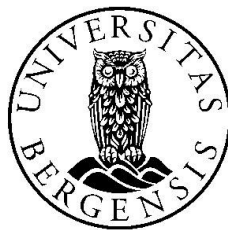


Hit ‘em where it hurts:  
Measuring and testing the impact of economic  
nonviolent strategies on democratization

Petter Kristiansen Arnesen



Masteroppgave

Vår/2019

Institutt for sammenliknende politikk

Universitetet i Bergen



## ABSTRACT

---

The literature on nonviolent political action has found that nonviolence far outpaces violence when it comes to winning political conflicts. Yet which actions nonviolent movements may perform to achieve success has rarely been studied. I argue that strategies which aim to limit the state's *economic capacity* are likely to be effective, and test whether such economic strategies are predictive of democratization. I build upon both recent and classic nonviolence- and democratization literature to craft a theoretical narrative of why I expect economic nonviolent strategies to be effective. I then construct a measurement model for economic strategies using a novel combination of the Nonviolent and Violent Campaigns and Outcomes 3.0 dataset and Bayesian item response theory methods. Using the resulting latent variable of economic strategies as an independent variable, I test whether it is predictive of transitions to democracy using Bayesian logistic regression. I find that nonviolent political campaigns that use economic strategies are significantly more likely to cause a transition to democracy than those which do not – A one standard deviation-change in economic strategy corresponds to a doubling in the odds of democratization. My findings are relevant to the nonviolence- and democratization literature as well as for practitioners of nonviolent action and fill an important research gap in an innovative way.



## ACKNOWLEDGMENTS

---

I want to thank my supervisors Michael Alvarez and Håvard Mogleiv Nygård for excellent guidance, motivation, understanding, and help with the project; My fellow master's students at UIB and PRIO, particularly the 7-Eleven gang, for their friendship, help, and support; Everyone at PRIO for making the past year fun and exciting, particularly Cathrine for her support and warmth, and Marianne for her understanding; To everyone at the Securing the Victory-project who let me write and work with them and who gave me vital feedback on the project; Xavier Fernández-i-Marín for the methodological inspiration and help; All my friends for their support and love; Ida for being a great friend and roommate; Silje for her unwavering love and support even in the darkest times; My family for their love and help, and especially Mom, for making me the person I am and who was always there for me.

This would not have been possible without all of you.

# Contents

1.	Introduction .....	1
1.1.	The Kapp Putsch and the general strike .....	2
1.2.	East German emigration .....	5
1.3.	Resilient Panamanian narco-militarism.....	6
1.4.	Investigating economic nonviolent action .....	8
2.	Literature review .....	10
2.1.	Nonviolent strategies .....	10
2.2.	Democratization.....	13
3.	Theory .....	16
3.1.	Nonviolently attacking sources of power .....	16
3.2.	Targeting states' material resources .....	18
3.3.	Hypotheses – Economic strategies and sources of state wealth .....	20
4.	Measuring economic nonviolent action .....	23
4.1.	Measurement theory .....	27
4.2.	Nonviolent action as contentious politics.....	33
4.3.	Sharp's methods of nonviolent action .....	36
5.	Methods.....	39
5.1.	Bayesian item response theory .....	39
5.2.	Bayesian logistic regression .....	45
6.	Data .....	47
6.1.	Data on nonviolent mobilization – NAVCO 3.0 .....	47
6.2.	Data-concept coherence.....	49
6.3.	Data on democratization – Dictatorship and Democracy index .....	50
6.4.	Dataset description .....	52
7.	Results and discussion.....	56
7.1.	Bayesian item response model .....	57
7.2.	Bayesian logistic regression .....	61
7.3.	Discussion.....	65

7.4.	MCMC model diagnostics .....	69
7.4.1.	IRT model: Item discrimination parameter alpha .....	71
7.4.2.	IRT model: Ability score parameter theta .....	74
7.4.3.	Logistic regression: Regression coefficients beta .....	78
7.5.	Alternative approaches .....	80
8.	Conclusion.....	82
9.	References .....	85
10.	Appendix .....	91
10.1.	List of indicators for IRT model .....	91
10.2.	Frequentist logistic regression .....	92
10.3.	Frequentist PCA analysis and logistic regression .....	92
10.4.	Ability scores for Kenya, Tanzania, and the US year.....	94





## List of figures

<b>Figure 1:</b> Causal graph of campaign strategy's effect on democratization .....	16
<b>Figure 2:</b> Conceptualization of nonviolent contentious politics.....	35
<b>Figure 3:</b> JAGS model for Bayesian IRT .....	43
<b>Figure 4:</b> JAGS model for Bayesian logistic regression .....	47
<b>Figure 5:</b> Sharp and NAVCO's conceptualization of nonviolent strategies .....	49
<b>Figure 6:</b> Highest posterior densities for alpha parameter.....	58
<b>Figure 7:</b> Highest posterior densities for theta parameter .....	59
<b>Figure 8:</b> Mean beta coefficients for covariates .....	62
<b>Figure 9:</b> Geweke and R-hat tests.....	70
<b>Figure 10:</b> Density- and traceplot for alpha parameters .....	72
<b>Figure 11:</b> Running means and autocorrelation plots for alpha parameters .....	73
<b>Figure 12:</b> Density- and partial comparison plots for theta parameters .....	75
<b>Figure 13:</b> Running means and autocorrelation plots for theta parameters.....	76
<b>Figure 14:</b> Traceplot for theta parameters .....	77
<b>Figure 15:</b> Geweke and R-hat tests.....	78
<b>Figure 16:</b> Density- and traceplot for beta parameters .....	79
<b>Figure 17:</b> Running means and autocorrelation plots for beta parameters.....	80

## List of tables

<b>Table 1:</b> Descriptive statistics of dataset .....	54
<b>Table 2:</b> Odds for covariates.....	63
<b>Table 3:</b> Transition years and standardized economic ability scores .....	65



# Hit 'em where it hurts:

Measuring and testing the impact of economic nonviolent strategies on democratization

“But a new leadership can and must be created by the masses and from the masses. The masses are the crucial factor. They are the rock on which the ultimate victory of the revolution will be built.”

(Rosa Luxemburg 1919)

“We live in capitalism, its power seems inescapable – but then, so did the divine right of kings. Any human power can be resisted and changed by human beings.”

(Ursula K. Le Guin 2014)

## 1. Introduction

Popular movements are viewed as one of the most important driving forces of democratization and political change. No political development or alteration is initiated, formulated, or implemented without humans' agency. This is very clear in routine politics, where the process is bureaucratic, hierarchic, and modeled to fit some efficient ideal. In contentious politics, however, the impact of peoples' actions on political developments are harder to observe and to measure. Yet the most important political changes spring from contention: democracy, capitalism, communism, empires, states, and nations all rise, fall, and change with contention. And contention is comprised of the strategic interaction of people who use their available resources, skills, and ideas to outsmart and defeat their opponents (Jasper 2006).

Recent research has found that the most effective grand strategy in contentious politics is nonviolence (e.g. Chenoweth, Perkoski, and Kang 2017; Chenoweth and Stephan 2011; Kudelia 2018; Nepstad 2011; Stephan and Chenoweth 2008). Chenoweth and Stephan (2011, 8-9) find that roughly 50 percent of nonviolent movements succeeded in the nineties and about 70 percent in the early 2000s. The success rates for violent campaigns were roughly 25 and 15 percent for violent campaigns in the same periods. The reasons for the efficiency of nonviolence are that nonviolence makes the movement appeal to a broad audience, it provokes defection in the security forces, and it makes it more difficult to repress the movements without causing a backlash-effect (Chenoweth, Perkoski, and Kang 2017; Chenoweth and Stephan 2011; Croissant, Kuehn, and Eschenauer 2018; Kudelia 2018; Nepstad 2011, 2015; Sharp 1973a, 2012; Stephan and Chenoweth 2008). The security forces as a possible tool of repression are central to the state's endurance. Because control over the police and the military, the monopoly

of legitimate violence, is one of the core definitional features of the state and its most severe sanctioning mechanism, it is an attractive target for popular campaigns. Without it, the state's available responses to dissent are circumscribed. The previous literature centers heavily around the Weberian state as the organization with a monopoly on legitimate violence (Weber 1958).

These findings seem to apply to the entire spectrum of nonviolent action and do not discriminate between vigils, strikes, blockades, speeches, or marches. They say nothing about which strategies of nonviolence achieve security force defections or backlash, or whether different actions produce different results. It is improbable that all types of nonviolent action are equally effective. Therefore, I investigate whether nonviolent movements that use *economic strategies* of contention are more likely to achieve democracy. Because targeting and limiting the violence-capacity of a state is an efficient way to win a political struggle, I believe the same should be true for its financial capacity. Threatening violence is more severe than most financial threats and may inspire repression from the state. But the economic capacity of states lays the foundation for every state activity, including violence. Money is a central incentive for recruitment to the security forces, and therefore necessary for its existence. Limiting the financial capacity of states can be an effective way of producing regime transitions. I argue that nonviolent movements can utilize *economic strategies* to limit the *capacity* of the state to repress them, analogous to directly targeting the monopoly of violence.

However, different strategies may produce unequal results in different contexts. Campaigns vary in size, framing, organizational structure, and class composition. The particularities of the target state are also bound to alter the campaign strategies' efficiency. Whether states are financed by taxation or loans, natural resource rents or foreign aid may require campaigns to adopt novel strategies. Not all revenue sources are equally susceptible to economic coercion by citizens. It is necessary, therefore, to account for macroeconomic differences between states when assessing how they might be challenged.

### 1.1. The Kapp Putsch and the general strike

On the 13<sup>th</sup> of March 1920, nationalist militants led by Wolfgang Kapp and Walther von Lüttwitz attempted to seize power to advance monarchist, conservative, and nationalist policies (see for example Feldman 1971; Sharp 1973b). While the political tumults of the Weimar Republic, unfair restrictions on German military capacity imposed by the Treaty of Versailles, and elephantine reparations owed by Germany provided a political space for opposition, it was the order to disband important *Freikorps* which ultimately prompted the putschists to act. This of course was a military takeover, and its tools were threats of violence. However, the elected

center-left Müller government, which was forced to flee from Berlin to Stuttgart as the putschists marched into the capital, called for a general strike on March 14<sup>th</sup>. Workers were quick to join in, and the strike spread throughout the country.<sup>1</sup> The new, self-declared Chancellor Kapp and his provisional government met much resistance in the bureaucracy, which refused to enforce policies – banks refused loan-orders with Kapp’s signature and ministers refused to resign or otherwise to implement changes. Even the staunchly anti-communist business class, holding great power in Weimar Germany and many of whom supported the organizations from which the Kapp Putsch arose, were critical of the coup (Feldman 1971, 101-103; Raloff 1971). The country was very quickly brought to a standstill by the general strike, and the putschists were unable to claim or exercise power because the bureaucracy did not cooperate (Raloff 1971; Sharp 1973b, 277). On March 17<sup>th</sup>, the putschists surrendered and were exiled, and the Müller government returned to Berlin (Feldman 1971).

The general strike illustrates several moments that are central to my research question. First, it demonstrates the possibility that would-be dictators can be stopped in executing their policies through mainly economic strategies. The general strike froze the German economy and bureaucracy, which in turn ensured that the putschists had no capacity to implement policy. Second, it demonstrates the role which consent plays in governance. Without the consent of the people, rulers will have a hard time implementing policy. The illegitimate and unpopular putschists had not established sufficient relations of loyalty with the German people, and unlike some successful coup-makers did not have the resources they needed to ensure compliance without consent.

However, this is an easy case where the target is illegitimate, unconsolidated, and unpopular. The putschists were opposed from day one, whereas other dictators have years to solidify their position and become a taken-for-granted part of life by their subjects. The exceedingly unpopular putschists were nothing of the sort, as even industrial capitalists lamented Kapp and his allies, as illustrated by one chemical industrialist:

One holds one’s head and asks whether men with brains and understanding or fools and lunatics have taken over the new leadership. As a businessman, therefore, I condemn what has happened thoroughly and completely, and I hope that the military hotheads in Berlin will soon come to their senses (Carl Duisberg, quoted in Feldman 1971, 102).

---

<sup>1</sup> No doubt the speed and enthusiasm was amplified by the *Arbeitsgemeinschaft* guaranteeing impunity for the striking workers, resolving that “[...] strike days up to and including Wednesday, March 17, were not to be counted against leave time, guaranteeing that workers would not be disciplined for participation in the general strike” (Feldman 1971, 111).

The situation in 1920's Germany is not common. Rarely are both bourgeoisie and proletariat so uniformly opposed to a regime or leader, and rarely are general strikes tolerated by capitalists.<sup>2</sup> Many coups-makers are more long-lived, however, such as the 1976 Argentine military coup installed a regime lasting for seven years. The 2013 military coup in Egypt saw general Fattah el-Sisi oust president Morsi, and el-Sisi was himself elected in 2014, still incumbent at the time of writing. With the recent change in the presidential term length, el-Sisi can potentially remain in office until 2030 (Michaelson and Youssef 2019). The speed with which the Kapp Putsch was thwarted is not a testament to an intrinsic weakness of coup-installed government, but to the importance of withdrawing consent and of popular, nonviolent action in contentious politics.

My argument that political conflict can be won by limiting the capacity of the state is supported by this. The failure of the putsch was a product of the general strike, which resulted in a paralyzed state with no capacity for policy implementation. The importance of the Weberian violence-monopoly was not very central here, as the only real source of power the putschists had was military power. But the general strike against the Kapp Putsch did more than limit growth or slow down production or infrastructure – it completely froze the political apparatus. Cases of nonviolent campaigns limiting the state's economic capacity alone should also exist, without bureaucratic support.

For similarly to the Weberian state-definition, Joseph Schumpeter argues that the modern state is characterized by its authority- and need to tax (1991[1919]). Historically, the “tax-state” evolved from rising war expenditures and insufficient credit supply, which led the medieval prince to seek revenue from taxation. Tilly (1985, 172) synthesizes the definitions of Weber and Schumpeter:

[...] the quest [for more effective war-making] inevitably involved them in establishing regular access to capitalists who could supply and arrange credit and in imposing one form of regular taxation or another on the people and activities within their spheres of control.

Both historically and conceptually then, the essential traits or institutions of the modern state is its ability to legitimately use violence on and extract taxes from its subjects. I juxtapose the two pillars of power: seeing that targeting and limiting the violence-capacity of a state is an efficient way to win a political struggle, I argue that the same should be true for its financial capacity. The threat of violence is more severe than most financial threats, but the financial capacity of a state underpins every state activity including violence. Money is a central motivation for anyone employed to enforce policy and states therefore need capital to form a security apparatus in the

---

<sup>2</sup> Duisberg and other capitalists blamed the strikes on the putschists, and Duisberg was satisfied to call the strikes “[...]not necessary because nothing can be achieved by them” (Feldman 1971), which is illustrative of the relative calm with which capitalists saw the strike.

first place. Limiting the financial capacity of states can plausibly be a fruitful tactic for nonviolent movements.

This is supported by Acemoglu and Robinson (2001), who argue that elites fear that democracy brings redistribution. Or, more elaborately, that elites who profit from dictatorship prefer to stay in power and therefore refrain from expanding the rights to political participation, redistributing their capital, and losing political influence to allowing popular campaigns to overthrow them. When popular campaigns can harm the economy and with it the income of the elite, the costs of staying in power rises. Democracy comes about as a compromise when the people threaten revolution and the rich want to remain powerful and wealthy. Similar points are made by among others Przeworski et al. (2000), O'Donnell and Schmitter (1986) and Skocpol (1979).

The Kapp Putsch was a weak case, unconsolidated and opposed by everyone. Are there stories of more stable regimes that have changed because of economic strategies?

## 1.2. East German emigration

No-one expected the Berlin Wall or the Soviet Union or its satellites to collapse in the late 1980s. They were considered very stable at the time, and the failure of political scientists to predict the collapse of the Eastern Bloc was not surprising (Nepstad 2011). Yet collapse they did. For the socialist German Democratic Republic (GDR, or East Germany), a contributing factor in the collapse was the mass emigration that occurred throughout its existence, in which mainly young, educated people fled the country to seek better fortunes in capitalist countries (Nepstad 2011, 44-45).<sup>3</sup> The economy suffered as important labor was absent and many goods, services, and technological development became unavailable. Industrial growth slowed down significantly over the years, and hospitals, factories, and bureaucracies were understaffed. The emigration, lack of human capital and resources, and public protests caused the fall of the Berlin wall in 1989, and the GDR was reunited with the capitalist West in 1990.

Being motivated perhaps primarily by fear and improvement of personal situations, the effects of migration on the GDR economy are unintended but not necessarily unknown or unwanted by the emigrants. It does however serve to illustrate the importance of economic capacity for states to execute their policies and stay in power. Economic downturns provide an opening for the opposition to manifest in the first place, because as grievances grow so does the urgency of changing politics. Furthermore, economic downturns limit the capacity of states to

---

<sup>3</sup> Some 5.275.000 people emigrated from the GDR from 1950 to 1993 according to a Council of Europe report (Kaya 2002).

counter the opposition – without supplying goods and services, taxation and other forms of domination become illegitimate and violence is left as the only tool for repression. The impact of economic strategies on states may therefore be two-pronged: it may increase the motivation to protest in the first place and limit the capacity of the state to resist popular demands.

As an additional factor in the collapse, the Gorbachev administration in the Soviet Union announced a weakening of ties between Moscow and its satellites (Nepstad 2011, 43-44). Both economic and military support was reduced, and the GDR could not count on Soviet support for repressing protest. This policy change was perhaps crucial to the socialist government's fall, as foreign economic support may have alleviated many of the concerns which harried it. This was the case in case in the GDR in 1953, and in Hungary in 1956 (Nepstad 2011, 43). A closer-knit Eastern bloc may have been able to eschew protests and prolong its existence.

While emigration from East Germany may not be immediately thought of as nonviolent action, I believe it is just that. A person physically removing themselves from a polity to oppose it costs the regime the person's entire contribution to the economic, political, and sometimes even the social life of the country. Hirschman (1970) discusses the choice of whether to exit a polity or voice one's discontent with a measure of spinelessness,<sup>4</sup> implying that the exiting a polity is indeed a type of political nonviolent action.

Different types of nonviolent action which target the state's economy and financial sanctioning power can be used to challenge the state and its policies. The efficiency of such action should depend on the size of the protest, the framing of protests, the movement's organization, and many other idiosyncrasies. It should also vary with the qualities of the state. States get funding for their activities from many sources: taxation, loans, state-owned businesses, or just straight-up printing money. Not all revenue sources are susceptible to economic coercion by citizens. Which state characteristics affect the use of economic nonviolence?

### 1.3. Resilient Panamanian narco-militarism

In the late 1980s, Panamanian military dictator Noriega was struggling with civilian protest, foreign withdrawal of support, a declining economy, high unemployment, and growing

---

<sup>4</sup> Hirschman (1970, 103-104) discusses exiting public goods-producing organizations such as states and argues that a measure of "spinelessness" can help explain why some stay or leave. As a public good turns into an evil, those with backbone leave and improve their lot elsewhere, while the spineless stay and suffer. If the stayers later grow spines however, because of a worsening of the situation for example, they may voice their dissatisfaction and change the evil-producing organization to the better. This last point is analogous to Sharp (1973a, 2012) and his removal of consent as essential for the onset of contentious nonviolence.



resentment from within his own forces (Nepstad 2011). Noriega's attempts to repress the opposition by violence, electoral fraud, and purging the military forces of disloyal members resulted in increased adversity in the population, worsening international relations, and withdrawal of investment and assets by international financial actors. Demonstrations were at one point some 750 000 strong,<sup>5</sup> and their leaders were inspired by the recent ouster of Philippine dictator Marcos. Despite these seemingly advantageous circumstances and significant turnout, Noriega only lost power when the US invaded and virtually crushed the Panamanian military resistance.

Nepstad (2011) argues that the resilience of the dictatorship is attributable to financial support from international allies such as Cuba, Taiwan, and Libya, and to noncooperation with the US from Japan, among others. Additionally, Noriega was allegedly part of the drug trade and an accomplice of the Medellín drug cartel, which provided both financial support and a threat of retribution should he resign and potentially rat them out to the US. The external assistance and support is exactly what was lacking in the case of the GDR.

Nygård (2015) argues a similar point: Interventionist international governmental organizations (IGOs) can solve commitment problems for the nondemocratic regime by guaranteeing that the opposition commits to certain policies. He finds that membership in interventionist IGOs decreases the likelihood of regime transitions (Nygård 2015, 423). IGOs can also sanction autocrats which repress the opposition. Dictators know this and can block such sanctions and deals through non-compliance and alleviate their impact by cooperating with sympathetic foreign leaders.

The Panamanian case illustrates that states have an advantage over the opposition by possessing unique tools capable of remedying the damage caused by nonviolent strategies. While Noriega's involvement in drug trafficking is a tactic rarely used by the state apparatus, material support from allied states is indeed used. With the right allies and counterstrategies, dictators can stay in power despite extreme opposition.

These three cases illustrate that states sometimes transition to democracy when faced with economic hardship, caused by structural factors or social movements. States are also capable of resisting such economic pressure even when it is great and have access to capital which is unavailable to non-state actors. States are thus not only vulnerable to defections in the security forces but also to the economic power of the people when used against them.

---

<sup>5</sup> The population of Panama per 1988 was roughly 2 300 000 (United Nations 2017). 750 000 is about 33 per cent of the population.

#### 1.4. Investigating economic nonviolent action

The question I pose is this: *Do campaigns using economic, nonviolent strategies affect the likelihood of a transition to democracy?* I thus attempt to fill a knowledge gap in the literature, as no quantitative analysis of which I am aware has tested this particular relationship. Attacks on state capacity by social movements and nonviolent political campaigns are rarely studied, and drops in state capacity are often seen as openings for potential mobilization rather than a consequence of political action (i.e. Skocpol 1979). This is relevant to the nonviolence literature which has produced convincing answers to how successful nonviolent strategies are vis-à-vis violent ones, but which has afforded little attention to the *actions* of nonviolent activists.

To fill this knowledge gap and provide convincing, novel results, I take advantage of the recent nonviolent literature in combination with much of the classic democratization literature to craft a theoretical narrative explaining *why* I expect economic nonviolent strategies to produce transitions to democracy. I construct a measurement model to identify the use of economic strategies in a set of political conflicts, which to my knowledge has not been done before. The freshly available Nonviolent and Violent Campaigns and Outcomes 3.0-dataset (NAVCO 3.0; Chenoweth, Pinckney, and Lewis 2018) provides a great starting point for this novel take and lets me answer my research question in a relevant way with new data.

I first review some key literature on nonviolence and democratization in chapter 2. The main point I make is that it should be possible to use nonviolent action to affect state economic capacity, with reference to for example Dahl (1971) and his cost of tolerance- and cost of repression concepts. The nonviolence literature has found a strong connection between nonviolence and security force defections, which exemplifies the causal utility of the Weberian definition of a state and what is important to its power.

In chapter 3 I supply a theoretical argument for why I expect economic nonviolent action to have an impact on the likelihood of transitions to democracy and under what conditions. I have already sketched the main points of the theory: Because revenue is important to a state's capacity, I expect economic strategies to be effective; and because states have relatively great economic power vis-à-vis popular campaigns, I expect the effect of economic strategies to vary by the state's sources of revenue. For example, it is unlikely that dictatorships rich on natural resources should be accommodating toward popular demands for regime change because the cost of losing power is very large. Those largely financed by taxation may be more amenable, as the population and their economic activities are important to their income. Additionally, states financed by foreign aid may be more or less susceptible to popular

demands, depending on the benefactor. Democratic financiers will threaten to withdraw if mobilization is answered by repression, while autocratic ones will defend their ally's rule. The above case-examples demonstrate the plausibility of these expectations and the mechanisms by which nonviolent campaigns interact with regimes.

In chapter 4, I argue in favor of using an item response theory (IRT) measurement model to estimate nonviolent campaign strategies' "economicness," or the degree to which they employ economically directed strategies. The best available data on strategies and tactics is the NAVCO 3.0 data (Chenoweth, Pinckney, and Lewis 2018), which records about 200 specific types of action in nominal variables. I separate the relevant indicators into dummy variables and use an IRT measurement model to construct a metric scale of economic strategy. This allows for metric scales and for agnostic weightings of the measurement's indicators, both of which are desirable traits for subsequent hypothesis testing (Schrodtt 2014). Furthermore, it is important to be mindful of the coherence between the concept which is measured and the quantification of that concept. This improves the validity, reliability, and credibility of the quantitative analysis and inference. I argue that measurement models should be used more often in political science and devote ample time to the necessary concept-measurement discussion which such efforts must entail.

In chapter 5, I explain the methods I use to construct the measurement model and to test my hypothesis that economic strategies increase the likelihood of democratization. I use a Bayesian IRT model to score the different nonviolent campaigns, and subsequently use this score as a predictor of democratization in a Bayesian logistic regression. IRT is similar to factor analysis but is applicable to data consisting of binary indicators rather than metric ones. The main advantage of Bayesian methods is that there is no assumption of repeated sampling. I argue that it is therefore appropriate for the data I use, and also that they help communicate the uncertainty inherent in the data.

In chapter 6, I describe the data with which I form the measurement model and the process of preparing the data for analysis. I use the data from the NAVCO 3.0 dataset (Chenoweth, Pinckney, and Lewis 2018) for strategy-data. For data on democratization, I use the Democracy and Dictatorship index (Cheibub, Gandhi, and Vreeland 2010). Both these sources I argue are state-of-the-art and conform well to my concepts and have the causal utility I need to answer my research question.

In chapter 7.1, I present the results from the Bayesian IRT model based on the indicators in the NAVCO 3.0 dataset. In chapter 7.2 I test the hypothesis that economic strategies are predictive of democratization. I report the results of my two analyses and discuss the application

of IRT for my particular purpose of measuring economic strategies, and the variation in the likelihood for democratization explained by economic strategies. I also discuss alternative specifications of my models to highlight strengths and weaknesses in my analysis and to suggest other, potentially fruitful approaches.

Lastly, in chapter 8, I discuss the implications of my findings for the literature on nonviolence, and for popular campaigns wanting to challenge their dictator with nonviolence. I conclude that my measurement model is fruitful and well-specified both theoretically and empirically, although different methods for constructing measurement models are also applicable. The approach of using measurement models to summarize data on strategy seems good, and I recommend that my approach is explored and developed further by others. I also find that economic strategies are robustly linked to the likelihood of democratization. Increasing the level of economicness in a movement's strategy by one standard deviation increases the odds of democratization in that country by roughly two, i.e. the likelihood doubles. This is a large effect, and it does not change its sign by removing any of the control variables, although it does vary somewhat in size. I argue that this is an important finding which is relevant for activists in addition to filling a knowledge gap in the nonviolence literature.

## 2. Literature review

To explain why I expect economic nonviolent strategies to impact democratization, it is first necessary to review some of the literature concerning both democratization and nonviolence.

### 2.1. Nonviolent strategies

Apart from the violent-nonviolent binary, little research has been done on the strategies of democratization campaigns. This lacuna in the literature is problematic, as it limits the ability of campaigns to draw on systematic and rigorous evidence to back up their efforts. The efforts of Sharp (2012) testify to the powerful impact such works can have, his book *From Dictatorship to Democracy* having been translated to over 30 languages and being a source of inspiration to activists in many countries.

Research on the violence-nonviolence dichotomy has produced diverging answers to the question of how campaigns ought to direct their resources toward success in bringing about democracy. Some find that violence sometimes works and that selective use of violence at the right time can be productive. The Ukrainian umbrella organization Right Sector during the 2014 Maidan protests exemplify this. They reinvigorated a mass protest which was losing momentum by using violence against police, ultimately succeeding in forcing the president to resign

(Kudelia 2018).

Others find positive effects of nonviolence. Stephan and Chenoweth (2008) find that nonviolent strategies are significantly more effective than violent ones, and attribute this to nonviolent campaigns' legitimacy, and the backlash effect that sometimes results from government repression of them. Karatnycky and Ackerman (2005) argue that strong, broad, nonviolent coalitions produce great democratic gains, contrary to smaller, narrower coalitions. Croissant, Kuehn, and Eschenauer (2018), Nepstad (2011), and Stephan and Chenoweth (2008) speak to the importance of the military, and that producing loyalty shifts on some decision-making level of the armed forces is essential to successful regime challenges. This is the main finding in the strategic nonviolence-literature and explains much of the difference in success rates between violent and nonviolent, and successful and unsuccessful nonviolent campaigns.

Testing correlations between regime type and strategy efficiency, Cunningham et al. (2017) find that given maximalist claims on governments, nonviolent campaigns are more likely to emerge in autocracies than in anocracies (i.e. semi-democratic regimes), and more likely in anocracies than democracies. The authors theorize that this is because grievances are high in non-democracies, which incentivizes maximalist claims-making. The incidence of nonviolent mobilization, furthermore, is positively influenced by the number of NGO chapters in the country, and nonviolent campaigns in neighboring countries (Cunningham et al. 2017, 478). Nonviolent campaigns seem to spring from fertile soil, where opportunities and resources are available to activists.

These findings all center around imposing costs on the opponent, and on using available channels and resources to gain leverage. Or, in other words, they concern minimizing the costs of organizing. Violence imposes direct, coercive costs on the opponent, and signals perseverance and commitment (Kudelia 2018, 503) – sometimes to an extreme extent, as with suicide bombings – raising the expected future costs as well (Pape 2005). Nonviolent campaigns employ strikes, blockades, marches and a large array of similar tactics (Sharp 1973b, 2012), some of which directly impose costs by coercion (Aitchison 2018; Klein and Regan 2018). Others focus on persuasion, such as making statements and communicating with potentially sympathetic segments of society, or deterrence, for example by signaling increased costs should the opponent choose to act in a way the campaign does not want (Aitchison 2018; Jasper 2006; Sharp 1973a).

What strategies and tactics impose the highest costs on the regime? Klein and Regan (2018, 489) argue that size, geographical dispersion, and the nature of the demands made by protestors increase the cost of repression. Larger protests obviously have higher disruptive

potential than smaller ones. One of the main weaknesses (or strengths) of nonviolence contra violence is that nonviolence requires comparatively many participants to be effective, whereas violence does not (Sharp 1973b, 110). This does not mean that all nonviolent protests have many participants – there is considerable variation in protest sizes.<sup>6</sup> Large, geographically dispersed protests combined can be very effective (Klein and Regan 2018), especially if they target important infrastructure and institutions. Concentration is not necessarily counterproductive, however, for instance in the cases of Egypt’s Tahrir Square protest<sup>7</sup> and Ukraine’s Maidan revolution (Kudelia 2018) which were arguably geographically concentrated and both resulted in the ousting of governments. Demanding regime change is probably more effective in the capital than in the countryside. Large protests with many participants further lower the risk that any single person will be arrested face personal consequences, which again lowers the cost of participation (Olson 1967).

