* = \gamma^*$ due to $\sigma = \theta$, and the two exemplary sets F and J embrace γ - β combinations for which $\lambda_{MAX} > \underline{\lambda} + \overline{\lambda}$ and $\overline{\lambda} > \lambda_{MIN} + \underline{\lambda}$ hold such that the corresponding equilibrium tax rate $t = 0$.

Chapter 3

Income, Democracy, and Critical Junctures

This chapter is joint work with Matteo Cervellati from the University of Bologna and Uwe Sunde and Thomas Vischer from the University of St. Gallen.

3.1 Introduction

The strong positive cross-country relationship between income and democracy is one of the few robust stylized facts in social sciences. This correlation has been interpreted in different ways. Institutions, in particular political institutions, have been identified as one of the fundamental determinants of long-run development, suggesting that the causality behind the relationship runs from democracy to income, see the discussion in Acemoglu, Johnson, and Robinson (2005) for example. On the other hand, economic development and "modernization", as reflected by income, has been argued to be the main driver behind the emergence and stabilization of good institutions, such as democracy. This is the so-called modernization hypothesis, or Lipset hypothesis, following the seminal article of Lipset (1959).[1]

In two related articles, Acemoglu, Johnson, Robinson and Yared (2008, 2009) try to isolate the direction of causality proposed by modernization hypothesis against the competing hypothesis of critical junctures. According to the latter, differences in the institutional and economic development of countries are jointly determined at certain critical historical junctures that affect the subsequent political-economic development path. A prime example for such a critical juncture is the colonization of a country and the conditions and strategies under which it was colonized. Whether a country became a colony purely for exploitation purposes or whether it was considered a place of residence, a so-called settler colony, has had a huge impact on its subsequent development.[2] The two articles

[1]The modernization hypothesis was assessed empirically in numerous studies, see Bollen and Jackman (1985), Arat (1988), Diamond (1992), Burkhart and Lewis-Beck (1994), Barro (1999), Przeworski et al. (2000), Boix and Stokes (2003), Glaeser et al. (2004), Epstein et al. (2006), and Przeworski et al. (2006) as the most prominent examples, see also Cheibub and Vreeland (2010) for a recent survey.

[2]The consequences of institutions that were set up at critical historical junctures in the distant past for institutional quality and consequently for long-run development have been shown by Acemoglu, Johnson, and Robinson (2001, 2002). Focusing on the case of former colonies they also show that the nature of these institutions was endogenous and varied substantially depending on natural circumstances like the disease environment or the density of the indigenous population. They find that clusters of bad, i.e., extractive (political, economic and social) institutions were more likely to emerge in colonies where exogenous circumstances provided a hostile environment for European settlers, whereas good institutions

by Acemoglu et al. (2008, 2009) provide evidence that the effect of income on democracy disappears once controlling for country-specific, time-invariant factors that might reflect historical conditions. Specifically, the empirical results show that within-country development in incomes does not explain political development in terms of variations in the quality of democracy. The authors interpret this finding as evidence against the modernization hypothesis and in favor of the critical junctures theory. Moreover, the results imply that the correlation between income and democracy must be driven either by third factors such as critical historical junctures, or by causality running from institutions to economic development.

This article presents new evidence suggesting that the income-democracy nexus might be heterogeneous rather than non-existent, and thus provides an alternative interpretation of the empirical findings of Acemoglu et al. (2008, 2009). The analysis on which this interpretation is based follows closely that of Acemoglu et al. (2008), but takes a scrutinizing look at the implications of the critical junctures hypothesis. According to Acemoglu et al. (2008), p. 812, as a consequence of critical historical junctures countries might "... embark on divergent political-economic development paths, some leading to relative prosperity and democracy, others to relative poverty and dictatorship." The finding that the positive partial correlation between income and democracy disappears when accounting for time-invariant country fixed effects, i.e., when looking at the correlation between within-country changes in income and democracy, therefore supports the critical junctures hypothesis and rejects the modernization hypothesis *on average*. What this analysis potentially conceals, however, is that the nexus between income and democracy might differ systematically between the "divergent political-economic development paths".

To illustrate the implications of this observation, consider a model with different political economic regimes, an oligarchic and a democratic one. Depending on the particular political regime increases in income might have completely different effects on the observable institutional quality. For instance, in the framework of Acemoglu and Robinson (2001, 2006), an increase in aggregate income tends to benefit the entire population under democracy. This is because incomes are more equally distributed in democracies due to

were implemented in countries that were more favorable for large-scale settlements.

the popular support for redistribution. One would therefore expect little, or, if anything, a positive effect on institutional quality since additional income is used to please all parts of society, including those that might not directly benefit from the income increase, or those that are discontent with the existing political order. Under oligarchy, on the other hand, the effect of an increase in income on institutional quality might be completely different. In such a regime, an increase in income might be seized by the ruling elite which is capable of expropriating its opponents, channeling resources in their pockets, and repressing the population. As a consequence of the appropriation, cleptocracy and repression associated with an increase in aggregate income, the measured quality of institutions is likely to deteriorate. Such differences in political-economic development paths have already been described by earlier models on the political Kuznets curve, see for example Acemoglu and Robinson (2002).[3] Following this reasoning, we test the hypothesis that income improvements have systematically different effects on subsequent institutional development depending on the particular political-economic development path a country embarked on at a critical historical juncture.

This hypothesis is also in line with evidence put forward by the political science literature according to which economic development might make established democracies more stable and improve their institutional quality, but might not necessarily lead to an increase in the probability of a transition from autocracy to democracy, see, e.g., Przeworski and Limongi (1997), Przeworski, Alvarez, Cheibub, and Limongi (2000), Benhabib and Przeworski (2006), and Epstein et al. (2006). It is also worth noting that, in his seminal article, Lipset (1959) explicitly restricted attention to self-governing states only, suggesting that other historical events, especially those restraining the independence of countries by third parties, might interfere with the relationship between income and democracy. Despite the fact that the empirical analyses is based on a sample of independent countries, it is impor-

[3]According to Acemoglu and Robinson (2002), economies with relatively low inequality in wealth are more likely to democratize as a consequence of economic growth than economies that are highly unequal and where income increases lead to a further increase in inequality possibly due to "... the absence of a well-developed civil society or other factors making it hard for the poor to organize", Acemoglu and Robinson (2002), p. 196. Such economies are likely to end up in an 'autocratic disaster': economic development is slow, and any increase in income tends to increase inequality, and to reduce the possibilities for institutional improvements, by ways of a revolution for example. Thus, institutional quality deteriorates because of the diversion of resources by the ruling elites and their successful repression of any revolutionary movements.

tant to note that "... many successor nations continued to be ruled in essentially colonial ways, enabling established leaderships to maintain a type of domestic neo-colonialism (Brown (1999), p. 710)." Accordingly, the colonial experience of a country has potentially long-lasting effects on the income democracy nexus.

Our analysis begins with a detailed investigation of the main correlates of the country fixed effects in the estimations of Acemoglu et al. (2008, 2009) to get a grasp of what is behind the historical junctures that are subsumed in the fixed effects. As suggested by Acemoglu et al. (2009) at the end of their paper, the fixed effects are related to various proxies for different development paths that countries have taken at critical junctures. Among these proxies are colonial status, the period of self-governance measured by years since independence, and early institutions proxied by settler mortality. A re-examination of the data used by Acemoglu et al. (2008, 2009) indeed delivers substantial heterogeneity in the income effect across different groups of countries. We find evidence for a significant negative income effect on democracy in colonies, in particular in colonies with a bad long-run institutional environment, as reflected by high settler mortality or late independence. On the contrary, we find evidence for a positive income effect in countries that were never colonized. This result is robust to different measures of the democracy score, alternative sources of income data and different panel frequencies.

Our results complement several recent papers that criticize the findings and interpretation of Acemoglu et al. (2008, 2009) on conceptual and econometric grounds. From a conceptual point of view, Gundlach and Paldam (2009) argue that the fact that accounting for fixed effects removes any effect of income on institutions does not provide conclusive evidence for the critical junctures hypothesis, and does not rule out other potential explanations. Rather than historical junctures, the fixed effects might capture geographical features as fundamental determinants of long-run development which might provide more relevant variation in economic development in light of the modernization hypothesis than short-run variations in income over time. This interpretation also challenges the view that institutions, rather than geography, are the fundamental determinant of long-run development. Our findings lend support to the view that critical junctures have long-run effects on institutional and economic development, as well as to Lipset's

own dictum that the modernization hypothesis might be valid for particular countries, but not for all. Recent work by Benhabib, Corvalan, and Spiegel (2011), Corvalan (2011) and Fayad, Bates, and Hoeffler (2011) has criticized the result of Acemoglu et al. (2008, 2009) concerning the income effect on democracy from an econometric perspective. Using models that account for censoring of the dependent variable as well as different data sets, or employing a binary coding of democracy and estimating a conditional logit model, respectively, Benhabib, Corvalan, and Spiegel (2011) and Corvalan (2011) find a significant positive income effect. In contrast, Fayad, Bates, and Hoeffler (2011) find a negative effect with dynamic panel estimation techniques that account for short-run cross-country heterogeneity and cross-sectional error dependence. While not addressing these criticisms, our central finding of a heterogeneous effect of income on democracy holds up under almost all alternative estimation methods, specifications, and sample compositions used in these studies.

The findings have several important implications. First, they suggest that the modernization hypothesis and the critical junctures hypothesis should not be seen as competing and mutually exclusive explanations of long-run political development. Second, the findings are difficult to explain on purely geographical grounds, lending credence to the view that institutions have important long-run implications for economic as well as institutional development. Third, the findings reconcile the diverging estimates of the income effect on democracy that have been reported using different econometric techniques and data sets by suggesting that the true effect of income on democracy might be heterogeneous. This implies that different sample compositions or techniques that give different weights to particular subsamples or parts of the distribution of observed outcomes are likely to generate different estimates. Yet, these estimates are likely to capture only part of the underlying heterogeneity in the relationship between income and democracy.

This paper is structured as follows. Sections 3.2 and 3.3 briefly elaborate on the data and the econometric specification we employ in our study. Section 3.4 presents the results that replicate and extend the earlier findings, followed by a thorough investigation of the robustness of the main finding. Section 3.5 concludes.

