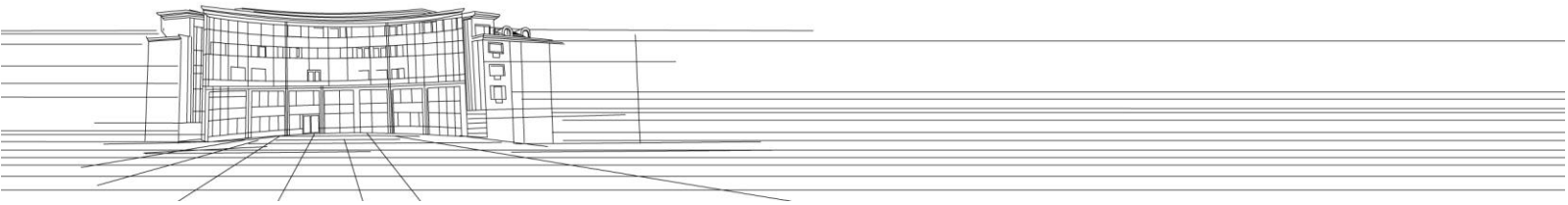


MASTER'S THESIS
INTERNATIONAL ADMINISTRATION AND GLOBAL GOVERNANCE

The Inequality of Fraud

Exploring the effect of societal inequality on electoral misconduct

Author: Valeriya Mechkova
Advisor: Eitan Tzelgov

26 May 2014

Words: 12,227

UNIVERSITY OF GOTHENBURG

Abstract:

The relationship between inequlity and democracy has been a subject to many academic studies. Yet no rigorous explanation has been offered about the connection between them. The present thesis engages in this debate by analyzing the effect of societal inequality on the democratic quality of elections. The hypothesis being tested is that the more the resources in society are unequally distributed, the greater the incentives and opportunities for the incumbent are to use illicit tactics to retain a privileged position. At the same time marginalized groups will be more willing to break the democratic norms in order to defend their rights better and access more power. The new V-Dem data allows for the first time to test these arguments in a comprehensive comparative analysis covering 113 years of history for 139 countries. Using a time-series cross-sectional regression model, the study tests whether and how social, economic and/or political inequality affects the level of electoral misconduct. One contribution of this study is that a new measure for electoral fraud is proposed that encompasses all legal and illegal tactics used by competitors to distort the electoral outcome. The empirical findings corroborate that on average inequality based on social group differences is associated with electoral misconduct. The frequency of fraud is higher when the underlying social differences are translated into the political life or affect civil liberties access.

Key words: inequality, electoral misconduct, elections, democratization

Table of Contents

Abstract: 2
1 Introduction 4
2 Theoretical framework 7
2.1 Democracy, elections and electoral misconduct 7
2.1.1 What is electoral misconduct? 7
2.2 Why do competitors cheat in elections? 8
2.3 Why inequality should cause electoral fraud? 9
2.4 Linear or inverted "U" effect 13
2.5 Research question and hypotheses 15
3 Data and methodology 16
3.1 Advantages of using the V-Dem data 17
3.2 Potential problems and how they are addressed 17
3.3 Operationalization of the dependent variable: electoral misconduct 19
3.4 Operationalization of the independent variable: inequality 22
3.5 Model specification 24
3.5.1 Control variables 27
4 Empirical analysis 28
4.1 Regression analysis 29
4.1.1 Robustness checks 33
4.2 Discussing linearity vs. inverted "U" effect 37
4.3 Limitations 39
5 Conclusion 40
References: 43
Appendix: 46

1 Introduction

The rising inequality in the world is becoming a central topic of discussion for academic scholars, practiocioners and politicians. Global leader meetings like the World Economic Forum[1] warn that social instability, weak institutions and bloody revolutions are likely an outcome of the increasing inequality. Both for developed and developing countries the tendency after the 1980s is that the top 1% wealthiest people have increased sharply their share of the overall income (Piketty, 2014). In the U.S. for example, the top percentile owned around 20% of the total income in 2010 (ibid). Similar facts about unequal distribution of power have driven a number of political science studies focused on answering important questions like: who really rules; is it possible to be a full democracy in conditions of very skewed distribution of social, political and economic resources; also, can a democracy develop in such unequal conditions.

Although the relationship between inequlity and democracy has been a central subject of many academic studies, no rigorous explanation has been established about the connection between the two. By contrast, scholars have found conflicting results. For example, Ansell and Samuels (2010) argue that historically while democratizing, countries have experienced increasing inequality. That is because economic development has led to a bigger gap between classes, as only certain groups accumulated more wealth leaving the larger masses behind. Boix (2003), on the other hand, reasons that democratization is more likely in more equal societies, whereas Acemoglu and Robinson (2006) argue that democratic transitions tend to happen when societal inequality is at middling levels.

This thesis engages in the debate about the relationship between the social structure and democracy by exploring specifically the effect of societal inequality on the election quality. An important assumption made is that elections are a key instrument of democratization,

[1]See the report for the current main Global Risks by WEF:http://www3.weforum.org/docs/WEF_GlobalRisks_Report_2014.pdf. Comments from world leaders on the report http://america.aljazeera.com/opinions/2014/1/davos-inequalityeconomicsinstability.html

and, therefore, understanding and predicting why electoral competitors would engage in fraud has important practical implications for the democratic governance of a country.

In the existing literature the structural conditions underpinning electoral manipulation and the precise causal mechanisms leading to misconduct are understudied. The research that connects the two phenomena is restricted mainly to case studies, captures only a few countries and has a limited time frame. [2] Also, the majority of the studies using inequality as a main explanatory variable focus on economic background solely, rather than consider a multivariate framework encompassing social, economic and political inequality.

The case study of Nineteenth' Century Germany by Ziblatt (2009) has very similar theoretical assumptions to the one employed in this thesis for a positive relationship between societal inequality and incidences of electoral fraud. Ziblatt's starting point is that the unequal distribution of social and economic power generates *opportunities* for the subversion of the democratic institutions that are supposed to isolate politics from pre-existing resource asymmetries (Ziblatt, 2009:3). In effect, socio-economic inequality can impede the "institutionally transformative effect of elections" (ibid). While Ziblatt's findings are both relevant and important; the generalizability of his empirical results may be limited. One may legitimately wonder whether the results will hold for an increased time and geographical span, and if the measures for the outcome and explanatory variables are developed more broadly.[3]

Therefore, there is a need for further investigation on the relationship between inequality and electoral misconduct, which is the research aim of this thesis. The release of the new Varieties of Democracy (V-Dem) data (Coppedge et al, 2013) allows for the first time to test the derived theoretical assumptions in a larger comparative analysis covering 113 years of history for 139 countries. Such a large sample will give reasons to draw generalizable conclusions for the posed research questions.