Additionally, different types of demands carry different costs to the regime (Klein and Regan 2018). Demanding alleviation of grievances or minor policy changes imposes small costs while calling for regime change and democracy is antithetical to the core interests of the regime. Maximalist demands are therefore inarguably costlier to nondemocratic regimes than more limited demands like taxes and social rights (Klein and Regan 2018). This is the core idea of Dahl’s (1971) cost of toleration-concept, wherein the likelihood of democratization is negatively correlated with the cost of accommodating popular demands. Conversely, the cost of repression is positively correlated with the likelihood of democratization – and increasing repression-costs is what nonviolent contentious action is all about. Klein and Regan (2018) find that different patterns of state response to mobilization depend on the degree of concession costs and disruption costs, which are analogous to Dahl’s (1971) costs of toleration and -repression (see also Acemoglu and Robinson 2001; Przeworski 1991).

The finding of Cunningham et al. (2017) that nonviolent mobilization is more likely to emerge in non-democracies suggests that despite an increased risk of repression relative to democracies (Klein and Regan 2018), the potential reward of democratization outweighs the high costs of repression. This should, according to Olson (1967), scarcely happen. Individual costs are high when challenging repressive regimes, and the individual contribution to the achievement of the public good diminishes with the size of protests. Protestors have “[...] much

---

<sup>6</sup> The NAVCO 3.0 estimate of participants (variable *num\_partic\_event*) for protests (variable *verb\_10* == 14) ranges from 1 to 5.000.000 participants with a mean of 28790 and median of 500 participants.

<sup>7</sup> Both 2011 and 2012 have a mean of 0.94 on NAVCO 3.0’s “nv\_concentration” variable. Means for less protest-dense years have insufficient variation for comparison, as many years have less than 10 protest-events.

more to lose than their chains” (Elster 1988, 223), yet maximalist nonviolence is more prevalent in regimes whose core interest is maintaining a stable ruler-subject relationship (Cunningham et al. 2017; Elster 1988, 223). This might however be an artifact of democracy itself because even maximalist claims are allowed to compete in the political routine, at least in many systems (Cunningham et al. 2017, 471). Additionally, several solutions to the participation paradox have been suggested, ranging from ideological commitment and moral imperative to social pressures and large grievances (Cunningham et al. 2017; Olson 1967; Ostrom 1990; Sharp 1973a; Sweezy 1972; Whiteley and Seyd 2002; Yashar 2005). The issue of nonviolent campaigns’ emergence, while interesting, I leave to others.

These findings and theories center around the related concepts of removing the legitimacy of autocrats and subsequently imposing costs of repression so large that the regime cannot viably resist accommodating the opposition’s demands. The above historical examples demonstrate these mechanisms at work. In the Kapp Putsch-case, the costs of repressing the constitutionally elected government were sky-high, in that Germany and its bureaucracy ceased to function almost completely during the coup. This is perhaps the most extreme display of a regime’s illegitimacy of which I have heard. In East Germany, the toll of mass emigration contributed significantly to the decline of production and economic growth, and at the same time demonstrates the illegitimacy of the socialist government. This, along with the diffusion of protest from neighboring SSRs and satellites, helped bring about the seemingly super-stable GDR. In Panama, despite enormous protests and blatant illegitimacy, the Noriega government stayed in power by relying on material aid from other sources than its own population. And while drug trafficking is not a common business for states to get into, it serves to illustrate that alternative sources of funding can sustain even the most detested regimes for a while. Thus, the most severe threat which an opposition campaign can bring an autocrat is the removal of their main sources of power, namely their legitimacy, their security forces, and their revenue.

## 2.2. Democratization

Processes of democratizing nondemocratic regimes have historically happened through three idealized paths, according to Dahl (1971, 7). From being closed hegemonies with varying degrees of military, monarchic, religious, or imperial qualities, many states have increased the level of inclusion and competition in their political systems. With inclusion, these rights are extended to larger parts of the population than those already privileged, such as to the unpropertied, non-men, non-whites, and so on. Subsequent or concurrent processes of liberalization and inclusion constitute democratization and contribute to the legal and political

equality of citizens. With contestation come free and fair elections, the right to run for office, protection from electoral fraud, and the equal weighting of votes.

These processes have occurred either by expanding the franchise to larger parts of the citizenry before the competition, as in the Soviet systems, by first liberalizing competition as in the late nineteenth century Europe, or both simultaneously as in recently democratic states (Dahl 1971). Tilly, Tarrow, and McAdam (2003) tie democratization to citizenship with their concept *protected consultation*, which conforms to Dahl's definition but with additional emphasis on the rule of law. These rights are to be guaranteed to and protected for the citizenry, particularly minorities (Tilly, Tarrow, and McAdam 2003, 13-14). Przeworski (1991, 54-55) describes liberalization as the toleration of autonomous groups, particularly political ones.

Hegemons rarely embark on these paths voluntarily, however. Democratization makes sense for a dictatorial regime when the cost of repressing the population's demands rises above the cost of tolerating their participation in the political system (Dahl 1971; O'Donnell and Schmitter 1986; Przeworski 1991). From the American Revolution to the Arab Spring, the collapse of the Soviet Union to the end of Apartheid, processes of democratization are contested. When by some mechanism the cost of denying rights of participation and contestation to the populace are increased, the rewards of dictatorial power are comparatively smaller (Dahl 1971, 15-16). If opposed, a dictatorial regime can choose to accommodate their demands, repress them, or ignore them (Klein and Regan 2018). The choice depends on both the severity of the costs imposed by the opposition's activities as they try and achieve their demands and on the gravity of their demands, which Klein and Regan (2018, 490) call disruption costs and accommodation costs. The combination of these in a particular situation forces the regime and the opposition to compete for their preferred outcomes.

All contentious activity between regime and opposition do not end in regime transition, however. Great disruption costs coupled with non-threatening costs of accommodation, such as liberalization of access to abortion, may be accommodated by the regime. The cost of tolerance is small, while the cost of repression or dismissal can be great. Inversely, demands for leader resignation or regime change by fringe groups with little political clout or a broad support base are likely to be repressed (Klein and Regan 2018, 517-518).

Repression too may be inspired by the costliness of demands and disruption. Klein and Regan (2018, 508) find that increased disruption costs lead to increases in the likelihood for accommodating opposition demands on average and that increases in concession costs, for example from single-issue policies to maximalist demands of leader resignation, is likely met with repression.



The same logic is applied by O'Donnell and Schmitter (1986). They argue that the key to lasting democratization is to guarantee material or political security to the regime actors, for instance in the form of pacts offering seats in the national assembly for junta members, or guarantees against large-scale redistributive policies (O'Donnell and Schmitter 1986, 45-46, 52-54). By leveraging disagreements or “splits” in the regime, the opposition can negotiate pacts with regime soft-liners and exclude hardline authoritarians from the process of democratization. Substantive guarantees and limits on the legitimate domain of politics can be undemocratic themselves, but lower concession costs and help increase the likelihood of democratic transition (Klein and Regan 2018; O'Donnell and Schmitter 1986). The approach of O'Donnell and Schmitter (1986) is an actor-centered one, which is an important perspective when dealing with regime-altering contention and often bloody conflicts that threaten lives on all sides, a threat perhaps strongly felt by dictators losing control of their population.

Similarly, Przeworski (1991) posits that democratization begins with splits in the regime or with popular mobilization. A regime split between hard- and soft-liners precedes democratization when soft-liners believe that a limited opening of the political system will have low or even negative concession costs, such as producing a broader power-base to legitimize their continued rule in exchange for some increased protected consultation. Upon liberalizing, both regime and population assess their opponent's preferences and the probability of succeeding and strategize how best to affect the interaction to their favor.

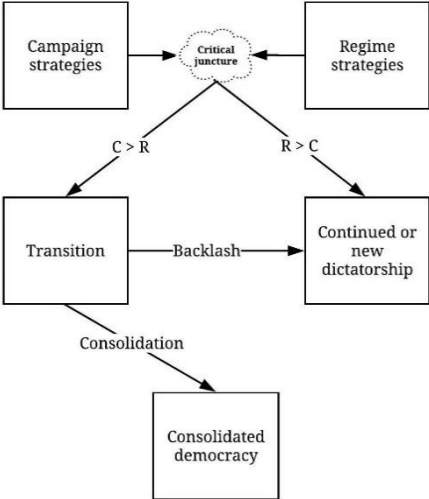
While Przeworski (1991, 60-66) uses a game theory framework and argues that no soft-liners will liberalize and no popular campaign will mobilize for transition unless their calculations about their opponent's perceived costs of oppression and tolerance are mistaken, he remarks that the presence of transitions prove that such miscalculations are common (1991, 60-61). Arguably then, assuming bounded rationality for both sides allows for opposition strategies to influence the course of a potential transition to democracy.

The goal of chapter 3 is to formulate hypotheses about the importance of campaign strategies as stylized in Figure 1. This I confine to hypotheses about the impact of movements' use or non-use of economic strategies. The research question of the thesis asks *whether using economic nonviolent strategies affect the likelihood of a transition to democracy*. Several other hypotheses can and should be derived from Figure 1, but I do not investigate the impact of economic strategies on backlash risk or democratic consolidation, or how the state acts to counter economic strategies.

There are probably many antecedent factors that affect campaign- and regime strategies. Structural conditions, the class composition of campaign and regime, ideologies, religion and

so on. I provide a link between such antecedent factors to the actual political contestation which takes place between regime and campaign and thus fill an important knowledge gap in the democratization literature.

**Figure 1:** Causal graph of campaign strategy’s effect on democratization



### 3. Theory

While the previous literature has identified some theoretical and empirical points – most concerning the Weberian state and the costs- and effects of popular mobilization – I flesh out a theoretical narrative for why I expect economic nonviolent strategies to cause transitions to democracy.

#### 3.1. Nonviolently attacking sources of power

A state’s ultimate source of power is its ability to impose sanctions by threat or use of violence (Nepstad 2011; Sharp 1973a). Attacking and effectively hampering this ability is an important predictor of success by nonviolent action and consequently a recurrent finding in the literature, usually by provoking defections from the security apparatus (e.g. Binnendjik and Marovic 2006; Chenoweth and Stephan 2011; Croissant, Kuehn, and Eschenauer 2018; Kudelia 2018; Nepstad 2011, 2015; Stephan and Chenoweth 2008). Yet while the monopoly of violence is the foremost source of state power, material resources may also be important strategic targets for activists. To be successful, a democratization campaign must disrupt the state’s core sources of power (Sharp 1973a, 2012). Provoking security force defections, recruiting huge numbers of protestors, or organizing general strikes are ways of damaging states’ repressive capabilities, legitimacy, and economic performance (Sharp 1973a, b, 2012).

Nepstad (2011) identifies the withholding of material resources as a possible strategy for civil resisters and finds an indeterminate relationship between the technique and success

using her comparative framework. East German democrats succeeded in utilizing material power (among other things) to oust the communist regime, while the Chilean overthrow of Pinochet and the Philippine “bloodless revolution” removing Marcos did not. Yet they ushered in democracy anyway. During unsuccessful campaigns in China, Panama, and Kenya, only the Chinese protestors failed to withhold material resources (Nepstad 2011, 127), yet all three campaigns failed to bring about democratization.

Similarly, Chenoweth and Stephan (2011) and Nygård (2015) argue that international sanctions, while potentially damaging to the civilian population as well as the regime, can help anti-regime campaigns. Chenoweth and Stephan (2011) show that international sanctions are more likely to be imposed in the presence of large and nonviolent campaigns than small and violent campaigns (Chenoweth and Stephan 2011, 52). Conversely, sponsorship and material aid from foreign states are more likely to go to violent than nonviolent campaigns. Nygård (2015) however, finds that membership in interventionist international governmental organizations (IGOs) decreases the likelihood of violent transitions, but that those autocracies with such memberships scarcely liberalize at all because they expect the IGO’s involvement. Both sanctions and state support are potentially good and bad for democratization campaigns. The South African campaign to end apartheid is an example where international boycotts, divestment, and sanctions helped put pressure on the regime. A similar strategy has only partially affected the Israeli occupation of Palestine, and the importance of foreign aid for the Noriega regime in Panama shows that sanctions can be effective if there are no ways of circumventing them.

Are nonviolent, economic strategies always positively related to democracy? Surely not. Nepstad (2011) and Chenoweth and Stephan (2011) present inconclusive results; older macro-theorists such as Moore (1966) argue that the bourgeoisie are essential for democratization, while Bernhard (2016) finds that this relationship does not hold after 1989; Wood (2008) argues that economic sanctions often produce unintended consequences for the people; Nygård (2015) finds that IGO membership is good for the transition process, but makes them unlikely in the first place. At the juncture of economics and politics, few things are certain or simple, but everything is important.

Moore’s theory and Bernhard’s findings suggest that *who the activists are* matters for the outcome. Dahlum (2018) finds that the coalition size and social segment of origin of nonviolent campaigns are positively related to democratization. In the pre-breakdown phase, broad coalitions are able to utilize diverse methods and resources thereby increasing disruption costs (Dahlum 2018, 6). Campaigns of working- and middle-class origins are more likely to

produce large coalitions, which suggests a relationship between economic class and democratization. These groups, along with peasants and the *lumpenproletariat*, are generally neither wealthy nor in control of direct influence over the state or state policy, which makes nonroutine collective action a potentially efficient tool for practicing contentious politics. When these classes are also the most likely democratizers, investigating their use of economic nonviolence may be interesting.

However, *what is done* by the opposition is potentially as important as who they are. And while the bourgeoisie and the proletariat possess very different motivations and resources it is the aggregate strategies and consequences of strategies that may bring a dictator to resign.

### 3.2. Targeting states' material resources

Different sources of material resources support the regime by different mechanisms. Sharp's theory of exercising power by withdrawing consent is focused on the vertical relationship between ruler and ruled (1973a), but horizontal support exists for both states and campaigns. Foreign states support or oppose each other by refraining from or engaging in trade, treaties, sanctions, and war, among other things. Interstate relations are not usually about domestic issues, because national self-determination is perhaps the key rule of the state game. The exceptions are however a relevant concern here: When states take a stand on others' domestic politics, what are the consequences? Limiting the inquiry to situations of nonviolent mobilization should reveal interesting relationships, such as whether foreign support for nonviolent movements makes repression of the movement more likely, and whether economic sanctions to support movements harm or help.

Nygård (2015) finds that regime transitions are less likely to occur when the state is a member of an interventionist international governmental organization (IGO, such as the International Monetary Fund (IMF), North Atlantic Treaty Organization (NATO), or the Organization of America States (OAS)). This is because IGOs can support or punish the new regime depending on whether the agreements of the transition are upheld. This is highly related to the protected consultation and rule of law-arguments of Tilly, Tarrow, and McAdam (2003) and Przeworski et al. (2000), and solves the commitment problem and uncertainty of the Przeworski (1991) game-theory model. With external support for domestic policy, a regime can be stabilized somewhat.

Popular campaigns often utilize international network and recruit other states to support their cause, too. State-movement relations are perhaps most obvious in territorial disputes, as in the cases of Palestine, Western Sahara, Transnistria, Kurdistan, and other would-be states.

Here, a major strategy for statehood-activists is to be acknowledged by other states, often requiring bottom-up mobilization in foreign countries to pressure governments. Movements also utilize foreign states' sanctioning- and political power, such as the Argentine campaign to oust the *Proceso* military dictatorship in the early 1980s (e.g. Brysk 1993), and the Boycott, Divest, Sanction-campaign (BDS), advocating against the Israeli occupation of Palestine. BDS also advocates for consumer and business boycotts of Israeli goods and lists country-specific products to avoid for activists. The campaign thus employs both vertical and horizontal transnational networks.

What accounts for the different legacies of economic nonviolent strategies? Surely, the autonomous economic capacity of a state is determined by the composition of its income. The relative importance of exports, aid, taxes, and resource wealth determines a state's ability to act independently of international and domestic actors.

Export-dependent states should be sensitive to external economic pressure to some degree, conditional on its market share and general demand. A monopolistic actor in a certain market may be in a stronger bargaining position than an atomistic one because importers have nowhere else to turn for the supplied good, which can make international boycott a less feasible method of resistance. The demand for a good may vary with the types of goods supplied – raw materials with utility for many industries such as oil or steel may be harder to boycott than luxury goods such as caviar or furs.

Boycotting industries on a large scale would however be damaging to the population and not just the regime's tax base (Afesorbor and Mahadevan 2016). Layoffs, bankruptcies, shortages, and wage cuts are some potential results of downturns in companies' fortunes, auxiliaries which may dissuade activists from advocating economic strategies (Allen and Lektzian 2012; Wood 2008). US trade-restrictions on Venezuela are good examples of strategies with large humanitarian costs, in which cutting off consumer-goods export and oil imports limits the availability of jobs and necessities to discredit and destabilize the regime (e.g. Meredith 2018; Rodríguez 2018). In capitalist economies with limited state control of ownership, the effects of sanctions on trade may be unpredictable and difficult to alleviate, and the effect on state capacity may only work indirectly through shrinking the tax base. Wood (2008) even finds a positive relationship between sanctions and state repression, suggesting that the instrument is a double-edged sword.

Aid- or loan-recipient states should be similarly sensitive to international economic action. Different types of aid and loans may relate differently to protests, however. Those types which are conditional on some policy development, such as liberalization of the economy or

efforts to curb corruption, may be withdrawn easily if rights are infringed upon or political skullduggery is discovered. This can plausibly happen alongside mobilization against the regime; uncovering corruption can provoke resistance, so too violation of rights. Conditionality may also prevent transgression in the first place, as is the intended purpose. Whether this happens is a different question, and the effectiveness of conditional aid and loans has produced a diverse literature (see Doucouliagos and Paldam 2009 for an overview). Additionally, Peksen and Woo (2018) find that economic sanctions, particularly human rights-related ones, or those imposed by the US, make participation in IMF loan programs less likely. This suggests a vicious cycle of stagnation for countries plagued by dictatorship, human rights issues, and insolvency, which would make external economic pressure an unstable tool for democrats to employ.

Directly targeting tax revenue is a rare method for activists to use, which is unsurprising. Refraining from paying taxes is always illegal and often easily trackable, especially in states with sophisticated bureaucracies and digitalized and automated tax-collecting procedures. It is however a powerfully symbolic method of resistance; refusing the government its revenue marks strong disapproval and harms state finances directly and carries direct financial incentives for activists (Sharp 1973b, 240-243). The method is perhaps most efficiently used by the wealthy, as they possess disproportionate amounts of money and can impose great costs on a government without organizing collectively. This does not mean that the working classes cannot utilize the method to their benefit.

General tax evasion and capital flight, which is mainly apolitical and motivated by maximizing profit, should perhaps be included because it may often be in response to changes in monetary- or fiscal policy. For example, the capital flight that occurred in Mitterrand's France in the early 1980s was a response to the Keynesian policies pursued by the elected government and led to a turn away from redistributive policies to inflation-control and privatization (e.g. Pisani-Ferry 2011, 24). The East German example is similar. Despite the political potency of capital, this thesis is limited to collective action and the effects of it on the likelihood for democratization. I do not explicitly include it in my empirical analysis.

### 3.3. Hypotheses – Economic strategies and sources of state wealth

Several factors should contribute to how well economic strategies work. First, I expect the impact of economic nonviolent strategies to differ with the state's relative dependence on tax revenue. If a protest is mainly domestic and the regime is tax dependent, then sizeable, economic mobilization should be able to hit state finances hard. Taxation is among the most visible and important evidence of states' dependence on the population, and this connection

makes demands for participation, rights, and redistribution just. Persson and Rothstein (2015) illustrate this with some qualitative interviews from Uganda, in which the respondents have few links to the state and little confidence in their ability to influence policy as a result. The effect of economic strategies will likely be contingent on the specific target institutions, the type of methods employed, and the goals of the campaign. In general, I expect that countries in which economic strategies are employed have a higher chance of becoming democratic than those in which economic strategies are not used.

The capacity of a state to impose taxation demonstrates its strength, however, and may also be a measure of citizen trust in the state, or of the state's ability to alleviate the concerns of citizens before they develop into contention. Hendrix (2010) argues that the tax-to-GDP rate is indicative but insufficient as a measure of state capacity – while it demonstrates the monetary capacity of a state, it does not differentiate between the bureaucratic apparatuses required to collect taxes. A rentier economy may have relatively high tax revenue but be unable to collect “difficult” income- or property taxes efficiently. Conversely, an economy that relies on difficult taxes will tend to have a greater capacity for monitoring and demanding payment from its citizens (Hendrix 2010, 278-279). It is perhaps more reasonable then to expect that economic strategies for democratization are likely to be effective in low-capacity states, where the regime is unable to monitor its population, but may still retain the ability to alleviate concerns with rent-wealth. I expect that states with a high tax-to-GDP ratio are less likely to be democratized by economic nonviolent strategies.

Second, I expect that states in which revenue from natural resources is comparatively high will not concede to maximalist demands. I rather expect that they will be more likely to repress maximalist demands and concede to minimalist ones. The findings of Wood (2008), that economic sanctions are related to state repression, are suggestive; arguing that when stability is threatened the state opts for repression, Wood's argument may travel to revenue-loss from taxes as well as sanctions. With resource-rich states, this effect may be strengthened because resource-income can pay for the increased cost of repression (monetary costs, that is – not the broader Dahlian term). However, it may also be that resource wealth is sensitive to strikes in the industry – miners, oil rig workers and other critical parts of the workforce may be accommodated in their demands, save for maximalist ones. In such cases, I would expect to observe repression of campaigns akin to that which Wood finds in the presence of sanctions. In short, when single, important sources of cash are disturbed, the state represses those who threaten its income. Whether that resource is trade or oil does not matter when it is being taken away abruptly.

Third, and already touched upon, I expect sanctions or discontinuations of foreign aid and loans to have a similarly diffuse relationship with the likelihood of a state to democratize. When trade and dependence on revenue from taxes on imports, exports, aid, and foreign capital are substantial, states have little bargaining power in an international conflict-situation. North Korea is an extreme example: the hostile international environment and its sanctions have contributed to a nuclear-militarist state ideology, which is leveraged by North Korea to increase international cooperation and improve the domestic situation. Another path chosen by many countries is conformity with neoliberal hegemonic rules-of-the-game, comparable to Mitterrand's turnaround, often implemented painfully fast but with positive effects on democracy scores and aggregate finances.

The theoretical uncertainty around whether economic strategies produce democracy necessitates a probabilistic and quantitative approach to illuminate relationships. The general research problem is clear, however: Are economic nonviolent strategies positively related to transitions to democracy? Based on the above theoretical discussion, I do expect this to be the case and to find a strong relationship contingent on the state's sources of revenue. I expect that nonviolent movements can use economic strategies to effect regime change, at least probabilistically.

This is because the economy is a core source of power for the state, and by effectively challenging it, a popular movement can gain influence over the state. I also expect this relationship to vary with state finance. States with a large economy are likely to be stable. So too those that receive aid from foreign states, and those with access to natural resource rents. Taxation too may be an expression of state capacity and resilience, but perhaps also of responsiveness. I expect higher taxation incomes to the state to be negatively linked to democratic transition. These hypotheses are listed below as a summary of my expectations.

- H1: Higher scores on economic nonviolent strategies are related to a higher likelihood of democratization in the same country-year.
- H2: States with a high tax-to-GDP ratio are less likely to be democratized by economic nonviolent strategies.
- H3: States with a high resource rent-to-GDP ratio are less likely to be democratized by economic nonviolent strategies.
- H4: Aid-financed states are less likely to be democratized by economic nonviolent strategies.
- H5: States with high GDPs are less likely to be democratized by economic nonviolent strategies.



## 4. Measuring economic nonviolent action

How should different nonviolent strategies be defined, conceptualized, and measured? What level of analysis is appropriate, which traits are important about different strategies, and which indicators can be used to measure them empirically? Several comparative-, case-, and statistical studies have addressed different aspects of nonviolent strategy and found interesting relationships between strategies, structural factors and outcomes (e.g. Ackerman 2007; Binnendjik and Marovic 2006; Chenoweth, Perkoski, and Kang 2017; Chenoweth and Uldfelder 2017; Kudelia 2018; Nepstad 2011; Stephan and Chenoweth 2008). In accordance with Goertz' second law (2006, 2), some of these authors have largely foregone explicit conceptualization in favor of spending their resources on measures, data collection, and hypothesis building and -testing. These studies are fruitful endeavors, as systematically collecting and analyzing data make generalizable investigations of social phenomena possible, from which much can be learned. It has however left the question of what constitutes a "strategy" undiscussed, which makes classification and fine-grained, quantitative analyses of nonviolent campaigns difficult. An overview of the different efforts to conceptualize nonviolent action is therefore necessary before I suggest a different approach.

Sharp (1973a, 65-67) distinguishes six classes of action in conflict situations. They are nonviolent action, persuasion, material destruction, physical violence against people, material destruction in combination with physical violence against people, and peaceful institutional procedures backed by threats and use of violence (i.e. ordinary politics in a state). These classes form the positive pole of the higher-order *conflict-action* concept, the negative pole being inaction (Goertz 2006, 31-32; Sharp 1973a, 64-65). An example of the same classification of a positive-pole concept is Lijphart's distinction between consensus-democracies and majoritarian democracies, both subtypes of democracy, both opposites of autocracy (Goertz 2006, 32; Lijphart 2012). This negative concept space is an important distinction, and it embeds Sharp's idea of whence power originates into the concept structure; for Sharp, change in the relations between regime and subject comes from the withdrawal of consent. When consent is present, you have no conflict with the regime and do not participate in conflict-action. The acknowledgment that states are not monolithic and that "Any human power can be resisted and changed by human beings" (Le Guin 2014) is the first step from consent to oppression and inaction toward action and change (Nepstad 2015, 4-7; Sharp 1973a, b, 2012). Nested in Sharp's concept of nonviolent action then, are mechanisms for its emergence, and the mechanisms by which change is produced in political systems.

Simultaneously, Sharp does not provide a clear-cut definition or an actual concept structure to form the basis of a measurement.<sup>8</sup> The closest thing to a definition of nonviolence Sharp gives is the following:

[...] nonviolent action is a technique by which people who reject passivity and submission, and who see struggle as essential, can wage their conflict without violence. Nonviolent action is not an attempt to avoid or ignore conflict. It is *one* response to the problem of how to *act* effectively in politics, especially how to wield power effectively (Sharp 1973a, 64, italics in original).

So, nonviolent action is 1) a technique for 2) nonviolently 3) wielding power effectively 4) when engaging in political struggle 5) for those who both do not consent to submission and 6) reject inactivity. Some clarification on these points is necessary.

First, and least problematic, is the characteristic of nonviolence as a *technique*. Nonviolence is simply a set of ways to engage in political conflict. It is an important characteristic because it separates nonviolent action from random but similar acts. A technique is consciously used and purposefully thought out to change something, while other forms of action may lack direction. For example, a reverse strike is different to working overtime because the former aims to change something other than the amount of wage-payment and work that is realized on a given work-day, namely changing power-relations in the workplace (Sharp 1973b, 402-403). By deviating from the expected pattern of behavior, namely going home at four o'clock, the workers display insubordination and organization, which alters the perceived power-relations in a negotiation or conflict.

Second, the “nonviolently”-part of the definition is murkier. Often, nonviolence is conflated with pacifism, which can indeed be a motivation behind nonviolence, but the two are unidentical. Nepstad (2015, 4-6) expounds upon the distinction divides nonviolence into a pacifist and a pragmatic type, the former describing the Gandhian school of moral nonviolence, the latter Sharp’s instrumental nonviolence. The main distinction between them according to Sharp (1973a, 82-83) is the emphasis on moral superiority and persuasion in the Gandhian strain; while pragmatists aim to out-manuever their opponents, pacifists aim to persuade them, a class of conflict-action which Sharp distinguishes from nonviolent action (1973a, 68). Furthermore, nonviolent action is not synonymous with pacifism, but may be motivated by practical, instrumental, ethical, religious or other grounds for preferring of nonviolence over other techniques (Gleditsch 1971; Nepstad 2015, 12-22; Sharp 1973a, 67-68, 70-71). These distinctions may have little influence on the performance of the actual methods of nonviolence but can supply different motivations for joining, supporting, or suppressing nonviolent

---

<sup>8</sup> Martin (2001, 23) states that “[...] nonviolence is easier to explain through examples than definitions or theory,” a statement accommodated with references to a list of case-studies on nonviolent action.

movements.

Third, nonviolent action is one of several ways to wield power effectively, or to cause change with the resources available. Everyone knows what power means, but the term is vacuous in analytical terms (Jasper 2006, 9). What is usually meant by power is a combination of resources and strategies, that is, directing available skills and means to get someone to act the way you want. The Weberian definition of power as being able to do something despite opposition or make someone do something they would not otherwise do thus grasps at the close relationship between the terms power and strategy (Jasper 2006, 9). Effective wielding of power must mean a successful, strategic use of the available skills and resources.

Sharp (1973a, 37) captures this by dividing the *sources* of power into authority, human resources, skills and knowledge, intangible factors (culture, ideology, predisposition to obedience, etc.), material resources, and sanctions. Outcomes of strategic social interactions depend on the amount of the different resources available to the actors and how they direct them to persuade, coerce, or deter the other toward some course of action (Jasper 2004, 2006; Sharp 1973a, 37). This is what “wielding power effectively” means. Importantly, it distinguishes nonviolent action from routine politics, which do not require wielding power at all but simply confirming the legitimacy of those who *do* wield it. In the context of attempting regime change, exercising power means heightening the cost of repression and lowering the cost of tolerance (Dahl 1971, 16). Previous findings on the efficiency of nonviolence suggest that this is indeed the case, particularly by undermining states’ repressive capabilities by provoking loyalty shifts in the security forces (Chenoweth and Stephan 2011, 46-49; Nepstad 2011, 128-131; 2015; Stephan and Chenoweth 2008).

Fourth, nonviolence is a way of engaging in conflict. When a disagreement over rights, duties, distribution, or priorities cannot be solved through routine political procedures, then nonroutine techniques of engaging with an opponent become viable alternatives (Sharp 1973a, 64). This too is an uncontroversial point, but research into how and when nonviolent campaigns emerge may pay attention to sequences of claims-making, claim salience, and escalation. There are arguments to be made about inefficient governance causing frustration and subsequent nonroutine mobilization, but it is not central to this thesis.