3.2 Data

Our analysis mainly draws on the data compiled by Acemoglu et al. (2008). We first replicate their results on the effect of income on democracy, using the same data sets and data sources as well as empirical specifications in order to demonstrate our point in a controlled and fully comparable way.[4] The primary source for data on GDP per capita is the Penn World Table and democracy data are taken from the Freedom House as well as the Polity IV project. Both indices are normalized to the range between zero and one, with higher values indicating higher levels of democracy. Acemoglu et al. (2008) also provide information on the former colonial status of a country, which is equal to one if a country is a former colony and zero otherwise.[5] To document the consistency of our results with the earlier findings and to demonstrate the relevance of the correct specification by accounting for heterogenous development dynamics, the baseline analysis does not introduce any changes to the data obtained from Acemoglu et al. (2008). The benchmark estimates are based on a five-year panel over the period 1960-2000 to replicate their estimates.

To document the robustness of our findings, we also consider ten-year panels, as well as alternative data sources for the democracy variable and for the coding of colonies. As alternative democracy measures, we use the index by Vanhanen (2000) which has the advantage of factually not being censored from above. Additionally, we employ the binary coding of democracy developed by Cheibub, Gandhi, and Vreeland (2010) which has been used by Benhabib, Corvalan, and Spiegel (2011) as well.[6] Moreover, since for a number of countries which are initially excluded from the analysis the colonial status in the original data of Acemoglu et al. (2008) is not identified,[7] we use data provided by the *French Center for Research and Studies on the World Economy* (CEPII) as an alternative data source.[8] The advantage of this data source is a comprehensive classification of countries

[4]The data is taken from `http://pubs.aeaweb.org/doi/pdfplus/10.1257/aer.98.3.808`.

[5]Table 3.1 provides an overview of the countries present in the sample as well as their colonial status.

[6]The source of the data is `https://netfiles.uiuc.edu/cheibub/www/datasets.html`.

[7]The discrepancy stems from the fact that the data is taken from a 500-year panel, where some countries are not present. Issues of sample selection are not very serious, however, because many of these countries lack GDP data for time periods before the fall of the Soviet Union.

[8]The source is `http://www.cepii.fr/`. The corresponding classification of countries with regard to their former colonial status is provided in Table 3.2 in the Appendix.

which allows us to identify the former colonial status of all countries in the sample of Acemoglu et al. (2008).[9] In addition, the data set also provides information on the most recent colonizing power and the major Western European colonizers for further analyses and additional robustness checks.

The summary statistics of the main variables that we use for our estimations in the different samples are reported in Table 3.3 for the observation period 1960-2000.

3.3 Econometric Specification

Consider the econometric model estimated by Acemoglu et al. (2008):

$$d_{i,t} = \alpha d_{i,t-1} + \gamma y_{i,t-1} + x'_{i,t-1}\beta + \mu_t + \delta_i + u_{i,t} \tag{3.1}$$

where $d_{i,t}$ is the democracy score of country i at time t. The estimated model is dynamic as it captures persistence as well as mean-reverting dynamics by including the lagged value of the dependent variable $d_{i,t-1}$. The coefficient of interest γ reflects the effect of the lagged value of log income per capita $y_{i,t-1}$ on democracy, and $x'_{i,t-1}$ contains additional covariates and a constant. To mitigate issues of endogeneity and unobserved heterogeneity, the specification includes a full set of country dummies δ_i and time fixed effects μ_t. Additional transitory shocks to democracy and other omitted factors are captured in the error term $u_{i,t}$.

The present specification posits a linear relation between log income and democracy, thereby assuming a homogenous effect of income on democracy over a heterogenous set of countries. As a first step, we replicate the results of Acemoglu et al. (2008, 2009) and estimate the specification (3.1) on the entire sample. However, in light of the above discussion and the emerging consensus for the importance of critical junctures and early institutions, the question arises whether this specification is suitable to identify the effect of income per capita, as a proxy for the level of development, on democracy. For this reason, in a next step we scrutinize the potential determinants and covariates of the

[9] In consequence, the sample of non-colonies is increased by three countries, the sample of colonies by thirteen, which leaves us with a total sample size of 150 countries, see Tables 3.1 and 3.2.

estimated fixed effects δ_i. We analyze whether they are correlated with proxies for long-run institutions, such as former colonial status or historical settler mortality.

To account for heterogeneous effects of income on democracy, we then modify equation (3.1) by allowing income per capita to partially interact with one of the aforementioned indicators of potentially different political-economic development paths that a country might embark on after a critical historical juncture (C_i) such that

$$d_{i,t} = \alpha d_{i,t-1} + \gamma y_{i,t-1} + \phi y_{i,t-1} C_i + x'_{i,t-1}\beta + \mu_t + \delta_i + u_{i,t} \ . \tag{3.2}$$

In this respect, equation (3.2) allows to identify a heterogeneous effect of per capita income on democracy for different groups of countries, such as colonies and non-colonies. The advantage of this specification is a relatively efficient estimation of heterogeneity since the full sample can be used. At the same time, the model nests the specification (3.1) estimated by Acemoglu et al. (2008, 2009) as a special case. The fact that proxies for long-run institutions are time invariant renders the estimation of the main effect of C_i in specification (3.2) unnecessary since within the given panel data framework it is already included in the country fixed effect. As a more flexible alternative, we finally estimate specification (3.1) separately for different groups of countries that are likely to be on different long-run political-economic development trajectories due to different historical legacies or different routes taken at critical historical junctures. Such a specification then also allows for heterogeneity in the effect of lagged democracy. We apply different proxies for C_i, including the former colonial status, as well as attached characteristics like the origin of the colonizers, the date of independence, or settler mortality which are supposed to account for differences in institutions.

3.4 Empirical Results

3.4.1 Main Results

As first step of our analysis, we replicate the main results of Acemoglu et al. (2008, 2009) by estimating specification (3.1) on the full sample of countries for five-year periods from

1960 to 2000 with a full set of country and period dummies. The results are presented in Table 3.4 and confirm the finding of no significant effect of lagged log income per capita on democracy.

We then regress the country dummy estimates from this estimation on different variables that might capture the distinct underlying development paths of countries in the two presumed subsamples. These variables include a binary indicator of whether a country has ever been a colony or not, the year of independence of a country, the log of settler mortality as well as indicator variables for colonies of the major colonizing powers of the 20th century Great Britain, France and Belgium.[10] The results are presented in Table 3.5. The estimates in columns (1)-(6) confirm the findings of Acemoglu et al. (2009) that the fixed effects indeed capture historical variables that are related to early institutions when regressed in isolation. The results imply that colonies in general, and especially colonies with a relatively high settler mortality, as well as countries that became independent relatively recently, display lower levels of democracy conditional on lagged income and democracy. The same is true for countries colonized by Great Britain, France and Belgium. Countries that were colonized by other Western European colonizing powers do not differ substantially in terms of the level of democracy conditional on lagged income and democracy. The overall picture is the same in multivariate regressions, as is shown in column (7), even though it is evident that in particular high settler mortality and late independence are associated with low levels of democracy. Similar results have been reported by Acemoglu et al. (2009).

These significant differences in the estimates of country fixed effects suggest that there might also be systematic variation in the development dynamics across countries. To test the hypothesis that this variation materializes in systematic differences of the effect of income on democracy across distinct groups of countries, we estimate partially interacted models as described by equation (3.2) that allow for heterogeneous income effects for countries with different historical institution proxies. This extends the previous analysis by accounting for a heterogeneous income effect while retaining the assumption that

[10]Note that former German colonies are implicitly included as well, since they are recoded with regards to the last colonial power (e.g., Namibia is coded as a former British colony).

institutional dynamics in terms of the effect of past democracy are homogeneous and allowing for country-specific time-invariant effects driving the level of democracy.

As already explicated above the colonization of a country represents a prime example of a critical juncture. Earlier work in political science suggests that a colonial past generally diminishes a democracy's prospects for survival, see Bernhard, Reenock, and Nordstrom (2004) and the references therein.[11] Table 3.6 presents a similar analysis as before, but now using information on whether or not a country ever was a colony as an indicator. The results for the Freedom House index of democracy in Panel (A) suggest that the income effect on democracy is positive for countries that have never been a colony, while the income effect is negative for former colonies. For the Freedom House data, this difference is significant. When estimating the model separately for non-colonies and former colonies, a similar picture emerges, with a significant positive effect of income on democracy for non-colonies.[12] The results are qualitatively identical, see columns 4 and 5, when employing the more appropriate difference GMM estimator as specified by Acemoglu et al. (2008) which mitigates potential concerns regarding biased estimates in dynamic panels with fixed effects.[13] When using a ten-year panel instead the relevant effects for both democracy measures are statistically even more significant and also quantitatively larger than in the five-year panel as can be seen in Table 3.7. Also, the reduction in the panel frequency avoids the issue of instrument proliferation in the non-colony sample for the difference GMM estimates due to the reduced number of available

[11] Bernhard et al. (2004), p. 229, cite a number of studies that "... argue that colonial economic development distorted the social structure in ways that (a) increased the power of classes that have been resistant to democracy while (b) weakening those classes whose struggles for political influence and incorporation have been historically associated with the establishment of democracy." On p. 230, they conclude that "... the vast preponderance of studies that take up the subject of the relationship between colonialism and democracy argue that its effects were negative. They offer an array of evidence in support of this and specify several different, not necessarily exclusive, mechanisms for why this is so."

[12] The positive income effect is robust to the respective inclusion of average school years, the log-transformed population size, the median age in the population as well as the percent of the labor share of gross value added. A regression simultaneously controlling for all these factors still delivers a positive and significant income effect. Detailed results are available upon request.

[13] A potential caveat with this result, however, is the issue of too many instruments, as pointed out by Roodman (2007). Accordingly, if the number of instruments exceeds the number of groups, the estimator may fail to expunge the endogenous component of the instrumented regressor. This issue is particularly relevant for the sample of the 35 non-colonies, where the difference GMM estimator generates 55 internal instruments. But when reducing the number of available lags for the generation of instruments in a common way, such that only lags three to seven are used, the same qualitative and statistically significant pattern for the sample of non-colonies is revealed. Additionally, neither the AR(2) test, nor the Hansen J test indicate evidence for further serial correlation or a rejection of the overidentifying restrictions.

lags. Table 3.8 presents comparable estimates with an alternative coding of colonies according to the CEPII data. With this classification, which delivers a full coding of all countries in the base sample, the results are qualitatively identical but quantitatively even stronger. Here also, the ten-year panel further strengthens our point, see Table 3.9.