[2]See for example the case studies of 19th century Germany (Ziblatt, 2009) and Costa Rica (Lehoucq and Molina).
[3]Ziblatt's measure for inequality is focused on difference in the land possession, while fraud is measured dichotomously whether elections were disputed or not.

In this thesis I argue that when resources in the society are unequally distributed, elections as key instrument to power, tend to matter more for the contesters. The incumbent rulers have greater incentive and opportunities to use even illegal tactics to retain their privileged position.The marginalized groups in turn will be willing to invest more, including to engage in fraud, in order to access more power or defend their rights better. However, the opportunities and incentives to conduct electoral misconduct are expected to be low at extreme levels of inequality. Thus, if power is concentrated in one group, they would not need to involve in electoral manipulations, and in the same time the powerless groups would not be able to challenge the status quo. More or less equal distribution of power is also not predicted to trigger high levels of electoral misconduct.

To test my main theoretical predictions, I utilize the disaggregated character of the V-Dem data, and create a new measure for electoral fraud. It encompasses all legal and illegal tactics employed by competitors to distort the electoral outcome in their favor, in a way that is violating the democratic norms. As main predictor variables for the occurrence of electoral misconduct I propose four measures that account for the extent to which social and economic discrepancies affect political power distribution and access to civil liberties. My results show that inequality based on *social group* differences is associated with more instances of electoral misconduct on average, regardless of the other country's characteristics.

The thesis is organized as follows: first, I discuss relevant theoretical findings from the existing literature on democracy, elections and fraud over several different research agendas; then the main arguments, research question and hypotheses are outlined. This is followed by an introduction and discussion of the data, the methodological strategy employed and a presentation of the main findings. Finally, conclusions and policy implications are discussed.

2 Theoretical framework

2.1 Democracy, elections and electoral misconduct

Elections are nearly universal in the contemporary world (Schedler, 2002:38) but scholars of political science agree that holding elections is not enough to call a country democratic (Schedler, 2002; Diamond, 2002; Lindberg, 2006, 2009). Yet elections are a necessary condition for democratic governance as they are the main mechanism that should ensure institutions are accountable to citizens (Stokes et al. 1999). Even when autocratic regimes conduct polls just to gain more internal or external legitimacy, elections are still important as they can introduce uncertainty about the final outcome (Hafner-Burton et al, 2013:155). Elections may well serve as critical turning points to an open democratic system (McFaul, 2002) or lead to "liberalizing electoral outcomes" in authoritarian systems (Howard and Roessler, 2006). Democratization becomes possible when the exercise of repeated elections itself brings about political liberalization, broader civic engagement, and improved political accountability (Lindberg, 2006, 2009; Roessler and Howard, 2006).

Yet, although elections are spread worldwide and can serve an important democratic function, more than half of the current elections in the world violate the democratic principles of basic freedom and fairness, and the respect for human rights (Hafner-Burton et al, 2013:152). It should be noted that electoral irregularities do not occur only in autocratic regimes. Ballot rigging, violence, and collation irregularities occur in established democracies as well (Breunig, Goerres 2011; Alvarez et al, 2011). That is one of the reasons that electoral fraud has become one of the central themes in the research about democracy.

2.1.1 What is electoral misconduct?

Most broadly, electoral fraud includes all tactics that violate the two main criteria for democratic elections described by Dahl (1971) as free and fair. "Clean" elections in this sense require impartial administration in charge of the conduct and control of the whole election process, reasonable and unbiased media coverage, opportunities for a broad spectrum of parties to compete, and, for citizens to vote without the threat of intimidation

and restrictions (Bermeo, 2010:1125,Elkit and Svensson, 1997:35). In fair elections all significant parties accept and respect the election process and its outcome (Pastor, 1998:160).

Therefore, the whole range of illegal and legal actions (Ziblatt, 2009) that breach democratic norms and violate human rights can be considered to be instances of electoral misconduct. Specific examples include fraudulent tactics – any unlawful activity before or during elections; electoral manipulation – bending the rules or legislation in someone's favour; irregularities – using flawed voting registry; violence used to intimidate voters or restrict the access to the polls. In this thesis, I typically use the term electoral misconduct as an umbrella term to encompass all types of tactics used by electoral competitors to influence the outcome of elections in an unfair way. However, since electoral fraud, manipulation and irregularities are terms widely used in the literature to describe the same phenomenon, I use them as synonyms.

2.2 Why do competitors cheat in elections?

Previous research has looked at different competing explanations for the occurrence of fraud, and the debate is ongoing. Naturally, a reason for incumbent rulers to engage in fraud and electoral violence is that they fear an unfavourable electoral outcome (Hafner-Burton et al, 2013). An incumbent uncertain about victory is more likely to use illegal means in order to stay in power than one who feels secure about winning (ibid: 150). In addition, fraud is more likely to be deployed by a highly unpopular incumbent ruler also because there is less to lose in terms of initial support (Collier and Vicente, 2012:119).

One alternative explanation to the one deployed in this thesis is from authors like Birch (2007) and Hicken (2007) who reason that electoral institutions are a mediating factor to the relationship between manipulation and level of political competition. They argue that fraud is more likely to occur in majoritarian or plural single member district systems than in proportional systems because of the more direct, intense, personal competition in a winner-takes-all situation.

Previously, it has been also argued that opposition parties can learn how to prevent fraud when they gain experience and the "democratic quality" of elections improves over time, even if the first elections were not free and fair (Lindberg 2006). Critically, an opposition capable of mobilizing a strategic coalition that can pose a credible challenge to the incumbent in national elections is more likely to be in a position to avert blatantly rigged elections (Howard and Roessler, 2006:370). At the same time, such a situation makes the incentives for an incumbent to circumvent the fairness of the electoral process even stronger.

According to another perspective the distribution of institutional power and the power of the imagined is more important than the distribution of wealth and class actors (Bermeo, 2010:1122). The main motivations to hold fair elections would be the perceptions of cross-class political leaders that fair elections will be beneficial to their organizations' interests, and the assessment that the citizens will expect clean elections.

Another group of authors have found that socioeconomic structures affect the likelihood of holding of free and fair elections. Significant differences in the access to political, economic and social resources between groups is often portrayed as one of the main reasons blocking the development of and consolidation of democratic institutions (Boix, 2003, Acemoglu and Robinson 2006). Since elections are a core practice of democracy (Ziblatt 2009:2), it seems reasonable to build on the findings in this existing literature and to expect that socioeconomic disparities – inequality – undermine the fairness and freedom of elections too.

2.3 Why inequality should cause electoral fraud?

All societies exhibit some degree of inequality and certain groups are wealthier and more politically powerful than others. If differences in socioeconomic conditions and influence are small, they might not have substantial political implications. However, we can expect the situation to be aggravated if income and material capital is translated into political

power such that some groups dominate the political development. The extreme situation occurs when political power is monopolized by a minority social or economic group that can shape policies to benefit their interests only, while the other groups are disconnected from the political process and their interests are not well represented.