Fifth, rejecting submission is essential to Sharp’s theory of nonviolent power. By realizing that obedience is unnecessary, withdrawing consent, and actively challenging unjust rule can oppression be eliminated and a fairer regime be constructed (Sharp 1973a, 2012). This is relatively easy in a dictatorial context – oppression is enacted by the dictator and suffered by the population. The target is clear and identifiable, and maximalist goals are easily formulated

as a desire to depose the autocrat. Martin (2001, 37-39) points out that this is often a simplification of suppression, and that Sharp's theory has a harder time of dealing with more complex dominance, such as that of workers and consumer-activists under oligopolistic capitalism, complicated, multi-layered bureaucracies, or more nebulous phenomena such as patriarchy and racism. It might not be clear who or which institutions are exerting dominance, especially in systems of exchange, and the technique of withdrawing consent is harder to leverage for change. Opposition to states and dictators is a most likely-case for observing successful nonviolent mobilization against a system of dominance (Cunningham et al. 2017).

Sixth, and perhaps most importantly, the rejection of inactivity and passivity is not equal to the withdrawal of consent. In fact, grudging cooperation is functionally equivalent to any other form of consent to oppression. How one feels or thinks about injustice is of zero consequence to a dictator unless it is acted upon.<sup>9</sup> Changing one's behavior by refusing expected tasks or performing unwanted ones is the primary weapon of any resistance campaign. Inaction is the conceptual negative of Sharp's action-classes. What motivates the shift from passive to active withdrawal of consent is not a question for this thesis but has produced a large and diverse literature, particularly the attention paid to the collective action problem (Olson 1967). What motivates the choice of *nonviolent* action is a different question but probably based on a pragmatic assessment; activists simply believe it will work. Structural factors may play into this assessment, such as the degree of political space (e.g. regime-type effects on the emergence of nonviolence, see Cunningham et al. 2017).

These six definitional criteria are all essential to the concept of nonviolent action, and the definition hints at a classical necessary-and-sufficient concept structure (Gerring 2012b; Goertz 2006). The concrete observation of a nonviolent campaign is then relatively unproblematic: a group of people actively utilizing nonroutine, nonviolent techniques of influencing an actor in power toward some course of action satisfies all the criteria.

There can however be no pretense that all campaigns utilizing nonviolent action plan, utilize, and succeed in identical ways. Nonviolence is varied; different goals, opponents, and contexts shape how successful a campaign is (Ackerman 2007; Ackerman and DuVall 2000; Chenoweth and Stephan 2011; Martin 2001; Nepstad 2011; Stephan and Chenoweth 2008). Efficient power-wielding depends on the power-sources propping up an opponent. How this

---

<sup>9</sup> The popular quote by Desmond Tutu is illustrative: "If you are neutral in a situation of injustice, you have chosen the side of the oppressor. If an elephant has his foot on the tail of the mouse, and you say you are neutral, the mouse will not appreciate your neutrality." (Quoted in Brown (1984, 19)).

definition may best be operationalized and measured is therefore an important issue in this thesis, to which I turn next.

#### 4.1. Measurement theory

Properly measuring social phenomena such as nonviolent strategies is what makes political science a *science*; it is “disciplined summary” of social phenomena (Hooghe et al. 2016, 3) or the act of “establishing a metric of equivalence so that diverse observations can be [...] compared” (Gerring 2012a, 726). Measurement links observation to theory by systematically ordering and analyzing observations. Rigorous, falsifiable, and replicable empirical treatments of concepts and their interrelations permit advancement of the communal understanding of human interaction, which requires some form of measurement, usually in the form of numbering observations for statistical testing (Gerring 2012b, 156; Sartori 1970). Importantly, this does not mean that investigations of unnumbered phenomena, description, or idiosyncratic events are unscientific on their own (Gerring 2012a); but proper description makes theory, classification, and conceptualization possible, and measurement links description to analysis (Gerring 2012b, 156).

Necessarily for valid measurements, “concept formation stands prior to quantification” (Sartori 1970, 1038). To assign numbers to something, it is essential to have a clear idea of what is being described by the numbers, and what the numbers mean in relation to the concept at hand (Adcock and Collier 2001). Sartori (1970, 1034) warns against the “unconscious thinker” whom he argues substitutes complex techniques of analysis for conceptual logic and measurement. Ignoring the ontology of concepts, what a thing *is*, diminishes the chance of measuring it correctly and learning something meaningful about it. These are not empty admonitions. Improperly specified regressions, under-theorized proxies, confounding variables, garbage-can models, and bad aggregations potentially affect the results of causal analyses tremendously (Achen 2005; Jackman 2008; Schrodt 2014). Indeed, these issues can hurt the advancement of quantitative social science and make it difficult to see what is gained from doing it (Johnson 2006). Consequently, great care and thought must be put into creating valid and reliable measures of social phenomena.

Adcock and Collier (2001) have a more iterative view of the joint processes of concept formation and measurement than Sartori, although they discuss much of the same issues. They argue that while clear concept and properly specified definitions are important for measurement, so too measurement and indicators are important for refining concepts and definitions (Adcock and Collier 2001, 531). An evaluation of measurement validity cannot be

detached from the conceptual definition, nor can conceptual discussions be separate from empirical observations. Arguing for a shared framework between the quantitative and qualitative camps, they propose a continual circle of evaluating concepts, definitions, indicators, and measurements. Their process is roughly equivalent to the steps I take in defining, operationalizing and measuring economic nonviolent strategy.

I have explained Sharp's definition of nonviolent action above. The definition is operationalizable and touches on the essential traits of the concept and guides the process of forming a measure of *economic* nonviolent action. Importantly, having a clear definition makes the relationship I want to investigate a bit plainer: by actively using the economic nonviolent technique to reject submission and passivity wield power over the state, political campaigns can force regime transitions. However, the task of measuring and testing this relationship quantitatively remains.

Obviously, some concepts lack true values because they are directly unobservable. A strategy is such a concept and numbers attached to it mean little without a comparative reference. Inferring meaningful numbers based on observable indicators is still possible and often the best option available. An illustrative example of such a latent concept is social class, which can be thought to characterize a person's societal status. To measure it, some combination of indicators of wealth and income, work and leisure, cultural habits, education and so on must be decided upon and weighted, the result of which can be said to measure social class. Using a proxy variable such as income to equal the latent social class-concept may cause problems in subsequent analyses (Jackman 2008, 126-129), for instance, if investigating whether social class affects voting patterns. A plausible finding is that higher class causes one to vote conservatively. Using a composite measure of the above potential indicators of social class might however illuminate other relationships and suggest that medium-level income combined with high education makes one likely to vote center-left, while low-income, low-education people vote for populist right or socialist parties. The results may furthermore be attenuated by using only a proxy variable in place of a composite measure because the measurement error present in the proxy weakens the estimated coefficient, causing misleading inferences (Jackman 2008, 126-129).

Multifaceted concepts make measurement complicated. Identifying what the different parts of a concept are and which of them are most important is almost always debatable, exemplified by the ceaseless debate over what "democracy" is (e.g. Bollen 1979; Cheibub, Gandhi, and Vreeland 2010; Ghandi 2015; Munck and Verkuilen 2002; Przeworski et al. 2000; Treier and Jackman 2008). The pragmatist answer to these debates is that a diverse environment

of concepts and operationalizations facilitates multiple approaches and refinements of hypotheses and theory. Investigating relationships between phenomena using different conceptualizations of the variables can reinforce findings that are common across approaches and help identify which definitional traits are central to different approaches. Acemoglu and Robinson (2012) for example, find that inclusive institutional characteristics are conducive to growth, while others find that democracy is (Diamond 1992; Lipset 1959; Przeworski et al. 2000; Teorell 2010). Maximalist democracy-definitions can be hard to disentangle from “inclusive institutions,” so using a minimalist definition in addition to test the same hypothesis could help identify which democracy-characteristics drive the relationship.

Working with novel concepts, identifying which indicators are or can be part of a concept is both a deductive and inductive task (Adcock and Collier 2001). Deductively, a definition identifying different core traits is the first step to arriving at a set of indicators (Gerring 2012b, 108; Goertz 2006, 3), an inference that requires a solid theoretical foundation. Inductively, looking at extant data thought to represent a concept to derive the indicators is potentially fruitful, paying particular attention to the dimensionality of the concept and the relationships between potential indicators (Gerring 2012b, 173-175). Concept formation rarely happens in a wholly deductive or inductive manner but is most often an iterative back-and-forth process between theory and observation (Adcock and Collier 2001; Goertz 2006, 140).

This inductive-deductive joint process is particularly useful in the context of economic nonviolent strategy. The data which exist on tactics, most notably from the NAVCO 3.0 dataset (Chenoweth, Pinckney, and Lewis 2018) which I use, are highly detailed and collect a list of about 200 different tactics. Selecting indicators that apply to the idea of economic strategies, i.e. those that aim at targeting the state financially, enlisting IGO assistance or otherwise exercising economic power over the state, is simple. The inclusion criteria for the dataset assure that the observations are examples of political action in one form or another, and by selecting those cases which target the state, most of Sharp’s definitional criteria are met. I explain the data in more detail in chapter 6.

Beyond the identification of appropriate indicators of a multifaceted concept, indicators must be joined together to form the concept. Such aggregation schemes depend on ontology and concept structure (Gerring 2012b, 167; Goertz 2006, 27). Necessary-and-sufficient type concepts, or classical concepts, are indicated using Boolean logic and aggregated with multiplication or logical AND-operators (Goertz 2006, 40). Each indicator is considered a necessary requirement for inclusion in a class, and all the indicators combined are considered sufficient for inclusion. This results in a binary aggregation – presence or absence – of as many

indicators as needed to satisfy the sufficiency criteria (Gerring 2012b, 167; Goertz 2006, 39-40). The definition of democracy promulgated by Przeworski et al. (2000) is an example. Lack of variation and loss of information are important criticisms of many uses of this concept structure. A relic of the simple aggregation is the danger that dissimilar observations may be classified as identical and similar ones as different. Another aggregation scheme is characterized by indicator substitutability and additive aggregation, resulting in graded measures (Goertz 2006, 59-62). This is the family resemblance concept structure, in which indicators' values are summed to a total score, and thresholds may be imposed to delineate membership of classes (Goertz 2006). The Polity measure of democracy is an example, in which weighted sums of ordinal variables above six on a scale from negative ten to positive ten are counted as full democracies (Marshall and Jaggers 2007; Munck and Verkuilen 2002). This concept structure provides variation but may be difficult to interpret because different combinations of indicator-values give identical aggregate values. Here too there is a risk of grouping different things together and similar things apart.<sup>10</sup>

Aggregation of classical concepts is given by definition – either all criteria are satisfied or not. Family resemblance aggregation is more flexible, as it allows the combination of necessary and unnecessary indicators, as well as different weightings of indicators based on their importance to the concept (Gerring 2012b, 167; Goertz 2006). This weighting is often somewhat arbitrary, however, and there are strong arguments for letting the data decide the weightings inductively through data summarization techniques (Hair et al. 2014; Jackman 2008; Munck and Verkuilen 2002, 24). Having few theoretical expectations for the relative importance of indicators of a concept often ends in the assumption that they should be weighted equally, which is a strong and sometimes unfounded assumption. Allowing the distribution of the data to determine how they relate to the concept empirically and using factor scores or difficulty parameters as weights is a transparent and reproducible alternative, simultaneously yielding more valid and reliable measures of some concepts.

I achieve an agnostic aggregation scheme by constructing a measurement model for economic strategies, which I explain in chapter 5. The main reason for aggregating the indicators with a measurement model is that it does not require me to make judgments about the importance of each variable. Figuring out which variables are most representative of economic strategy is left to the statistical computations, and not given arbitrary weights *a priori*.

---

<sup>10</sup> For example, Suriname in 1986 and Tajikistan in 1992 receive the same Polity-score of -6 but differ very much on the indicators. Suriname gets a score of 0 on the variable *xropen*, registering the openness of executive recruitment, while Tajikistan gets 4, the maximum possible score for the indicator (Marshall and Jaggers 2007).



An important aspect of the ontology of concepts is to consider what the essential parts of a concept are (Goertz 2006, 27-28). This is similar to what Gerring (2012b, 130-131) calls causal utility, namely the extent to which a conceptualization captures the traits relevant for solving a problem or demonstrating a relationship. The utility of a concept is not fixed but depends on the problem at hand. Descriptively, every quality of a thing is a part of the thing itself, but some traits have different uses; the color and texture of a metal might be important to the artist, while the engineer is concerned with its ability to lead electricity or its reaction to a chemical. Likewise, the political scientist analyzing democracy and wealth may be interested in different traits than the one analyzing democracy and terrorism.

Furthermore, causal utility in science depends on exogeneity (Gerring 2012b, 130). If a concept is thought to explain a phenomenon in some way – say, democracy affects the quality of governance – it makes little sense to have “democracy” and “governance” overlap by defining both with a “rule of law”- trait, for instance. Just as one moves all instances of  $x$  to one side of an algebraic equation to solve it for  $x$ , the explained and explanatory variables must be separate to make meaningful statements about the relationships between them. While encompassing definitions of a concept can be appropriate in some contexts, minimalist definitions utilizing only the core essentials are often better for crafting exogenous explanations. However, some mechanisms linking the explanatory with the explained concept should at least be grounded in the former to make theories of a causal relationship convincing (Johnson 2006). For instance, the minimalist democracy of Przeworski et al. (2000) contains only criteria about elections and turnover, and is exogenous to governance, rule of law, national income and so on, and can easily be used to test relations between them and democracy. Why this relationship exists, meanwhile, requires some formulation of mechanisms by which democracy affects the dependent variable.

When constructing measures, particularly those intended for use as independent variables in a regression, several considerations and trade-offs must be made. First, the measurement of a variable must be valid both theoretically and empirically. A valid measurement measures what the creator claims it measures, resonates with the commonly understood meaning of the measured concept, and is comprised of indicators that tap into the important aspects of it (Gerring 2012b, 161; Jackman 2008). It is similar to unbiasedness and linked to the data collection process, wherein what is believed to be observations of interest in describing and analyzing a phenomenon are documented. Here the simultaneous deductive-inductive nature of measurement is evident, as some assumption or theory determines what data are collected, with consequences for hypothesis testing. Validity is largely a theoretical exercise

– achieving it requires an understanding of the concept and how it maps onto the empirical world, as well as an overview of previous attempts at measuring the concept. It can however be tested empirically by comparison with other operationalizations of the concept (Jackman 2008, 121-123).

Second, the measure must be reliable or accountable to be convincing. Given multifaceted or latent concepts, the procedure of choosing and aggregating indicators must be transparent and replicable (Gerring 2012b, 167; Jackman 2008). Reliability, assuming that the concept of interest has a true value, can be measured by the variance of the indicators of the concept (Jackman 2008, 123-126). Precision is a synonym to reliability, and it is largely an empirical issue; a concrete measurement of a phenomenon is reliable if an observation is measured identically over time or by different observers (Jackman 2008, 124). This can be estimated by the test-retest method or an inter-rater reliability test. Given unobservable phenomena, inter-item reliability is a measure of the internal consistency of a composite measure, where the quantity of interest is the correlation between indicators of a concept (Jackman 2008, 124).

Achieving causal utility for a variable summarizing the extent to which nonviolent campaigns use economic strategies is not very hard. There is unlikely to be a problem of endogeneity in two concepts as distinct as a regime transition and a mobilization strategy. It is worth mentioning that there could still be considerable bias in the sample, such that those cases which see democratization were more likely to do so despite the presence of economic strategies. This represents yet another trade-off, that between parsimony and omitted variable bias (Gerring 2012b; Schrodtt 2014). The validity and reliability of the indicators I have chosen are discussed in chapter 6.

To summarize: measurements are structured descriptions which, by assigning numbers to phenomena, permit analysis of hypotheses. The phenomenon in question may be observable or latent and may be one- or multidimensional. Both multidimensional and latent phenomena require aggregation, and aggregation depends on the structure of the concept. To be useful beyond description, the measured phenomenon must have causal utility, which is affected by the research question. For arguments building on measurements to be convincing, they must be valid and reliable representations of the systematized concept. How phenomena are represented by measured indicators have consequences for causal analysis; proxy-variables, measurement error, and endogeneity can introduce significant bias and error in causal analyses.

## 4.2. Nonviolent action as contentious politics

Keeping the lessons about conceptualization, measurement, and aggregation in mind, it is also important to ask what an economic nonviolent strategy is *a case of*. Sartori (1970, 1036) argues that to meaningfully compare phenomena they must be examples of the same higher-level concept or class. This Aristotelianism is in part mimicry of the natural sciences pervading the social sciences, which inspires measurement of everything measurable, without foundation in theory (1970, 1036-1038). Before quantification, Sartori writes, we must know “what it is that we are measuring” (Sartori 1970, 1038). Consequently, an effort to define what a concept is and is not, how it is distributed, what traits are essential to it, and how the traits relate to another is important.

To compare nonviolent strategies the higher-level concept must be defined. My subject of analysis can be described as nonviolent collective action, because I am interested in groups, acting more-or-less coordinated to achieve some goal. Olson (1967) describes collective action as a group of individuals acting together to achieve some shared goal or interest. This definition is extensive and contains nation states, interest organization, black metal bands, and elementary school classes simultaneously. Collective action is a classical concept with necessary and sufficient-inclusion criteria, as removing any of the three traits no longer qualifies for inclusion in the concept. A group of people acting together is not a case of collective action unless they have some shared interest they want to achieve; capitalists in the marketplace, for instance, want to reduce the size of the group to only themselves, achieving monopoly and maximization of profit (Olson 1967, 37-38). Similarly, a group of people waiting in line at the supermarket would be better off if there was no group at all, and they were the sole customer. This Olson calls *exclusive goods* as opposed to public ones, and while they are strictly shared (every member of the group wants more profit or shorter wait times), they are not collective interests.<sup>11</sup> If capitalists merge businesses or form cadres, or customers collectively plan and divide themselves efficiently into equal lines, one may speak of collective action, but groups competing for scarce resources or interests can hardly be described as acting collectively but rather as acting simultaneously.

---

<sup>11</sup> Schumpeter (1954, 550-551) argues a similar point when discussing the term *class*: “[...] when we speak of working-class movements, we are indeed referring to masses of individuals but of individuals that rally around a group standard and form, as it were, a psychological corporation [...]. When we consider the group of all the people who derive their incomes from selling services (personal efforts), we find that we are combining social types that have very little in common and hardly ever feel and act in unison. [...] in short, we are considering a category that we [researchers] have formed ourselves.” His intention is to illustrate that classification must be valid and refer to real-world phenomenon to be analytically fecund.

In addition, collective action can be apolitical. Metal bands and school classes are not necessarily so,<sup>12</sup> but the arbitrariness and breadth of the definition is a problem (Collier and Adcock 1999, 548-550; Sartori 1970, 1035). Only the collective action geared toward public goods, perhaps, should be considered. This way political concerts and reading groups can be included while groups without explicit political aims or activities are excluded.

Collective action does not capture the informal or noninstitutionalized nature of protest, however, and an apter higher-level category might be *contentious politics* (Tarrow 1994), itself a subcategory of collective action. Tarrow defines this as “what happens when collective actors join forces in confrontation with elites, authorities, and opponents around their claims or the claims of those they claim to represent” (1994, 4). This largely conforms to Sharp’s conflict-action (1973a). Compared to collective action, contentious politics includes outsiders to the political system challenging the insiders, moving down the conceptual ladder to a more intensive concept. What contentious politics is *not*, is apolitical action, institutionalized politics, professional persuasion, or elites’ strategic interaction (Tarrow 1994, 8).

The unit of observation for Tarrow is the social movement, which is defined by four characteristics: *a common purpose*, a goal which the movement is aiming to achieve; *collective challenges*, in which groups act against their opponent to achieve their common purpose, directly or indirectly; some *social solidarity*, a common identification with other participants or their shared interest; and *sustained contention*, some continuity in their activities and goals, separating social movements from one-off spontaneous collective acts (Tarrow 1994, 8-12).

Contentious politics, then, is the phenomenon in which social movements consisting of outsiders to a political system collectively act against insiders to challenge some aspect of politics (Tarrow 1994). Whether the substance of the challenge is grandiose or not is definitionally unimportant; contentious politics can be small-scale and local, concerning some town environmental policy, or large-scale and global, touching on imperialism, wars or capitalism. For my thesis, I add the two movement characteristics with which I am occupied, namely nonviolence, and that the movements must have goals concerning the national political system. The concept is illustrated in Figure 2.

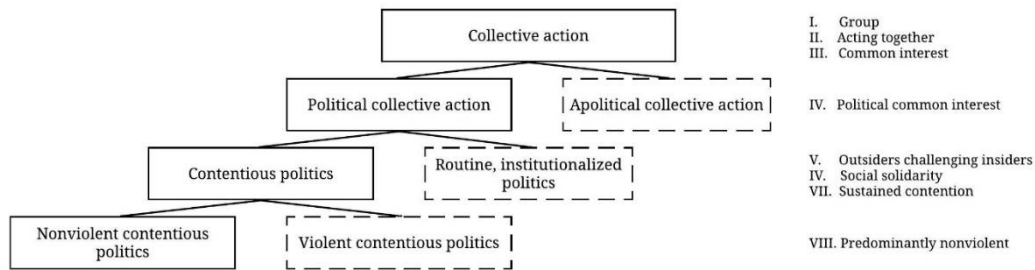
Figure 2 illustrates a necessary-and-sufficient type concept, which I believe is accurate, causally fecund, and which resonates with Sharp, Tarrow and Olson’s concepts (Olson 1967;

---

<sup>12</sup> The socialization of children in schools and the political potential of syllabi are often subjects of debate, however. See for example Sheehan (2017).

Sharp 1973a; Tarrow 1994, 8-12). The eight conditions of inclusion speak to the universe of potential cases I address.

**Figure 2:** Conceptualization of nonviolent contentious politics



*Solid-edge boxes contain concepts of interest to the thesis, dashed-edge boxes do not. Characteristics I-VIII apply to solid-edge boxes and do not to dashed-edge boxes.*

I make three decisions when phrasing nonviolent contentious politics in this terminology. First, I divide political and apolitical collective action. I wish to explicate that while many acts of collectives have the potential to become political, they are not necessarily so. Music groups, social clubs, sports teams and such are examples of this. *Political* collective actions are those concerned with supplying some sort of good, duty, or reward to a manifestly politicized group in society. Second, I separate contentious- from routine, institutionalized politics to highlight that political action sometimes happens outside the arenas of legislatures, courts, governments, parties, and NGOs (Tarrow 1994; Tilly, Tarrow, and McAdam 2003). The omnipresence of such noninstitutionalized struggles for rights, duties, and goods make up much of the grand events political scientists are committed to studying, e.g. revolutions, democratization, extensions of suffrage, terrorism, riots – in short, it is when ordinary people without routine access to the political sphere occupy it that contentious politics happen. And while phenomena like voting and petitioning are often nonviolent and political, they are outside the domain of this thesis. This distinction accords with Sharp (1973a) and his six-fold separation of social and political action in conflicts, wherein nonviolent action is separated from peaceful, institutional conflict-solving (Sharp 1973a, 66).

Third, I partition nonviolent- from violent contentious politics. This is not a normative statement on violence, but motivated by the literature which finds different causes, mechanisms, and results depending on the strategies chosen by contentious actors to further their interests (e.g. Chenoweth and Uldfelder 2017; Croissant, Kuehn, and Eschenauer 2018; Cunningham et al. 2017; Sharp 2012; Stephan and Chenoweth 2008). I build on these findings by theorizing an

economic dimension of nonviolent action which possesses particular causal traits, at a lower level of abstraction “nonviolent action.”

A predominantly nonviolent campaign needn't be entirely peaceful and civil, and nonviolence is not synonymous with pacifism, however (Sharp 1973a, 68). I would not consider damage to property or other non-person things violence, for one thing. There are strong reasons to call this type of behavior “violent” in some cases – hoarding food during famines is arguably a violent act, as is damaging agricultural land or people's houses. Striking and damaging a company's profits or blocking a road, while potentially causing multiple times the damage in monetary terms, are scarcely violent in themselves, however. This is because violence is a question of structural relations (Aitchison 2018). When used to challenge “certain objectionable forms of political domination” (2018, 667), Aitchison writes, it can be both morally defensible and be employed as a surrogate-, remedial-, or mobilizational tool in lieu of other channels of contestation. Coercion, meaning the interference in others' choices by disincentivizing, forcing, or tricking them (Aitchison 2018, 668), is not necessarily violent. Simultaneously, persuading someone to voluntarily pursue a course of action can also be violent, for example by communicating the seriousness of a demand, which is the case with terrorism (Aitchison 2018; Pape 2005). The perpetrator and target of seemingly violent or nonviolent action is a critical contextual factor to adjudicate the justness of it. Empirically, few campaigns are completely violent or nonviolent, and adding the qualifier *predominantly* to nonviolence is necessary. Some margin of tolerance is needed to delineate between predominantly nonviolent and violent campaigns, although each concrete event can be strictly one or the other.

#### 4.3. Sharp's methods of nonviolent action

What characterizations are shared among different strategies of nonviolent contentious politics? Which courses of action distinguish them from each other? Sharp (1973b, xi-xviii) structures his list of tactics under the headings *protest and persuasion*; *social noncooperation*; *economic noncooperation*: (1) *economic boycotts*, (2) *the strike*; *political noncooperation*; and *nonviolent intervention*.

Firstly, Sharp (1973b, 117-119) defines nonviolent *protest and persuasion* as arguing for or against something; it may speak to an opponent or the public, and it is mainly communicative. They are distinguished from noncooperation and intervention by being primarily about expressing opinions or calling for action, but not exerting direct pressure or costs on an opponent or bystander. Yet they are distinctly political, as they are *collective* efforts rather than personal communication, and aimed at achieving a political goal and/or targeting a

politically potent audience (Sharp 1973b, 118). Sharp lists 54 methods of nonviolent protest and persuasion and points out that their efficacy differs according to their prevalence in a particular context, and that their prevalence is determined by sanctions and opportunities for collective and expressive action and speech (Sharp 1973b, 119). Examples of this type of action are picketing, wearing symbols or colors of political salience, protest meetings, public speeches, and parades (Sharp 1973b, 119-172).

Second, *noncooperation* is a refusal to continue particular relations. Sharp divides the methods of noncooperation into social-, economic-, and political noncooperation, based on what kind of relationships are discontinued (1973b, 184-185). Social noncooperation targets relations between people or groups of people. While the aim of social noncooperation is political, the arena wherein action is taken is social, for instance refusing to socialize or behave normally with strikebreaking scabs or Nazi collaborators (Sharp 1973b, 185-189). Individuals or groups who have done something disagreeable become symbols of the political disagreement, and costs imposed upon them to sanction and deter such actions. The experiences of Norwegian women who were (allegedly) romantically involved with the occupying Nazi soldiers provide an extreme example, as they had their hair shaved, were labelled *tyskertøs* (lit. “German-slut”), and were ostracized for many years, as were with their children, called *tyskerunger* (lit. “German-kids,” pejoratively).

*Economic noncooperation* encompasses boycotts and strikes, which impose economic costs on a target. Boycotts are discontinuations of purchase, sales, or handling of goods from certain suppliers, while strikes involve ceasing labor activity, usually collectively and simultaneously, to disrupt production (Sharp 1973b, 219-221, 257-259). The common denominator is the economic impact imposed on an objectionable target, while the goal of the action may be economic, social or political (1973b, 219-221). Examples include the boycott of Israeli-produced goods by individuals and institutions to protest and deter the occupation of Palestine, the Chinese state boycott of Japanese goods in the early 1900s to protest Japanese policies, and the Icelandic Women’s Strike in 1975 to protest pay gaps and gender inequality.

*Political noncooperation* is the discontinuation of political activity, for example with police, the state, foreign states, parties and the like (Sharp 1973b, 285-286). The isolation of the Sweden Democrats from government negotiations in Swedish parliament after the 2014 election and the expulsion of Russian diplomats from multiple countries following the poisoning of former Russian spy Sergej Skripal and his daughter Julia in 2018 are examples of political noncooperation. Political noncooperation imposes costs in different ways, for instance by creating legitimacy costs for parts of the state if citizens do not cooperate with them, or

economic and political costs on foreign governments when suspending diplomatic activity.

Lastly, *nonviolent intervention* is a category of methods directly effecting change, akin to direct action, commonly used by anarchists. Instead of sanctioning, deterring, or persuading others to pursue some course of action, nonviolent intervention directly hinders it, for instance by destroying anti-homeless architecture, squatting in unused homes to alleviate homelessness, or arranging lunch-ins in racially segregated restaurants (Sharp 1973b, 357-359). Sharp partitions the category into social, political, economic, psychological, and physical intervention (Sharp 1973b, 358). These subdivisions speak to the different targets and spheres in which intervention can be employed: social intervention by establishing new social norms, like treating Black people as equal to Whites in the 19<sup>th</sup> century US; political intervention by establishing parallel institutions like courts and parties; economic by creating cooperatives to lower prices; psychological by verbally harassing opponents; and physically by invading or trespassing politically salient institutions or companies (Sharp 1973b, 359-435).

The conceptualizations offered all possess causal utility for my research question, as no category or subcategory contains traits endogenous to democracy or democratic transitions (Gerring 2012b, 130-131). Several authors have argued and tested the hypothesis that the organization of civil society contributes to the evolution of social capital within a society which subsequently aids democratization and democratic consolidation (e.g. Diamond 1999; O'Donnell and Schmitter 1986; Putnam 1993; Teorell 2010). This is however not an issue of conceptualization, but formulations of the hypotheses I wish to test. Those before me who have argued the importance of civil society have not deconstructed nonviolent action and tested its components systematically but have often used independent variables like “civil society strength” and “organizational density” to correlate with the probability of democratization. The causal utility of Sharp, NAVCO’s and my own conceptualization of economic nonviolent action contain no elements of regime types and are exogenous to these concepts.

One issue with using the subdivisions of Chenoweth, Pinckney, and Lewis (2018) and Sharp (1973b) is that they are not so much categorizations of strategy as they are classes of methods by which strategies can be executed – tactics in military language. I have no expectation that campaigns will align to only one of these categories but hypothesize that a combination of different methods from different subcategories of nonviolent contentious action will be utilized for each struggle. Some combinations will probably be more common than others, and some more efficient. There is no reason to believe that there is an ideal “protest and persuasion” campaign which for some reason only use these methods to achieve their goals and refuse to employ noncooperation and intervention-techniques. However, I do believe that there



are distinct qualities possessed by *economic* strategies that alter the likelihood of democratization by inflicting costs on the regime. By theorizing these qualities and mechanisms, and by conceptualizing and measuring the concept of economic nonviolence, I hope to find a probabilistic relationship that answers the research question and provides relevant information to the research literature and to activists.