3.4.2 Robustness

To test the robustness of our results with respect to alternative measures of democracy than those used by Acemoglu et al. (2008), we follow Benhabib, Corvalan, and Spiegel (2011) and employ the Vanhanen democracy index as well. This index has the advantage of factually not being bounded from above which reduces problems of censoring in the dependent variable.[14] The results are reported in Table 3.10 where Panel A presents results for the five-year and Panel B for the ten-year panel. The results in both panels are qualitatively identical. As before, income changes appear to improve the democracy score particularly for non-colonies. The effect is quantitatively much smaller and statistically weaker for the subsample of colonies compared to non-colonies. Note also that the positive effect vanishes completely for former colonies once dynamic panel estimation techniques are employed. In the ten-year panel, the income effect for former colonies is always statistically insignificant whereas it is also significantly positive on a 99% confidence level for non-colonies for both estimation techniques.

An alternative robustness check follows the analysis of Corvalan (2011) and uses a binary classification of democracy rather than a continuous or multivariate index. There is a debate in the political science literature about the appropriate measurement of democracy as a binary or multivariate concept, see Przeworski et al. (2000), Epstein et al. (2006), and Cheibub, Gandhi, and Vreeland (2010). We apply three different binary measures of democracy. Based on the benchmark sample, we code countries as democracies that score 0.6 or higher on either the Freedom House or Polity IV index, following common practice in the applied literature, see for example Persson and Tabellini (2006). As an alternative binary measure, we follow Corvalan (2011) and apply the binary measure proposed by

[14] For comparability, the index has also been normalized to lie between zero and one. By construction, there might still be censoring from below at zero, but the censoring is much less pronounced than in the Freedom House and Polity IV indices.

Cheibub, Gandhi, and Vreeland (2010). For comparability with the previous results, we estimate linear probability models with fixed effects as in specification 3.1. The estimation results are presented in Table 3.11. For each of the three binary democracy measures, we find no income effect in the full sample. However, once the sample is split into colonies and non-colonies we again find evidence for heterogeneous effects, with income if anything exerting a negative effect on the propensity of observing democracy in former colonies, while the income effect is positive in non-colonies. This pattern is consistent throughout all datasets.[15]

A last robustness check refers to Benhabib, Corvalan, and Spiegel (2011) who proposed the use of the Tobit estimator in the given context as a substantial mass of observations lies on the lower and upper bound of the democracy score. Thus, the estimation of the regression parameters might suffer from data censoring. Table 3.12 shows the corresponding results. Estimations in Panel A use a Tobit estimator for two-sided censoring. Here, only the partially interacted models are highly significant with the regression parameters having the expected signs whereas in the split samples the results are not significant. This changes however, when employing a one-sided Tobin estimator for censoring from above where most of the mass lies. In this case, the partially interacted models as well as the split samples reveal a positive and statistically significant effect of income on democracy, see Panel B.

3.4.3 Heterogeneity within the Sample of Colonies

While the comparison between colonies and non-colonies delivers pronounced variation in long-run development experiences, it might also involve some selection in terms of which countries were colonized and which were not, and typically were colonizers instead. This is not necessarily a problem in light of the hypothesis to be tested, namely whether the income effect might differ systematically across groups of countries that follow different political-economic development paths. It is nevertheless interesting to investigate the income effect on democracy when considering different groups of countries within the

[15] Additional estimates using conditional logit models exhibit qualitatively similar patterns and largely confirm the results from linear probability models. Also the use of ten-year panels or the coding of democracies at a score of 0.5 or higher delivers similar results, details are available upon request.

sample of colonies. Moreover, one might expect substantial heterogeneity among colonies depending on their colonial history. While essentially all colonizing powers implemented regimes of elite rule in which no popular participation in government was intended, these regimes differed in their repressiveness and thus in terms of how well the populations were prepared for democracy, see for example Diamond (1988) or Barro (1999).[16] Also, Acemoglu, Johnson, and Robinson (2001) show that economic development varies greatly within the sample of colonies and depends, in particular, on the early institutions that were implemented by the colonial powers which reflect different colonization strategies. Depending on whether countries were colonized for exploitation purposes only or were considered for large-scale settlements by the colonizing power subsequent development paths differed substantially.

Table 3.13 presents results for the same sample and coding as in Table 3.8, but now colonies are partitioned with regard to the respective colonizing power. In particular, columns (1) and (4) contain fixed effects and GMM estimates for the subsample of colonies of the main colonizing powers of the early 20th Century, namely Great Britain, France, Germany and Belgium. Columns (2) and (5) contain results for the other colonies, and columns (3) and (6) replicate the estimates for the subsample of non-colonies. The results are striking and strikingly different across subsamples. The income effect for the colonies that belonged to the major late colonial powers is significantly negative, whereas the income effect is statistically insignificant for the remaining colonies, and significantly positive for the non-colonies. The same qualitative pattern with statistically significant regression parameters is revealed when considering a ten-year panel instead, see Table 3.14. Having in mind that colonies that were in the hands of the late colonizers were held for exploitation rather than long-term settlements, this result is perfectly in line with the arguments put forward by Acemoglu et al. (2001).

Along similar lines, the year of independence is another useful measure to distinguish between colonies. Typically, colonies in which colonizers settled and implemented strong institutions became independent relatively early compared to colonies that were

[16]Brown (1999), p. 710, notes that "... many successor nations continued to be ruled in essentially colonial ways, enabling established leaderships to maintain a type of domestic neo-colonialism."

essentially set up for exploitation only, see Acemoglu et al. (2001 and 2009) for example. Moreover, there appears to be a consensus in the political science literature that the effect of colonial history is diminishing with time, i.e., with the years since independence, see Rueschemeyer, Huber Stephens, and Stephens (1992), Lipset, Seong, and Torres (1993), or Bernhard, Reenock, and Nordstrom (2004).[17] Table 3.15 presents the corresponding estimation results. The qualitative pattern is again similar: the income effect is positive, albeit insignificant, but the interaction between income and year of independence is negative, indicating that a later independence, which is likely associated with worse institutional quality from the colonial time, reduces the income effect on democracy in former colonies. Sample splits distinguishing between colonies that became independent before and after 1900, or before and after 1945 deliver a similar picture. The income effect is negative for countries that gained independence later. This holds for both democracy measures. In the ten-year panel the effect is even more pronounced, see Table 3.16.

Another proxy for different long-run development paths of colonies is the settler mortality variable that has been used by Acemoglu et al. (2001) as an instrument for differences in institutions across colonies. Table 3.17 reports the results from estimates within the sample of former colonies that account for differences in settler mortality. In Panel A, the dependent variable is the Freedom House democracy score, in Panel B it is the Polity IV data. Within each panel, columns 1 to 3 refer to the five-year panel, whereas columns 4 to 6 show the results for the ten-year panel. To investigate the presence of heterogeneity a binary variable is constructed which takes the value of one for countries whose settler mortality is within the 75th-percentile.[18] The partial interactions indicate substantial heterogeneity within the sub-sample of colonies for which the settler mortality data is available. Even though the estimates for the Freedom House data are insignificant, the general and consistent pattern is a more pronounced negative effect of income on democracy for countries with high settler mortality. With the Polity IV data,

[17] This view is also supported by recent findings of Bertocchi and Guerzoni (2010) who analyze the causes of state fragility in Sub-Saharan Africa.

[18] We employ a binary measure to circumvent the discussion whether to use settler mortality data in logs or levels. As noted by Acemoglu et al. (2001), p. 1383, there is no theoretical justification for preferring one over the other. Though a log specification mitigates the impact of outliers in the sample it postulates a non-linear relationship between the variables under consideration.

on the other hand, the results are also statistically significant, in particular the partial interactions.[19]

Even though the diminished sample size, due to availability of settler mortality data, reduces the precision of the obtained estimates, all results follow a consistent pattern which indicates that higher settler mortality, reflecting lower incentives to set up settler colonies with good institutions, exhibits lower or more negative income effects. Together with the results on the income effect conditional on the year of independence this provides additional pieces of evidence on the importance of institutions within colonies. This not only corroborates the argument of Acemoglu et al. (2001), p. 1383, but is also suggestive of different political-economic development paths.

3.5 Concluding Remarks

This paper revisits the relation between income and democracy and takes into account the crucial role of critical junctures in the history of countries. Focusing on colonization as one of the prime examples of critical junctures we find a hetereogenous effect of income on democracy which qualitatively depends on the nature of institutions set up by the respective colonizing power. Once controlling for the colonial past of a country we reveal a positive statistically significant effect of income on democracy for all non-colonies, and a statistically insignificant or negative one for former colonies, depending on the employed data or estimation technique. Our main result is obtained by using the same dataset as the previous literature despite controlling for country and time fixed effects. Moreover, it is robust against variations in the lags or measures of the democracy score, the use of common proxies for institutional quality like years of independence or settler mortality, and the employment of various estimation techniques.

Our results provide a leeway for reconciling certain findings in the literature. In general, income seems to be correlated with changes in political institutions even when considering variations within countries. But the nature of these institutional changes appears itself to be affected by the quality of institutions that were implemented at certain

[19]The results are qualitatively similar when continuous levels of settler mortality are used. Detailed results are available upon request.

historical junctures in the particular country; in the case of our study at times of colonization. However, as our results suggest, this influence of the institutional environment weakens over time. Thus, in independent countries a positive income effect on democracy might prevail in the (very) long run.