In essence, I predict that when distribution of political, social and economic power is skewed both the incumbent ruler and the marginalized groups will have stronger incentives to employ even illegal tactics to win an election. First, I will look at the motivation and opportunities for the more resourceful to engage in electoral misconduct. Presumably more powerful groups (politically, economically and/or socially) would prefer to avoid relinquishing their advantage. If they have already achieved unproportional access to power winning the next elections is important to not lose that advantage. Therefore, using all means possible, including fraud, is justified as the stakes are higher (Lehoucq and Molina 1999, 2002).

In cases of existing unequal distribution of political power, which happens in autocratic regimes, the ruling parties have many opportunities to influence the outcome of elections. Politicians in incumbent regimes have asymmetrical access to state resources that they can use to their advantage compared to the opposition (Greene 2007; Magaloni 2006). Institutions such as courts, electoral management bodies, and prosecutors are more easily manipulated to influence the organization, conduction, monitoring and certification of elections (Magaloni 2010). For example, the media can be used for propaganda purposes to affect public opinion. Bending electoral rules in advantage for the incumbent party in order to manipulate and divide the opposition parties is another possibility to keep the asymmetrical political power distribution (Lust-Okar 2005).

However, even in non-autocratic regimes, the existence of democratic principles and political equality (e.g. universal suffrage) can be effectively weakened by economic inequality (Boix 2003, Ziblatt 2009). If money can be used to influence political actors, its impact is likely to be greater in more unequal societies (Rosset et al 2013:820). In this

environment, the role of the material base is greater because wealthier citizens will have better access to resources to gain political influence including through enhancing their performance in elections for example by having more expensive and better designed electoral campaigns. Therefore, even under the conditions of formal democratic rules, the procedures and outcome of elections are challenged by the in-built asymmetry in resources and the possibility to replicate the socioeconomic gaps into the institution of elections (Ziblatt 2009:3). For instance, a recent study on U.S. politics showed that economic elites and organizations representing business interests are exerting significantly greater impact on the outcome policies compared to the independent influence of average citizens and mass-based interest groups (Gilens and Page, 2014:4).

Social status can be another source of inequality in terms of distribution of power. Top-level positions in political parties and key government institutions can be occupied by groups defined by specific social characteristics (race, language, religion, region etc.). Based on the common features these groups achieve unity through their common background, interests and social interactions (Gilens and Page, 2014:6). Thus, for example authors like Mills (1959) argue that historically, politics was shaped largely by elite groups whose status is not defined solely by their wealth but other coinciding interests and social characteristics.

In unequal societies, the marginalized groups have more reasons to involve in electoral misconduct as well. First argument is that inequality leads to underrepresentation of the poor in the political system which effectively leads to poor defence of their preferences (Rosset et al 2013:819). Therefore, poorer people can be expected to invest a lot in an election if it will lead to guarding their interests better and increasing their share of the power. The same argument is valid for social groups that are excluded from the decision-making process. Rigging elections is similarly rationalized as the only way to protect their interests.

Unequal access to civil liberties can be another trigger to violate some of the election rules. Since in modern history repressive state institutions were the main violator of civil liberties (Møller, Skaaning, 2013:84), protecting those liberties will require a change on the

incumbent ruler. Therefore, extreme tactics in election could be easily justified as the only way to oust the ruling regime and thus, protect the marginalized groups' rights. It is important to account for civil liberties protection also because its unequal distribution might undermine other characteristics of equality. Thus, for example, if private property rights are not defended accordingly for a certain social group, the economic wealth they have accumulated might be jeopardized by coercive state behaviour.

Furthermore, independent institutions like legislatures, other strong political parties, the armed forces or the judiciary can serve as accountability groups and constrain attempts to conduct electoral misconduct (Hafner-Burton et al, 2013:154). If perpetrators of fraud, both the incumbent and opposition parties, realistically face a response and some kind of penalty on part of powerful accountability groups, the motive to engage in fraud decreases (ibid, p. 156). Yet this causal mechanism works only if the mentioned groups are not significantly weaker than the violators, which tends not to be true in societies with much skewed distribution of political power. Thus, in closed authoritarian systems, the opposition could not rely on impartial reaction from the state apparatus to instances of fraud. In addition, if the state institutions are used by the incumbent to protect their own interest in elections, the opposition will be discouraged from participating peacefully and lawfully because of the low expectations that their votes will be counted fairly.

Another argument why we should expect a relationship between inequality and fraud is that less resourceful people can be more vulnerable against perpetrators. Groups with lower income also tend to have lower levels of education and knowledge about their rights as citizens and how to defend them (Converse, 1972; Verba, Schlozman and Brady 1995:305). As Converse argues (p.324) the better educated a citizen is, the more knowledgeable he/she is about their rights and politics as a whole, and the more they are motivated to participate in political activities. This is because formal education brings about a stronger interest in politics, a better understanding of the importance of elections, and not the least, education nurtures the commitment to being an active citizen (ibid).

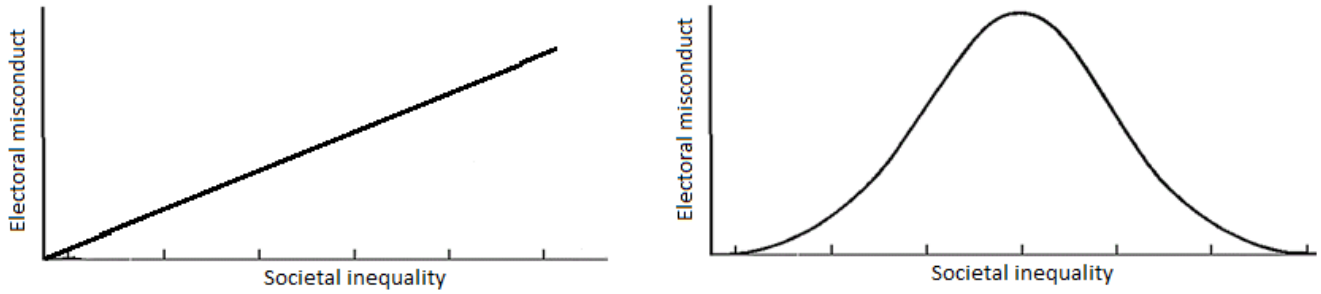
In addition, the lack of financial resources can be expected to hinder the consolidation and organization of poorer groups in practical terms as well. For instance, financing political parties and their activities like meetings or public campaigns becomes less viable. We can also expect that especially one of the strategies of fraud – vote buying will be more widespread in poorer regions, where due to economic difficulties people might accept to sell their votes (Kitschelt and Wilkinson, 2009). An illustration of the last arguments is the study on elections in Costa Rica by Molina and Lehoucq (1999). Their findings support the hypothesis that the incidences of fraud were concentrated in the poorest, rural and least populated areas of the country. The citizens from these regions were not able to defend themselves against the violations of electoral law compared to their counterparts from richer urban areas.