## 5. Methods

To test the impact of economic strategies on the likelihood of democratization, I first construct a measurement model of economic strategies. I then use the resulting measure as my main independent variable to test whether movements using more economic strategies are more likely to achieve their goals than those who use them less. The advantages to using a measurement model compared with binary predictors<sup>13</sup> from the NAVCO dataset are many, relating to reliability, validity, model fit, and parsimony. I use Bayesian item response theory to explore the relationships between the chosen variables for economic strategies and summarize them into a standardized, metric score. To test my hypotheses, I then use the score to estimate the causal effect of it on the likelihood of democratization in a logistic regression model, with appropriate control variables.

### 5.1. Bayesian item response theory

Measurement models take indicators of observed data that are thought to represent a multifaceted or unobservable concept and merge them together to form a new, metric variable representing the concept. This use of measurement models is called data reduction and the objective is to summarize the variation in the predictor variables with fewer numbers and thus achieve a more parsimonious regression model (Hair et al. 2014, 96-97). Treier and Jackman (2008) for example, take indicators from the Polity IV project and apply a measurement model, which allows them to assess the measurement error of the score, and to agnostically weigh the indicators – in short, an approach preferable to the arbitrary one used by Polity IV. My approach is similar: I use my chosen indicators of nonviolent tactics from NAVCO 3.0 and apply a measurement model which results in an ability-score for nonviolent campaigns' use of economic strategies, or a summated scale of their strategy's 'economicness'.

---

<sup>13</sup> The variables *verb\_10*, *verb\_100* and *verb\_1000* are coded nominally, and simply using them as predictors in a regression would be equivalent to making dummy variables for each value. This is at best a profuse approach – having close to two hundred dummy variables as predictors would obfuscate the relationship between the variables and make it nearly impossible to make sense of the resulting coefficients, all controlled for one another. I therefore make dummy variables for each nominal category and make a measurement model using the 42 indicators of economic strategy.

Several techniques for data reduction are available which can explore how variables related to economic nonviolent strategies covary. One way to reduce the number of variables used to measure a concept is item response theory (IRT; see Hair et al. 2014; Jackman 2008, 2009; Jordana, Fernández-i-Marín, and Bianculli 2018). IRT-models require binary coding of a variables' presence or absence and are mainly employed in psychometrics and related fields (Jackman 2008). The most common use in political science is for analyzing roll call- or judicial decision-data to reveal the latent ideological positions of senators and judges (see Clinton, Jackman, and Rivers 2004; Jackman 2001; Jackman 2008, 2009). Despite the omnipresence of binary variables and a multitude of potential uses, the method is unpopular in political science. The technique is particularly useful to alleviate some of the problems highlighted by Achen (2005) and Schrodtt (2014), who lament the use of binary independent variables and its consequences for model fit.<sup>14</sup>

The core idea of a measurement model is to estimate a latent dimension from the observed data. A latent dimension as discussed above is a variable that is not directly observable, such as class or ideology. Obviously, observing the latent characteristics directly would be preferable, but this is often not possible. The benefits of a measurement model compared to creating an additive index or using proxy variables are many. First, measurement models allow the data to weight themselves, relieving the researcher of creating arbitrary weighting schemes (Achen 2005; Goertz 2006, 46-50). Weights must be justified with reference to the concept structure, and the default weighting schemes of taking the mean, median, or sum of the indicators constitute strong assumptions that are rarely theorized. While seemingly neutral and agnostic with regards to the distribution of the latent dimension, it can affect the results greatly (see Hooghe et al. 2016; Treier and Jackman 2008). This is a point made by Treier and Jackman (2008), that while the aggregation rules of the Polity IV index are simple, they are also arbitrary. Using a measurement model is both more transparent and more neutral.

Second, given some theoretical validity of the chosen indicators of a concept, measurement models enhance the confidence of subsequent regression models compared to those using proxy variables or ill-specified indexes (Achen 2005; Jackman 2008). Achen shows that small misspecifications, bad proxies, or adding or removing control variables can have huge impacts on regression analyses, making them almost useless and often uninterpretable. Achen (2005) and Schrodtt (2014) both argue for parsimony in regressions, i.e. using few

---

<sup>14</sup> The same can be achieved by factor analysis, although this method usually requires non-binary coding of variables, and treats model parameters differently in some respects (e.g. Fox 2010; Jordana, Fernández-i-Marín, and Bianculli 2018; Reckase 2009).

independent variables in a single regression model. By reducing indicators of a concept into a single measure with statistically determined weightings one is better equipped to perform hypothesis-testing analyses because much of the same variation is explained by the reduced measure and the model fit can be assessed more easily with fewer explanatory variables.

Measurement models like IRT are very similar to common regression analyses in that the objective is to estimate unknown parameters based on known data. Linear regressions are usually formalized as

$$\gamma = \alpha + \beta X + \sigma$$

where  $\gamma$  is the observed data, the values of which are predicted using the estimated intercept  $\alpha$ , and the estimated coefficient  $\beta$  of the observed predictor variables  $X$ . The residuals or unexplained variance are represented by  $\sigma$ . The key point is that the beta-coefficient is estimated based on observed data  $X$  and  $\gamma$  but is itself unobservable. In measurement models, the latent  $X$  is also estimated from the observed data  $\gamma$ .

For IRT models, imagine a class of students taking a maths test on addition. Their answers ( $\gamma$ ) are marked as either correct (1) or incorrect (0) and recorded in a student-by-question matrix ( $i$  by  $v$ ). Based on this matrix of observed characteristics  $\gamma$ , it is possible to calculate each question's difficulty parameter ( $\beta$ ),<sup>15</sup> discrimination parameter ( $\alpha$ ), and each student's ability ( $\theta$ ).<sup>16</sup> The difficulty of a question ( $\beta_v$ ) tells how likely it is that students get it right, given their ability ( $\theta_i$ ). A student that gets everything right will have high ability score, and questions that only the brightest get right are difficult.<sup>17</sup> The discrimination parameter  $\alpha_v$  describes how good a question is at distinguishing between students; those questions that everyone gets right or wrong have a low discriminatory value, whereas those that separate students well are good discriminators. The students' ability-score will represent their skill in addition, assuming that other skills are not required or are evenly distributed, such as reading and understanding text and numbers.

For my measurement model of economic strategies, I want to extract the abilities of different nonviolent campaigns to use economic strategies. Because each event in the NAVCO 3.0 dataset is assigned only one strategy-indicator I need to aggregate temporally to produce

---

<sup>15</sup> The name "difficulty parameter" is simply the test based IRT nomenclature and does not mean that variables with high difficulty scores are necessarily harder to "perform."

<sup>16</sup> For those familiar with factor analysis, discrimination is equivalent to loading or weights, and ability is equal to factor scores, i.e. the measurement of the latent dimension in which I am interested.

<sup>17</sup> Some IRT models allow for a guessing-parameter, i.e. factoring in the likelihood that some respondent randomly get the question right, despite lacking the ability to do so.

variation in the data – a test with only one question taken by one student cannot reveal much about the class’ relative abilities and the questions’ difficulty or discriminatory value. I aggregate by the strategy-indicators’ maximum count by campaign-year. That is, I record whether they are present in a given year no matter how many events. This retains the dichotomous coding and allows the use of IRT.<sup>18</sup>

The worst consequence of this is that I lose a lot of data and observations. The number of observations shrinks from some 112.000 distinct events to 442 country-years, and because I aggregate by the maximum count of each dummy variable each year, I also lose the number of times a strategy is used. There are available methods that do not require all of this, which would be equally suited to my analysis, which I discuss in chapter 7.4.

Each nonviolent campaign thus takes a “test” in economic tactics each year. Campaigns either do or do not use a particular economic nonviolent tactic. Based on the data matrix of tactic presence or absence, the ability-score of each campaign-year is estimated, a score which communicates how “good” a campaign was at using economic strategies that year. The tactics which are common will have a low difficulty parameter and the campaigns which use few or only easy tactics will have a low ability score. I expect that the ability of a nonviolent campaign to use economic strategies helps explain transitions to democracy, and the ability score thus forms my main independent variable.

Why am I using a *Bayesian* framework for my measurement model, rather than the conventional frequentist one? The key reason is that my data are not representative of the population to which I would want to generalize my findings. As the NAVCO 3.0 data are explicitly not chosen randomly or with any pretense at randomness or representativeness of other countries, but rather because they are particularly eventful, this limits the applicability of frequentist confidence intervals and significance levels. Additional sampling from the population would probably yield very different results and alter the relationship between the variables (Jackman 2009, xxxi-xxxii; Schrodt 2014, 293). Breaking this particular assumption of representativity should always be a reason to at least evaluate the choice of frequentist or Bayesian methods (Fox 2010; Jackman 2009; Kruschke 2010; Schrodt 2014).

Using a Bayesian measurement model can improve the measurement model additionally by including the uncertainty of measurements in the analyses. Missing values for example, of which there are many in datasets on social movements, are utilized in estimating the uncertainty

---

<sup>18</sup> An alternative approach would be to summarize the periodic occurrences of the indicators, so that a tactic used in five events counted as 5 rather than 1. This would require factor analysis for data summation, which is basically equivalent to IRT.

of the posterior probability distribution, which is a more appropriate response than the habitual listwise exclusion resulting in fewer observations (Fox 2010; Jackman 2009; Jordana, Fernández-i-Marín, and Bianculli 2018, 3).

My two-step plan of first formulating a measurement model and then using one of its parameters as an independent variable makes uncertainty very important. By stating my results as probability distributions, this uncertainty can be carried on through the process and inform the result better than a ninety-five percent credible interval can. Communicating the uncertainty of an estimate is an important part of evaluating the credibility of a piece of evidence. And as frequentist methods often do so in a limited way, for instance by deleting observations with missing values, I use Bayesian methods which do not hamper the communication of uncertainty. This way, the uncertainty of my measurement model (which is considerable, given the high degree of missingness from my variables) is carried on to the hypotheses testing phase. My item response model looks like Figure 3:<sup>19</sup>

**Figure 3: JAGS model for Bayesian IRT**

```

1      'model {
2      for (rowIdx in 1:Ntotal) {
3          y[rowIdx] ~ dbern(pCorr[rowIdx])
4          pCorr[rowIdx] <- ilogit(itemDisc[itemID[rowIdx]]
5              *(subjAbil[subjID[rowIdx]] - itemDiff[itemID[rowIdx]]))
6      }
7      for (subjIdx in 1:Nsubj ) {
8          subjAbil[subjIdx] ~ dnorm(muAbil, 1/sigmaAbil^2)
9      }
10     for (itemIdx in itemDiffEstIdx ) {
11         itemDiff[itemIdx] ~ dnorm(muDiff, 1/sigmaDiff^2)
12     }
13     for (itemIdx in 1:Nitem) {
14         itemDisc[itemIdx] ~ dnorm(muDisc, 1/sigmaDisc^2)
15     }
16     muDisc      ~ dnorm(0.1,0.1^2)
17     sigmaDisc   ~ dunif(0.001,0.1)
18     muAbil      ~ dnorm(50,1/25^2)
19     sigmaAbil   ~ dunif(1,50)
20     muDiff      <- 50
21     sigmaDiff   <- 50
22     }'
```

---

<sup>19</sup> The reason for this tiresome mathematical display is that there are no packages available in R that I am aware of with enough options and specification for me to execute this part of my research design. Furthermore, a few scholars have used similar models before me, which allowed me to learn a great deal about the maths behind the model (notably Curtis 2010; and most importantly Fernández-i-Marín 2019, who introduced me to IRT and JAGS in the first place; Jackman 2001, 2009; Reckase 2009). This particular model is adapted from the online appendix to Kruschke (2010), available at <http://doingbayesiandataanalysis.blogspot.com/2015/12/bayesian-item-response-theory-in-jags.html>

What follows here is a restatement of the model in, hopefully, more understandable terms. First recall that the item response model is performed on a matrix of binary data, and estimates the parameters based on this matrix. I write the model in JAGS (Just Another Gibbs Sampler) and run it from R using the *rjags* package (Plummer 2018).

Line 3 signifies that the data ( $Y$ ) for each cell are generated from a Bernoulli-distributed<sup>20</sup> probability parameter ( $p_{\text{Corr}}$ ). This means that the values of the cells take the value 1 and 0 with the frequency they observably do. It is also a way of restricting the model to only consider outcomes between 0 and 1, as Bernoulli distributions only have two outcomes.

Line 4 and 5 state the regression model itself. The probability parameter for each cell equals the logit of the discrimination for the variable times the observation's ability minus the variable's difficulty. The logit is the natural logarithm of the odds rate for each value. The probabilities of the data, which are Bernoulli distributed, are given by the relations between parameters in the matrix, and these parameters are thus estimated based on the data. The ability is determined by how many ones are in the observation's row, and the difficulty by how many are in the column. Rows with many ones have high ability, and columns with many ones have low difficulty. The ability of a column is based on how well the columns contribute to separating observations from one another.

Lines 7 and 8 specify that the ability scores are normally distributed with the estimated mean and standard deviations specified in lines 18 and 19, the mean is normally distributed around 50 and with a low precision parameter, the standard deviation uniformly distributed, with high precision. The precision specified in the priors is a measure of how certain the user is about the parameter taking the mean value and is interpretable similarly to variance. High precision values, such as the 50 used in line 18, means that I am quite uncertain that the mean is correct.

Lines 10 and 11 fix the difficulty parameter between 0 (easiest indicator) and 100 (hardest indicator). The mean and standard deviation for the difficulty is fixed to 50. This is simply a constraint on the model which allows the other parameters of interest, discrimination and ability, to vary in relation to each other. The discrimination of each indicator is affected by the others' discrimination, and each campaign-year's ability does the same. The shrinkage which this relativity implies makes outliers less influential (Kruschke 2010, 175-176). Fixing the difficulty in this way has no negative consequences because the difficulty parameter itself is not an interesting quantity to the research question.

---

<sup>20</sup> A Bernoulli distribution is one with a single experiment with two possible outcomes, where the probability of an event occurring ( $p$ ) is between 0 and 1, and the probability of it not occurring ( $q$ ) is  $1 - p$ .

Lines 13 and 14 specify the discrimination of the indicators and are similar to lines 7 and 8, except that the priors for the mean and standard deviations in lines 16 and 17 are fixed at other locations with different strengths.

The priors are meant to be vague to allow the data to dominate the model because I have no similar measurement models to base my beliefs on. This is very much an explorative application and biasing the output with an unjustifiably strong prior would not be proper.

## 5.2. Bayesian logistic regression

The abilities of the campaigns to utilize economic strategies are then appended to my dataset to be used as my main independent variable in testing whether economic ability is related to the likelihood of democratization. I do this by estimating a posterior distribution for the logit of the economic ability variable, statistically controlling for GDP, tax revenue-to-GDP ratio, resource rent-to-GDP ratio, and net aid received.

Logistic regression is not very different from the standard linear regression model as discussed above. The main difference is that the dependent variable  $Y$  is dichotomous and not continuous. Rather than reporting the average amount of change in  $Y$  with a one-unit increase of  $X$ , as the coefficient in linear regression does, the output of interest is the average change in the natural logarithm of the odds rate-change in  $Y$  with a one-unit increase in  $X$ , i.e. the change in probability of observing a positive response one, controlled for the other covariates' effects. This is a very difficult metric to interpret, and visual aids such as density plots of the coefficients are helpful. The logistic regression model furthermore limits the possible values of  $Y$  to be 1 and 0, which eliminates the possibility that the outcome can occur more than all the time, which is logically impossible – a likelihood cannot be greater than one or less than zero.

The Bayesian method of hypothesis testing works as such: I inform the model of my prior beliefs about the location, shape, and precision of the posterior distribution, and the model subsequently incorporates the data so that the posterior distribution reflects a mixture of my beliefs and the actual data (Jackman 2009, xxvii). The posterior distribution is a probability statement about the parameter of interest, which allows me to answer my research question as an intuitive likelihood, for instance by saying that I am seventy percent certain that the logit of my ability score is positive, given my prior beliefs and data, and sixty percent sure that it lies in a certain interval.

The prior is a subjective statement on what I believe the truth of the parameter to be. Subjective probability refers rather to a state of mind than of the world, which means that different priors will be used by different researchers, and the results of an analysis will differ

based on it (Jackman 2009, 13-19). If a prior strongly states that a coefficient is exactly 5 and the data equally strongly say it is 15, the posterior distribution will center around 10 and be somewhat uncertain. A prior of 15 for the same data will however be reinforced and made more certain. Furthermore, the confidence lent to the prior matters. The stronger the prior, the less influence the data have on the posterior distribution. A weak prior, on the other hand, where little is known or believed about the location and distribution of the parameter, will influence the posterior distribution comparatively little – this is appropriate for hypothesis testing such as mine, where previous research is scarce.

Therefore, my prior confidence in the hypothesis is weak. I incorporate that weak confidence into the analysis with the prior. Only after reviewing the results of the analysis do I update my beliefs and answer the hypothesis. I will not, however, be certain that the resulting coefficient is an objective truth in the universe. Rather, I will be more certain about my hypothesis' level of veracity. There is no objective law that determines how economic strategies work or under which circumstances, but there are trends or patterns which can be summarized to a regression coefficient. This is not the same as saying that the result of my analysis will be undeniably true, but that the data I have viewed should make me update my subjective idea of how important economic strategies are for regime transitions. A different, more convincing analysis can change it again by falsifying the hypothesis or changing the coefficient, for instance by incorporating a relevant covariate I have omitted.

The effect of something so unscientific as a researcher's subjective beliefs on the result of a statistical analysis seemed to me at first irresponsible and arbitrary, but it touches on core ideals of the scientific method. To infer something about the world, we must compare our initial idea with observable data, and rationally update our beliefs to fit those data. Strong beliefs require strong observations to be disproven. Compared to the frequentist approach of requiring a p-value of  $< 0.05$  to discard an improbable null hypothesis, the Bayesian incorporation of prior beliefs is rational and in many situations more appropriate than the assumption of repeated sampling (Jackman 2009, 6-8).

Furthermore, stating the estimated regression coefficient as a posterior probability communicates the uncertainty of the estimation and provides information with which to update beliefs about the coefficient. The key output to interpret is the logit for every covariate, which when positive indicates that the likelihood of a transition to democracy increases with each unit increase in the covariate, while the others are held constant. Graphical depictions of the posterior distributions also communicate the shape, thereby allowing readers to evaluate the results in detail. The logistic regression model I use looks like Figure 4:



**Figure 4:** JAGS model for Bayesian logistic regression

```
1      "model{
2        for(i in 1:n){
3          Y[i] ~ dbern(q[i])
4          logit(q[i]) <- beta[1] + beta[2]*X[i,1] + beta[3]*X[i,2] +
5                      beta[4]*X[i,3] + beta[5]*X[i,4] + beta[6]*X[i,5]
6        }
7        for(j in 1:6){
8          beta[j] ~ dnorm(0,1)
9        }
10     }"
```

Line 3 states that the data  $Y$ , the dependent variable, are distributed the way they are. As in the IRT model, they are Bernoulli distributed into ones and zeroes, with the frequency  $q[i]$ , i.e. as it is in the data. Lines 4 and 5 state the regression model, in which the logit of the probability with which the data take their values is determined by the coefficients of the covariates,  $\beta[2]$  through  $\beta[6]$ , multiplied by the value of the observation.  $\beta[1]$  is the intercept, and not multiplied by a cell value.

Lines 7 and 8 specify the prior distribution for the betas, stating that they are normally distributed with location zero and precision one, which allows the data to dominate the posterior distribution. This is the point, as my actual prior knowledge of the relationship is weak. While I expect the coefficient to be positive and strong, this is more of a guess than a prior belief based on evidence, and I do not use a prior to represent that expectation.

## 6. Data

To investigate whether economic strategies are better at producing democratic transitions than non-economic strategies, I need variables that reveal variation in tactics and strategies, and ones that document whether transitions to democracy occur. I use what I consider state-of-the-art data-resources which fit the conceptual discussions above, and which allow me to answer the research question in a convincing and reproducible manner. Unsurprisingly, these qualities are not present in one ready-made dataset, and I therefore merge several datasets to supply control variables. Furthermore, a lot of data cleaning, aggregation, and amending is necessary. This process, along with the reasons for- and limitations of my choice of data and the fit between the data and the concepts they measure, are all detailed in the following sections, along with descriptive statistics. R-scripts for replication can be supplied on request.

### 6.1. Data on nonviolent mobilization – NAVCO 3.0

The very best quantification of strategic nonviolent mobilization currently available is, I think, the NAVCO 3.0 database (Chenoweth, Pinckney, and Lewis 2018). The dataset consists of

112,381 event-days, covering the years from 1991-2012 for 26 countries.<sup>21</sup> My reason for choosing this dataset is primarily its relevance to the research question: It contains detailed information on the tactics used by political movements, which can be used to create a measurement model of strategy.

The data are sourced from Agence France Press (AFP), coded by research assistants under supervision, and tested for inter-coder reliability (see Chenoweth, Pinckney, and Lewis 2018 for more details on the dataset). While I have a high degree of confidence in the coding itself (having found only a handful of errors in the attractive verb-variables), the fact that only one newspaper is used is a cause for concern. AFP is a large newspaper and highly regarded for quality journalism, but this does not necessarily remedy the potential bias in relying on one provider of stories to code. This is perhaps my greatest concern with using the NAVCO 3.0 data, as the absence of data triangulation by using different sources may bias the collected data in some way, for instance by over- or underestimating numbers of participants, economic consequences, deaths, injuries and so on. The reliance on AFP therefore introduces a source of uncertainty to the data that needn't have been there.

The sample is also unrepeatably. According to Chenoweth, Pinckney, and Lewis (2018, 527-528) the dataset includes “[...] all events in a country for the entire period of study,” which means that all relevant events reported by AFP are recorded. Doing the data collection again would not alter the result save for some measurement error, and to speak of repeated sampling does not make much sense. Collecting data from other newspapers, time periods or countries would of course also be valuable but would represent a different population than NAVCO 3.0.

Additionally, the actual news stories are not recorded by NAVCO 3.0, and only the title is present in the dataset. As the stories may be updated online after the coders have finished documenting them, this makes replication of the measurement difficult. More than once have I looked for articles to investigate a suspected coding mistake and failed to find the source. However, it is understandable that keeping track of and updating over 112,000 news stories is an enormously impractical and near-Sisyphean task.

I choose to use NAVCO version 3.0 rather than for example earlier NAVCO datasets, the European Protest and Coercion Data (EPCB), or the Minorities at Risk Organizational Behavior dataset (MORAB) because of NAVCO 3.0's detailed coding of actors and actions. It

---

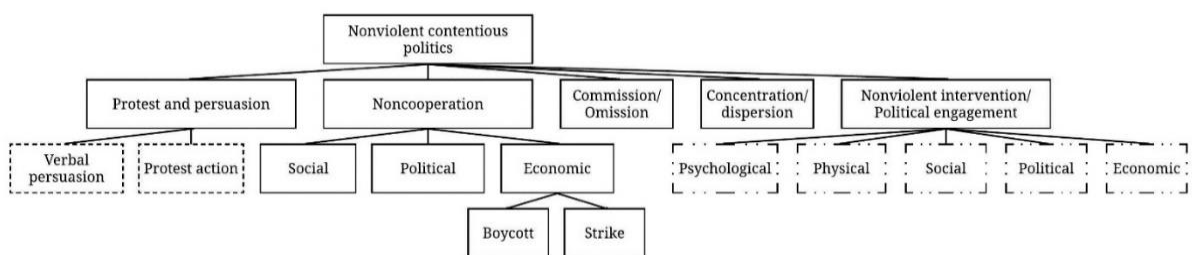
<sup>21</sup> The authors describe the country sample as potentially “more ‘eventful’ than the global average,” a source of bias they discuss in an online appendix. The data furthermore do not cover the full twenty one-year stretch for all the countries (Chenoweth, Pinckney, and Lewis 2018).

is relatively easy to summarize, separate, and manipulate the data. Furthermore, it is not regionally bound to a region like the EPCB<sup>22</sup> or MORAB<sup>23</sup> datasets, and it ends relatively recently, making it more likely that findings from the data are still relevant.

## 6.2. Data-concept coherence

A concept-illustration of Sharp's categories alongside the variables of NAVCO 3.0 (Chenoweth, Pinckney, and Lewis 2018, 529) for the corresponding concepts is depicted in Figure 5.<sup>24</sup>

**Figure 5:** Sharp and NAVCO's conceptualization of nonviolent strategies



*Dotted edges are only present in NAVCO 3.0, dotted-and-dashed edges only in Sharp's typology.*

Sharp and NAVCO 3.0's variables have some differences, marked by the edges of the boxes in Figure 5. In NAVCO 3.0, events are characterized by the category of nonviolent action, divided into protest and persuasion (subdivided into verbal persuasion and protest actions); non-cooperation (social, political, economic-strike, economic-boycott); and political engagement. Both Sharp (1973b, 68) and Chenoweth, Pinckney, and Lewis (2018, 529) allow for variation differences in *concentration or dispersion*, or whether geographic distribution is a goal of the event, and for *commission or omission*, signifying whether an event is a performance of unusual or unexpected action, or a refusal to perform expected activities, respectively. The two sources thus have near-identical ideas on how best to conceptualize nonviolent action.

In addition, NAVCO 3.0 has several variables for more detailed description of the concrete actions performed in campaign events, which adds specificity and potential for accuracy and increased intension in the use of indicators or for analysis. These, unfortunately, do not accord perfectly with Sharp's descriptions of nonviolent tactics but are adapted from the CAMEO codebook.<sup>25</sup> While Sharp (1973b) describes detailed actions like sit-ins, protest

<sup>22</sup> While the EPCB has detailed descriptions of the actual events similar to NAVCO 3.0 (verbs, what is being done to achieve something), its records are slightly older, documenting events from 1980-1985.

<sup>23</sup> MORAB also has an ethno-political inclusion criterion, which is not the focus of this thesis.

<sup>24</sup> While my overarching concept and contrast space differ somewhat from Sharp's, as discussed in the previous section, the subdivisions conform. I argue that material destruction could fit into the physical- and economic nonviolent intervention-category.

<sup>25</sup> <http://eventdata.parusanalytics.com/data.html>

disrobing and homages at burial places, Chenoweth, Pinckney, and Lewis (2018) have more general subcategories of action, such as “demonstrate for leadership change” or “obstruct passage to demand rights.” This is a challenge for the concept-measurement consistency, as the preferred deconstruction of the categories of nonviolent action drawn in Figure 5 would be one identical to the prime theoretician’s, Sharp. Chenoweth, Pinckney, and Lewis (2018) might however have more breadth in their measures than Sharp does, and that they are less sensitive to mobilizational culture and regional variations, which is advantageous for clustering and measuring success (Collier and Levitsky 1997; Gerring 2012b, 61-64; Sartori 1970).

Despite the sub-optimal fit between Sharp and NAVCO 3.0’s action-categories, I believe that the latter’s operationalization is more apt for data reduction. Sharp’s methods of nonviolent action are potentially culture-bound and are restrictively fine-grained. The first point is a problem, but the latter not necessarily so; with a proper aggregation-scheme, fine-grained data would be great. An issue with Sharp’s descriptions of methods is that they are inductively formed and therefore hard to systematically quantify, especially in retrospect. Newspapers for instance, which is the data source of NAVCO 3.0 (and many other projects) do not necessarily report that a strike was a “lightning strike” or a “limited strike.” The NAVCO 3.0 verb-variables which are theoretically representative of economic strategies and use them in my measurement model to form a variable for the degree of economic strategy used by democratization campaigns.

The indicators I use are listed in Appendix 10.1. They are chosen based on the NAVCO 3.0 variable on economic noncooperation as described in Figure 5, and, because the economic noncooperation variable has a high degree of missingness, on reviewing the list of tactics recorded by NAVCO 3.0 and choosing those which seem relevant. I believe the chosen indicators are all relevant for the concept of economic nonviolent action conform to the conceptual discussions above.

### 6.3. Data on democratization – Dictatorship and Democracy index

For my dependent variable, I use the Dictatorship and Democracy (DD; Cheibub, Gandhi, and Vreeland 2010) index’s measure of whether a transition to democracy happened in a certain year. The variable is binary and records a minimalist-procedural concept of democracy concerning the way representatives are chosen for legislatures and executives, and whether turnover in government has occurred since democratic elections were introduced. It is a somewhat controversial measure as many believe the concept of democracy cannot be summarized sufficiently a dichotomous measure. (For various views, see Bollen 1979;

Cheibub, Gandhi, and Vreeland 2010; Ghandi 2015; Munck and Verkuilen 2002; Przeworski et al. 2000; Treier and Jackman 2008) As there is neither space nor need to delve into the discussion here, I only explain my reasoning for choosing the dichotomous measure briefly.

Compared to metric measures of democracy, such as V-Dem, Polity IV or the Vanhanen measure, DD is appropriate for measuring changes from one regime to another. What concerns me and the research question for this thesis is the effect of economic nonviolence on the likelihood of transitions to democracy. Smaller changes in the *level* of democracy can be significant too – citizens in an anocracy may be pleased with policies that increase their democracy level from three to four despite their elections not being completely free. However, I prefer to consider political systems as bounded wholes and qualitatively different from non-democracies for ontological reasons: political systems that do not have observably free elections do not meet the minimal requirements for the procedural conception democracy. It therefore makes sense to treat them as two distinct categories, although both democracies and dictatorships have varying degrees of inclusion and competition. I adhere to the advice of Sartori (1970) and consider democracies and autocracies as taxonomically different, rather than to choose one of the many continuous measures which exist. This is not a hindrance for investigating the effects of economic strategies however, as I compare the different effects of strategies on the *likelihood* for transitions. A continuous measure of democracy would require me to measure the average change in the *level* of democracy rather than the *probability* of transition and to make an arbitrary cut-off point for what constitutes a transition, whereas the DD measure does not.

For the purposes of testing my novel measure of economic strategies, however, a dichotomous measure of democracy is probably a poor strategic move. Metric, fine-grained measures carry more variation and finding a substantial average effect would be easier with for example Vanhanen's 100-point measure. A dichotomous measure thus represents something of a hard test for the hypothesis. The advice of Sartori is however relevant again, as comparisons between qualitatively different classes of things cannot be meaningful (1970, 1036). Dictatorships and democracies cannot be meaningfully compared on a scale of "democraticness" because the former does not possess the quality, and as such belongs to a different class of political regime. Furthermore, the investigation into whether a variable increases the likelihood of a holistic transformation of political life rather than only incremental change is a far more interesting question and augmenting the research question in favor of a more easily discoverable finding seems problematic and, frankly, boring.

## 6.4. Dataset description

The final dataset contains data for 26 countries, 12 of which have data for the full 1991-2012 period (Chenoweth, Pinckney, and Lewis 2017). They are Algeria, Egypt, Kenya, Mexico, Morocco, Pakistan, Sierra Leone, Sudan, Syria, Turkey, Uzbekistan, and Yemen.<sup>26</sup> Aggregated by year, this gives a total of 442 observations.<sup>27</sup> The coverage is somewhat geographically biased – no South American countries are included, and the Middle East and North Africa (MENA) are overrepresented, comprising half of the countries in the dataset.<sup>28</sup> This means that I cannot reliably generalize my from this sample, as they are unrepresentative of the population of countries in the world.