3.6 Appendix

Table 3.1: Classification of Countries (Acemoglu et al. (2008))

Non Colonies	*Colonies*	*Colonies (ctd.)*	*Not Classified*
Albania	Algeria	Jamaica	Armenia
Austria	Angola	Kenya	Azerbaijan
Belgium	Antigua	Lesotho	Belarus
Bulgaria	Argentina	Madagascar	Croatia
China	Australia	Malawi	Czech Republic
Cyprus	Bangladesh	Malaysia	Ethiopia 1993-
Denmark	Barbados	Mali	Kazakhstan
Estonia	Belize	Malta	Kyrgyz Republic
Ethiopia -pre 1993	Benin	Mauritania	Macedonia, FYR
Finland	Bolivia	Mauritius	Moldova
France	Botswana	Mexico	Namibia
Germany	Brazil	Morocco	Slovakia
Greece	Burkina Faso	Mozambique	Slovenia
Hungary	Burundi	New Zealand	Ukraine
Iceland	Cambodia	Nicaragua	Uzbekistan
Iran	Cameroon	Niger	Yemen
Ireland	Canada	Nigeria	
Israel	Cape Verde	Pakistan-post-1972	
Italy	Central African Republic	Pakistan-pre-1972	
Japan	Chad	Panama	
Jordan	Chile	Papua New Guinea	
Korea, Rep.	Colombia	Paraguay	
Latvia	Comoros	Peru	
Lebanon	Congo, Dem. Rep.	Philippines	
Lithuania	Congo, Rep.	Rwanda	
Luxembourg	Costa Rica	Sao Tome and Principe	
Nepal	Cote d'Ivoire	Senegal	
Netherlands	Cuba	Seychelles	
Norway	Dominica	Sierra Leone	
Poland	Dominican Republic	Singapore	
Portugal	Ecuador	South Africa	
Romania	Egypt, Arab Rep.	Sri Lanka	
Russia	El Salvador	St. Kitts and Nevis	
Spain	Equatorial Guinea	St. Lucia	
Sweden	Fiji	St. Vincent and the Grenadines	
Switzerland	Gabon	Tanzania	
Syrian Arab Republic	Gambia, The	Togo	
Taiwan	Ghana	Trinidad and Tobago	
Thailand	Grenada	Tunisia	
Turkey	Guatemala	Uganda	
United Kingdom	Guinea	United States	
	Guinea-Bissau	Uruguay	
	Guyana	Venezuela, RB	
	Haiti	Vietnam	
	Honduras	Zambia	
	India	Zimbabwe	
	Indonesia		
N = 41		N = 93	N = 16

Table 3.2: Classification of Countries (CEPII (2011))

Non Colonies	*Colonies*	*Colonies (ctd.)*	*Colonies (ctd.)*
Albania	Algeria	Guinea-Bissau	South Africa
Armenia	Angola	Guyana	Sri Lanka
Austria	Antigua	Honduras	St. Kitts and Nevis
Azerbaijan	Argentina	Iceland	St. Lucia
Belarus	Australia	India	St. Vincent and the Grenadines
Belgium	Bangladesh	Indonesia	Syrian Arab Republic
Bulgaria	Barbados	Ireland	Tanzania
China	Belize	Israel	Togo
Denmark	Benin	Jamaica	Trinidad and Tobago
Estonia	Bolivia	Jordan	Tunisia
Ethiopia -pre 1993	Botswana	Kenya	Uganda
Ethiopia 1993-	Brazil	Lebanon	United States
Finland	Burkina Faso	Lesotho	Uruguay
France	Burundi	Luxembourg	Venezuela, RB
Germany	Cambodia	Madagascar	Vietnam
Greece	Cameroon	Malawi	Yemen
Haiti	Canada	Malaysia	Zambia
Hungary	Cape Verde	Mali	Zimbabwe
Iran	Central African Republic	Malta	
Italy	Chad	Mauritania	
Japan	Chile	Mauritius	
Kazakhstan	Colombia	Mexico	
Korea, Rep.	Comoros	Morocco	
Kyrgyz Republic	Congo, Dem. Rep.	Mozambique	
Latvia	Congo, Rep.	Namibia	
Lithuania	Costa Rica	New Zealand	
Macedonia, FYR	Cote d'Ivoire	Nicaragua	
Moldova	Croatia	Niger	
Nepal	Cuba	Nigeria	
Netherlands	Cyprus	Pakistan-post-1972	
Norway	Czech Republic	Pakistan-pre-1972	
Portugal	Dominica	Panama	
Romania	Dominican Republic	Papua New Guinea	
Russia	Ecuador	Paraguay	
Slovakia	Egypt, Arab Rep.	Peru	
Spain	El Salvador	Philippines	
Sweden	Equatorial Guinea	Poland	
Switzerland	Fiji	Rwanda	
Taiwan	Gabon	Sao Tome and Principe	
Thailand	Gambia, The	Senegal	
Turkey	Ghana	Seychelles	
Ukraine	Grenada	Sierra Leone	
United Kingdom	Guatemala	Singapore	
Uzbekistan	Guinea	Slovenia	
$N = 44$		$N = 106$	

Table 3.3: Summary Statistics for 5- and 10-Year Panel (1960-2000)

	Panel A: Freedom House					
	5-Year Panel			**10-Year Panel**		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
$Democracy_t$	0.568	0.363	945	0.591	0.362	457
$Income_{t-1}$	8.164	1.018	945	8.16	0.989	457
	Panel B: Polity IV					
	5-Year Panel			**10-Year Panel**		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
$Democracy_t$	0.574	0.381	854	0.603	0.375	419
$Income_{t-1}$	8.145	1.023	854	8.125	1.004	419
	Panel C: Vanhanen (2000)					
	5-Year Panel			**10-Year Panel**		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
$Democracy_t$	0.292	0.293	946	0.307	0.295	463
$Income_{t-1}$	8.179	1.011	946	8.166	0.984	463
	Panel D: Cheibub, Gandhi, and Vreeland (2010)					
	5-Year Panel			**10-Year Panel**		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
$Democracy_t$	0.493	0.500	940	0.529	.500	461
$Income_{t-1}$	8.169	1.009	940	8.153	0.982	461

Table 3.4: Replication of Main Results from Acemoglu et al. (2008)

	Dependent Variable:	
	Freedom House	Polity IV
	(1)	(2)
$Democracy_{t-1}$	0.379***	0.449***
	[0.047]	[0.058]
$Income_{t-1}$	0.01	-0.006
	[0.032]	[0.036]
Observations	945	854
Adj. R^2	0.234	0.389
Sample	150 Countries	136 Countries
Estimator	Fixed Effects	Fixed Effects

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively. All regressions include country and time fixed effects. The results in column (1) replicate those in Acemoglu et al. (2008), Table 2 Col. (2); results in column (2) replicate those in Acemoglu et al. (2008), Table 3 Col. (2);

Table 3.5: Correlates of Country Fixed Effects

	Panel A						
	Dependent Variable: Fixed Effects from Freedom House						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Colony	-0.131***						
	[0.036]						
Year of Independence		-0.059***					-0.073*
		[0.014]					[0.039]
Log Settler Mortality			-0.082***				-0.071***
			[0.011]				[0.015]
Colony of GBR, FRA, BEL				-0.105***		-0.144***	0.005
				[0.032]		[0.038]	[0.054]
Colony of other					-0.031	-0.108**	
					[0.036]	[0.043]	
Constant	-0.105***	0.924***	0.144***	-0.162***	-0.200***	-0.123***	1.503**
	[0.031]	[0.263]	[0.054]	[0.023]	[0.020]	[0.030]	[0.712]
Observations	134	137	79	149	149	149	76
Adj. R^2	0.088	0.062	0.338	0.058	-0.003	0.089	0.33
	Panel B						
	Dependent Variable: Fixed Effects from Polity IV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Colony	-0.144***						
	[0.032]						
Year of Independence		-0.060***					-0.073**
		[0.014]					[0.035]
Log Settler Mortality			-0.071***				-0.058***
			[0.011]				[0.015]
Colony of GBR, FRA, BEL				-0.135***		-0.168***	-0.016
				[0.029]		[0.034]	[0.048]
Colony of other					-0.011	-0.096**	
					[0.033]	[0.039]	
Constant	-0.048*	0.979***	0.150***	-0.102***	-0.152***	-0.068**	1.504**
	[0.027]	[0.275]	[0.053]	[0.020]	[0.019]	[0.027]	[0.640]
Observations	120	123	76	135	135	135	73
Adj. R^2	0.143	0.087	0.312	0.123	-0.007	0.155	0.339

Robust standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively.
The fixed effects have been obtained from estimating specification 3.1.

Table 3.6: Acemoglu et al. (2008) Colonies with 5-Year Panel

Panel A

Dependent Variable: Freedom House

	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	0.365***	0.314***	0.588***	0.388***	0.649***
	[0.045]	[0.050]	[0.093]	[0.107]	[0.168]
$Income_{t-1}$	0.064*	-0.052	0.135***	-0.225*	0.132*
	[0.036]	[0.039]	[0.043]	[0.136]	[0.070]
$Income_{t-1}*$ $Colony$	-0.108***				
	[0.041]				
Observations	927	638	289	573	263
Adj. R^2	0.24	0.188	0.465		
Hansen J Test				0.101	0.954
AR(2) Test				0.546	0.37

Panel B

Dependent Variable: Polity IV

	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	0.444***	0.359***	0.683***	0.520***	0.654***
	[0.057]	[0.063]	[0.093]	[0.120]	[0.142]
$Income_{t-1}$	0.014	-0.054	0.121**	-0.340*	0.141
	[0.043]	[0.043]	[0.055]	[0.182]	[0.112]
$Income_{t-1}*$ $Colony$	-0.046				
	[0.048]				
Observations	837	566	271	500	246
Adj. R^2	0.389	0.346	0.586		
Hansen J Test				0.311	0.929
AR(2) Test				0.564	0.294

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively. All regressions include country and time fixed effects.