2.4 Linear or inverted "U" effect

Building on the arguments presented so far, the intuition is that the relationship between inequality and electoral fraud/irregularities is linear. That is, the larger inequality is, the more fraudulent elections will be. Alternatively, we can predict that the probability of fraud will be lower at the extreme levels of inequality. If this is true, we would then expect a relationship between societal inequality and electoral fraud that looks like an inverted "U".

Similarly, Acemoglu and Robinson predict an inverted "U" relationship in which democratization will be possible at middle levels of inequality (2006). Applying this logic to the research question examined in this paper, we could expect that at moderate levels of inequality, elections will be more significant for all participants, and therefore, the tendency for fraud should increase. Figure 1 helps to visualize the comparison between the two predictions for the relationship between the explanatory and outcome variables.

Figure 1. Linear versus inverted "U" relationship between societal inequality and electoral misconduct



The expectation in the second graph is that in the most hegemonic and authoritarian regimes, where power is concentrated in one group, electoral manipulation should not be necessary because the opposition is too weak to contest the elections. It is more likely that rulers choose to engage in fraudulent tactics and repression when they feel insecure in their victory in elections (Magaloni 2010, Diamond 2002).

Regarding the opposition, investing in fraud could be justified only if victory is believed to be within reach. This is possible only in a system that provides an opportunity to change the status quo. Whereas in a rigid hegemonic society, where much of the power is concentrated in one group and civil liberties are fundamentally violated, change might be viewed as unfeasible. Thus, in a more open and equal system, opposition parties perceive the incumbent regime as the key obstacle to achieving their goals. By contrast, in closed authoritarian regimes the opportunity for change – elections, does not exist (Howard and Roessler, 2006:369) and severe civil liberties restriction might not allow real competition. Similarly, when we consider the material background, if the opposition possesses little economic resources, financing election campaigns is more difficult. Also, conducting electoral fraud requires a certain amount of resources as well, for example for vote buying, bribing officials, acquiring weapons to intimidate voters/opponents etc.

On the other end of the equality spectrum, in a system of more or less equal power distribution, the incentives for fraud should be smaller as well. In democratic regimes, where political power distribution is relatively equal, groups have accepted the rules of the game and the institution of elections. The introduction of free and fair elections "institutionalizes uncertainty" (Dahl, 1971). That is, the way democratic institutions are

created ensures that the political process is neutral by separating as much as possible political outcomes from the pre-existing social structure (Tilly 2007: 117–20, Dahl 1971). In addition, if there is no gross inequality by socio-economic groups present, mobilizing resources to manipulate the electoral outcome can be expected to be more difficult to justify. Similarly, if civil liberties are denied to whole groups, mobilization would be easier.

2.5 Research question and hypotheses

The broad research question this thesis aims to answer is *whether and how societal inequality influences the occurrence of electoral misconduct.* Drawing from the theory discussed in the previous sections, two hypotheses will be tested to answer the main research question.

> *Hypothesis 1: Societal inequality is positively related to the instances of electoral misconduct.*
>
> *Hypothesis 2: Instances of electoral misconduct are more frequent at moderate levels of societal inequality.*

While in Hypothesis 1 the relationship between outcome and predictor variables is expected to be linear, the second one suggests a relationship that looks like an inverted "U". That is, both at the extreme levels of inequality, with concentration of power in one group, and in relatively equal societies, the incentives to use fraudulent tactics will be fewer.

The first two hypotheses will provide evidence regarding the main theoretical questions raised in this thesis whether there is a relationship between societal inequality and manipulations during elections. The question remains, however, which aspect of equality is most detrimental to the process of clean elections. By utilizing the disaggregated character of the V-Dem data, I include in my analysis measures that take account for two specific types of societal inequality. The last two hypotheses will articulate the more specific

arguments that electoral misconduct is caused by inequality based on social or economic grounds.

Hypothesis 3. Instances of electoral misconduct are more frequent in societies with inequality based on socio-economic position.

Hypothesis 4. Instances of electoral misconduct are more frequent in societies with inequality based on social group.

3 Data and methodology

The new data on different dimensions of democracy that V-Dem has produced allows empirical tests of many theoretical arguments in the field of democracy studies. Below the core ideas of the project are reviewed, as well as some of the advantages and disadvantages of using V-Dem data for the purposes of testing the above hypotheses.

The main goal behind the V-Dem project is to produce transparent and measurable indicators capturing various aspects of democratic systems and practice[4]. The data collection covers all countries in the world starting from 1900 to the present. The dataset is compiled by gathering factual information from existing data sources, and by expert coding for questions that require evaluation. The majority of experts are nationals of the country they are coding, which is one of the biggest strengths of V-Dem. That is, V-Dem incorporates "deep, local knowledge" about the history of a country and by standardized measurement matches this knowledge to a global understanding of what democracy is.

[4] www.v-dem.net

3.1 Advantages of using the V-Dem data

Below is a summary from the V-Dem project description (Coppedge et al 2013), of the main features that distinguish the V-Dem data in comparison to other indices[5] that motivated my decision to choose this dataset instead of other existing measures.

First, V-Dem seeks to create quantitative measures that capture as precisely as possible the different dimensions that make a country more or less democratic. To this aim, V-Dem distinguishes among seven main principles of democracy[6]. Each index is disaggregated into a number of constituent component parts, in total almost fifty, and each component is measured by several indicators. For example, one of the core components of electoral democracy – the quality of elections is assessed by combining 38 different indicators. The disaggregated nature of the V-Dem data allows selecting the indicators that capture most accurately the theoretical concept of electoral fraud motivated in this thesis. Having many disaggregated measures will also allow designing my own indexes and explore relationships between specific elements for the purposes of the study– including how different aspects of inequality relate to electoral manipulation tactics. In addition, as the V-Dem coding starts in 1900 for all countries in the world, it is possible to investigate systematically the relationship between inequality and electoral misconduct by utilizing both variations across time and polities.

3.2 Potential problems and how they are addressed

Quantifying phenomena like political equality and electoral misconduct is challenging, because it can be argued that these concepts are 'latent.' That is, while we can all agree that political equality is greater in contemporary Sweden than in 1930s Germany, individuals will tend to disagree on the degree to which cases differ and might have different understandings of what "inequality" is. In addition, coders' thresholds vary for where meaningful "big shifts" occur on a scale from maximal inequality to non-existent. Different individuals may simply have different intuitions of what for example the midpoint between

[5]The most widely used indices now are Freedom House (www.freedomhouse.org) and Polity IV (http://www.systemicpeace.org/polity/polity4.htm).
[6] The seven principles are electoral, liberal, participatory, majoritarian, consensual, deliberative, and egalitarian democracy.

these two extremes look like. This means that the probability they will assign a given numerical score to a given case can depend on various individual attributes like education, cultural background but also on their interpretation of "inequality" is as a concept. Finally, one might say that coders would provide correlated ratings even for different indicators as their answers will be influenced by their general perspective on the development of the country.