Furthermore, I have amended the data substantially. The amendments consist largely of corrections for obvious coding mistakes, which were revealed by tabulating the *verb*-variables in the NAVCO 3.0 dataset. Mistakes such as observations coded with values that are absent from the codebook and therefore obviously coding mistakes, or categories with very few observations, were checked and re-coded where appropriate to values which I believe are correct.<sup>29</sup> It should have no negative impact on the data quality, as I think I have been conservative when deciding what to change and what to leave as-is.

The actual variables on economic nonviolent strategies that I use are those which can, in some way or another, theoretically represent an economic focus. Reviewing the list of the *verb*-variables in the codebook, I selected those that were somehow linked to material economic activity, like strikes, sanctions, property destruction, and aid. Additionally, I referenced the economic noncooperation-variable in the NAVCO 3.0 dataset and used it as a guideline for the selection. The 42 indicators are listed in Appendix 10.1.

The data available from DD originally stop in 2008, which limits the applicability of the NAVCO 3.0 strategy data. Many protest events take place after this, particularly in the

---

<sup>26</sup> Those not fully covered have data for these years: Bahrain (1994-2012), China (1991-92, 2010-12), Estonia (1991, 1999, 2002-2012), India (1991, 2011-12), Iraq (1999-2000, 2009-12), Jordan (1991, 1993-2012), Libya (1991-2000, 2002-2012), Madagascar (1991-93, 1995, 1997-98, 2000-06, 2009-12), South Korea (1992-96, 2012), South Sudan (2011-2012), Tanzania (1991-2006, 2008, 2010-12), Tunisia (1994-2012), Ukraine (1992-2012), and the USA (2007-11).

<sup>27</sup> It has been pointed out to me by one of the dataset authors that this aggregation does the many coders a great disservice. I agree but maintain that a considerable loss of information is a necessary trade-off when using a measurement model on event data.

<sup>28</sup> More than half if Sudan, South Sudan, and Pakistan are counted among them.

<sup>29</sup> The process of finding mistakes and recoding them produced an R-script of roughly 1000 lines of repetitive code which provides explanations for my changes and the code to repeat them. The process took more than a month to complete, as the original dataset is very large and the coding detailed. However, for the sake of veracity, causal utility, and reliable measurement, the process was necessary to get reliable results. The script with the suggested amendments has been sent to the authors of the dataset for review and will hopefully feature in a future errandum.

overrepresented MENA region. This is easily amended, however, as the coding rules for the DD dataset are commendably simple and easy to apply and so I have coded the four missing years myself.<sup>30</sup>

The variables which are interesting are those on regime transitions from autocracy to democracy, and vice versa. The coding of the years 2009-2012 supplied three additional democratic transitions, namely those of Madagascar in 2010, Tunisia in 2011, and Egypt in 2012 (the latter was reversed by a military coup in 2013<sup>31</sup>). This brings the number of transitions to democracy in the sample up to nine, and the number of regime transitions in total up to 12 (Madagascar 2009, Pakistan 1999, and Sierra Leone 1997 are reversals of previous democratizations).

The scarcity of transitions is not an issue for the analysis. As I am interested in whether those campaigns which utilize have a higher likelihood of becoming democratic than those who do not, I want to infer merely the direction of the coefficient (although quantifying the effect of economic ability in probabilistic terms is also interesting). With logistic regression analysis, I can estimate the probability of a democratic transition given high or low levels of economic ability, enabling me to answer the research question satisfactorily.

The four covariates which are not constructed using my measurement model (*gdp*, *aid*, *tax2gdp*, and *rent2gdp*) are all taken from the World Bank's Databank (World Bank 2019a, b, c, d). The reason for this is the accessibility and open access licensing of their data. One issue is, again, the geographically biased presence of data, as well as the degree of missingness itself. Most worrisome is the *tax2gdp* variable, with nearly half of the observations missing values, and *rent2gdp*, with a quarter of the cases missing information. However, the variables with the highest degree of missingness are also those which are relatively hard to come by. The information is not missing at random, according to a few quick tests from the *MICE* imputation package in R (van Buuren and Groothuis-Oudshoorn 2011), meaning that the missingness is correlated with the other variables in the dataset. To fix this I have run a quick imputation script to investigate whether the missing values matter very much. They do not, and the small differences in impact are discussed in chapter 7.2.

---

<sup>30</sup> I was assisted by a fellow MA student who showed me a coding device of his own making, which greatly simplifies R data entry. Anyone who would like to try it should download the function, available here: <https://github.com/Peder2911/ChildGotReeds>

<sup>31</sup> The Egyptian transition is also debatable as a proper transition to democracy according to the rules of the DD democracy definition. I include it because it exemplifies a transition and the fall of a long-standing dictator, and because it is difficult to say whether the Morsi government would have yielded in the face of an opposition election victory. Its inclusion does not affect the results of my hypothesis testing in any significant degree.

Table 1 describes the variables I use in my regression models. The first thing to note in Table 1 is that all the independent variables are standardized to have a mean of zero and a standard deviation of one.<sup>32</sup> This affects the substantive interpretation of the coefficients so that they are less intuitive but more easily comparable. Unstandardized covariates can be interpreted directly – an increase in GDP of one dollar increases the likelihood of democratization by some value of the coefficient, or a decrease of one percentage point in tax revenue-to-GDP lowers the likelihood of democratization by some coefficient. With standardized coefficients, this is no longer possible (unless the variables are transformed back). Now, the interpretation of coefficients is that an increase of *one standard deviation* in the covariate changes the likelihood by the value of the coefficient. Because the ability-score is not directly interpretable anyway, and it is the independent variable of interest, this loss of intuitive interpretability is inconsequential. For the control variables it is important, however, to point out that the coefficients no longer correspond to changes in the units, but in the number of units corresponding to a standard deviation.

**Table 1:** Descriptive statistics of dataset

<b>Variable name</b>	<b>Min / max</b>	<b>Mean</b>	<b>Median</b>	<b>SD</b>	<b>NAs</b>	<b>Variable description</b>
<i>tt</i>	0 / 1	0.020362	0	0.006	0	Dependent variable. Measures transition to democracy.
<i>ability</i> <sup>33</sup>	-1.219 / 3.43624	0	0.06079	1	0	Independent variable. Ability score for economic strategies-IRT.
<i>gdp</i>	-0.1960 / 8.9439	0	-0.1776	1	8	Control variable. Annual gross domestic product, current USD.
<i>aid</i>	-1.6938 / 6.2704	0	-0.3660	1	23	Control variable. Annual net aid received, current USD.
<i>tax2gdp</i>	-1.88314 / 4.02552	0	-0.05941	1	215	Control variable. Percentage of tax revenue to GDP.
<i>rent2gdp</i>	-0.72419 / 5.18779	0	-0.32815	1	110	Control variable. Percentage of natural resource rents to GDP.

The upside is that the coefficients are directly comparable. Unstandardized variables of GDP may make sense on their own, but which is the more important predictor: a 0.00002 size coefficient for a one-dollar increase in GDP, or a 0.006 coefficient for a one-percentage-point

<sup>32</sup> That is, I have subtracted the mean and divided by the standard deviation. A standard deviation is the average distance to the mean, or variance squared.

<sup>33</sup> This is the output of the Bayesian IRT model. The indicators used to estimate the measure are too many to summarize in the text here but can be viewed in Appendix 10.1.



increase in the tax-to-GDP ratio? Difficult to say. Standardized coefficients can be compared directly, as an increase of one standard deviation means the same for every variable. The coefficients come to represent the average change in the likelihood for democratization with an increase of a standard deviation, and thus weights the variables equally in a regression, rather than overestimating the importance of variables with high values, such as GDP or net aid received.

The *values* presented in Table 1 are worth discussing, as they reveal how they are distributed. The dependent variable, transition to democracy (*ttd*) is straightforward – it is a binary variable, recording whether a country-year saw a transition to democracy. It is right-skewed with only nine positive observations in the dataset and thus has a low mean and standard deviation.

*Ability*, my main independent variable and output theta-parameter from the IRT model, is skewed slightly to the left but is close to normally distributed. A high number of observations have *ability*-values near zero, however, and a few observations have very high or low scores – six observations have ability scores over three standard deviations from the mean, and 329 observations – roughly 74 percent – have scores which are equal to- or less than one standard deviation from the mean. A perfect normal distribution would have 68.27 percent within one standard deviation of the mean, which means that *Ability* is crowded in the middle and has thin tails.

Next, all the control variables are skewed to the right, i.e. that they cluster below the mean. This affects the *gdp* variable most severely, whose median is close to the minimum value. The maximum value is close to nine standard deviations from the mean, meaning the richest countries in the sample are extremely rich compared to the rest. Unsurprisingly, those countries with GDP above three standard deviations are China and the US in 2010-12 and 2007-11, respectively. The same pattern is clear in the *aid* variable, albeit less severely. Nine country-years have values over three standard deviations from the mean, and Egypt in 1991 is over six standard deviations from the mean. *Rent2gdp* is similar also, but the high numbers are spread out more, resulting in a less steeply sloped distribution. *Tax2gdp* is closer to a normal distribution than the other control variables but does have a high number of missing values (215) as nearly half of the observations do not have data on the percentage of tax revenue compared to GDP.

These skewed distributions could mean that removing outliers would be beneficial to the analysis. Removing extreme values such as the GDP of China and the US may reveal previously obscured and patterns in the data and accentuate interesting effects.

Additionally, it is worth mentioning that there are no considerable correlations between the variables in the dataset. The strongest correlation is between the *resource rent-to-GDP* ratio and the *tax-to-GDP* ratio, which is at .388. This is not a worrisome value but excluding one of the variables could potentially alter the results and illuminate other relationships in the analysis. The second strongest correlation is between *ability* and *aid*, at .237.

I further content myself with these five explanatory variables (at the most) in an attempt to limit the garbage can model-problems emphasized by Achen (2005) and (Schrodt 2014). They both phrase the issues better than I can, so I do not reiterate their points here. Suffice to say that overstuffing a regression model with covariates makes the results nearly impossible to analyze and investigate and that three covariates are usually neither too few nor too many to present a credible result.<sup>34</sup>

## 7. Results and discussion

My expectations for my models are as follows. First, I expect that the Bayesian item response model will yield enough variation with which to summarize the variables for economic tactics into a single variable. The undesirable result of no variation is equivalent to a regression coefficient being indistinguishable from zero. It is however possible to use a low-variation measurement model for data reduction; the central tendency of the estimates can vary a great deal despite high uncertainty. It is of course preferable to construct a variable in which the indicators have discrimination parameters which are distinguishable from each other, as this makes the specification of the model easier. If the discrimination parameters overlap, how can one separate those which contribute to the ability score and those who do not?

Second, I expect that economic ability is positively related to the likelihood of democratization. I believe the dependent variable will be negatively related to all my control variables: *gdp*, *aid*, *tax2gdp*, and *rent2gdp*. All are indicative of state capacity, albeit in different ways. This means that as I add control variables to the regression model, the effect of economic ability on the likelihood for democratization should shrink. I will investigate the relationships between the variables to better interpret the changes and supply various specifications of the model to tap it for information. Furthermore, I discuss model diagnostics which aid in the interpretation of the models.

---

<sup>34</sup> According to Schrodt, I would probably be better off if I combined the variables for tax-ratio and GDP into a stability-measuring variable and the rent-ratio- and aid-variables into a variable for economic dependence or something like that. I will however test models with only three covariates and hope that it will suffice for the methodologists.

## 7.1. Bayesian item response model

In this section, I report the findings on two of the parameters from the IRT model, namely the item discrimination alpha for every variable, and the ability score theta for every observation. The item discrimination of a variable describes how much it contributes to the resulting ability scores, i.e. how much of the variance in the data is explained by that single variable (Fox 2010, 114-115; Jackman 2009, 455). Higher item discrimination values therefore signify more important variables. Given that I have fixed the item *difficulty* parameter delta, both the alpha- and theta scores are constrained to a certain interval. This means that the numeric scores will be low for the alpha scores, but the size matters less than the slope. The expected result for the alpha parameters is that they are distinct from zero, and that they contribute differently to the variance explained by the latent ability score theta (Fox 2010; Jackman 2009; Kruschke 2010).

For the ability score parameter theta, which is the latent variable which I argue determines the likelihood of using economic strategies, I also expect a slope. Some nonviolent campaigns should be more skilled in economic strategies than others or have access to resources others do not. Higher theta values represent greater “economic ability.” Again, because of the fixed item difficulty parameter delta, the values are unimportant, and I have standardized the resulting values to have a mean of zero and a standard deviation of one in any case, so numerical values do not matter for the evaluation of the model.

I run the model as specified in chapter 5.1 above, with sampling parameters as follows: 1000 adaptive- and burn-in iterations 1.000.000 iterations, thinning every ten iterations, with two simulation chains,<sup>35</sup> and I use the Mersenne Twister pseudorandom number generator for sampling from the posterior distribution. The resulting alpha-parameter estimates are displayed in the caterpillar plot in Figure 6.<sup>36</sup>

The distribution of the discrimination parameter alpha shows that all indicators contribute to the model, and the indicators are distinguishable from each other to some degree. The seemingly low values on the x-axis result from the fixed difficulty-parameter. There is a significant deal of overlap between the indicators, which means that each variable does not contribute very much and that there is uncertainty about the location of the values. However, none of them have highest probability distributions (HPDs) which cross zero, which is striking

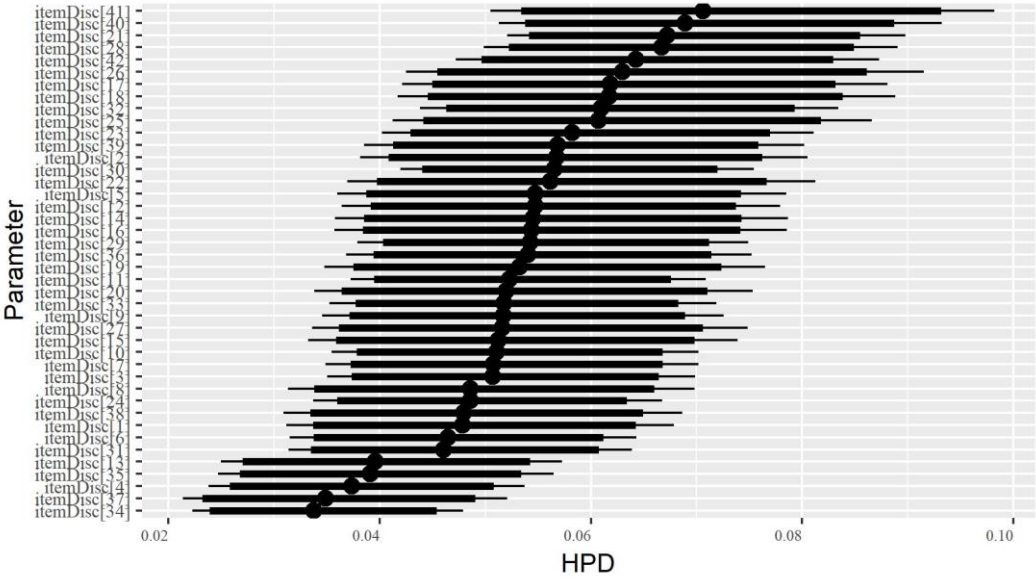
---

<sup>35</sup> Using two chains makes the process somewhat more reliable when the chains with different starting points converge on the same value. It does however require more computational power and time. The other sampling parameters for the simulation are chosen based on trial and error where I have found a combination which is not too slow, but which delivers the expected convergence.

<sup>36</sup> The labels for the parameters are prohibitively difficult to alter, and I have therefore listed the indicators and the corresponding number in Appendix 10.1. For example, itemDisc[41] is for the variable recording “confiscate property”-actions.

given that I specified the prior with the most likely location being very close to zero. The scores tend to become more uncertain the farther they are from zero, which may reflect that those who contribute more to the model are simultaneously those who deviate from the prior. Their posterior probability distributions may be “flattened” by the prior and given too high credible intervals relative to the strength of the data.

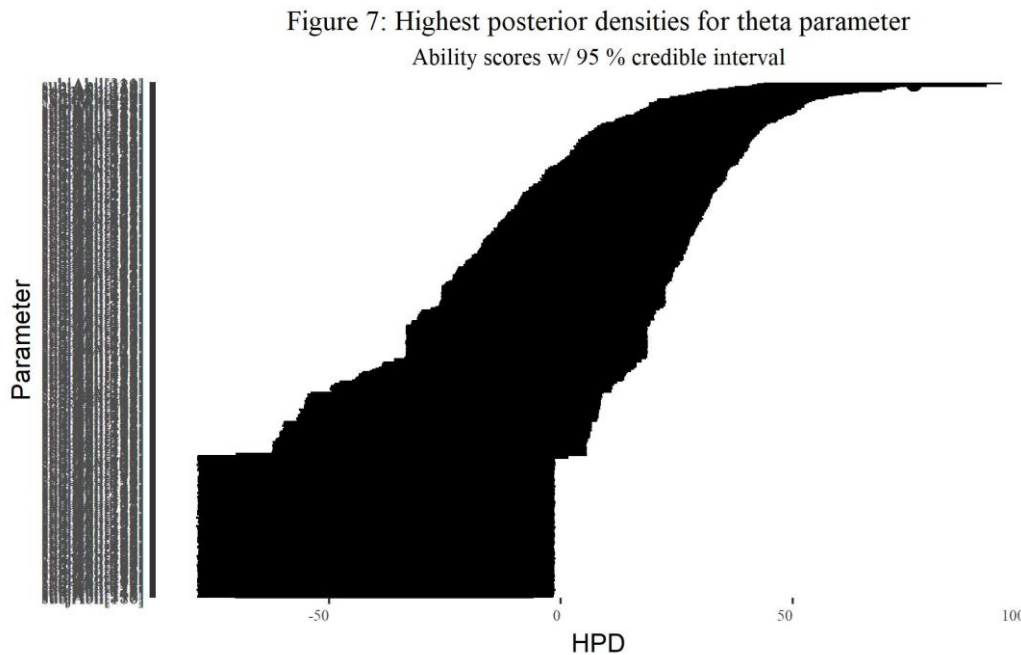
Figure 6: Highest posterior densities for alpha parameter  
Item discrimination scores w/ 95 % credible interval



It is worth noting that the Bayesian credible intervals are analogous to frequentist confidence intervals. The difference is that the credible interval incorporates information from the prior and that the credible intervals are fixed, not random (Jackman 2009, 26-28). The interpretation is similar, however, as the credible intervals represent the probable region where the value is located based on the data but does not assume that repeated sampling is possible or realistic.

The posterior distribution of the theta parameters is very interesting. It is displayed as a caterpillar plot in Figure 7 below. First, the slope of the scores is good, and the estimates are in fact different from one another. A uniform buzz around zero would mean that the scale was invariant, and therefore useless as an independent variable for hypothesis testing. Variation is required to explain variation, and while the number of transitions to democracy is low, it is all the more important to have some separation between low- and high-ability campaign-years to

explain that scarce variation. Note that the y-axis is unreadable because over 500 parameters are estimated.



Second, the fact that the lower values are more uncertain than the high ones is not odd, given the prior. It is possible that the same “flattening” by the prior occurs here, and that values that would be closer to the prior given stronger data are disadvantaged, and that the very low values are simply less certain in the data. Indeed, this is not unlikely at all, given the type of data I am using. There are many campaigns which use few economic tactics, and as such are difficult to place in relation to the others who do the same. Additionally, my aggregation scheme of counting the mere presence or absence of a tactic in a given year entails a loss of information and therefore contributes to the uncertainty in the ranking and scoring of the parameters.

The most important point to take away from the ability scores from the IRT model is that it is statistically feasible to construct a measure of strategy in this way and get meaningful results from the posterior. The resulting scale measures the degree to which campaigns use economic strategies in political contestation. Figure 7 shows that most receive low scores, and as such do not use or use only low-item discrimination tactics. Some campaigns use a large degree of economic strategies, and it is reasonable to infer that these possess skills and resources which allow them to do this, but which other campaigns lack. It is highly possible to use this constructed measure of economic strategy as the dependent variable in a regression to figure out which phenomena contribute to the differences. For example, are movements’ class composition linked to economic strategy? My measure of economic strategy may indeed be a

link relevant for Dahlum (2018), who investigates whether class composition and size of movements are linked to democratization. Likewise, my measure could, given more expansive data, constitute a causal, action-centered variable between the structural antecedents of Skocpol (1979) and Moore (1966), or an antecedent to the cost of repression-concept of Dahl (1971).

It remains to see whether there is a relationship between this score and the likelihood of democratization before these bold claims can be made. First, however, the scale should be mapped to some cases to evaluate the validity of the measure. Does the model make sense as an inference from observation to measurement?

I have argued above for the increased use of measurement models in political science, and for the NAVCO 3.0 variables being appropriate for data reduction through IRT. The validity of the measure should however be tested (Adcock and Collier 2001; Jackman 2008). An obvious approach for evaluating the validity of the resulting variable is to compare it to cases in which economic strategy has been used, and simply seeing if the numbers match the qualitative description.<sup>37</sup> However, few cases of prominent economic strategy exist on which to test the limited findings. This is a feature of the knowledge gap I attempt to fill: my research question is motivated by the lack of literature which treats attacks on state financial capacity.

A couple of examples can illustrate the validity, however. First, Nepstad (2011, 127) argues that the Kenyan campaign against the Moi dictatorship in the late eighties and early nineties succeeded in withholding material resources, a concept which is close to my concept of economic strategies. If the Kenyan campaign prohibited the state from using a set of material resources, some spike in economic strategy should be observed. However, the years are too early to be included in the dataset, as Moi announced elections in 1992 after facing resistance for several years, and most significantly so in 1990. The 1991 economic ability for Kenya is however not distinctively high. The peak years are 1997, 1998 (the year of transition to democracy), and 2008 (a year of intense protest after a disputed presidential re-election), ability scores being 1.5, 1.2, and 2.7 standard deviations above the mean of the dataset, respectively. This is however a weak test, as Nepstad (2011, 101-102) puts the height of resistance before the dataset starts recording.

A second example which provides a better test is the Occupy movement in the US in 2011 and 2012. Here, thousands gathered all over the country to protest inequality, faux democracy, and capitalism. The goals of the campaigns were famously vague, and their methods directionless: Gene Sharp (2011) criticizes the movement for this, and stated that

---

<sup>37</sup> Comparing it to previous quantitative measures, i.e. a test of convergent validity (Jackman 2008), would be better. I am however not aware that any such measures exist.

”They don’t have any specific demands or a clear objective” and “If they think they will change the economic system by simply staying in a particular location, then they are likely to be very disappointed.” This is hardly controversial; the Occupy movement has so far had little impact on the politics or economy of the US. However, the year 2011 stands out as a particularly strong year for economic strategies in the US, scoring 1.5 standard deviations above the mean. This is a large increase from the other years and provides a stronger test of validity.

Thirdly, Tanzania is an interesting case in which protest has usually been high in election years. My measure of economic strategies counts 1 and 1.1 standard deviations above the mean in 2000 and 2005, which would seem to weakly validate my measure. Every other observation for Tanzania in the dataset has values either close to zero or below the mean.

A line plot of the economic ability scores  $\theta$  for Kenya, Tanzania, and the US is visible in Appendix 10.4, showing the trends described above. These three cases of convergence between my measure and the qualitative examples confirm the validity of the measure weakly. Despite wanting to test the validity of the measure more formally with other measures, these convergence-tests and the content validity of the concept as discussed in chapter 4 is all that is available to me (Jackman 2008). Closer knowledge of the countries covered would supplement the convergent validity, but I believe the content validity is sufficiently strong to justify the limited tests of convergence-reliability (see Adcock and Collier 2001).

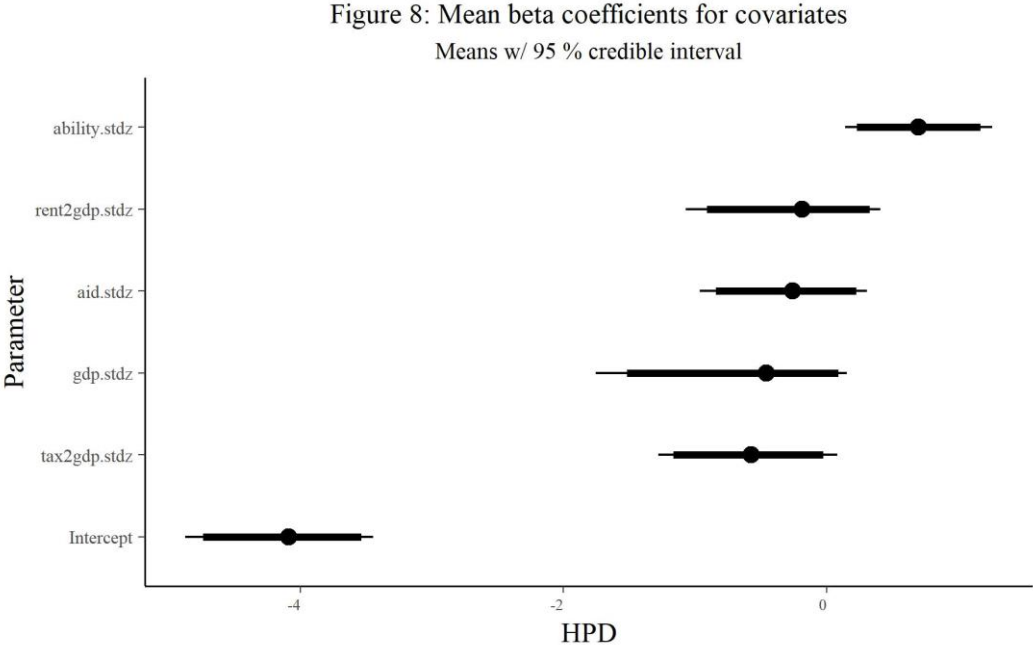
It is important to keep in mind that the scores are relative to every other observation and variable included in the dataset. Removing one of either makes the entire score different because every cell is involved in estimating the parameters. An expected high value relative to a single country’s history may not be high on the score at all, because it may be comparatively low in the entire dataset. This is because the IRT takes the entire data matrix as its starting point and does not account for variation in time and space; the country name- and year-labels are not part of the IRT analysis. Each observation does get its appropriate score which likely *does* make sense, but they are not absolute and do depend on the rest of the matrix.

## 7.2. Bayesian logistic regression

My expectation is that economic strategies should be positively related to changes in the likelihood of democratization. To test it, I use a JAGS model for Bayesian logistic regression. The powerful graphics packages which take JAGS-outputs as their input is the main reason for this, as several packages provide basic Bayesian logistic regression. They demand much less work but are harder to manipulate and often come with sub-par graphical functions. The *ggmcmc* package from Fernández-i-Marín (2016) makes great use of the *ggplot2* package and

supplies brilliant and easily interpretable graphics.

I run the model with 100.000 iterations, 1000 adaptive- and burn-in iterations, thinning every five iterations, two chains, and with the Mersenne-Twister random number generator. The parameters of interest for the hypothesis testing are all the beta-parameters in the analysis, i.e. the coefficients for each of the variables plus the intercept, along with their standard deviations as a measure of uncertainty. The results for the beta parameters are displayed in Figure 8.<sup>38</sup>



The intercept at -4.1 shows that there is very little chance of a country democratizing in the first place. This is expected because there are only nine transitions to democracy recorded in the 442 country-year dataset. All the control variables are negative at the mean, but all reach positive values by the 97.5<sup>th</sup> percentile; the standardized *rent-to-gdp* ratio is just positive at the 75<sup>th</sup> percentile. Increases in the control variables then decrease the likelihood of democratization, in accordance with the hypotheses. The fact that the *tax-to-GDP* ratio is the most acutely negative control variable may indicate that it is a proxy for government capacity rather than responsiveness to citizen demands. The same theoretical expectations are true for the other variables; increases in resource rents, foreign aid, and GDP decreases the likelihood of democratization on average.

Notably, the richest countries in the sample are the US and China, the former of which is already a democracy, the latter famously stable (although the same was true of GDR in the

<sup>38</sup> The shapes of the posterior distributions of the beta-parameters are visible in Figure 16, chapter 7.4.3.



late 1980s). Removing the ten observations in which the US and China are recorded changes the sign of the GDP coefficient and gives a mean coefficient size of 0.335. The variable still has credible intervals that cross zero, and the removal of the outliers does not affect the ability score's coefficient noticeably (from .7003 with outliers to .7063 without). The other covariates are similar, too. The outlier-free model does however conform to the Most General Finding (*sic*) in political science, namely that democracy and wealth are positively correlated (e.g. Acemoglu and Robinson 2012; Bernhard 2016; Bollen 1979; Diamond 1992; Fukuyama 2016; Inglehart and Welzel 2005; Lipset 1959; Przeworski et al. 2000).

The ability-score from the IRT model, which measures economic strategy, is consistently positive, however. The credible interval does not cross zero, which means that any amount of economic strategy increases the likelihood of democratization. The mean coefficient is 0.7, and it reaches 1.26 at the 97.5<sup>th</sup> percentile. This is a considerable size, which confirms the hypothesis and answers the research: *There is a positive relationship between economic strategies and the likelihood of democratization.*

**Table 2:** Odds for covariates

	<b>Econ. Ability</b>	<b>Rent-to-GDP</b>	<b>GDP</b>	<b>Aid</b>	<b>Tax-to-GDP</b>
<i>Odds</i>	2.0983736	0.8560520	0.6413993	0.7999677	0.5943638
<i>(SD)</i>	(0.6118742)	(0.3024607)	(0.2690986)	(0.2529224)	(0.2088025)

Translated into odds, the coefficients look like in Table 2. Economic ability contributes significantly to the likelihood of democratization, with a change in odds of nearly 2.1. This means that an increase in economic strategy of one standard deviation more than doubles the chance of a country experiencing a transition to democracy in the same year.

The odds of the control variables are all below one, meaning that increased control variable values lead to decreased likelihoods of democratization. Most of the odds for the control variables are not very strong however: the tax-to-GDP ratio, the strongest of the covariates, reduces the likelihood of democratization by forty percent; GDP by thirty-five percent.

The impact of imputation is small.<sup>39</sup> The odds rate for the economic ability variable changes from 2.1 to 2.04, which is negligible. The credible intervals of the control variables do shrink however and make all consistently negative. The means do not change noticeably, but simply explain more of the variation which in the non-imputed version is accounted for by the

<sup>39</sup> The imputation scheme I used is the standard *mice* command, with m=16, maximum iterations = 100, and the predictive mean matching (pmm) method. The imputation process could be specified very much more appropriately but suffices to demonstrate that the missing values matter little.

ability score.