Table 3.7: Acemoglu et al. (2008) Colonies with 10-Year Panel

	Panel A				
	Dependent Variable: Freedom House				
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	-0.06	-0.136*	0.230*	0.146	0.379**
	[0.067]	[0.072]	[0.123]	[0.139]	[0.181]
$Income_{t-1}$	0.149**	-0.091	0.385***	-0.257	0.448**
	[0.066]	[0.056]	[0.100]	[0.285]	[0.208]
$Income_{t-1} * Colony$	-0.207***				
	[0.064]				
Observations	455	311	144	229	109
Adj. R^2	0.139	0.116	0.382		
Hansen J Test				0.02	0.118
AR(2) Test				0.685	0.266
	Panel B				
	Dependent Variable: Polity IV				
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	0.046	-0.088	0.347***	0.267	0.572***
	[0.075]	[0.081]	[0.112]	[0.179]	[0.160]
$Income_{t-1}$	0.044	-0.091	0.237*	-0.106	0.052
	[0.072]	[0.063]	[0.127]	[0.343]	[0.179]
$Income_{t-1} * Colony$	-0.101				
	[0.074]				
Observations	418	280	138	198	104
Adj. R^2	0.251	0.259	0.364		
Hansen J Test				0.02	0.195
AR(2) Test				0.832	0.82

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively. All regressions include country and time fixed effects.

Table 3.8: CEPII (2011) Colonies with 5-Year Panel

	Panel A				
	Dependent Variable: Freedom House				
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	0.363***	0.327***	0.568***	0.401***	0.620***
	[0.044]	[0.048]	[0.107]	[0.105]	[0.159]
$Income_{t-1}$	0.080**	-0.048	0.169***	-0.231**	0.203***
	[0.039]	[0.035]	[0.047]	[0.112]	[0.074]
$Income_{t-1} * Colony$	-0.118***				
	[0.042]				
Observations	945	706	239	631	207
Adj. R^2	0.242	0.199	0.46		
Hansen J Test				0.25	0.998
AR(2) Test				0.605	0.574
	Panel B				
	Dependent Variable: Polity IV				
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	0.439***	0.366***	0.659***	0.517***	0.533***
	[0.056]	[0.063]	[0.096]	[0.120]	[0.135]
$Income_{t-1}$	0.03	-0.063	0.129**	-0.273*	0.327***
	[0.044]	[0.040]	[0.059]	[0.140]	[0.117]
$Income_{t-1} * Colony$	-0.072				
	[0.048]				
Observations	854	619	235	545	202
Adj. R^2	0.391	0.345	0.578		
Hansen J Test				0.204	0.994
AR(2) Test				0.539	0.523

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively. All regressions include country and time fixed effects.

Table 3.9: CEPII (2011) Colonies with 10-Year Panel

	Panel A				
	Dependent Variable: Freedom House				
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	-0.069	-0.127*	0.212	0.118	0.369*
	[0.065]	[0.069]	[0.137]	[0.127]	[0.212]
$Income_{t-1}$	0.190***	-0.095*	0.490***	-0.294	0.443**
	[0.070]	[0.052]	[0.109]	[0.229]	[0.217]
$Income_{t-1} * Colony$	-0.247***				
	[0.067]				
Observations	457	344	113	252	86
Adj. R^2	0.153	0.121	0.46		
Hansen J Test				0.03	0.268
AR(2) Test				0.468	0.497
	Panel B				
	Dependent Variable: Polity IV				
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	0.033	-0.069	0.287**	0.223	0.501***
	[0.072]	[0.082]	[0.108]	[0.178]	[0.138]
$Income_{t-1}$	0.085	-0.114**	0.299**	-0.25	0.097
	[0.076]	[0.056]	[0.128]	[0.264]	[0.159]
$Income_{t-1} * Colony$	-0.170**				
	[0.077]				
Observations	419	306	113	217	85
Adj. R^2	0.263	0.261	0.375		
Hansen J Test				0.034	0.272
AR(2) Test				0.611	0.712

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively. All regressions include country and time fixed effects.

Table 3.10: Acemoglu et al. (2008) Colonies

	Panel A: 5-Year Panel				
	Dependent Variable: Vanhanen				
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	0.295***	0.234***	0.430***	0.328***	0.520***
	[0.050]	[0.063]	[0.071]	[0.104]	[0.105]
$Income_{t-1}$	0.094***	0.038**	0.141***	-0.039	0.299***
	[0.024]	[0.019]	[0.028]	[0.070]	[0.091]
$Income_{t-1} * Colony$	-0.054**				
	[0.026]				
Observations	940	645	295	576	274
Adj. R^2	0.363	0.303	0.466		
Hansen J Test				0.129	0.998
AR(2) Test				0.331	0.918
	Panel B: 10-Year Panel				
	Dependent Variable: Vanhanen				
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	GMM Colonies (4)	GMM Non Colonies (5)
$Democracy_{t-1}$	-0.077	-0.096	-0.011	0.101	0.073
	[0.058]	[0.067]	[0.114]	[0.134]	[0.158]
$Income_{t-1}$	0.163***	0.023	0.268***	0.337	0.511***
	[0.042]	[0.035]	[0.074]	[0.246]	[0.135]
$Income_{t-1} * Colony$	-0.128***				
	[0.046]				
Observations	468	318	150	227	116
Adj. R^2	0.33	0.245	0.447		
Hansen J Test				0.244	0.327
AR(2) Test				0.82	0.03

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively. All regressions include country and time fixed effects.

Table 3.11: Binary Democracy Measures with 5-Year Panel

	Dependent Variable:								
	Freedom House			**Polity IV**			**Cheibub, Gandhi, and Vreeland (2010)**		
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects Non Colonies (3)	Fixed Effects Full Sample (4)	Fixed Effects Colonies (5)	Fixed Effects Non Colonies (6)	Fixed Effects Full Sample (7)	Fixed Effects Colonies (8)	Fixed Effects Non Colonies (9)
$Democracy_{t-1}$	0.247***	0.192***	0.395***	0.404***	0.355***	0.561***	0.337***	0.271***	0.520***
	[0.051]	[0.055]	[0.099]	[0.053]	[0.061]	[0.101]	[0.055]	[0.063]	[0.089]
$Income_{t-1}$	0.014	-0.105	0.225**	-0.027	-0.07	0.136	0.04	-0.026	0.271**
	[0.060]	[0.072]	[0.105]	[0.054]	[0.067]	[0.087]	[0.059]	[0.067]	[0.104]
Observations	945	638	289	854	566	271	940	645	277
Adj. R^2	0.116	0.0998	0.275	0.299	0.277	0.418	0.241	0.204	0.439

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively. All regressions include country and time fixed effects.
The p-value in column 6 is 12.7%.

Table 3.12: Acemoglu et al. (2008) Colonies with 5-Year Panel

Panel A: Two-Sided Censoring

	Dependent Variable: Freedom House			**Dependent Variable: Polity IV**		
	Tobit Full Sample (1)	Tobit Colonies (2)	Tobit Non Colonies (3)	Tobit Full Sample (4)	Tobit Colonies (5)	Tobit Non Colonies (6)
$Democracy_{t-1}$	0.429***	0.372***	0.705***	0.410***	0.349***	0.648***
	[0.057]	[0.060]	[0.137]	[0.060]	[0.064]	[0.125]
$Income_{t-1}$	0.180***	-0.038	0.111	0.119**	-0.039	0.085
	[0.051]	[0.050]	[0.070]	[0.051]	[0.044]	[0.060]
$Income_{t-1}*$	-0.225***			-0.156**		
$Colony$	[0.060]			[0.062]		
Observations	927	638	289	837	566	271
Pseudo R^2	0.91	0.87	1.004	1.232	1.403	1.133

Panel B: One-Sided Censoring (From Above)

	Dependent Variable: Freedom House			**Dependent Variable: Polity IV**		
	Tobit Full Sample (1)	Tobit Colonies (2)	Tobit Non Colonies (3)	Tobit Full Sample (4)	Tobit Colonies (5)	Tobit Non Colonies (6)
$Democracy_{t-1}$	0.398***	0.344***	0.662***	0.408***	0.348***	0.645***
	[0.050]	[0.052]	[0.127]	[0.061]	[0.065]	[0.126]
$Income_{t-1}$	0.175***	-0.029	0.093*	0.115**	-0.039	0.101**
	[0.045]	[0.041]	[0.055]	[0.050]	[0.044]	[0.046]
$Income_{t-1}*$	-0.213***			-0.149**		
$Colony$	[0.054]			[0.061]		
Observations	927	638	289	837	566	271
Pseudo R^2	1.151	1.217	1.138	1.251	1.411	1.171

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively.
All regressions include country and time fixed effects.

Table 3.13: Partitioned Colony Sample with 5-Year Panel

Panel A

	Dependent Variable: Freedom House					
	Fixed Effects Colonies (UK, FRA, GER, BEL)	Fixed Effects Colonies (other)	Fixed Effects Non Colonies	GMM Colonies (UK, FRA, GER, BEL)	GMM Colonies (other)	GMM Non Colonies
	(1)	(2)	(3)	(4)	(5)	(6)
$Democracy_{t-1}$	0.329***	0.299***	0.568***	0.259**	0.387***	0.620***
	[0.068]	[0.059]	[0.107]	[0.105]	[0.092]	[0.159]
$Income_{t-1}$	-0.101**	0.076	0.169***	-0.297***	0.029	0.203***
	[0.043]	[0.070]	[0.047]	[0.115]	[0.177]	[0.074]
Observations	482	223	240	424	207	207
Adj. R^2	0.178	0.276	0.46			
Hansen J Test				0.424	0.995	0.998
AR(2) Test				0.211	0.593	0.574

Panel B

	Dependent Variable: Polity IV					
	Fixed Effects Colonies (UK, FRA, GER, BEL)	Fixed Effects Colonies (other)	Fixed Effects Non Colonies	GMM Colonies (UK, FRA, GER, BEL)	GMM Colonies (other)	GMM Non Colonies
	(1)	(2)	(3)	(4)	(5)	(6)
$Democracy_{t-1}$	0.327***	0.356***	0.624***	0.297**	0.472***	0.533***
	[0.090]	[0.095]	[0.102]	[0.141]	[0.120]	[0.133]
$Income_{t-1}$	-0.128***	0.112	0.117*	-0.389***	0.021	0.316***
	[0.036]	[0.085]	[0.064]	[0.142]	[0.118]	[0.121]
Observations	424	193	237	368	175	204
Adj. R^2	0.301	0.48	0.546			
Hansen J Test				0.603	1.000	0.996
AR(2) Test				0.810	0.520	0.510

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively. All regressions include country and time fixed effects. The colonies in column 1 and 4 were colonized by Great Britain, France, Germany or Belgium. The colonies in column 2 and 5 were colonized by Spain, Portugal, Netherlands, Austria or Denmark.