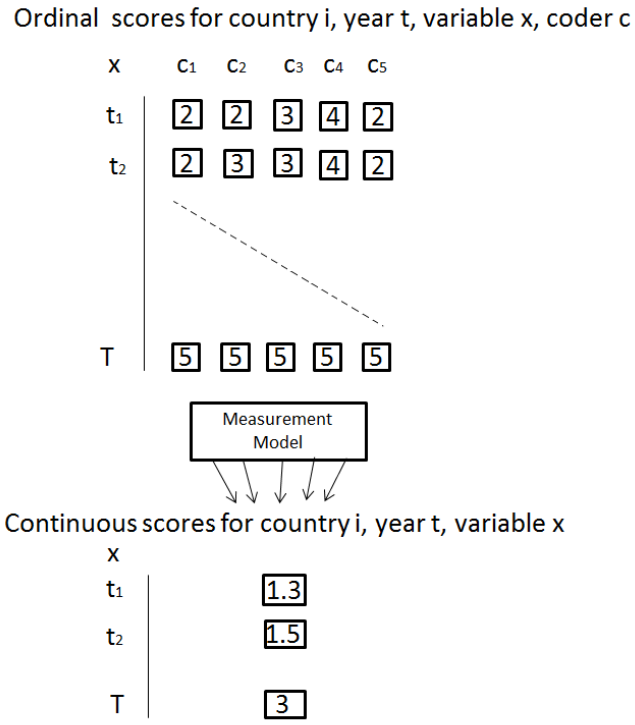
To address these issues and increase the validity of the data, the aggregation of the V-Dem data is done using a statistical measurement model[7] designed to test and correct inter-coder reliability. The V-Dem's measurement strategy is to build on the 'Item Response Theory' model, commonly used in educational and psychometric testing.[8] Figure 2 below illustrates visually how the measurement model works. In essence the measurement model is based on the ordinal scores multiple coders provide for a single variable X, country *i* and year *t*. A single continuous score for each case (question, country, and year) is produced by calculating a point estimates on a newly constructed latent scale. The model takes into consideration how reliable the individual coders are and what their threshold is to move from one category to another for the variable of concern (e.g. *Election vote buying*) [9].

[7] The measurement model is designed and implemented by Dan Pemstein, Eitan Tzelgov and Yiting Wang.

[8] In the field of political science IRT is developed by authors like Jackman, 2001; Cox and Poole, 2008 mainly to estimate legislators' ideology using their recorded votes.

[9]To get a further intuition regarding the way the model works, consider the way these models are used in education studies. Test designers would like to write a questionnaire that would be able to estimate the test taker's IQ, but also know how good the questions in estimating the 'latent' concept - intelligence. Thus, the measurement model estimates three parameters. First, the model estimates the IQ level for each test taker based on their wrong/correct answers. Second, the model estimates two question level parameters. The first is generally called a 'difficulty' parameter, and thus reflects the probability that test takers will choose the correct answer. The second parameter is named a 'discrimination' parameter, and reflects the degree to which a given question provides information on the latent concept (IQ) being measured. In this regard, questions with high discrimination parameters are considered to be better.
V-Dem uses a similar model, in which country experts provide ratings regarding various aspects of democracy. Based on these ratings, every case (specifically, country year) is assigned a continuous score on the newly constructed latent scale, and raters are assigned discrimination parameters (essentially, how good a coder is) and difficulty parameters (i.e thresholds between different levels of the variable).

Figure 2. Measurement model structure



Lastly, one might assume that coders from country A are different from coders from country B. In order to ameliorate this problem, V-Dem uses 'bridge coders'. These are coders that code multiple countries and the information they provide is essential to guarantee cross-national comparability. By introducing a number of control variables and performing a series of robustness checks I seek to further address some of the issues mentioned above.

3.3 Operationalization of the dependent variable: electoral misconduct

The aim of this thesis is to investigate how structural conditions in society affect the tactics electoral competitors employ during elections. Specifically, the outcome variable on which the analysis focuses is electoral misconduct.

There is no consensus in the literature on a measure for fraudulent tactics used during elections. Similarly to measuring corruption, quantifying electoral manipulation as a shadow activity is particularly difficult because the subjects of fraud want to remain hidden

and are unwilling to reveal their actions. Scholars have used measures such as the number of filed petitions to nullify elections and the number of contestations (Ziblatt, 2009, Lehoucq and Molina, 1999) to approximate the level of electoral misconduct. However, accusations of fraud are often used by opposition parties to justify their defeats and in order to try to reduce the political legitimacy of the winners (Lehoucq and Molina, 1999, Lindberg 2006) putting a question-mark to the validity of such measures.

To create a measure for electoral misconduct, I have selected ten V-Dem indicators that capture different aspects of irregularities conducted during national elections. They are presented in brief in Table 1 with the main tactics they account for. The last column summarizes considerations from the existing literature on how each tactic affects the quality of elections, and this is the justification for the inclusion in the measure[10].

Table 1: Measuring Electoral Misconduct

	Variable name	Main tactics included in the measure	How does the tactic affect the quality of elections
1	Election vote-buying	Distributing money or gifts to influence decision to vote or whom to vote for	Violates the freedom of choice in elections (Schedler 2010:40, Ziblatt 2009); affects particularly the economically disadvantaged
2	Elections multiparty	A few parties are legally allowed to stand for elections but they are all strongly influenced by the incumbent party	Elections are not meaningful unless citizens can choose between substantially different options; freedom of choice is restricted otherwise (Schedler 2010:40)
3	Election voter registry	Intentionally manipulation of the registry by adding/deleting names of citizens entitled to vote	Manipulations of the registry and its flaws might lead to disenfranchisement of voters, double-voting and impersonation (Coppedge et al, 2013)
4	Government intimidation	Violent harassment and intimidation of the opposition by the government or its agents	Voters or opposition parties could be intimidated and discouraged to vote or continue their participation in elections (Collier, Vicente; 2010)
5	Other electoral violence	Election-related violence conducted from and between citizens/non-governmental agents	Intimidation of citizens/parties, at the extreme levels, could lead to taking over power by violence (Collier, Vicente; 2010)

[10] The variable names, and the description of the tactics and their effects are drawing from the V-Dem Codebook (Coppedge, Gerring, Teorell, Lindberg; 2013)

6	EMB Autonomy restricted	EMB is prevented from applying election laws and administrative rules impartially	EMB is a central institution exercising constraints on the opportunities for fraud(Hafner-Burton et al, 2013)
7	Free campaign media	Access to media and campaign coverage is restricted to the ruling parties and candidates only	Citizens should have access to information and learn about available political alternatives; free elections include freedom of opinion formation (Schedler, 2002:39)
8	Eligibility restricted	Legal provisions prevent the eligibility of candidates for national office restricted by ethnicity, race, religion, or language	Equal opportunities on individual level to stand for office makes elections "fair"
9	Other voting irregularities	Other intentional irregularities: e.g. using double identities, intentional lack of voting materials, ballot-stuffing, misreporting of votes etc.	Intentional irregularities might distort the will of the electorate and steal the purpose of elections
10	Elections free and fair	Comprehensive measure of the overall election process encompassing all tactics compromising elections	Impairs the opportunity to effectively exercise the democratic right to select the rulers

Eight of the ten indicators focus on manipulations and irregularities conducted right before or during national elections. The two indicators measuring institutionally designed factors that might affect the quality of elections are *elections multiparty* and *eligibility restricted*.