Furthermore, ANOVA-testing several specifications of the regression model with different combinations of the covariates yield only slight differences in explained variance. The model with all five variables has 70.71 percent residual deviance, while every other specification has a little more. None of the differences are significant in a frequentist ANOVA test between the models however, which suggests that the control variables do not contribute much to the explained variance. This means that the most parsimonious and appropriate regression model for the research question may be the one with *only the ability-variable* and that the control variables are only complicating the result. This does seem to be the case, as the coefficient and credible intervals change very little both for the intercept and the *ability* variable.

In the model with only *ability* as a predictor, the coefficient for the variable takes a value of 0.6573, which is a small decrease in size. This means that the control variables make the relationship stronger, signifying that some of the variation in the dependent variable is explained by them. The fact they strengthen the effect of *ability* accords with their negative coefficients: the control variables contribute mainly to predicting non-transitions, as they correlate negatively with transitions to democracy. They thus predict mainly zeroes, while *ability* is positively correlated with transitions and is more closely related to the nine transitions in the dataset.

Furthermore, I have performed a frequentist logistic regression of my model, presented in Appendix 10.2. It displays precisely the same trends as the Bayesian one, albeit with slightly different values for the coefficients. In the frequentist regression, the *ability*-coefficient is stronger than in the Bayesian one, and it is significant at the five-percent level. Additionally, the Hosmer-Lemeshow goodness-of-fit-test provides no evidence of poor fit for the Bayesian regression model, with a p-value of .77 and  $\chi^2$ -value of 4.9. The test is vulnerable to the number of groups specified (Hosmer et al. 1997), but altering the number of groups makes no substantial difference to the result.

It seems that the logistic regression has yielded results which are both stronger and more robust than I expected them to be. The research question is answered positively, and for the limited post-Cold War data sample available there has been a marked effect of economic strategies on the likelihood for democratization. Increased use of economic strategies equal to one standard deviation more than doubles the likelihood of a transition to democracy.

By viewing the cases in which transitions have occurred, it is easy to see that none of the cases display very low scores on the economic strategy measure. Table 3 lists the country-

years in which a transition to democracy has occurred. Every transition except that in Madagascar in 1993 has economic ability scores above the standardized mean of zero. Five of the nine transitions have scores above the 3<sup>rd</sup> quartile or 0.730, and none are below the 25<sup>th</sup> quartile of -1.19.

**Table 3:** Transition years and standardized economic ability scores

Country	Year	Economic ability
<i>Madagascar</i>	1993	-0.740
<i>Sierra Leone</i>	1996	0.356
<i>Kenya</i>	1998	1.189
<i>Sierra Leone</i>	1998	1.627
<i>Mexico</i>	2000	0.273
<i>Pakistan</i>	2008	1.568
<i>Madagascar</i>	2010	0.347
<i>Tunisia</i>	2011	2.057
<i>Egypt</i>	2012	2.568

Obviously, this relationship is a probabilistic one. Nothing in the analysis suggests that a high level of economic strategic action necessarily makes a state transition to democracy. But the high increase in likelihood still testifies to the potency of the relationship, and that economic strategies may be a good strategic choice for social movements.

These findings suggest that an action-centered approach to democratization may provide a link between structural or actor-centered approaches and regime characteristics. The effect of economic strategy on the likelihood for democratization is on a less abstract level than many previous analyses such as Dahl (1971), Moore (1966), Skocpol (1979), Przeworski et al. (2000), and may provide links between democracy and its antecedents previously unexplored quantitatively. My analysis thus provides valuable insights and fills a knowledge gap for both the democratization literature at large and the nonviolence literature in particular.

### 7.3. Discussion

My findings are important for several reasons. First, they speak to the literature on nonviolent strategy and democratization and fill an important knowledge gap. Second, my use of item response theory exemplifies the applicability of measurement models for data reduction in situations where the available data are very fine-grained. Third, the findings show that Bayesian methods for hypothesis testing are accessible, easily interpretable, and more appropriate than frequentist methods when dealing with inherently unrepresentative data whose repeated

collection is impossible. Fourth, I supply potentially valuable information for activists by providing a probabilistic framework for what to expect when using different strategies to democratize a country. Fifth, my conceptual discussion in chapter 4 enriches the interpretation of the results by situating them in a broader context and provides grounds for new veins of research.

First, the literature on nonviolent strategy has found important and reliable relationships between strategy and outcomes of contentious politics. The best of these findings is that provoking defections from the security forces is an almost sure-fire way to win such a conflict (e.g. Chenoweth, Perkoski, and Kang 2017; Chenoweth and Stephan 2011; Nepstad 2011, 2015; Stephan and Chenoweth 2008). In essence, this speaks to the causal utility inherent in Weber's definition of a state as the organization with a monopoly on legitimate violence (Weber 1958). This violence-monopoly is the state's foremost source of power. I have argued that it is likely effective to target its other sources of power. With reference to Schumpeter's definition of a state as the organization with a monopoly on legitimate taxation (Schumpeter 1991[1919]) – arguably the source of violent power in the first place – I have found that states are also vulnerable to attacks on their economic capacity. Nepstad (2011), using a comparative design, found no deterministic pattern between the removal of material resources from a state and nonviolent success. I have found a strong probabilistic one.

This is indeed a novel finding. No analysis of which I am aware has tested a similar hypothesis with quantitative methods as I have, and none have attempted to summarize the characteristics of nonviolent campaign strategies with measurement models. These approaches are fruitful, reproducible (with enough computing power...), and relevant. The NAVCO dataset is indeed new, and so the level of detail in the data collection efforts and the work it must have taken cannot be exaggerated. The new dataset has obviously not been analyzed to exhaustion yet and provides a great point of departure for just the kind of novel, knowledge-producing efforts political science must produce.

Crucially, the nonviolence literature should increasingly focus on strategy and the concrete actions which promote democratization. I have made one step in that direction, but scarcely enough to rely upon for anyone risking death or severe repression. The relationship between strategy and state action must be vigorously investigated both quantitatively and qualitatively to be able to write blueprints for democratization. Sharp (2012) has done most in this regard. His advice is more general than that which I can give. However, the nonviolence research community should pay attention to strategy both in the aggregated case, as I have, and in the very concrete sense, investigating individual cases, interrelations between different

strategies and tactics and their efficiency.

Additionally, my findings identify a possible vein for new research, namely causes of different strategies. Different strategies may spring from different structural factors or actor compositions, so my findings may identify a link between these antecedents and democratization. This is not trivial and can inform future research on both nonviolence and democratization.

Second, I have used the NAVCO 3.0 dataset in a novel and interesting way (Chenoweth, Pinckney, and Lewis 2018). Despite the various shortcomings of the data (which the authors themselves highlight in the accompanying article, and which I have amended to a degree), it is unequalled in detail. While I have made short work of that detail with my aggregation schemes, I believe that such fine-grained event data are very attractive for use in measurement models. The verb-variables from the NAVCO 3.0 dataset provide a great starting point for quantitative analyses of strategy, and I believe my approach is a fecund one to be developed further with other techniques, datasets, and hypotheses.

Third, making quantitative inferences about macro-level phenomena with data which are neither random samples nor experiments with Bayesian is both appropriate and accessible. Many have pointed this out before me and highlighted the potential gains of being armed both with frequentist and Bayesian tools of analysis (e.g. Collier, Brady, and Seawright 2010; Jackman 2001, 2009; Schrodt 2014). In my analysis, it is meaningless to ask how likely it is that the data are anomalous compared to the population. The NAVCO 3.0 authors write themselves that the data are drawn from particularly eventful countries and make no pretense to generality (Chenoweth, Pinckney, and Lewis 2018). It is also uninteresting to formulate a null hypothesis with which to compare such an anomalous sample. While the nonviolence-literature is methodologically sophisticated compared to much of contemporary political- and social science and we are acutely aware of these points, it is worth reiterating to justify my choice of methods.

Fourth, my findings are potentially important for people wishing to overthrow dictators. The empirical fact that in my dataset the odds of democratization doubles when using economic strategies is valuable information to democrats. Given that they have the resources and skills to do so, it is advisable to attempt to target the state's economy in one way or another. Strikes, blockades, divestment, economic sanctions, and property destruction are, in some here-unexplored combination, tactics which increase the likelihood of democratization. What the resources and skills needed to perform these types of actions are, I have not explored. There are no notable correlations between economic ability and my control variables save that net aid

received has a 23.5 correlation coefficient. What causes economic ability can be explored by future research.

I also think it is very likely that the efficiency of economic strategies of nonviolence is contingent upon other factors which I have not accounted for. For one thing, which strategies the state employs to counter its opposition is bound to be important. As mentioned in the introduction, the Panamanian dictator Noriega resisted protests consisting of a third of his subjects. This is probably not attributable only to his involvement with drug smugglers and foreign autocrats, but to some more concrete action taken. It is unlikely that the Kapp putsch could withstand protest for very long given the enormous opposition from all parts of society, but perhaps the coup could have lasted longer with some state-strategy. And no doubt the East German regime could have done more to dissuade emigration, even without abandoning its socialist principles entirely.

Moreover, different structural factors are likely to affect my findings. Many things other than the revenue sources of states differ in my sample of states. They all have different political regimes, religious and ethnic make-ups, climates, foreign relations, histories, colonial heritage, and many other differences. (This is partly the reason why a subset of states is not representative of the state population.) Different analyses of different such relationships are necessary to create a reliable theory (*sic*) or even a generalizable relationship of economic nonviolent strategy and democratization. My results should be interpreted with caution by anyone seeking to practice nonviolent action in an autocratic context.

Fifth, my findings provide support for an action-focused vein in quantitative political science which bridges much of the theory and empirical findings on democratization, nonviolence, and regime stability. My conceptual discussion has situated the findings in a contentious politics-context and my theoretical discussion has provided arguments for why and how economic nonviolent strategies should affect democratization. Combined with the theoretical discussion on state behavior- and stability, my approach thus takes a step toward an action-perspective to connect structural, institutional, strategic, and actor-centered explanations of transitions to democracy. This advances the logical economy in the field and encourages researchers to move a step down the ladder of abstraction and look more directly at *what is being done* by activists. It also brings the quantitative knowledge closer to the qualitative literature by focusing on variables which are particular for every case yet generalizable, and which lie close to the concrete causal mechanisms of change.

## 7.4. MCMC model diagnostics

It is necessary to investigate whether the simulation-based methods have converged and produced reliable results. The most important form of model diagnostics for Markov Chain Monte Carlo (MCMC)-based simulation models is convergence. The model specified above samples from the posterior distribution and estimates the likelihood that the data exist in a certain range. This simulation of samples uses a random number generator to a ‘random walk’ through the posterior distribution, and therefore needs to run for many iterations in order to converge on a likely HPD. Badly specified models may take a very long time to converge, although what constitutes a sufficient number of iterations depends on the model, data, and parameters to be estimated (Jackman 2009, 187-190; Kruschke 2010, 108-109).

Convergence is usually evaluated with plots, and if there is no evidence of non-convergence the model is thought to have run for long enough. Showing the convergence plots for every parameter of the analysis is unfeasible however, as there are over 532 different parameters, two chains estimating each parameter, and several types of convergence plots to evaluate them with.<sup>40</sup> I will show a subset of these to illustrate the general finding that there is no evidence of non-convergence in my model, and every parameter seems to have found its mark. The rest I have saved and can be supplied upon request. The R-package I use to evaluate convergence, *ggmcmc*, is versatile and user-friendly, and communicates very well with the popular *ggplot2*-package, the go-to plotting tool for many R-users.

I first evaluate all the parameters with a Geweke diagnostic and an R-hat test, both of which are formal tests of convergence. Formal, in this context, means that they are based on actual calculated values from the MCMC chains, not on a visual inspection of them. Neither formal nor informal convergence tests are sufficient to test convergence on their own but are complimentary. Second, I evaluate the alpha- and theta parameters with informal tests, such as running means, trace plots, autocorrelation plots, and density plots.

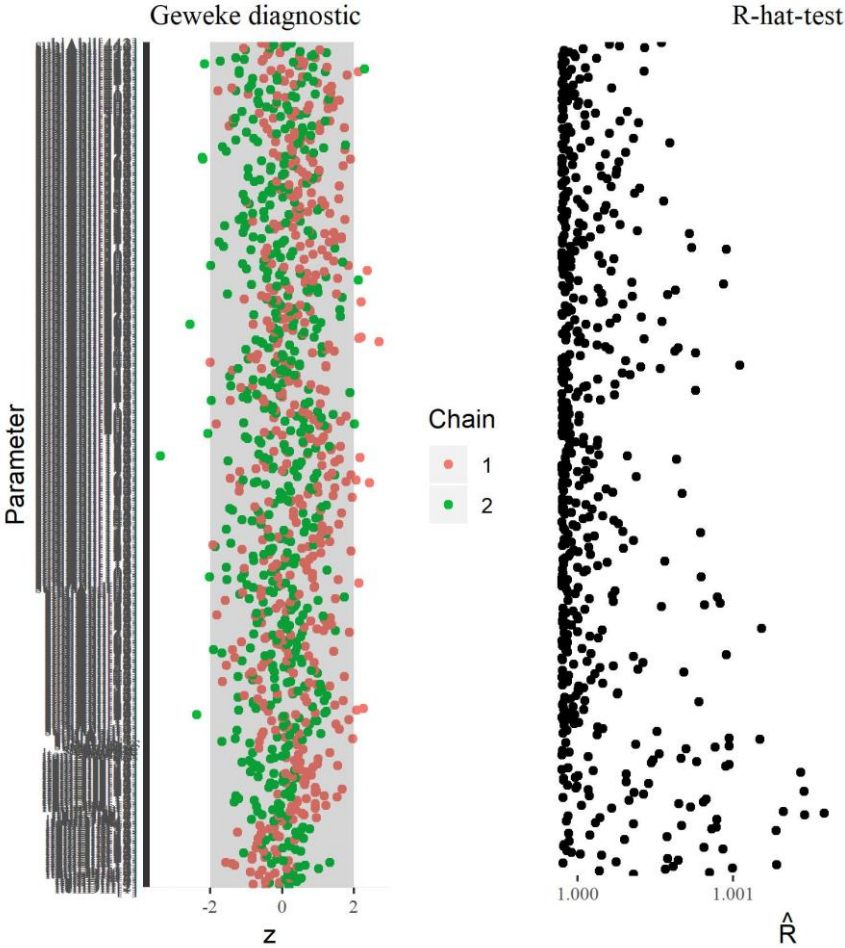
The Geweke z-score diagnostic (Fernández-i-Marín 2016, 9-10) in Figure 9, left panel compares the first and last parts of the MCMC chains and reports the z-score for the estimated parameter means. The goal is to have no more than 5 percent of the estimates be outside the shaded  $-2/2$  interval. This helps identify troublesome chains which do not conform to the

---

<sup>40</sup> The *ggmcmc*-package potentially produces a pdf of 637 pages containing multiple graph for every parameter. It is tremendously helpful, but it requires a selective presentation. Furthermore, because this selection demands a specification of which parameters to display in the form of a “regular expression” in R, it is necessary to show ten at a time. When showing only five, I can unfortunately only pick from single-digit parameters, i.e. the first nine. When showing ten parameters at once, I can select whichever set of parameters I want which start with the same two digits (e.g. `theta[350-359]`). This somewhat limits the presentability of the results, and diagnostics for the first nine (e.g. `alpha[1-9]`) parameters will be somewhat overrepresented.

expectations (Fernández-i-Marín 2016, 9-10). The plot shows thirteen red dots and eleven green outside the preferred interval, which means that only about 2.25 percent of the 1064 parameter estimates are off.

Figure 9: Geweke and R-hat tests



The Geweke-diagnostic is sensitive to specifications, however. Over several model trials,<sup>41</sup> I encountered values that did not conform to the expectations, usually with a family or two of parameters showing diverging results in one of the chains. The specifications in question are those of model run-time, such as the number of iterations in the simulation. The final model, which I ended up using because of the promising convergence diagnostics is the one shown in the graphs and whose iteration-values are described above in Chapter 6.1.

Similarly, the potential scale reduction factor ( $\hat{R}$ , or R-hat) compares the between-chain with the within-chain variation. If the result is 1 or very close to 1, all is well. A dot plot of R-

<sup>41</sup> While running multiple models to see which one performs best is often a sign of bad science, this is not the case with MCMC simulations (Jackman 2009). It is necessary to evaluate the performance of the simulation post-hoc to be confident in the results, and the practice is expected and not equivalent to running “[...] dozens – or more likely hundreds – of alternative formulations of the estimation” which Schrodt laments (2014, 287).



hat for every parameter is shown in Figure 9, right panel. Note that the scale of the X-axis is fitted to the data so that the largest  $\hat{R}$ -value is only about 1.003. With the default scaling (x-axis goes from 1 to 1.5), the dots make an almost straight line at  $\hat{R} = 1$ , which is the expected result. Based on these diagnostics, it is safe to say that there is no dire risk of the simulation not having converged. The R-hat test is very much satisfactory, and the Geweke shows that the estimates' z-scores conform to expected values. It is also useful to informally investigate whether a sub-selection of parameters demonstrate the expected patterns. In the next two sections, I check whether some of the alpha- and theta parameters display evidence of non-convergence or if they seem fine.<sup>42</sup>

#### 7.4.1. IRT model: Item discrimination parameter alpha

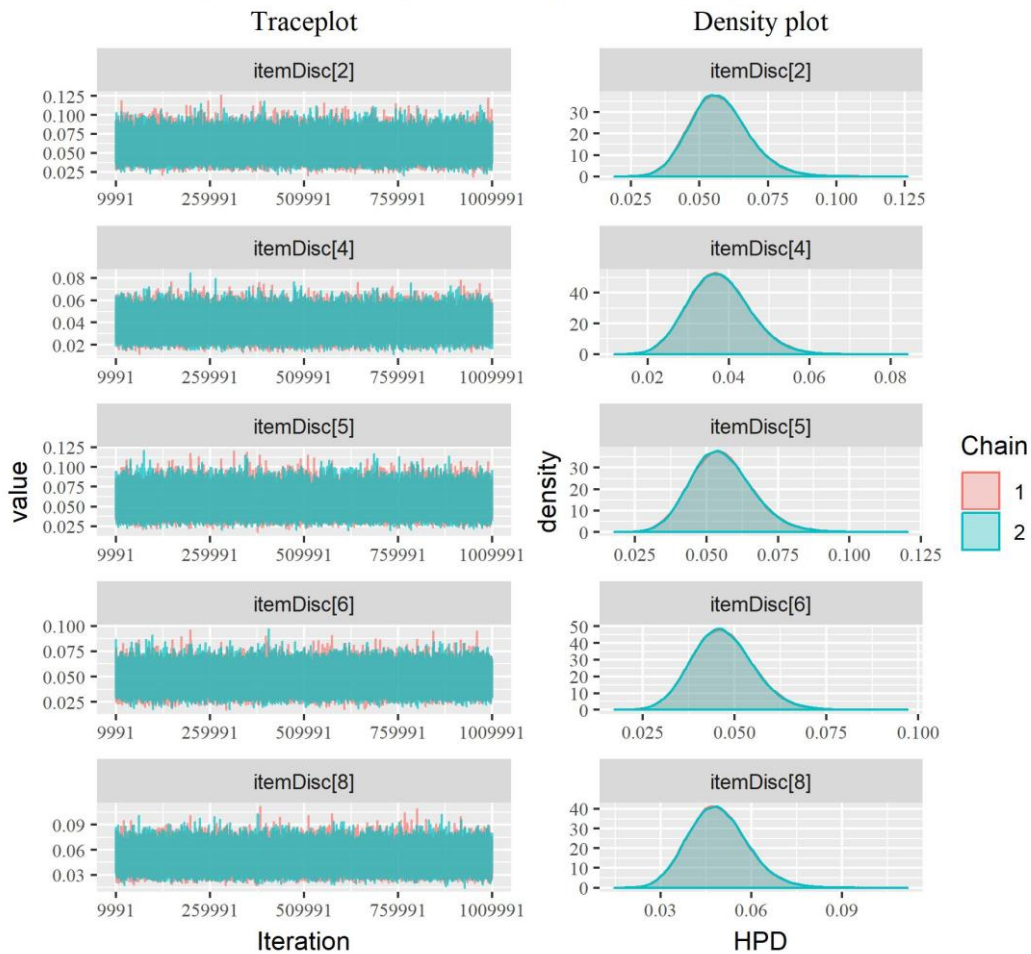
There are 42 indicators of the economic ability measure, each of which has an estimated discrimination parameter alpha. Figure 10, left panel shows a time series of the simulation process by iteration for five of the parameters, or a trace plot, with their estimated value on the Y-axis. It shows the mixing of the five selected parameters for the two chains and reveals that the mixing is quite good. The goal here is to have a uniform, pattern-less buzz, akin to a scatterplot to test for heteroskedasticity in the residuals of a regression analysis. (Or the look of a hairy caterpillar, as Funk, Camacho, and Johnson (2018) put it). If the chains are flat lines or move in the same direction over many iterations, it is a sign of bad mixing and an ill-specified model. There is no evidence of that here. (If Markov chains enter a “closed state” from which it cannot leave, i.e. a flat line in the trace plot, it is a violation of the irreducibility-assumption (Jackman 2009, 179-180).) Note that the first 9,990 iterations are gone – these are the burn-in- and adaptive iterations. The fact that the visible iterations start and end at approximately the same values means that the number of discarded iterations is not too low.

The right panel of Figure 10 shows the density plots for the same parameters. The expected result of the density plots is to have the chains overlap as much as possible. Large deviations and mismatched shapes of the posterior distribution between the chains is a sign of model misspecification and non-convergence in the chains. Neither deviations nor mismatched shapes are visible here, however. Plots for items 4, 5, and 6 show some *very* slight differences between the chains, but this is far from worrisome. It is clear from the graphs that there is no evidence of non-convergence in the simulation chains.

---

<sup>42</sup> Here I want to thank Martin Tegnander at PRIO, who gave me a whopping 96 GB of memory to create these plots from my huge dataset. The convergence diagnostics would not have been so complete without his help.

Figure 10: Density- and traceplot for alpha parameters



The discrimination parameter's running means for five of the indicators are shown in Figure 11, left panel. The figure indicates that there is little evidence of non-convergence in the alpha parameters. Every red and blue line showing running means stabilizes around the black line signifying the overall mean. Additionally, the speed with which the running means center around the mean is quite high, signifying that there are few issues with the model specification. Some parameters take longer than others notably number seven, but none show large deviations after about 26,000 iterations.

Figure 11: Running means and autocorrelation plots for alpha parameters

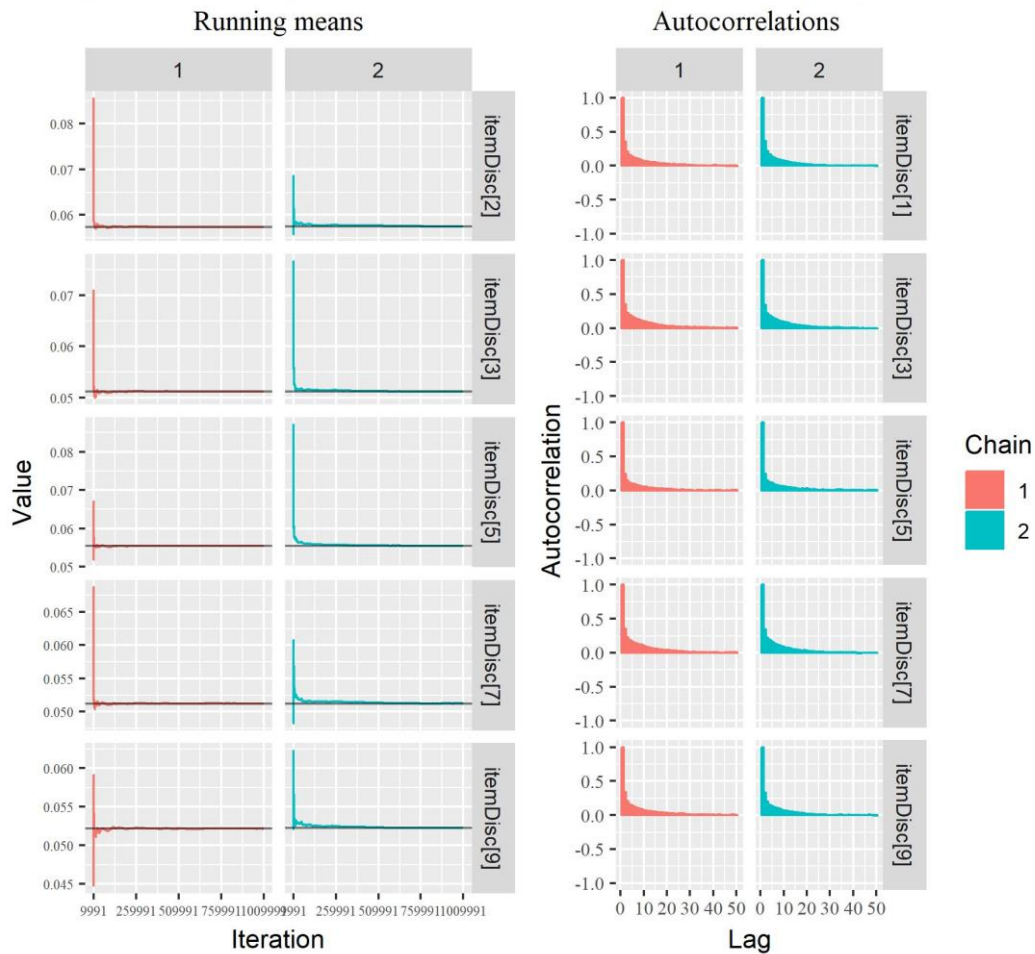


Figure 11, right panel shows autocorrelation plots of the simulation process for the same five alpha parameters. The values in the plots are the average autocorrelation between the parameter and the parameter  $x$  lags later. The goal is decreasing autocorrelation as the lags increase, ideally with no autocorrelation after lag 1 (Fernández-i-Marín 2016, 8; Jackman 2009, 268). These are all pretty quick to reach very low values of autocorrelation. None have notable autocorrelation values after about 15 lags. This means that the chains are mixing reasonably fast and that the correlation between the individual samples is low. It is not ideal, but autocorrelation is not a devastating problem in MCMC simulation theory (Jackman 2009).

It is possible to get even greater reductions in the autocorrelation after lag 1, for example by increasing the number of samples discarded by the thinning parameter. This hurts the estimation somewhat, as all discarded information does. It can however help with reducing the storage space and memory used. Thinning is very much a trade-off between efficiency and sample size and should not be done without caution. I therefore leave my sample as it is, with thinning every ten iterations and 1.000.000 iterations. Any increased number of iterations is

unfeasible because of the time it takes, and any additional thinning is unnecessary.

Based on these graphs is it safe to say that the alpha-parameters have been simulated thoroughly enough. The MCMC process is credible and satisfactorily executed to reflect the posterior distribution of the alpha parameter. Except for the autocorrelation test, which could be a bit better, the informal convergence-tests for the alpha parameters display no evidence of non-convergence. While there are arguments to be made for a more selective choice of indicators, I believe that my decision to use every indicator concerning economics of some sort is valid for the exploratory and novel way of synthesizing strategy.

#### 7.4.2. IRT model: Ability score parameter theta

For each of the 442 campaign-years in the aggregated sample, a theta parameter is estimated. Figure 12<sup>43</sup> left panel shows density plots for the two chains for ten of the theta parameters. As was the case with the alpha parameters the differences between the chains are small. None of the ability scores show divergence which is worth mentioning. This again means that no sign of non-convergence is discernible and that the model has run for long enough.

The right panel of Figure 12 compares the densities of the entire chain for each parameter with the last ten percent of the chain. The parameters are the same as in the left panel but are divided into columns by which chain is examined. It is very easy to see which parts of the chain affect the posterior distribution in which direction: for instance, for theta parameter 354 it is the entire first chain that creates the little red bump at approximately value -35. The second chain tilts slightly to the right of this value, as does the last ten percent of the first chain. When large deviations occur, these easily comparable visual aids can help identify errors in different parts of the simulation.

With both the density plot and the partial comparison plot however, the case is the same: there is no evidence of serious non-convergence. None of the graphs show serious divergence from its comparable companion. All the parameters seem to have converged nicely, independently of which chain and which part of the chain is examined.

---

<sup>43</sup> The ticks and labels on the Y axes are removed from the entire figure to save space. The densities themselves are not important however, only the overlapping shapes of the figures. This is because the two plot-types show the same information, partitioned differently. The densities across the graph are therefore an uninteresting quantity for evaluating convergence.