Table 3.14: Partitioned Colony Sample with 10-Year Panel

Panel A

Dependent Variable: Freedom House

	Fixed Effects Colonies (UK, FRA, GER, BEL)	Fixed Effects Colonies (other)	Fixed Effects Non Colonies	GMM Colonies (UK, FRA, GER, BEL)	GMM Colonies (other)	GMM Non Colonies
	(1)	(2)	(3)	(4)	(5)	(6)
$Democracy_{t-1}$	-0.124	-0.200*	0.212	0.036	0.016	0.369*
	[0.105]	[0.098]	[0.137]	[0.150]	[0.164]	[0.212]
$Income_{t-1}$	-0.139*	0.016	0.490***	-0.473**	0.263	0.443**
	[0.073]	[0.102]	[0.109]	[0.193]	[0.194]	[0.217]
Observations	229	115	113	165	87	86
Adj. R^2	0.0922	0.247	0.46			
Hansen J Test				0.919	0.256	0.268
AR(2) Test				0.379	0.892	0.497

Panel B

Dependent Variable: Polity IV

	Fixed Effects Colonies (UK, FRA, GER, BEL)	Fixed Effects Colonies (other)	Fixed Effects Non Colonies	GMM Colonies (UK, FRA, GER, BEL)	GMM Colonies (other)	GMM Non Colonies
	(1)	(2)	(3)	(4)	(5)	(6)
$Democracy_{t-1}$	-0.197	-0.07	0.287**	0.034	0.104	0.501***
	[0.148]	[0.101]	[0.108]	[0.244]	[0.162]	[0.138]
$Income_{t-1}$	-0.178***	0.047	0.299**	-0.544**	-0.003	0.097
	[0.049]	[0.156]	[0.128]	[0.233]	[0.248]	[0.159]
Observations	206	99	114	146	71	85
Adj. R^2	0.262	0.45	0.376			
Hansen J Test				0.323	0.170	0.272
AR(2) Test				0.417	0.434	0.712

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively. All regressions include country and time fixed effects. The colonies in column 1 and 4 were colonized by Great Britain, France, Germany or Belgium. The colonies in column 2 and 5 were colonized by Spain, Portugal, Netherlands, Austria or Denmark.

Table 3.15: Year of Independence with 5-Year Panel

Panel A

	Dependent Variable: Freedom House					
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects After 1900 (3)	Fixed Effects Before 1900 (4)	Fixed Effects After 1945 (5)	Fixed Effects Before 1945 (6)
$Democracy_{t-1}$	0.345***	0.308***	0.333***	0.278***	0.325***	0.282***
	[0.045]	[0.049]	[0.070]	[0.075]	[0.078]	[0.065]
$Income_{t-1}$	1.417**	1.807	-0.080*	0.019	-0.079*	0.003
	[0.577]	[1.105]	[0.043]	[0.096]	[0.044]	[0.085]
$Income_{t-1}*$	-0.001**	-0.001*				
$Year$	[0.000]	[0.001]				
Observations	842	638	472	166	424	214
Adj. R^2	0.246	0.192	0.214	0.218	0.203	0.205

Panel B

	Dependent Variable: Polity IV					
	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects After 1900 (3)	Fixed Effects Before 1900 (4)	Fixed Effects After 1945 (5)	Fixed Effects Before 1945 (6)
$Democracy_{t-1}$	0.399***	0.344***	0.310***	0.332**	0.289***	0.394***
	[0.058]	[0.064]	[0.085]	[0.122]	[0.091]	[0.104]
$Income_{t-1}$	1.275**	3.137**	-0.093**	0.076	-0.098**	0.052
	[0.607]	[1.510]	[0.040]	[0.155]	[0.041]	[0.136]
$Income_{t-1}*$	-0.001**	-0.002**				
$Year$	[0.000]	[0.001]				
Observations	757	566	411	155	363	203
Adj. R^2	0.387	0.358	0.312	0.474	0.31	0.456

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively.
All regressions include country and time fixed effects.

Table 3.16: Year of Independence with 10-Year Panel

Panel A

Dependent Variable: Freedom House

	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects After 1900 (3)	Fixed Effects Before 1900 (4)	Fixed Effects After 1945 (5)	Fixed Effects Before 1945 (6)
$Democracy_{t-1}$	-0.091	-0.147**	-0.138	-0.172	-0.129	-0.199**
	[0.067]	[0.071]	[0.112]	[0.106]	[0.125]	[0.090]
$Income_{t-1}$	1.954**	2.741**	-0.098	-0.14	-0.103	-0.116
	[0.829]	[1.356]	[0.071]	[0.124]	[0.073]	[0.113]
$Income_{t-1} * Year$	-0.001**	-0.001**				
	[0.000]	[0.001]				
Observations	404	311	222	89	196	115
Adj. R^2	0.151	0.126	0.12	0.221	0.118	0.199

Panel B

Dependent Variable: Polity IV

	Fixed Effects Full Sample (1)	Fixed Effects Colonies (2)	Fixed Effects After 1900 (3)	Fixed Effects Before 1900 (4)	Fixed Effects After 1945 (5)	Fixed Effects Before 1945 (6)
$Democracy_{t-1}$	-0.016	-0.113	-0.295**	-0.034	-0.352***	0.029
	[0.075]	[0.080]	[0.122]	[0.106]	[0.127]	[0.100]
$Income_{t-1}$	1.747*	4.757***	-0.121*	-0.086	-0.139**	-0.065
	[0.894]	[1.653]	[0.063]	[0.204]	[0.063]	[0.185]
$Income_{t-1} * Year$	-0.001*	-0.003***				
	[0.000]	[0.001]				
Observations	369	280	198	82	172	108
Adj. R^2	0.282	0.292	0.28	0.427	0.31	0.383

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively.
All regressions include country and time fixed effects.

Table 3.17: Settler Mortality

Panel A

Dependent Variable: Freedom House

	5-Year Panel			**10-Year Panel**		
	Fixed Effects Full Sample	Fixed Effects ≤ 75th Percentile	Fixed Effects > 75th Percentile	Fixed Effects Full Sample	Fixed Effects ≤ 75th Percentile	Fixed Effects > 75th Percentile
	(1)	(2)	(3)	(4)	(5)	(6)
$Democracy_{t-1}$	0.272***	0.280***	0.275	-0.207***	-0.166**	-0.424***
	[0.059]	[0.056]	[0.169]	[0.074]	[0.074]	[0.133]
$Income_{t-1}$	-0.047	-0.044	-0.213**	-0.103	-0.14	-0.108
	[0.051]	[0.056]	[0.072]	[0.077]	[0.085]	[0.15]
$Income_{t-1}$*	-0.183**			-0.042		
$Mortality$	[0.091]			[0.154]		
Observations	464	353	111	231	178	53
Adj. R^2	0.152	0.149	0.153	0.119	0.122	0.131

Panel B

Dependent Variable: Polity IV

	5-Year Panel			**10-Year Panel**		
	Fixed Effects Full Sample	Fixed Effects ≤ 75th Percentile	Fixed Effects > 75th Percentile	Fixed Effects Full Sample	Fixed Effects ≤ 75th Percentile	Fixed Effects > 75th Percentile
	(1)	(2)	(3)	(4)	(5)	(6)
$Democracy_{t-1}$	0.345***	0.354***	0.267*	-0.086	-0.045	-0.371**
	[0.072]	[0.082]	[0.15]	[0.088]	[0.095]	[0.157]
$Income_{t-1}$	-0.013	-0.04	-0.257***	-0.076	-0.176	-0.257**
	[0.065]	[0.077]	[0.072]	[0.094]	[0.115]	[0.088]
$Income_{t-1}$*	-0.274***			-0.19		
$Mortality$	[0.1]			[0.124]		
Observations	447	340	107	225	173	52
Adj. R^2	0.338	0.382	0.256	0.259	0.321	0.326

Clustered standard errors are in brackets. ***, **,* indicate significance at 1-, 5-, and 10-percent level, respectively.
All regressions include country and time fixed effects.

Bibliography

Acemoglu, D. (2002): "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, Vol. 40, No. 1, 7–72.

——— (2009): *Introduction to Modern Economic Growth*. Princeton University Press.

Acemoglu, D., G. Egorov, and K. Sonin (2008): "Coalition Formation in Non-Democracies," *Review of Economic Studies*, Vol. 75, No. 4, 987–1009.

Acemoglu, D., S. Johnson, and J. Robinson (2005): "Institutions as the Fundamental Cause of Long-Run Growth," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf, chap. 6. Elsevier, Amsterdam.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, Vol. 91, No. 5, 1369–1401.

——— (2002): "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, Vol. 177, No. 4, 1231–1294.

Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2008): "Income and Democracy," *American Economic Review*, Vol. 98, No. 3, 808–842.

——— (2009): "Reevaluating the Modernization Hypothesis," *Journal of Monetary Economics*, Vol. 56, 1043–1058.

Acemoglu, D., and J. A. Robinson (2000): "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective," *Quarterly Journal of Economics*, Vol. 115, No. 4, 1167–1199.

——— (2001): "A Theory of Political Transitions," *American Economic Review*, Vol. 91, No. 4, 938–963.

——— (2002): "The Political Economy of the Kuznets Curve," *Review of Development Economics*, Vol. 6, No. 2, 183–203.

——— (2006): *Economic Origins of Dictatorship and Democracy*. Cambridge University Press.

——— (2008): "Persistence of Power, Elites and Institutions," *American Economic Review*, Vol. 98, No. 1, 267–293.

——— (2010): "The Role of Institutions in Growth and Development," *Review of Economics and Institutions*, Vol. 1, No. 2, Article 1.

Alesina, A., and G.-M. Angeletos (2005): "Fairness and Redistribution," *American Economic Review*, Vol. 95, No. 4, 960–980.

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

Alesina, A., and E. L. Glaeser (2004): *Fighting Poverty in the US and Europe. A World of Difference*. Oxford University Press, Oxford.

Alesina, A., and D. Rodrik (1994): "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, Vol. 109, No. 2, 465–490.

Alesina, A., and E. Spolaore (1997): "On the Number and Size of Nations," *Quarterly Journal of Economics*, Vol. 112, No. 4, 1027–1056.