After selecting the relevant aspects of electoral misconduct, my goal is to transform these variables into a general index of the underlying latent variable. To this aim, I use factor analysis to reduce the ten indicators quantifying different aspects of electoral misconduct from Table 1 to one single indicator. The scores from the factor analysis[11] are applied in the regression analysis as dependent variable, retaining essentially the variation from the original data (Rummel, 1967).[12] Lower values for that index correspond to more incidences of fraud while higher values will mean "cleaner" elections. [13]

[11]Factor analysis output is presented in the Appendix, Table A.1.

[12]This is done with the following steps (Rummel, 1967): first, the loadings from the factor analysis matrix determine whether there is a pattern in the variables variation. Every constituting variable is weighted according to its involvement in the pattern and, hence, variables with less involvement in the pattern will have lower weight in the final score and respectively, variables with more involvement in the pattern will have bigger weight. Subsequently, the initial score from the data (country-year-variable) is multiplied by the weight of that variable in the pattern. The score derived for all variables is then summed to produce a final factor score on electoral fraud for each country and year.

[13] The exact formulation of the indicators questions and answers is presented in the Appendix.

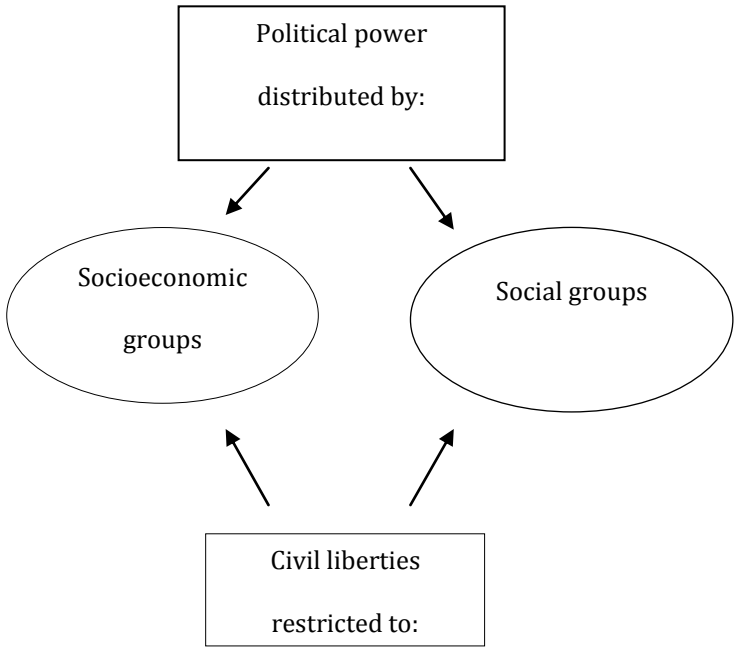
3.4 Operationalization of the independent variable: inequality

The explanatory variables in my analysis account for levels of inequality in a country. In the existing literature inequality is typically perceived as a homogenous phenomenon that can be measured by a single variable – often this is Gini coefficient. As it was mentioned, the previous study on the relationship between inequality and electoral fraud uses the difference in possession of land – a major source of wealth and power in the past, as a proxy for inequality. The unequal distribution of land can undermine the fairness of elections because landlords were able to influence electoral outcomes by using their social power (Baland and Robinson, 2006). Ziblatt extends that argument by adding that land elites were also able to exercise control over local institutions, thus acquiring the institutional base and coercive resources to rig the election conduct and outcome (2009:9).

Ziblatt's measure of inequality– differences in the holding of lands provides a good mediation of the distribution of wealth and power in the nineteenth century Germany when land was a key source of political influence. However, the importance of landholding for the distribution of wealth and power has changed significantly over time. In addition, inequality has other important dimensions that can be expected to affect the exercise of democratic rights differently. Since it is possible that various aspects of inequality affect the incidence of electoral misconduct, a multivariate framework is used in the following analysis. This makes it possible to assess the degree to which different aspects of inequality affect the dependent variable.

Specifically, I am interested in the extent to which social group and economic differences are detrimental to other key characteristics of a society – political power distribution and access to civil liberties. Figure 3 visualizes the aspects of inequality as they are specified with the main independent variables of interest, extracted from the V-Dem dataset.

Figure 3.Aspects of societal inequality as captured by the main explanatory variables.



Four indicators capturing different levels of social, political and economic inequality are available in the V-Dem data. The first two indicators measure whether *political power* is distributed according to *socioeconomic position* or *social groups*. Specifically, they focus on the extent to which wealth and the class structure are transformed into political power. A social group is termed as individuals that identify themselves as having common ethnicity, caste, language, race, religion, come from the same region or define themselves with some combination of the mentioned (Coppedge et al, 2013)[14].

The third and fourth V-Dem indicators measure whether all *social* and *socio-economic groups* enjoy the same level of *civil liberties*. That is, whether all people have equal access to justice, private property rights, freedom of movement, freedom from forced labor (ibid).

Lower values for the inequality variables describe a gross unequal distribution of resources, while the highest values correspond to more or less equal societies. These four indicators together capture in a relative comprehensive way the different aspects of inequality in a society. Analyzing them will give a sense of whether societal inequality in general affects the instances

[14] The exact formulation of the indicators questions and answers is presented in the Appendix.

of electoral misconduct, while the individual indicators included in the analysis will allow more fine-grained tests of how exactly the mechanism works.

3.5 Model specification

To test the assumed hypotheses on the relationship between societal inequality and the freedom and fairness of elections time-series cross-sectional (TSCS) regression model is applied, using ordinary least square (OLS) estimation procedure and has the following form:

$$y_{i,l} = \beta_0 + \beta_k X_{k\,i,l} + \varepsilon_{i,l} \qquad (1)$$

In equation (1) the measure β_k is the predicted effect that one unit of change in k number of independent variables $X_{i,l.}$ will produce in the dependent variable $Y_{i,l}$. The equation also includes a common intercept β_0 and an individual error term $e_{i,\,l.}$ The observations are indexed by unit (country) "i" and time "l", which signifies election year in my models.