Figure 12: Density- and partial comparison plots for theta parameters

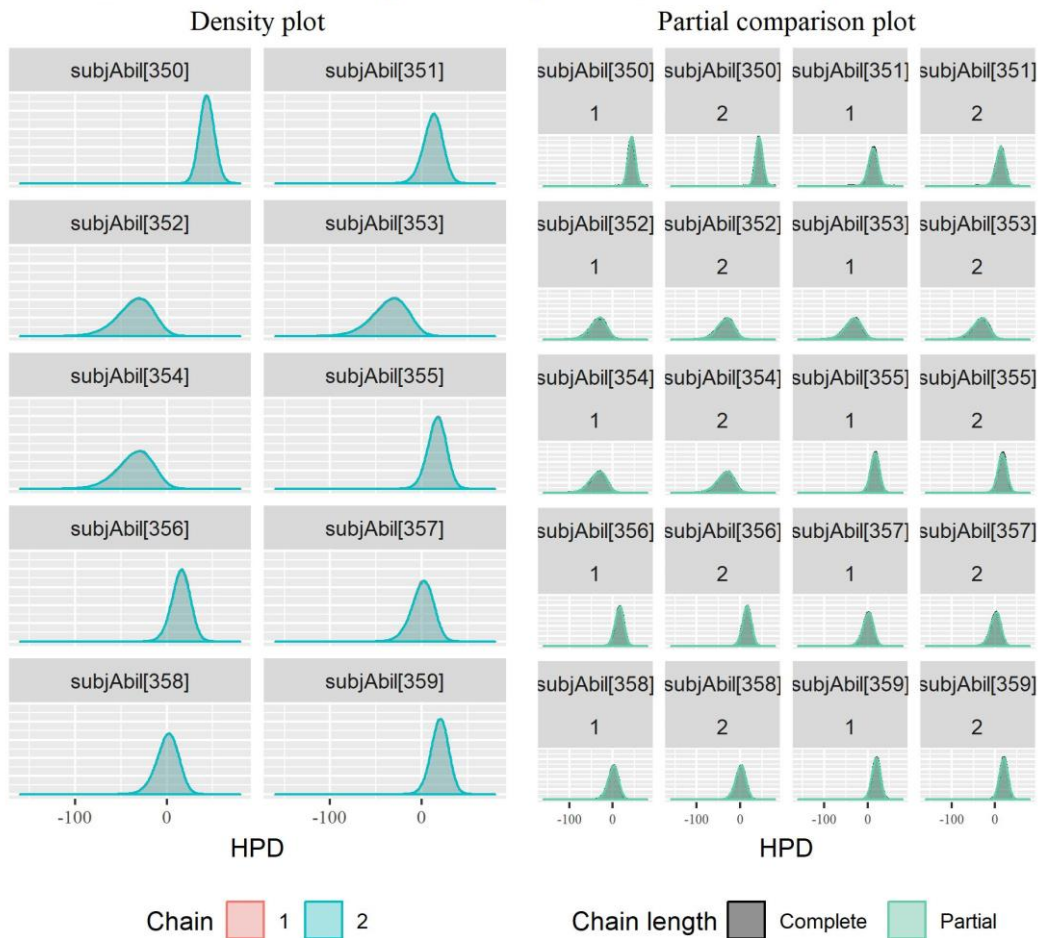


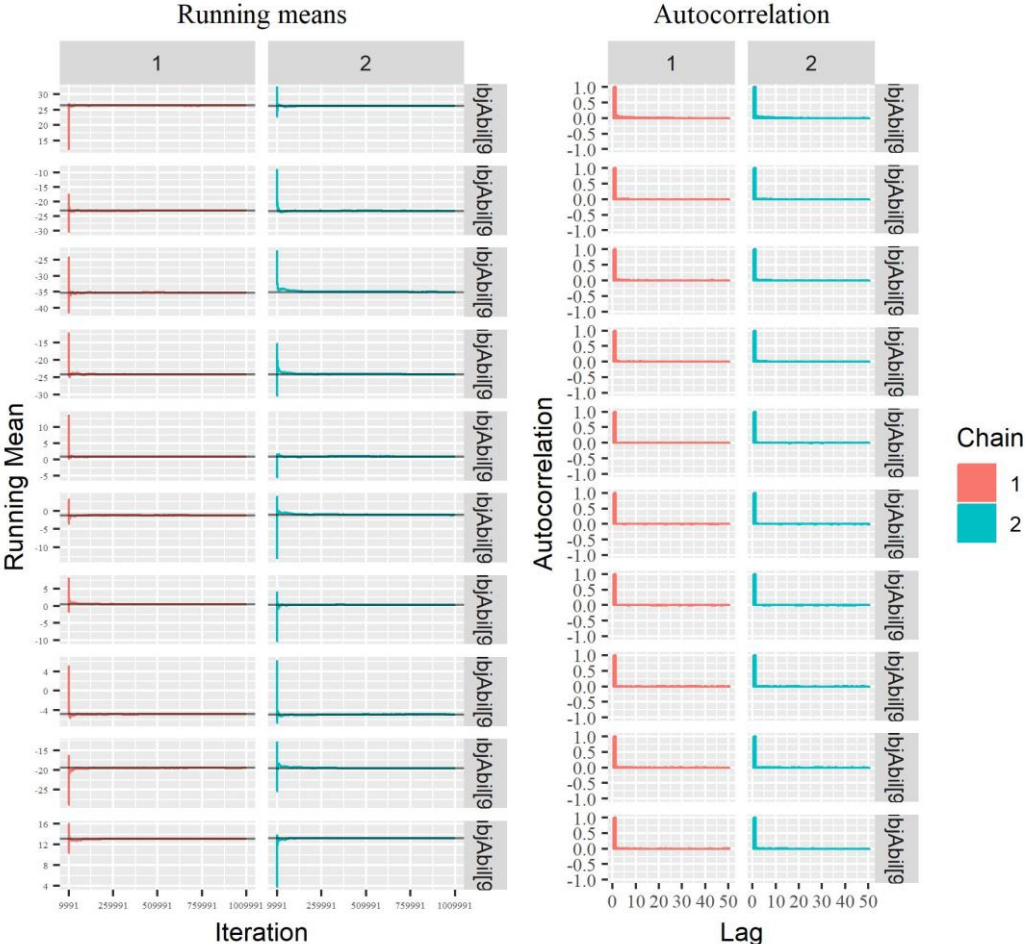
Figure 13, left panel displays the running means of the ability score theta for observations 90 to 99. The scores all converge comfortably after about 12,000 iterations. None of the individual ability scores display evidence of non-convergence in this plot. The parameters which are not displayed here are mostly similar to those shown, with a few exceptions where it takes longer time to converge.<sup>44</sup> It is also good to see that the different chains converge on approximately the same values. This is not necessarily the case, and when it is, something is usually wrong with the model specifications. The phenomenon would of course also be visible in the density plot, where two very distinct distributions would have emerged.

The running means plot is nice for between-chain comparisons because the panels for chain one and two are equally scaled. Together with the density- and partial comparison plots above, they also give a very clear idea of which chain, and which part of a chain, is causing the posterior to have its particular distribution. When the running means are smooth and show no

<sup>44</sup> Observations 35, 60, 67, 136, and 153 are some examples of this, for those who want to check for themselves. Replication material will be supplied on request.

signs of non-convergence as here, that is difficult to do, but when part of a chain diverges very much, it is easily identifiable. These plot diagnostics thus make it easier to correct such errors, either by increasing thinning, the number of iterations the simulation runs, or increasing the burn-in phase of the simulation. Usually, the latter option is preferred, as the first part of the chain is probably the most problematic.

Figure 13: Running means and autocorrelation plots for theta parameters



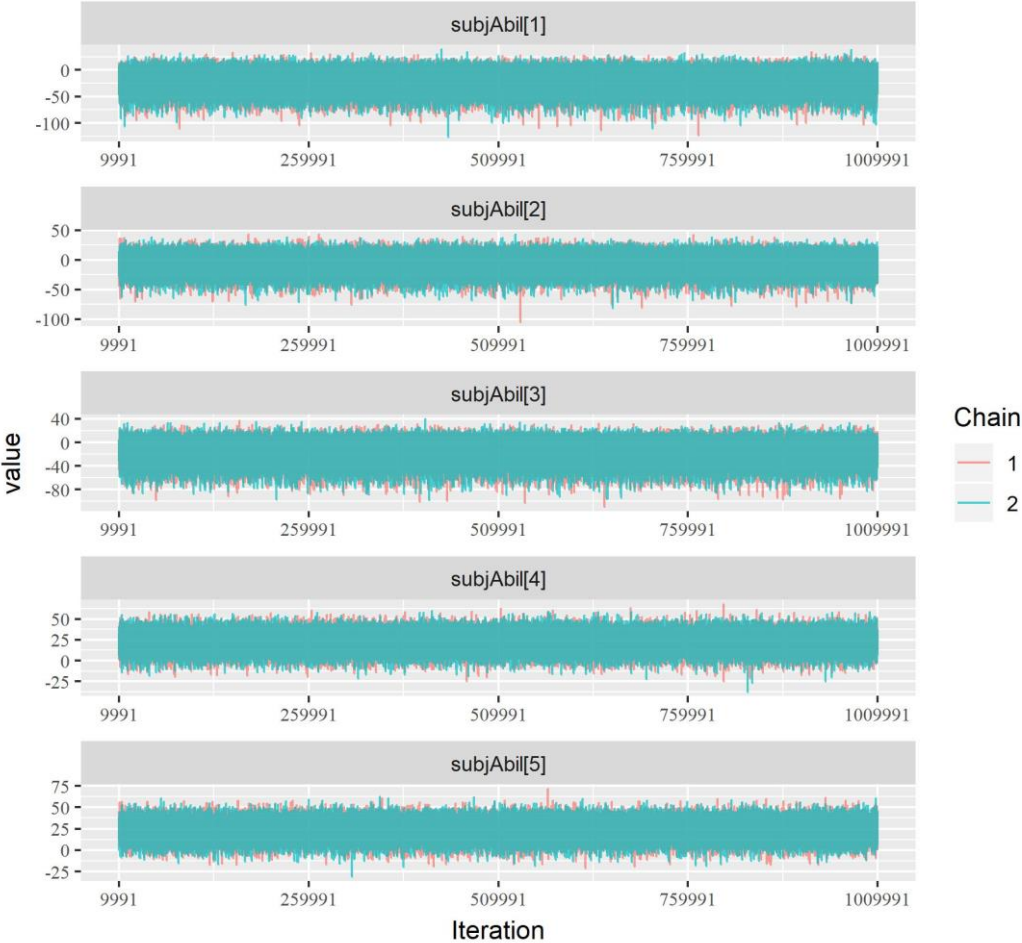
The autocorrelation plot in Figure 13, right panel shows that for the theta parameters,<sup>45</sup> the problem of autocorrelation is nonexistent. While the alphas displayed some autocorrelation, which could have been troublesome, none of the thetas do. The autocorrelation drops to nothing right after the first lag. The same is true for most of the thetas, although a few show some autocorrelation, but not as much as the alphas and not nearly enough to be a problem.

Figure 14 shows the trace plots for five theta parameters. As with the alpha parameters and with the previous theta-diagnostics, there are no patterns in the distribution other than a

<sup>45</sup> These are theta parameters 90-99.

uniform buzz which stays around the same values as the iterations increase. Here too the starting value of the iterations are similar to the final ones, indicating that the burn-in period is well specified. The scale of the ability scores as visible in all the preceding figures vary between approximately -75 and 100. After standardization of the variable to have a mean of zero and a standard deviation of one, the scale shrinks to that displayed in Table 1, to be between -1.219 and 3.43624. This only enhances the comparability of the covariates for the hypothesis-testing and does not impact the strength or interpretation of the ability score.

Figure 14: Traceplot for theta parameters

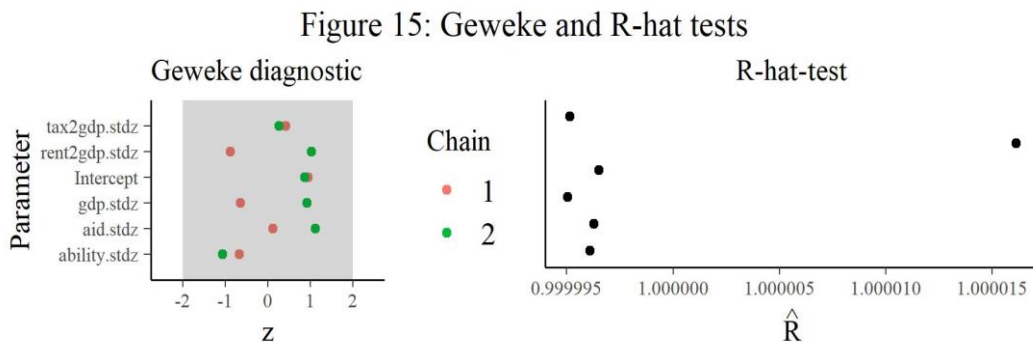


The model is generally successfully applied, with some slight weaknesses in the convergence of the alpha parameters, specifically concerning autocorrelation. This is however not very important. The values of the theta parameters show variation and are therefore useable as an independent variable for hypothesis testing. Neither set of parameters shows evidence of non-convergence in any other test, and the encompassing formal test (Geweke and R-hat) provide no ground for changing or discarding the model.

### 7.4.3. Logistic regression: Regression coefficients beta

To assess the convergence of the MCMC simulations for the regression model, I use many of the same tools as I did for the IRT model above. The logistic regression model has fewer parameters than the IRT model, and as such is much easier to interpret visually. The six beta coefficients for each right-hand variable and the intercept can easily be gathered in single plots and there is no need to choose only a few parameters to inspect.

The Geweke- and  $\hat{R}$ - tests can be seen in Figure 15, left- and right panel respectively. The chains display no signs of non-convergence in these tests. The Geweke test records no parameter for either chain outside the shaded  $-2/2$  interval, which means that the first and last parts of the chains are not too different. The most problematic parameter is the *rent2gdp* parameter, with z-scores of 1 and -1 for the chains. This is not a product of the significant number of missing values for this variable, as *tax2gdp* does not show similarly high z-values but has almost twice as many missing values. This may mean that the values of *rent2gdp* converge slower than the rest of the parameters.



For the  $\hat{R}$ -test, again the scaling on the X-axis is removed, and the *rent2gdp.stdz*-parameter is well within the bounds of what is expected for a converged chain. Here too, *rent2gdp* is most problematic, meaning that the two simulation chains differ somewhat. Again, the difference is negligible, and the chains display no evidence of non-convergence.

Figure 16 shows trace- and density plots for the beta parameters. The convergence here is very good, as there are no discernible patterns in the trace plots, nor any divergence between the two chains in the density plots. This is better than the convergence for the IRT model, even though that was already satisfactory. The solid blocks of “white noise” in the trace plots are exactly what the time-series of a Monte Carlo Markov Chain should look like, and the two chains’ density plots are barely distinguishable from one another. This means that there is no



evidence of non-convergence from the values discussed in chapter 6.3, and the simulation process seems to have produced reliable estimates.

Figure 16: Density- and traceplot for beta parameters

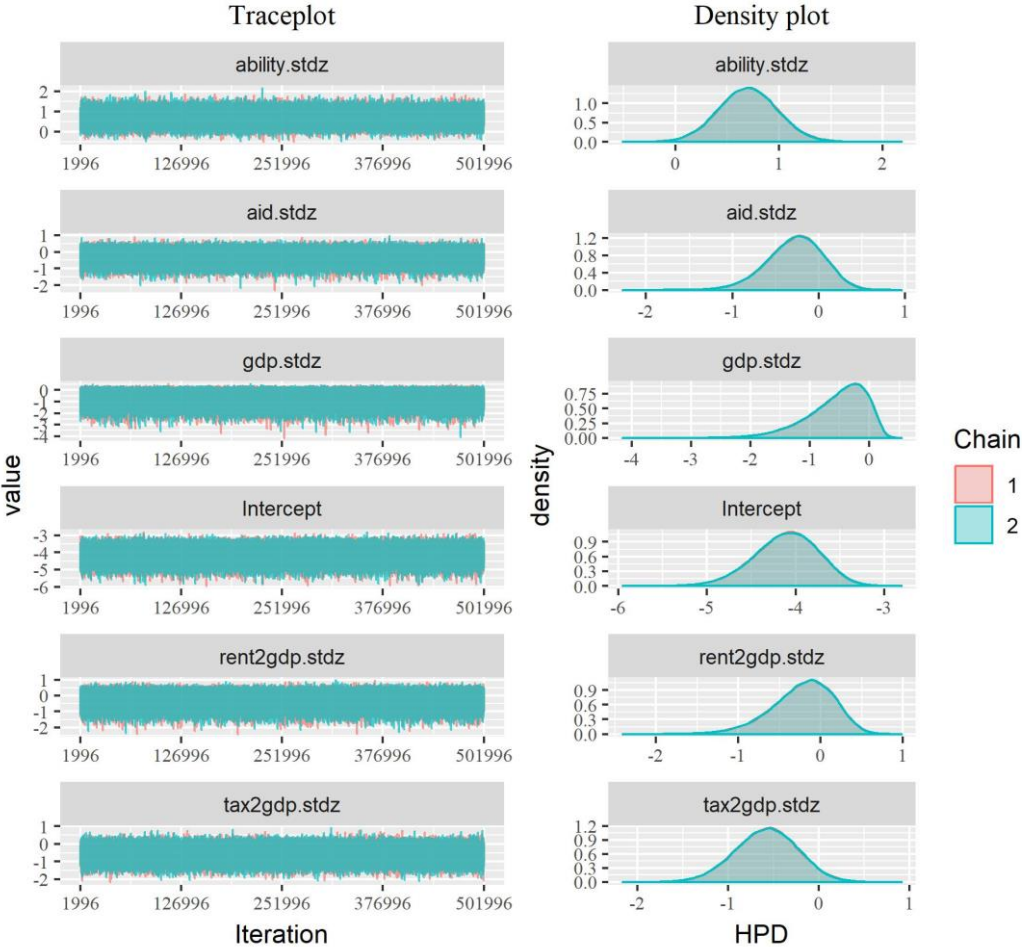
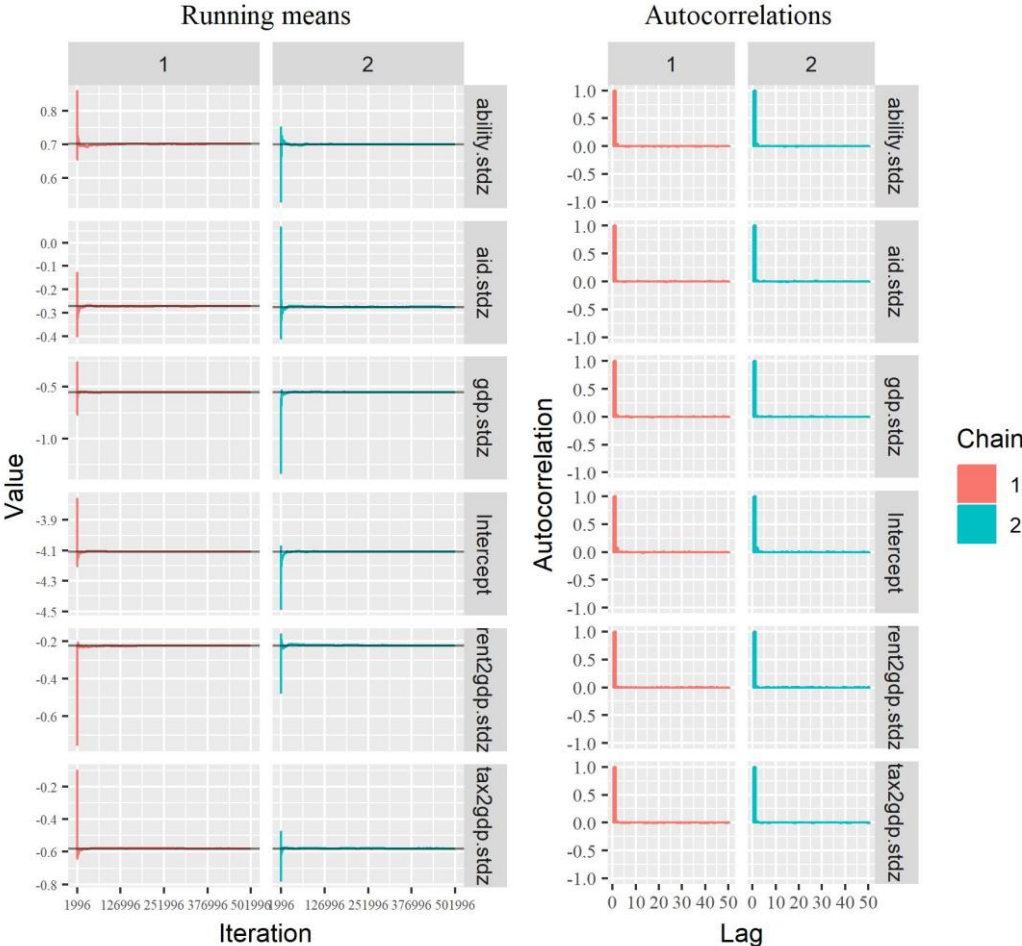


Figure 17 displays the running means and autocorrelations for the beta parameters in the left- and right panel, respectively. Again, the convergence seems to be very good, and there are no signs of bad mixing in the running means-plots or of iteration autocorrelation in the autocorrelation plots. There is virtually no autocorrelation beyond lag 1, and the running means center after about 2.000 iterations. The latter point may be a sign that I have chosen an unnecessarily high number of iterations, but the computer power and time required for the logistic regression model is very low, at least compared to the IRT model (25 minutes versus about 42 hours; This is because of the very high number of parameters needed for the estimation of the IRT model, whereas only six are estimated for the logistic regression). The number of burn-in- and adaptive iterations may also be unnecessarily high with 1.000 of each. Note also that the *coefficients* for the parameters are not the same as their odds; the fact that the mean of the *ability*-variable is around .7 is consistent with the odds presented in Table 2, chapter 7.2.

The running means-plot also displays the ease with which the between-chain variation can be assessed here – it is very simple to see whether the running means for each chain are identical or not. While some slight variation is to be expected, the fact that they look almost perfectly aligned makes the values more reliable.

Figure 17: Running means and autocorrelation plots for beta parameters



### 7.5. Alternative approaches

There are several other ways in which I could have solved my research problem. For example, as I have already pointed out, I could have used principal component factor analysis rather than IRT to construct my measure of economic strategies. That way, less information would have been lost in the aggregation process. Because factor analysis relies on the correlation matrix of variables rather than the data matrix of binary values, I could have aggregated every indicator of economic strategy by its sum each year, rather than the maximum count. My reason for choosing to use IRT, which requires binary variables in the data matrix, rather than principal component factor analysis is simply that I had never used it before and wanted to learn it. This is certainly no way to do science – novelty for novelty’s sake is, as Schrodtr (2014, 295) puts it,

dysfunctional. I have therefore performed a quick, frequentist principal component analysis and subsequent logistic regression to amend this problem. The results are in Appendix 10.3. and show how 17 different orthogonally rotated factors constructed from the verb-variables contribute the proportion of explained variance in the data, and how these together explain the likelihood of democratization. The effects while unrefined are messy, with coefficients signs in both directions, and are trivial in size. This may indicate that my IRT-aggregation of all 42 verb-variables is better for predicting democratization. My IRT model might however be overexplaining democratization as well by including too many indicators. Comparing the two is unfair, however – I’ve spent close to nine months on one model and an afternoon on the other.

The utility of factor analysis is accentuated by its inductive character, which allows for groupings of variables based on their shared variance rather than theoretical coherence. Of course, deduction and induction in measurement are complimentary, but it is advantageous when choosing indicators to first, investigate which are correlated and explain the same variation and second, to check whether they can be conceptually linked, for instance as indicative of an “economic” strategy.

This advantage is lost in my IRT model because I have no theory supporting an approach with several types of economic strategy. By selecting all those indicators with some connection to economic factors and excluding those without it, I have lost some room for interpretation of the item discrimination parameters. Surely, I could have limited the number of indicator variables to only those which discriminate campaigns above a certain threshold. The resulting ability score for a campaign’s economicness would be more parsimonious, and the MCMC simulations easier to run and evaluate. This may have an effect on the hypothesis testing.

Another possible approach is to use a clustering algorithm and find groups of observations which share some trait or another, and then test whether some are more likely to produce democratization by simple ANOVA test, for example. This approach would potentially result in a typology of groups as well as a test of my hypothesis. Jordana, Fernández-i-Marín, and Bianculli (2018) construct measurement models for four dimensions of institutional features of regulatory agencies using IRT and factor analysis and subsequently a clustering algorithm to identify six types of regulatory agencies. An identical framework for nonviolent campaigns would have allowed me to simply compare the occurrence of democratization in each cluster with ANOVA. This would indeed have been an elegant way of answering the research question, and with the interesting auxiliary output of a potential typology of nonviolent campaigns. Perhaps some combination of economic and violent strategy as particularly disadvantageous; or perhaps economic strategy without some other strategy in combination

with it is useless. I strongly recommend this approach as a continuation of my work in this thesis.

Machine learning and random forest models are also options which could satisfactorily answer the research question. While my knowledge of them is limited, such models can be useful for both classification and regression and are not prohibitively complicated to execute. Learning to use them properly is relatively easy with free information and dedicated people on the Internet, and particularly the R community is very helpful. In my own analyses, few problems I had were novel, and most had easily accessible solutions online.

As the source material from which the data are recorded consists of AFP newspaper articles, an interesting approach would also be to use a structural topic model to investigate the language in the articles, as well as the frequency of articles on a political conflict. Are higher volumes of press coverage, or some words or topics connected with democratization or repression? Sadly, the newspaper articles themselves are not currently available, but it would be interesting to see an analysis similar to Grimmer (2009), who analyzes latent topics in the text of US Congress press releases. The approach is similar to mine in that it uses latent topics as an interesting quantity as I use economic strategy but uses text as data rather than as source materials.

## 8. Conclusion

In this thesis, I set out to investigate whether economic nonviolent strategies are conducive to democratization. I was primarily motivated by the research gap on nonviolent strategy in the aggregate. The current literature on nonviolent movements has little to say about which strategies succeed and which fail, apart from the general “nonviolence is better than violence”-finding. Seeing that this finding is explained foremost by the importance of limiting the state’s access to means of violence – its primary defining characteristic and source of power – I argued that this is a testament to the Weberian definition of a state. A counterdefinition is the Schumpeterian one. Swapping Schumpeter’s “tax-monopoly” for the more general “economically capable state,” I then deduced that a similar relationship should exist with the state’s economic resources. I stated that states need financial as well as military capacity and argued that it should be possible to replace its regime by targeting state finances through economic nonviolent protest. More specifically, the capacity of a state to impose its policies depends on the financial capacity and consent of its population, and therefore can be limited with certain strategic foci. To situate the problem in the broader democratization literature, I connected these points to the Dahlian cost of repression/cost of tolerance-concepts and

narrowed the theoretical discussion toward my particular research question. This is a commensurable approach which “advance[s] the logical economy in a field” (Gerring 2012b, 60, 68-69) and merges the Dahlian democratization theory with Sharp’s theory of wielding power effectively. It further resonates well with most existing literature on democratization in that it touches upon concepts like stability, capacity, wealth, and class.

Then, to investigate the relationship between economic nonviolent strategy and democratization, I indulged in a discussion of what the core definitional traits of a ‘strategy’, ‘nonviolence’, and how best to operationalize the terms with resonance and commensurability. The result of this discussion was first that the data collected to not directly correspond to the conceptual discussions available. Secondly, I concluded that the data on nonviolent tactics and strategies which exist, and primarily the NAVCO 3.0 data, are well-suited to aggregation.

Thirdly, the conceptual discussion contributed to specifying the domain in which my findings apply. For studying mechanisms, causality, patterns, or trends in social behavior it is imperative to understand the context in which humans act. This is an epistemological point as well as a practical one: we cannot know much about a phenomenon without an idea of its context or domain; and in order to collect and compare findings about similar phenomena, we must identify likenesses and dissimilarities between phenomena, their contexts, and their defining traits. I have argued that for economic nonviolent action, the conceptual framework of contentious politics is appropriate, and it should be discussed as part of that domain. I also argued that it is appropriate to compare economic nonviolence to other types of nonviolence, as well as to violent contentious politics. The latter necessitates attention to different causal mechanisms and patterns of state- and protest behavior, but violent and nonviolent contentious politics are not opposites. Rather, they are different styles of action with different traits. As Kudelia (2018) and Aitchison (2018) argue, violence and nonviolence can complement and substitute one another.

Following the theoretical and abstract discussions, I explained my choice to aggregate and reduce the NAVCO 3.0 data by way of an IRT model. The method requires binary data and results in a considerable information loss when they are aggregated by its maximum value rather than their count. However, the results seem reliable and possessive of causal utility. The indicators chosen to represent economic strategies were deductively chosen and could have been selected with greater attention to their interrelationships.

Desiring to use my ability score for hypothesis testing, I then explained my choice to use logistic regression. Having chosen a binary indicator of democratization rather than a metric one for primarily ontological reasons but also as a hard test of my hypothesis, the logistic

regression follows logically. Both of these analyses were performed using Bayesian methods. Because I have not employed sample- or experimental data where potentially repeated sampling makes frequentist significance testing possible, this is appropriate.

The results of the logistic regression show that an increase of one standard deviation in economic strategy corresponds to more than a doubling in the likelihood for transitions to democracy. This is a large effect that speaks to both the nonviolence- and the democratization literature. My research design and -question, which have not been used or investigated quantitatively before in the nonviolence literature, contribute to filling a research gap which is relevant both theoretically, empirically, and practically. I suggest that subsequent research use more latent variable models such as item response theory, factor analysis, cluster analysis, structural topic modeling and similar techniques to find and test the efficiency of different nonviolent strategies. This will make researchers able to give relevant advice to activists by bringing the level of analysis down to a more concrete level. It will also contribute to the cumulative knowledge on democratization and nonviolence and as such better our understanding of these important and interesting phenomena.

The finding also speaks to the importance of conceptualizing and measuring phenomenon properly, and to theorizing how and how much different types of contentious political action impact politics. I have argued that the data I used are appropriate to definitionally represent Gene Sharp's concept of nonviolent action and that they possess causal utility for my research question. It is not obvious that variables recording very fine-grained types of action are appropriate without paying attention to the surrounding concepts, the inclusion criteria of the data, contrast-space, and theoretical discussions about contentious politics. My conceptual discussion allows the results of my analysis to be interpreted as a part of a larger set of phenomena from which many possible research questions and -designs spring, and they provide fertile ground for theorizing and testing new links between action and consequence.

I have thus taken a convincing step toward establishing a reliable, quantitative relationship between the strategies of nonviolent political movements and regime transitions. Much work remains to be done to craft a complete empirical narrative of the relationships in Figure 1, and this thesis has made significant contributions to the understanding of those relationships.

## 9. References

- Acemoglu, Daron, and James A. Robinson. 2001. "A Theory of Political Transitions." *The American Economic Review* 91 (4):938-963, <http://www.jstor.org/stable/2677820>.
- Acemoglu, Daron, and James A. Robinson. 2012. *Why Nations Fail: The origins of power, prosperity and poverty*, <http://norayr.am/collections/books/Why-Nations-Fail-Daron-Acemoglu.pdf>. New York: Crown Publishers.
- Achen, Christopher H. 2005. Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong. *Conflict Management and Peace Science* 22: 327-339. doi:10.1080/07388940500339167.
- Ackerman, Peter. 2007. "Skills or Conditions: What Key Factors Shape the Success or Failure of Civil Resistance?" Conference on Civil Resistance & Power Politics, St Antony's College, University of Oxford.
- Ackerman, Peter, and Jack DuVall. 2000. *A Force More Powerful: A Century of Nonviolent Conflict*. New York: Palgrave.
- Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *The American Political Science Review* 65 (2):529-546, <https://www.jstor.org/stable/3118231>.
- Afesorgbor, Sylvanus Kwaku, and Renuka Mahadevan. 2016. "The Impact of Economic Sanctions on Income Inequality of Target States." *World Development* 83:1-11. doi: <http://dx.doi.org/10.1016/j.worlddev.2016.03.015>.
- Aitchison, Guy. 2018. Domination and Disobedience: Protest, Coercion and the Limits of an Appeal to Justice. *Perspectives on Politics* 16 (3): 666-679. doi:10.1017/S1537592718001111.
- Allen, Susan Hannah, and David J. Lektzian. 2012. "Economic sanctions: A blunt instrument?" *Journal of Peace Research* 50 (1):121-135. doi: 10.1177/0022343312456224.
- Bernhard, Michael. 2016. "The Moore thesis: what's left after 1989?" *Democratization* 23 (1):118-140. doi: <http://dx.doi.org/10.1080/13510347.2015.1094796>.
- Binnendijk, Anika Locke, and Ivan Marovic. 2006. "Power and persuasion: Nonviolent strategies to influence state security forces in Serbia (2000) and Ukraine (2004)." *Communist and Post-Communist Studies* 39:411-429. doi: 10.1016/j.postcomstud.2006.06.003.
- Bollen, Kenneth A. 1979. Political Democracy and the Timing of Development. *American Sociological Association* 44 (4): 572-587. Accessed 11.09.2018, <https://www.jstor.org/stable/2094588>.
- Brown, Robert McAfee. 1984. *Unexpected news : reading the Bible with Third World eyes*. Philadelphia: The Westminster Press.
- Brysk, Alison. 1993. "From Above and Below: Social Movements, the International System, and Human Rights in Argentina." *Comparative Political Studies* 26 (3):259-285. doi: <https://doi.org/10.1177/0010414093026003001>.
- Cheibub, José Antonio, Jennifer Gandhi, and James Raymond Vreeland. 2010. Democracy and dictatorship revisited. *Public Choice*, 10.1007/s11127-009-9491-2 (143): 67-101. doi:10.1007/s11127-009-9491-2.
- Chenoweth, Erica, Evan Perkoski, and Sooyeon Kang. 2017. "State Repression and Nonviolent Resistance." *Journal of Conflict Resolution* 61 (9):1950-1969. doi: DOI: 10.1177/0022002717721390.
- Chenoweth, Erica, Jonathan Pinckney, and Orion Lewis. 2018. "Days of rage: Introducing the NAVCO 3.0 dataset." *Journal of Peace Research* 55 (4):524-534. doi: 10.1177/0022343318759411.