——— (2003): *The Size of Nations*. MIT Press.

Almond, G. A., and S. Verba (1963): *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton University Press.

Ansolabehere, S., J. M. Snyder, A. B. Strauss, and M. M. Ting (2005): "Voting Weights and Formateur Advantages in the Formation of Coalition Governments," *American Journal of Political Science*, Vol. 49, No. 3, 550–563.

Arat, Z. F. (1988): "Democracy and Economic Development: Modernization Theory Revisited," *Comparative Politics*, Vol. 21, No. 1, 21–36.

Aristotle (1943): *Politics*. Jowett, B., New York.

Bähr, J. (2004): *Bevölkerungsgeographie. Verteilung und Dynamik der Bevölkerung in globaler, nationaler und regionaler Sicht*. UTB Verlag, Stuttgart, 4 edn.

Balafoutas, L. (2011): "How Much Income Redistribution? An Explanation Based on Vote-buying and Corruption," *Public Choice*, Vol. 146, 185–203.

Barro, R. J. (1999): "Determinants of Democracy," *Journal of Political Economy*, Vol. 107, No. 6, 158–183.

Bassett, W. F., J. P. Burkett, and L. Putterman (1999): "Income Distribution, Government Transfers, and the Problem of Unequal Influence," *European Journal of Political Economy*, Vol. 15, 207–228.

Becker, G. S. (1983): "Theory of Competition Among Pressure Groups for Political Influence," *Quarterly Journal of Economics*, Vol. 98, No. 3, 371–400.

Bellettini, G., and C. Berti Ceroni (2007): "Income Distribution, Borrowing Constraints and Redistributive Policies," *European Economic Review*, Vol. 51, 625–645.

Benabou, R. (2000): "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, Vol. 90, No. 1, 96–129.

Benabou, R., and E. A. Ok (2001): "Social Mobility and the Demand for Redistribution: The POUM Hypothesis," *Quarterly Journal of Economics*, Vol. 166, No. 2, 447–487.

Benabou, R., and J. Tirole (2006): "Belief in a Just World and Redistributive Politics," *Quarterly Journal of Economics*, Vol. 121, No. 2, 699–746.

Benhabib, J., A. Corvalan, and M. M. Spiegel (2011): "Reestablishing the Income-Democracy Nexus," Federal Reserve Bank of San Francisco, Working Paper 2011-09.

Benhabib, J., and A. Przeworski (2006): "The Political Economy of Redistribution under Democracy," *Economic Theory*, Vol. 29, 271–190.

Bernhard, M., C. Reenock, and T. Nordstrom (2004): "The Legacy of Western Overseas Colonialism on Democratic Survival," *International Studies Quarterly*, Vol. 48, 225–250.

Bernheim, D., B. Peleg, and M. Whinston (1987): "Coalition-proof Nash equilibria: I Concepts," *Journal of Economic Theory*, Vol. 42, No. 1, 1–12.

Bertocchi, G., and A. Guerzoni (2010): "Growth, History, or Institutions? What Explains State Fragility in Sub-Saharan Africa," CEPR Discussion Paper No. 8406.

Bertola, G. (1993): "Factor Shares and Savings in Endogenous Growth," *American Economic Review*, Vol. 83, No. 5, 1184–1198.

Boix, C. (2003): *Democracy and Redistribution*. Cambridge University Press.

Boix, C., and S. C. Stokes (2003): "Endogenous Democratization," *World Politics*, Vol. 55, No. 4, 517–549.

Bollen, K. A., and R. W. Jackman (1985): "Political Democracy and the Size Distribution of Income," *American Sociological Review*, Vol. 50, 438–457.

Borck, R. (2007): "Voting, Inequality and Redistribution," *Journal of Economic Surveys*, Vol. 21, No. 1, 90–109.

Botzenhart, M. (1993): *Deutsche Verfassungsgeschichte 1806-1949*. Kohlhammer Verlag, Berlin.

Bresciani-Turroni, C. (1939): "Annual Survey of Statistical Data: Pareto's Law and the Index of Inequality of Incomes," *Econometrica*, Vol. 7, No. 2, 107–133.

Breyer, F., and H. W. Ursprung (1998): "Are the Rich Too Rich to Be Expropriated?: Economic Power and the Feasibility of Constitutional Limits to Redistribution," *Public Choice*, Vol. 94, No. 1, 135–156.

Brown, J. M. (1999): "Epilogue," in *The Oxford History of the British Empire*, ed. by J. M. Brown, and W. R. Louis, vol. IV, pp. 703–711. Oxford University Press, Oxford.

Bryce, J. (1921): *Modern Democracies*, vol. 1. Macmillan, New York.

Buchsteiner, I. (1993): *Grossgrundbesitz in Pommern*. Akademie Verlag, Berlin.

Bullock, A. (1964): *Hitler: A Study in Tyranny*. Harper & Row, New York.

Burkhart, R. E., and M. S. Lewis-Beck (1994): "Comparative Democracy: The Economic Development Thesis," *American Political Science Review*, Vol. 88, No. 4, 903–910.

Campante, F. R., and F. H. Ferreira (2007): "Inefficient Lobbying, Populism and Oligarchy," *Journal of Public Economics*, Vol. 91, 993–1021.

Cervellati, M., J. Esteban, and L. Kranich (2010): "Work Values, Endogenous Sentiments and Redistribution," *Journal of Public Economics*, Vol. 94, 612–627.

Cervellati, M., P. Fortunato, and U. Sunde (2008): "Hobbes to Rousseau: Inequality, Institutions, and Development," *Economic Journal*, Vol. 118, 1354–1384.

Cheibub, J. A., J. Gandhi, and J. R. Vreeland (2010): "Democracy and Dictatorship Revisited," *Public Choice*, Vol. 143, 67–101.

Cheibub, J. A., and J. R. Vreeland (2010): "Economic Development and Democratization," in *Dynamics of Democratization*, ed. by N. Brown. Johns Hopkins University Press.

Corneo, G., and H. P. Gruener (2002): "Individual Preferences for Political Redistribution," *Journal of Public Economics*, Vol. 83, 83–107.

Corvalan, A. (2011): "On the Effect of Income on Democracy during the Postwar Period," Mimeo, NYU.

Dahrendorff, R. (1965): *Gesellschaft und Demokratie in Deutschland*. Piper, München.

de Tocqueville (1864): *Democracy in America*. Harvard University Press.

Dederke, K.-H. (1996): *Reich und Republik: Deutschland 1917-1933*. Klett-Cotta, Stuttgart, 8th edn.

Diamond, L. (1988): "Introduction," in *Democracy in Developing Countries*, ed. by L. Diamond, J. Linz, and S. M. Lipset, vol. II. Lynne Rienner, Boulder.

——— (1992): "Economic Development and Democracy Reconsidered," in *Reexamining Democracy*, ed. by G. Marks, and L. Diamond. Sage Publications, London.

Dreze, J. H., and J. Greenberg (1980): "Hedonic Coalitions: Optimality and Stability," *Econometrica*, Vol. 48, No. 4, 987–1003.

Dumke, R. (1991): "Income Inequality and Industrialization in Germany, 1850-1913," in *Income Distribution in Historical Perspective*, ed. by Y. S. Brenner, H. Kaelble, and M. Thomas, pp. 117–149. Cambridge University Press.

Einy, E., and B. Peleg (1995): "Coalition-proof Communication Equilibria," in *Social Choice, Welfare, and Ethics*, ed. by W. A. Barnett, H. Moulin, M. Salles, and N. J. Schofield, chap. 13, pp. 289–301. Cambridge University Press.

Epple, D., and T. Romer (1991): "Mobility and Redistribution," *Journal of Political Economy*, Vol. 99, No. 4, 828–858.

Epstein, D. L., R. Bates, J. Goldstone, I. Kristensen, and S. O'Halloran (2006): "Democratic Transitions," *American Journal of Political Science*, Vol. 50, No. 3, 551–569.

Eschenburg, T. (1983): "Jahre der Besatzung: 1945-1949," in *Geschichte der Bundesrepublik Deutschland*, ed. by K.-D. Bracher. Vol. I. Brockhaus, Wiesbaden.

Esteban, J., and D. Ray (1994): "On the Measurement of Polarization," *Econometrica*, Vol. 62, 819–852.

——— (2006): "Inequality, Lobbying, and Resource Allocation," *American Economic Review*, Vol. 96, No. 1, 257–279.

——— (2008): "Polarization, Fractionalization and Conflict," *Journal of Peace Research*, Vol. 45, No. 2, 163–182.

Fayad, G., R. H. Bates, and A. Hoeffler (2011): "Income and Democracy: Lipset's Law Inverted," OxCarre Research Paper No. 61.

Frechette, G. R., J. H. Kagel, and M. Morelli (2005): "Gamson's Law versus Non-cooperative Bargaining Theory," *Games and Economic Behavior*, Vol. 51, 365–390.

Gamson, W. A. (1961): "A Theory of Coalition Formation," *American Sociological Review*, Vol. 26, No. 3, 373–382.

Georgiadis, A., and A. Manning (2007): "Spend It Like Beckham? Inequality and Redistribution in the UK, 1983-2004," CEP Discussion Paper No. 816, August.

Glaeser, E., R. LaPorta, F. Lopez-de Silanes, and A. Shleifer (2004): "Do Institutions Cause Growth?," *Journal of Economic Growth*, 9(3), 271–303.

Gouveia, M., and N. A. Masia (1998): "Does the Median Voter Model Explain the Size of Government?: Evidence from the States," *Public Choice*, Vol. 97, 159–177.

Gradstein, M. (2007): "Inequality, Democracy, and the Protection of Property Rights," *Economic Journal*, 117(516), 252–269.

Grossman, G. M., and E. Helpman (1996): "Electoral Competition and Special Interest Politics," *Review of Economic Studies*, Vol. 63, 265–286.

Gundlach, E., and M. Paldam (2009): "A Farewell to Critical Junctures: Sorting Out Long-run Causality of Income and Democracy," *European Journal of Political Economy*, 25(3), 340–354.

Harms, P., and S. Zink (2003): "Eating the Rich vs. Feeding the Poor: Borrowing Constraints and the Reluctance to Redistribute," *Public Choice*, Vol. 116, 351–366.