There are three important assumptions that need to be considered when applying OLS procedure with TSCS model. First, we have to take into account that the observations in TSCS are yearly observations for the same political units. This might violate the OLS assumption that the observations are independent. The problem that should be considered is that there is high probability that the independent variables in equation (1) are endogenous, or in other words correlated with the error term in ε_i. Secondly, all errors should have the same variance across units (homoscedasticity assumption). If the errors are not "spherical" in this sense, the standard errors in the model will be miscalculated affecting the significance of the results (Beck, Katz, 1995:4). The third issue to be taken into account when applying regression analysis is the direction of causality or the claim that one variable causes changes in the other.

In addition it should be noted that the variables included in vector X (independent variables) should not be correlated with each other. Otherwise they will cause multicollinearity problem, and produce larger standard errors which will make it difficult to

reject the null hypothesis. If this is the case, I should consider excluding one of the collinear variables.

In the literature there are a number of recommendations how to deal with violations of the described assumptions and achieve a more reliable model.

One of the methods widely applied in social sciences is to use country fixed effects (Greene, 2008:183). In order to make sure that the regression results are not caused by constant characteristics of the countries not included in the model such as geography, institutions and population size and in order to isolate the time-variant effects we are interested to measure. When we use fixed effects, we assume that we need to control for the individual time-invariant specifics of a country that might bias or impact the outcome or explanatory variables.

Lagging the independent and dependent variables is another adjustment recommended to overcome the endogenity bias and autocorrelation (Beck, 2001; Keele and Kelly 2004). In effect, lagged dependent variable serves as a proximate test for causality direction and also, controls for "history" or in other words is a proxy for the effect of other omitted variables in the model. These two specifications modify the model in the following way:

$$Y_{i,l} = \beta_0 + \beta Y_{i,l-1} + \beta_k X_{ki,l-1} + \gamma_i + \varepsilon_{i,l} \qquad (2)$$

Where γ_i specifies the inclusion of country fixed effects and the design l-1 denotes the lagging of the independent and dependent variables. For electoral misconduct, the lag will measure the levels of fraud in the previous elections, and for the explanatory variables – the levels of inequality one year before the election.

Theoretically, the lagging of the main variables also holds ground, since in reality the behavior of actors is affected by history. Thus, we could specify that equation (2) measures the current level of electoral fraud as a function of past levels of fraud modified by the perceived information on levels of inequality. The lagging of the dependent variable and

independent variables consequently adds dynamic interpretation to the model (Keele and Kelly 2004:5).

The lagging of the main variables, however, might create side effects since they will be highly collinear with the original variables, leading to imprecise estimates of the betas (Keele and Kelly 2004:6). While Keele and Kelly show that the bias is not serious as long as panel sizes are large enough, the inclusion of lags can artificially reduce the explanatory power of the theoretically motivated variables. To control for such bias, I will estimate and compare both models with and without a lag.

The last modification advised by the methodological literature aims to overcome heteroskedasticity and non-spherical errors, by applying panel-corrected standard errors (PCSEs) (Beck 1995, 2001). The advantage of this method is that it reflects closely the variability of the βcoefficients produced by OLS without distorting the data while correcting for problems that affect the measure of the standard errors (Beck 2001:13). The final model as described with the above specifications can be represented in the following way:

$$ElectoralMisconduct_{i,l} = \quad (3)$$
$$\beta_0 + \beta_1 Political\ power\ distributed\ by\ social\ group_{i,l}$$
$$+ \beta_2 Political\ power\ distributed\ by\ socioeconomic\ position_{i,l}$$
$$+ \beta_3 Civil\ liberties\ by\ social\ group_{i,l}$$
$$+ \beta_4 Civil\ liberties\ by\ socioeconomic\ position_{i,l}$$
$$+ \beta_5 X_{ki,l-1} + \beta\ Electoral\ Misconduct_{i,l-1} + \beta_6 x_{ji,l} + \gamma_i + \varepsilon_{i,l}$$

Where$X_{ji,l}$ is a vector including all control variables to be introduced in the model.

The interpretation of equation (3) is that the level of electoral misconduct depends on the levels of different aspects of societal inequality all other factors held equal. The results of this regression will give us arguments to support or reject the four specified hypotheses. The discussion of statistical significance and coefficients results estimates will allow further exploration of the relationship between the dependent and independent variables.

3.5.1 Control variables

When analyzing the effect of inequality on electoral misconduct, it is important to consider alternative explanations that may affect the behavior of electoral competitors. The introduction of control variable will thus reduce the likelihood of spurious findings on the relationship between the dependent and independent variables. To that aim, based on previous findings in the literature I supplemented the main model by adding other important factors that might affect the occurrence of electoral misconduct.

The first control variable is ethnic fractionalization since internal divisions in a country can affect the likelihood of violence. The variable chosen uses a definition of ethnicity involving a combination of racial and linguistic characteristics collected by Alesina et al (2003).

In addition, control is introduced for a standard measure for wealth in a country to make sure that the changes in the dependent variable are not result of differences in economic development across the countries. The indicator GDP/PPP is extracted from the World Bank Development Indicators and measures the gross domestic product using purchasing power parity rates (World Bank WDI, 2013).

A third control variable is an estimate of Gini-index of inequality. The Gini coefficient varies from minimal value 0 which corresponds to the theoretical possibility for perfectly equal income distribution in the society to maximum value 100 in which case the society's total income belongs to only one person (Teorell et al, 2013). The indicator comes from the United Nations University's World Income Inequality Database (Solt 2008), and is supplemented with the data on income inequality gathered by Ansell and Samuels mainly for years before 1960 (2010). Controlling for Gini coefficient tests for the theoretical assumption that economic inequality is the only motive to conduct fraud.

Since previous research (see for example Birch, 2007; Hicken, 2007) has found that the type of electoral system affects the dynamics of elections significantly, a control for this factor is introduced. The measure is extracted from QoG database and uses Golder's (2005)

Democratic *Electoral Systems* dataset to identify the type of electoral system – majoritarian, proportional or mixed.

Lastly, I will control for regime type to make sure that it is not the specific characteristics of democratic regimes only that prevent fraud (Lindberg, 2006; Howard and Roessler, 2006). To that aim, I will use the mean score from Freedom house and Polity for democracy (fh_ipolity2) designed by Teorell and Hadenius (2005)[15].

4 Empirical analysis

To test the observable implications of my arguments for the relationship between inequality and electoral misconduct, I perform regression analysis, using data on elections and inequality for 139 countries[16] for 2453 election years, held from 1900 until 2012. Before testing my main models, I discuss the preliminary analysis of my data. Table 2 below provides summary information of the descriptive statistics for all variables employed in the subsequent analysis.