- Chenoweth, Erica, Jonathan Pinckney, and Orion A. Lewis. 2017. Nonviolent and Violent Campaigns and Outcomes Dataset. University of Denver.
- Chenoweth, Erica, and Maria. J. Stephan. 2011. *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*. Edited by Hoffman, Bruce, *Columbia Studies in Terrorism and Irregular Warfare*. New York: Columbia University Press.
- Chenoweth, Erica, and Jay Uldfelder. 2017. Can Structural Conditions Explain the Onset of Nonviolent Uprisings? *Journal of Conflict Resolution* 61 (2): 298-324. doi:10.1177/0022002715576574.
- Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98 (2):355-370.
- Collier, David, and Robert Adcock. 1999. Democracy and Dichotomies: A Pragmatic Approach to Choices about Concepts. *Annual Review of Political Science* 1999 (2): 537-565. Accessed 16.10.2018. doi:<https://doi.org/10.1146/annurev.polisci.2.1.537>.
- Collier, David, Henry E. Brady, and Jason Seawright. 2010. "Sources of Leverage in Causal inference: Toward an Alternative View of Methodology." In *Rethinking Social Inquiry*, edited by Brady, Henry E. and David Collier, 161-199. Plymouth: Rowman & Littlefield Publishers, INC.
- Collier, David, and Steven Levitsky. 1997. "Democracy with Adjectives: Conceptual Innovation in Comparative Research." *World Politics* 49 (3):430-451, <http://www.jstor.org/stable/25054009>.
- Croissant, Aurel, David Kuehn, and Tanja Eschenauer. 2018. Mass Protests and the Military. *Journal of Democracy* 29 (3): 141-155. doi:<https://doi.org/10.1353/jod.2018.0051>.
- Cunningham, David E., Kristian Skrede Gleditsch, Belén González, Dragana Vidović, and Peter B. White. 2017. Words and deeds: From incompatibilities to outcomes in anti-government disputes. *Journal of Peace Research* 54 (4): 468-483. Accessed Jul. doi:10.1177/0022343317712576.
- Curtis, S. McKay. 2010. "BUGS Code for Item Response Theory." *Journal of Statistical Software* 36 (1):1-34, <https://www.jstatsoft.org/article/view/v036c01/v36c01.pdf>.
- Dahl, Robert A. 1971. *Polyarchy*. London: Yale University press, Ltd.
- Dahlum, Sirianne. 2018. "Joining forces: Class-coalitions, mass protest and democratization." Workshop on Civil Resistance and Democratization, Norwegian University of Science and Technology (NTNU), Nov 19th.
- Diamond, Larry. 1992. Economic Development and Democracy Reconsidered. *American Behavioral Scientist* 35 (4/5): 450-499. doi:doi.org/10.1177/000276429203500407.
- Diamond, Larry. 1999. *Developing Democracy: Toward Consolidation*. Baltimore: The John Hopkins University Press.
- Doucouliagos, Hristos , and Martin Paldam. 2009. "The Aid Effectiveness Literature: The Sad Result of 40 Years of Research." *Journal of Economic Surveys* 23 (3):433-461. doi: <https://doi.org/10.1111/j.1467-6419.2008.00568.x>.
- Elster, Jon. 1988. "Marx, Revolution and Rational Choice." In *Rationality and Revolution*, edited by Taylor, Michael, 206-228. Cambridge: Cambridge University Press.
- Feldman, Gerald D. 1971. "Big Business and the Kapp Putsch." *Central European History* 4 (2):99-130, <http://www.jstor.org/stable/4545598>.
- Fernández-i-Marín, Xavier. 2016. "ggmcmc: Analysis of MCMC samples and Bayesian inference." *Journal of Statistical Software* 70 (9):1-20. doi: 10.18637/jss.v070.i09.
- Fernández-i-Marín, Xavier. 2019. "Xavier Fernández i Marín." accessed 22.03.2019. <http://xavier-fim.net/>.
- Fox, Jean-Paul. 2010. *Bayesian Item Response Modeling: Theory and Applications*, DOI 10.1007/978-1-4419-0742-4. London: Springer.



- Fukuyama, Francis. 2016. Macro theory and the Study of Political Development. *Scandinavian Political Studies* 39 (3): 207-225. Accessed 11.09.2018. doi:10.1111/1467-9477.12064.
- Funk, Sebastian, Anton Camacho, and Helen Johnson. 2018. "Practical session: MCMC diagnostics." accessed 03.04.2019. [http://sbfknk.github.io/mfiidd/mcmc\\_diagnostics.html](http://sbfknk.github.io/mfiidd/mcmc_diagnostics.html).
- Gerring, John. 2012a. Mere Description. *British Journal of Political Science* 42: 721-746. doi:doi:10.1017/S0007123412000130.
- Gerring, John. 2012b. *Social Science Methodology: A Unified Framework*. Edited by Elman, Colin, John Gerring and James Mahoney. 2nd ed. Cambridge: Cambridge University Press.
- Ghandi, Jennifer. 2015. Elections and Political Regimes. *Government and Opposition* 50 (3): 446-468. doi:10.1017/gov.2015.11.
- Gleditsch, Nils Petter. 1971. "Inledning." In *Kamp Utan Vapen: Berömda exempel på ickevåldsmotstånd från olika länder och tider i urval av Nils Petter Gleditsch*, edited by Gleditsch, Nils Petter, 10-24. Oskarshamn: Bokförlaget Prisma Stockholm.
- Goertz, Gary. 2006. *Social Science Concepts A User's Guide*. Oxfordshire: Princeton University Press.
- Grimmer, Justin. 2009. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 2010 (18):1-35. doi: doi:10.1093/pan/mpp034.
- Hair, Joseph F. , William C. Black, Barry J. Babin, and Rolph E. Anderson. 2014. *Multivariate Data Analysis*, <https://is.muni.cz/el/1423/podzim2017/PSY028/um/ Hair - Multivariate data analysis 7th revised.pdf>. Harlow: Pearson Education Limited.
- Hendrix, Cullen S. . 2010. "Measuring state capacity: Theoretical and empirical implications for the study of civil conflict." *Journal of Peace Research* 47 (3):273-285. doi: DOI: 10.1177/0022343310361838.
- Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty - Responses to Decline in Firms, Organizations, and States*. Cambridge: Harvard University Press.
- Hooghe, Liesbet, Gary Marks, Arjan H. Schakel, Sara Niedzwiecki, Sandra Chapman Osterkatz, and Sarah Shair-Rosenfield. 2016. "Transparency in Measurement." In *Measuring Regional Authority: A Postfunctionalist Theory of Governance, Volume I*, 33. Oxford: Oxford University Press.
- Hosmer, D. W, T Hosmer, S Le Cessie, and S Lemeshow. 1997. "A Comparison of Goodness-of-Fit Tests for the Logistic Regression Model." *Statistics in Medicine* 16:965-980.
- Inglehart, Ronald, and Christian Welzel. 2005. *Modernization, Cultural Change, and Democracy: The Human Development Sequence*. Cambridge: Cambridge University Press.
- Jackman, Simon. 2001. "Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking." *Political Analysis* 9 (3):227-241. doi: <https://doi.org/10.1093/polana/9.3.227>.
- Jackman, Simon. 2008. "Measurement." In *The Oxford Handbook of Political Methodology*, edited by Box-Steffensmeier, Janet M., Henry E. Brady and David Collier. Oxford University Press.
- Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. Edited by Balding, David J., Noel A. C. Cressie, Garrett M. Fitzmaurice, Iain M. Johnstone, Geert Molenberghs, David W. Scott, Adrian F. M. Smith, Ruey S. Tsay, Sanford Weisberg, Harvey Goldstein, Vic J. Barnett, J. Stuart Hunter and Jozef L. Teugels, *Wiley series in probability and statistics*. West Sussex: John Wiley & Sons, Ltd.

- Jasper, James M. 2004. A Strategic Approach to Collective Action: Looking for Agency in Social-Movement Choices. *Mobilization: An International Journal* 9 (1): 1-16. doi:10.17813/mai.9.1.m112677546p63361.
- Jasper, James M. 2006. "Getting your way: Strategic dilemmas in the real world." In. London: University of Chicago Press, Ltd.
- Johnson, James. 2006. "Consequences of Positivism: A Pragmatist Assessment." *Comparative Political Studies* 39 (2):224-252. doi: 10.1177/0010414005282982.
- Jordana, Jacint, Xavier Fernández-i-Marín, and Andrea C. Bianculli. 2018. "Agency proliferation and the globalization of the regulatory state: Introducing a data set on the institutional features of regulatory agencies." *Regulation & Governance* 0 (0):1-17. doi:doi:10.1111/rego.12189.
- Karatnycky, Adrian, and Peter Ackerman. 2005. "How Freedom is Won: From Civic Resistance to Durable Democracy." In. New York: Freedom House. <https://freedomhouse.org/report/special-reports/how-freedom-won-civic-resistance-durable-democracy> (accessed 02.10.2018).
- Kaya, Bülent. 2002. The Changing Face of Europe – Population Flows in the 20th Century. In *Project "Learning and teaching about the history of Europe in the 20th century"*, edited by Studies, Swiss Forum for Migration and Population. Strasbourg: Council of Europe.
- Klein, Graig R., and Patrick M. Regan. 2018. Dynamics of Political Protests. *International Organization* 72 (2): 485-521. doi:doi:10.1017/S0020818318000061.
- Kruschke, John K. 2010. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS and Stan*. 2nd ed. London: Academic Press / Elsevier.
- Kudelia, Serhiy. 2018. When Numbers Are Not Enough: The Strategic Use of Violence in Ukraine's 2014 Revolution. *Comparative Politics* 50 (4): 501-521. doi:<https://doi.org/10.5129/001041518823565623>.
- Le Guin, Ursula Kroeber 2014. "Ursula K Le Guin's speech at National Book Awards: 'Books aren't just commodities'." *The Guardian*, 20. nov. Accessed 06.11.2018. <https://www.theguardian.com/books/2014/nov/20/ursula-k-le-guin-national-book-awards-speech>.
- Lijphart, Arend. 2012. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. Michigan: Yale University Press.
- Lipset, Seymour Martin. 1959. Some Social Requisites of Democracy: Economic Development and Political Legitimacy. *The American Political Science Review* 53 (1): 69-105. doi:10.2307/1951731.
- Luxemburg, Rosa. 1919. Order Prevails in Berlin. *Marxist.org*, <https://www.marxists.org/archive/luxemburg/1919/01/14.htm>. Accessed 02.10.2018.
- Marshall, Monty G. , and Keith Jagers. 2007. POLITY IV PROJECT: Dataset Users' Manual. [http://home.bi.no/a0110709/polityiv\\_manual.pdf](http://home.bi.no/a0110709/polityiv_manual.pdf): 1-87.
- Martin, Brian. 2001. *Nonviolence versus capitalism*. London: War Resisters' International.
- Meredith, Sam. 2018. "US embargo would be the 'death knell' for Venezuela's oil industry, analyst says." *CNBC*, May 1st 2018. Accessed 14.12.2018. <https://www.cnbc.com/2018/03/01/us-embargo-on-venezuelas-oil-industry-would-be-death-knell-analyst-says.html>.
- Michaelson, Ruth, and Adham Youssef. 2019. "Sisi wins snap Egyptian referendum amid vote-buying claims." *The Guardian*, 23.04. <https://www.theguardian.com/world/2019/apr/23/sisi-wins-snap-egyptian-referendum-amid-vote-buying-claims>.
- Moore, Barrington Jr. 1966. *Social Origins of Dictatorship and Democracy*. Boston: Beacon Press.

- Munck, Gerardo L., and Jay Verkuilen. 2002. Conceptualizing and Measuring Democracy: Evaluating Alternative Indices. *Comparative Political Studies* 35 (1): 5-34. Accessed 14.09.2018. doi:10.1177/001041400203500101.
- Nepstad, Sharon Erickson. 2011. *Nonviolent Revolutions - Civil Resistance in the Late 20th Century*. Edited by Bob, Clifford and James M. Jasper, *Oxford studies in culture and politics*. Oxford: Oxford University Press.
- Nepstad, Sharon Erickson. 2015. *Nonviolent Struggle: Theories, Strategies, & Dynamics*. New York: Oxford University Press.
- Nygård, Håvard Mokleiv. 2015. The role of international organizations in regime transitions: How IGOs can tie a dictator's hands. *Conflict Management and Peace Science* 34 (4): 406-430. doi:10.1177/0738894215599554.
- O'Donnell, Guillermo, and Philippe C. Schmitter. 1986. *Transitions from Authoritarian Rule - Tentative Conclusions for Uncertain Democracies*. Edited by O'Donnell, Guillermo, Philippe C. Schmitter and Laurence Whitehead. 5 vols. Vol. 5, *Transitions from Authoritarian Rule*. London: The Johns Hopkins University Press. Research.
- Olson, Mancur. 1967. *The Logic of Collective Action*. Harvard: Harvard University Press.
- Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.
- Pape, Robert A. 2005. *Dying to Win: The Strategic Logic of Suicide Terrorism*. New York: Random House.
- Peksen, Dursun, and Byungwon Woo. 2018. "Economic Sanctions and the Politics of IMF Lending." *International Interactions* 44 (5):681-708. doi: <https://doi.org/10.1080/03050629.2018.1429427>.
- Persson, Anna, and Bo Rothstein. 2015. "It's My Money: Why Big Government May Be Good Government." *Comparative Politics* 47 (2):231-249, <http://www.jstor.org/stable/43664141>.
- Pisani-Ferry, Jean. 2011. *The Euro Crisis and its Aftermath*. New York: Oxford University Press.
- Plummer, Martyn. 2018. "Package 'rjags' : Bayesian Graphical Models using MCMC." CRAN, <http://mcmc-jags.sourceforge.net/>.
- Przeworski, Adam. 1991. *Democracy and the market: Political and economic reforms in Eastern Europe and Latin America*. Edited by Elster, Jon and Michael S. McPherson, *Studies in Rationality and Social Change*. Cambridge: Press Syndicate of the University of Cambridge.
- Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub, and Fernando Limongi. 2000. "Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990." In. Cambridge: Cambridge University Press (accessed 05.09.2018).
- Putnam, Robert D. 1993. *Making democracy work : Civic traditions in modern Italy*. West Sussex: Princeton University Press.
- Raloff, Karl. 1971. "Icke-våldsmotståndet som kvävde Kapp-kuppen." In *Kamp Utan Vapen: Berömda exempel på ickevåldsmotstånd från olika länder och tider i urval av Nils Petter Gleditsch*, edited by Gleditsch, Nils Petter, 69-79. Oskarshamn: Bokförlaget Prisma Stockholm.
- Reckase, Mark D. 2009. *Multidimensional Item Response Theory*, DOI 10.1007/978-0-387-89976-3. London: Springer.
- Rodríguez, Francisco. 2018. "Why More Sanctions Won't Help Venezuela." *Foreign Policy*, Jan 12th 2018. Accessed 14.12.2018. <https://foreignpolicy.com/2018/01/12/why-more-sanctions-wont-help-venezuela/>.
- Sartori, Giovanni. 1970. Concept Misformation in Comparative Politics. *The American Political Science Review* 64 (4): 1033-1053. <http://www.jstor.org/stable/1958356>.

- Schrodt, Philip. 2014. Seven deadly sins of contemporary quantitative political analysis. *Journal of Peace Research* 51 (2): 287-300. doi:10.1177/0022343313499597.
- Schumpeter, Joseph A, ed. 1954. *History of Economic Analysis*. Edited by Schumpeter, Elizabeth Boody. 6th ed. London: George Allen & Unwil LTD.
- Schumpeter, Joseph A. 1991[1919]. "Crisis of the Tax State." In *Joseph A. Schumpeter: The Economics and Sociology of Capitalism*, edited by Swedberg, Richard, 99-140. Princeton: Princeton University Press.
- Sharp, Gene. 1973a. *The Politics of Nonviolent Action Part One: Power and Struggle*. 9th ed. 3 vols. Vol. 1. Boston: Extending Horizons Books.
- Sharp, Gene. 1973b. *The Politics of Nonviolent Acton Part Two: The Methods of Nonviolent Action*. 8th ed. 3 vols. Vol. 2, *The Politics of Nonviolent Action*. Boston: Extending Horizons Books.
- Sharp, Gene. 2011. "Q&A: Gene Sharp." *Al Jazeera*, December 6th 2011, Opinion: Latin America. <https://www.aljazeera.com/indepth/opinion/2011/12/201112113179492201.html>.
- Sharp, Gene. 2012. *From Dictatorship to Democracy*. 4th ed. Croydon: Serpent's Tail. Political writing.
- Sheehan, Eleanor. 2017. "Publisher to Recall Whitewashed Textbook Claiming First Nations People 'Agreed to Move' For Settlers " *Splinter News*. <https://splinternews.com/publisher-to-recall-whitewashed-textbook-claiming-first-1819121949>.
- Skocpol, Theda. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press. Comparative.
- Stephan, Maria. J., and Erica Chenoweth. 2008. Why civil resistance works - The strategic logic of nonviolent conflict. *International Security* 33 (1): 7-44. Accessed Sum. doi:10.1162/isec.2008.33.1.7.
- Sweezy, Paul M. 1972. "Marx and the Proletariat." In *Modern Capitalism and other essays*, 147-165. New York: Monthly Review.
- Tarrow, Sidney G. 1994. *Power in Movement: Social Movements and Contentious Politics*. 3rd ed. Cambridge: Cambridge University Press.
- Teorell, Jan. 2010. *Determinants of Democratization: Explaining Regime Change in the World, 1972-2006*. Cambridge: Cambridge University Press.
- Tilly, Charles. 1985. "War Making and State Making as Organized Crime." In *Bringing the State Back In*, edited by Evans, Peter B., Dietrich Rueschemeyer and Theda Skocpol, 169-91. Cambridge: Cambridge University Press.
- Tilly, Charles, Sidney Tarrow, and Douglas McAdam. 2003. *Contention and Democracy in Europe, 1650-2000*, <http://ebookcentral.proquest.com/lib/priono-ebooks/detail.action?docID=255200>. Cambridge, UNITED KINGDOM: Cambridge University Press.
- Treier, Shawn, and Simon Jackman. 2008. Democracy as a Latent Variable. *American Journal of Political Science* 52 (1): 201-217.
- United Nations, Department of Economic and Social Affairs, Population Division (2017). 2017. World Population Prospects: The 2017 Revision, Online Demographic Profiles.
- van Buuren, S., and K. Groothuis-Oudshoorn. 2011. "mice: Multivariate Imputation by Chained Equations in R." *Journal of Statistical Software* 45 (3):1-67.
- Weber, Max. 1958. "Politics as a Vocation." In *From Max Weber: Essays in Sociology*, edited by Gerth, H. H. and C. Wright Wright Mills. New York: Oxford University Press.
- Whiteley, Paul F., and Patrick Seyd. 2002. "Theories of Participation and High-Intensity Participation." In *High-Intensity Participation*, 35-58. Ann Arbor: The University of Michigan Press.

- Wood, Reed M. 2008. "'A Hand upon the Throat of the Nation': Economic Sanctions and State Repression, 1976–2001." *International Studies Quarterly* 52:489-513. doi: <https://doi.org/10.1111/j.1468-2478.2008.00512.x>.
- World Bank. 2019a. GDP (current US\$). edited by World Bank national accounts data, and OECD National Accounts data files.: World Bank.
- World Bank. 2019b. Net official development assistance and official aid received (current US\$). edited by Development Assistance Committee of the Organisation for Economic Co-operation and Development, Geographical Distribution of Financial Flows to Developing Countries, Development Co-operation Report, and International Development Statistics database. : World Bank.
- World Bank. 2019c. Tax revenue (% of GDP). edited by International Monetary Fund, Government Finance Statistics Yearbook and data files, and World Bank and OECD GDP estimates.: World Bank.
- World Bank. 2019d. Total natural resources rents (% of GDP). edited by), World Bank staff estimates based on sources and methods described in "The Changing Wealth of Nations 2018: Building a Sustainable Future" ( Lange et al 2018: World Bank.
- Yashar, Deborah J. 2005. *Contesting Citizenship in Latin America: The rise of indigenous movements and the postliberal challenge*. Cambridge: Cambridge University Press.

## 10. Appendix

### 10.1. List of indicators for IRT model

Variable code	Description	Label in MCMC figures, tables
v_0211	Appeal for economic cooperation	1
v_0231	Appeal for economic aid.	2
v_0233	Appeal for humanitarian aid	3
v_0311	Express intent to cooperate economically	4
v_0254	Appeal for easing of economic sanctions, boycott, or embargo	5
v_0331	Express intent to provide economic aid.	6
v_0333	Express intent to provide humanitarian aid	7
v_0354	Express intent to ease economic sanctions, boycott, or embargo	8
v_061	Cooperate economically	9
v_071	Provide economic aid	10
v_073	Provide humanitarian aid.	11
v_0842	Return, release property	12
v_085	Ease economic sanctions, boycott, embargo.	13
v_0863	Allow humanitarian access.	14
v_1011	Demand economic cooperation	15
v_1031	Demand economic aid	16
v_1033	Demand humanitarian aid	17
v_1054	Demand easing of economic sanctions, boycott, or embargo.	18
v_1211	Reject economic cooperation.	19
v_1221	Reject request for economic aid.	20
v_1223	Reject request for humanitarian aid.	21
v_1244	Refuse to ease economic sanctions, boycott, or embargo	22
v_1311	Threaten to reduce or stop aid.	23
v_1312	Threaten with sanctions, boycott, or embargo.	24
v_1314	Threaten economic consequences.	25
v_1381	Threaten blockade	26
v_1382	Threaten occupation	27
v_143	Conduct strike or boycott, not specified below	28
v_1431	Conduct strike or boycott for leadership change.	29
v_1432	Conduct strike or boycott for policy change.	30
v_1433	Conduct strike or boycott for rights	31
v_1434	Conduct strike or boycott for change in institutions, regime.	32
v_1441	Obstruct passage to demand leadership change.	33
v_1442	Obstruct passage to demand policy change	34

<b>v_1443</b>	Obstruct passage to demand rights.	35
<b>v_1444</b>	Obstruct passage to demand change in institutions, regime.	36
<b>v_1621</b>	Reduce or stop economic assistance.	37
<b>v_1623</b>	Reduce or stop humanitarian assistance	38
<b>v_1663</b>	Expel or withdraw aid agencies.	39
<b>v_171</b>	Seize or damage property, not specified below.	40
<b>v_1711</b>	Confiscate property	41
<b>v_1712</b>	Destroy property.	42

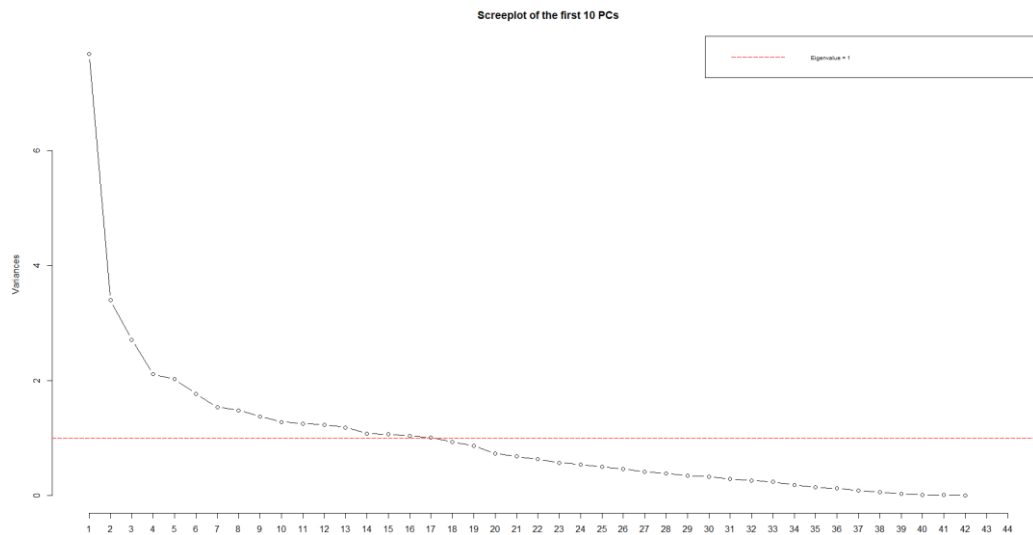
## 10.2. Frequentist logistic regression

<b>Covariates</b>	<b>Coefficient</b>	<b>SD</b>	<b>z-value</b>	<b>Pr(&gt; z )</b>
<i>ability</i>	1.1625**	0.3888	2.990	0.002
<i>gdp</i>	-0.4916	1.0267	-2.046	0.632
<i>aid</i>	-0.4307	0.4343	-0.992	0.321
<i>tax2gdp</i>	-1.1095*	0.5423	-2.046	0.040
<i>rent2gdp</i>	-0.2495	0.6048	-0.412	0.680
<i>intercept</i>	-5.0327***	0.6966	-7.224	0.000

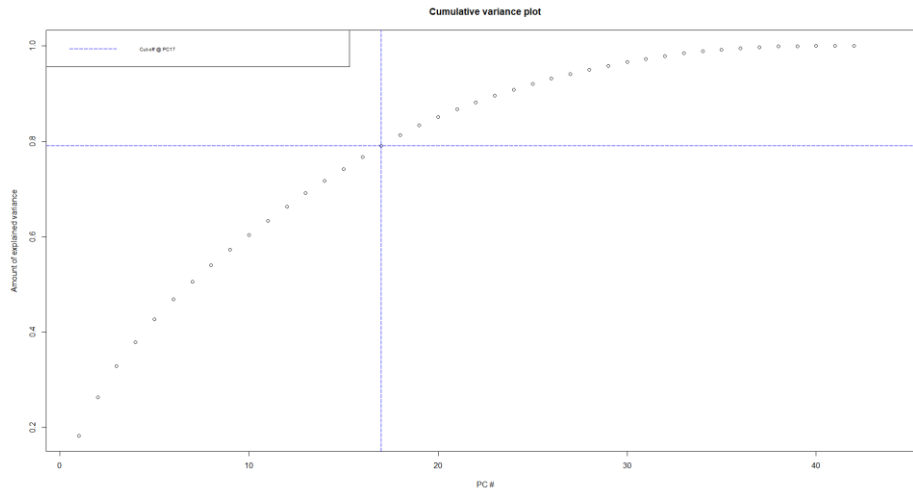
\*\*\* =  $p < 0.000$ , \*\* =  $p < 0.001$ , \* =  $p < 0.05$ . Dependent variable: Transition to democracy. N=442.

The frequentist regression shows that the trends are similar albeit with different coefficient sizes. To avoid deletion of cases with missing values on the control variables, I used the imputed dataset mentioned in chapter 7.2.

## 10.3. Frequentist PCA analysis and logistic regression



The first 17 indicators, sorted by eigenvalue, have eigenvalues  $> 1$  and explain more than one indicator's worth of the variation.



These same 17 indicators together explain 79 per cent of the variation. The results of the frequentist logistic regression can be seen in the image below. The principal components PC1-PC17 are uncorrelated, and they contribute unequally to the prediction of democratization, and in different directions. The results can help understand which indicators tug in what direction, and subsequently improve future measurement models for economic nonviolent strategy.

```

call:
glm(formula = ttd ~ PC1 + PC2 + PC3 + PC4 + PC5 + PC6 + PC8 +
PC9 + PC10 + PC11 + PC12 + PC13 + PC14 + PC15 + PC16 + PC17 +
tax + rent + aid + gdp, data = toreg)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.32584 -0.02623 -0.01418 -0.00368  0.97542

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0208890  0.0066013   3.164  0.001667 **
PC1          0.0038353  0.0024292   1.579  0.115129
PC2         -0.0091505  0.0036334  -2.518  0.012158 *
PC3          0.0147403  0.0041184   3.579  0.000385 ***
PC4         -0.0036059  0.0045755  -0.788  0.431082
PC5         -0.0086254  0.0047111  -1.831  0.067827 .
PC6         -0.0005513  0.0050243  -0.110  0.912676
PC8          0.0079810  0.0054833   1.456  0.146270
PC9          0.0008378  0.0056595   0.148  0.882393
PC10        -0.0025115  0.0058968  -0.426  0.670388
PC11        -0.0046622  0.0059287  -0.786  0.432091
PC12         0.0025267  0.0059764   0.423  0.672678
PC13         0.0043820  0.0061137   0.717  0.473919
PC14        -0.0095370  0.0063555  -1.501  0.134211
PC15        -0.0122188  0.0064066  -1.907  0.057171 .
PC16         0.0001342  0.0065228   0.021  0.983595
PC17         0.0001929  0.0066341   0.029  0.976819
tax          -0.0110309  0.0068474  -1.611  0.107936
rent         -0.0003484  0.0071916  -0.048  0.961384
aid          -0.0119746  0.0071451  -1.676  0.094499 .
gdp          -0.0047785  0.0069195  -0.691  0.490211
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.01917502)

Null deviance: 8.8167  on 441  degrees of freedom
Residual deviance: 8.0727  on 421  degrees of freedom
AIC: -470.91

Number of Fisher Scoring iterations: 2

```

#### 10.4. Ability scores for Kenya, Tanzania, and the US year.