Hoffmann, W. G. (1965): *Das Wachstum der deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts*. Springer Verlag, Berlin.

Hohorst, G., J. Kocka, and G. A. Ritter (1975): *Sozialgeschichtliches Arbeitsbuch II, Materialien zur Statistik des Kaiserreichs 1870-1914*. C.H. Beck, München.

Husted, T. A., and L. W. Kenny (1997): "The Effect of the Expansion of the Voting Franchise on the Size of Government," *Journal of Political Economy*, Vol. 105, No. 1, 54–82.

Iversen, T., and D. Soskice (2006): "Electoral Institutions and the Politics of Coalitions: Why Some Democracies Redistribute More Than Others," *American Political Science Review*, Vol. 100, No. 2, 165–181.

Jeck, A. (1968): "The trends of Income Distribution in West Germany," in *The Distribution of National Income*, ed. by J. Marchal, and B. Ducros. Macmillan, London.

Jung, F., and U. Sunde (2011): "Inequality, Development, and the Stability of Democracy – Lipset and Three Critical Junctures in German History," CEPR Discussion Paper No. 8406.

Karabarbounis, L. (2011): "One Dollar, One Vote," *Economic Journal*, Vol. 121, 621–651.

Krusell, P., and J.-V. Rios-Rull (1999): "On the Size of U.S. Government: Political Economy in the Neoclassical Growth Model," *American Economic Review*, Vol. 89, No. 5, 1156–1181.

Kuznets, S. (1955): "Economic Growth and Income Inequality," *American Economic Review*, Vol. 45, No. 1, 1–28.

Lind, J. T. (2007): "Fractionalization and the Size of Government," *Journal of Public Economics*, Vol. 91, 51–76.

Lindert, P. H. (1996): "What Limits Social Spending?," *Explorations in Economic History*, Vol. 33, 1–34.

Lipset, S. M. (1959): "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," *American Political Science Review*, Vol. 53, No. 1, 69–105.

Lipset, S. M., K.-R. Seong, and J. C. Torres (1993): "A Comparative Analysis of the Social Requisites of Democracy," *International Social Science Journal*, Vol. 45, 155–175.

Lizzeri, A., and N. Persico (2004): "Why did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform'," *Quarterly Journal of Economics*, Vol. 119, No. 2, 707–765.

Llavador, H., and R. Oxoby (2005): "Partisan Competition, Growth and the Franchise," *Quarterly Journal of Economics*, Vol. 120, No. 3, 1155–1189.

Lorenz, M. O. (1905): "Methods of Measuring the Concentration of Wealth," *Publications of the American Statistical Association*, Vol. 9, No. 70, 209–219.

Luttmer, E. F. P., and M. Singhal (2011): "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy*, Vol. 3, No. 1, 157–179.

Maddison, A. (2003): *The World Economy: Historical Statistics*. Development Centre of the Organisation for Economic Cooperation and Development, Paris.

Marshall, M. G., and K. Jaggers (2011): *Polity IV Project: Political Regime Characteristics and Transitions, 1800-2010*. University of Maryland.

Mattozzi, A. (2010): "Policy Uncertainty, Electoral Securities, and Redistribution," *International Economic Review*, Vol. 51, No. 1, 45–71.

Meltzer, A. H., and S. F. Richard (1981): "A Rational Theory of the Size of Government," *The Journal of Political Economy*, Vol. 89, 914–927.

Milanovic, B. (2000): "The Median-Voter Hypothesis, Income Inequality, and Income Redistribution: An Empirircal Test with the Required Data," *European Journal of Political Economy*, Vol. 16, 367–410.

Mitchell, B. (2003): *International Historical Statistics: Europe 1750-2000*. Palgrave-Macmillan, New York.

Moene, K. O., and M. Wallerstein (2001): "Inequality, Social Insurance, and Redistribution," *American Political Science Review*, Vol. 95, No. 4, 859–874.

Montalvo, J. G., and M. Reynal-Querol (2005): "Ethnic Polarization, Potential Conflict and Civil War," *American Economic Review*, Vol. 95, 796–816.

Moreno, D., and J. Wooders (1996): "Coalition-Proof Equilibrium," *Games and Economic Behavior*, Vol. 17, 80–112.

Morrisson, C. (2000): "Historical Perspectives on Income Distribution: The Case of Europe," in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon, vol. 1. North-Holland, Amsterdam.

Olson, M. (1965): *The Logic of Collective Action*. Harvard University Press.

Pareto, V. (1896): *Cours d'economie politique*. F. Rouge, Lausanne.

Perotti, R. (1996): "Growth, Income Distribution, and Democracy: What the Data Say," *Journal of Economic Growth*, Vol. 1, 149–187.

Persson, T., and G. Tabellini (1994): "Is Inequality Harmful for Growth?," *American Economic Review*, Vol. 84, No. 3, 600–621.

——— (2003): *The Economic Effects of Constitutions*. MIT Press, Cambridge, MA.

——— (2006): "Democracy and Development: The Devil in the Details," *American Economic Review*, Vol. 96, No. 2, 319–324.

Petzina, D., W. Abelshauser, and A. Faust (1978): *Sozialgeschichtliches Arbeitsbuch III, Materialien zur Statistik des Deutschen Reiches 1914-1945*. C.H. Beck, München.

Piketty, T. (1995): "Social Mobility and Redistributive Politics," *Quarterly Jorunal of Economics*, Vol. 100, No. 3, 551–584.

Przeworski, A. (2006): "Self-enforcing Democracy," in *Handbook of Political Economy*, ed. by B. Weingast, and D. Wittman, chap. 5, pp. 312–329. Oxford University Press.

Przeworski, A., M. E. Alvarez, J. A. Cheibub, and F. Limongi (2000): *Democracy and Development. Political Institutions and Well-Being in the World, 1950-1990*. Cambridge University Press.

Przeworski, A., and F. Limongi (1997): "Modernization: Theories and Facts," *World Politics*, Vol. 49, 155–183.

Putnam, R. D. (1993): *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press.

Ramcharan, R. (2010): "Inequality and Redistribution: Evidence from U.S. Counties and States, 1890-1930," *Review of Economics and Statistics*, Vol. 92, Nr. 4, 729–744.

Reif, H. (1999): *Adel im 19. und 20. Jahrhundert*. R. Oldenbourg, München.

Rodriguez, F. C. (1999): "Does Distributional Skewness Lead to Redistribution? Evidence from the United States," *Economics and Politics*, Vol. 11, No. 2, 171–199.

——— (2004): "Inequality, Redistribution, and Rent-Seeking," *Economics and Politics*, Vol. 16, No. 3, 287–320.

Roemer, J. E. (1998): "Why the Poor Do Not Expropriate the Rich: An Old Argument in New Garb," *Journal of Public Economics*, Vol. 70, 399–424.

Rogowski, R., and D. C. MacRae (2008): "Inequality and Institutions: What Theory, History, and (Some) Data Tell Us," in *Democracy, Inequality and Representation*, ed. by P. Beramendi, and C. Anderson, chap. 11, pp. 354–387. Russell Sage Foundation.

Roodman, D. (2007): "A Note on the Theme of Too Many Instruments," *Center for Global Development*, Working Paper, No. 125, 1–38.

Rueschemeyer, D., E. Huber Stephens, and J. D. Stephens (1992): *Capitalist Development and Democracy*. University of Chicago Press, Chicago.

Saint Paul, G., and T. Verdier (1996): "Inequality, Redistribution and Growth: A Challenge to the Conventional Political Economy Approach," *European Economic Review*, Vol. 40, 719–728.

Schulz, G. (1987): *Zwischen Demokratie und Diktatur: Deutschland am Vorabend der Krise*. Vol. II. Walter de Gruyter, Berlin.

——— (1992): *Zwischen Demokratie und Diktatur: Von Brüning zu Hitler*. Vol. III. Walter de Gruyter, Berlin.

Shayo, M. (2009): "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution," *American Political Science Review*, Vol. 103, No. 2, 147–174.

Shelton, C. A. (2007): "The Size and Composition of Government Expenditure," *Journal of Public Economics*, Vol. 91, 2230–2260.

Traxler, C. (2009): "Voting over Taxes: The Case of Tax Evasion," *Public Choice*, Vol. 140, 43–58.

Vanhanen, T. (2000): "A New Dataset for Measuring Democracy, 1810-1998," *Journal of Peace Research*, Vol. 37, No. 2, 251–265.

Warwick, P. V., and J. N. Druckman (2001): "Portfolio Salience and the Proportionality of Payoffs in Coalition Governments," *British Journal of Political Science*, Vol. 31, No. 4, 627–649.

Winkler, H. A. (1993): *Weimar 1918-1933: Die Geschichte der ersten deutschen Demokratie*. C.H. Beck, München.

Zollitsch, W. (1999): "Die Erosion des traditionellen Konservativismus. Ländlicher Adel in Preussen zwischen Kaiserreich und Weimarer Republik:," in *Parteien im Wandel, Vom Kaiserreich zur Weimarer Republik*, ed. by D. Dowe, J. Kocka, and H. A. Winkler, pp. 161–183. Oldenbourg, München.

Curriculum Vitae

Born 25th April 1980 in Gießen, Germany.

Education

2007 - 2011	Ph.D. Studies, University of St. Gallen, Switzerland,
2011	Visiting Assistant in Research, Yale University, USA,
2002 - 2007	Diploma, Business Administration, University of Passau, Germany,
2001 - 2007	M.A., Political Science, University of Passau,
2004	Exchange Student, University College Dublin, Ireland.

Working Experience

2007 - 2011	Research Assistant, Institute of Economics, University of St. Gallen,
2006	Intern, Indo-German Chamber of Commerce, Bangalore, India,
2005	Intern, Siemens, Beijing, China.

Conference and Seminar Presentations

2011	Society of New Institutional Economics, Stanford University, USA,
	Macro Lunch Series, Yale University,
	American Economic Association, Denver, USA,
2010	World Congress of the Econometric Society, Shanghai, China,
	Swiss Society of Economics and Statistics, Fribourg, Switzerland,
	Canadian Economics Association, Quebec City, Canada,
	Brown Bag Series, University of St. Gallen.