Table 2. Descriptive Statistics of the variables used

Variable name	*N*	*Mean*	*Std. Dev.*	*Min*	*Max*
Electoral misconduct	2453	0.08	0.96	-2.22	2.45
Civil liberties equality for social groups	2570	0.31	0.85	-1.67	1.87
Civil liberties equality for social class	2496	0.27	0.84	-1.89	2
Political power by socioeconomic position	2507	0.16	0.80	-1.74	1.94
Political power by social group	2467	0.43	0.74	-1.48	2.09
Ethnic fractionalization	1639	0.43	0.27	0	0.93
Gini coefficient	1781	26.58	19.61	0.28	71.33
GDP per capita, PPP	951	8489.61	9413.24	207.05	47626.3
Regime type (FH_polity)	1200	5.95	3.18	0.25	10
Electoral system	931	1.82	0.65	1	3

[15] The exact formulation of the indicators questions and answers is presented in the Appendix.
[16] Table A.2 in the Appendix lists the countries included in the analysis.

The first variable in Table 2 is the dependent variable – *Electoral misconduct*, derived using factor analysis; followed by the four main explanatory V-Dem variables, and lastly, the control variables are presented.

The dependent variable is continuous and displays a more or less normal distribution (Graph A.1 in the Appendix). Although it shows a slight negative skew and lightly tailed distribution (Table A.3 in the Appendix), the distribution of the observations are close to normal. In addition, the check for multicollinearity problem show that the independent variables are not a linear function of one another as the VIF-values (variance inflation factor) in Table A.5 in the Appendix demonstrate.

However, Pearson correlation coefficient[17] shows that the control variable measuring regime type (combined Freedom House and Polity IV measure) is highly correlated with the dependent variable (0.88). This signals that the variance of the left and right hand side variables is explained by similar factors. In essence, this means that the effect of regime type on electoral misconduct is correlated with the error term, and in the regression estimates this might lead to biased beta coefficients. To avoid the occurring endogeneity problem and still make sure that the type of regime is not the most important factor affecting electoral misconduct, I will re-estimate the main models and exclude from the data countries with *extreme levels* for the variable measuring regime type in the robustness checks. Thus, election years for full democracies with scores higher than 9, and full autocracies with scores lower than 2, will be excluded from the regression.[18]

4.1 Regression analysis

Table 3 reports the estimates of the models described in the previous sections.The discussion of the statistical significance and estimated coefficients from the regression will provide arguments to support or reject the hypotheses presented above.

[17] Pearson correlation coefficient table is attached in the Appendix, Table A.4. The remaining variables do not show strong correlation between each other.
[18] Similar method is used by Howard and Rossler (2006:368) for their purposes. They are excluding closed authoritarian systems (lowest scores) and full electoral and liberal democracies (highest scores) to get data only on competitive authoritarian regimes.

Table 3. Regression estimates of the effect of Societal Inequality on Electoral Misconduct

Variable name	Model 1 V-Dem Ind.V	Model 2 LDV	Model 3 LDV PCSE FE	Model 4 LDV, Control variables, PCSE	Model 5 LDV, Control variables, PCSE, FE
Power distributed by socioeconomic position	0.137***	0.005	0.045	-0.029	-0.002
	(0.03)	(0.02)	(0.02)	(0.04)	(0.06)
Power distributed by social group	0.368***	0.093***	0.200***	0.067	0.382***
	(0.04)	(0.02)	(0.03)	(0.07)	(0.09)
Social group equality in respect for	0.272***	0.052**	0.155***	0.069	0.363***
civil liberties	(0.03)	(0.02)	(0.02)	(0.05)	(0.09)
Social class equality in respect for	0.046	0.005	0.032	0.044	-0.002
civil liberties	(0.04)	(0.02)	(0.03)	(0.04)	(0.12)
Electoral misconduct lagged		0.874***	0.721***	0.680***	0.270***
		(0.01)	(0.03)	(0.05)	(0.06)
Ethnic fractionalization				0.016	.
				(0.07)	.
Gini coefficient				0.001	-0.011*
				(0)	(0)
GDP per Capita, PPP				0.000	0.000
(Constant International USD)				(0)	(0)
Electoral system				0.068**	0.188***
				(0.02)	(0.05)
Constant	-0.254***	-0.030**	-0.354***	-0.126	-0.594**
	(0.02)	(0.01)	-0.07	(0.15)	(0.22)
Observations	2173	2043	2043	510	510
ll	-2576.723	-911.858			
aic	5163.446	1835.715	.	.	.
R-squared	0.349	0.851	0.87	0.796	0.882

Legend: * p<0.05; ** p<0.01; *** p<0.001.Robust standard errors within parantheses.

The first basic model presented in the table includes only the four V-Dem main predictor variables measuring inequality and calculates that three of them are statistically significant predictors of electoral misconduct at .001 level. Model 2 augments the original by including in vector X a lag with one election of the dependent variable *Electoral misconduct*. By comparing the two models I test for two different theoretical assumptions – while the first model will predict how one unit change in the variables for inequality affects the dependent variable, the addition of the levels of electoral misconduct in the previous elections will test whether the *change* of occurrence of electoral fraud from one year to another is influenced by inequality. The lag of the dependent variable is also improving the model in methodological terms (e.g. serves as a proxy for omitted variables). The regression results in Model 2 support the hypothesis that the changes in the levels of electoral misconduct are explained by the variables measuring inequality based on social groups (*Power distributed by social group and Civil liberties equality by social group*).

I re-estimate the obtained results in Model 3 by including country fixed effects[19] in order to analyse the impact of the explanatory variables over time by controlling for constant country characteristics, and adjusting the standard errors with PCSE. The regression results only in a modification to the values of the beta coefficients without changing either the directions or the levels of significance in comparison to Model 2. The findings are thus robust so far.

Model 4 adds all control variables derived from the theoretical review as alternative explanations to my main arguments[20], and applies the method of panel-corrected standard errors (PCSE). It is interesting to note that Model 4 does not show statistically significant results for any of the main independent variables. The statistical significance returns in the last model (Model 5) which applies country fixed effects with country dummies, used to

[19] Conducting Hausman test to decide between fixed or random effects produced results advising to use the first one as the null hypothesis that the error terms are not correlated with the regressors was rejected (Greene 2008, Chapter 9).
[20] Table A.6 in the Appendix presents the regression results when the control variables are added step by step without PCSE method and country fixed effects.

isolate the time-invariant country effects, as well as all main explanatory and control variables included on the right hand side of the equation. That change of the significance levels can be explained with large variance of the observations in the variables *across* countries which do not allow getting significant results in Model 4. The isolation of country-specific characteristics, however, tests the effect of the independent variables on the outcome variable *within* country. The substantive implication from these results is that the significant results obtained are valid *at country level*. In other words, this means that higher levels of inequality in country X than in country Y is not associated with the variation in levels of electoral misconduct that these two countries display. Yet, a change from higher to lower inequality within country X over time significantly affects and lowers the frequency of electoral misconduct in that country. This is exactly the type of substantive effects the theory predicted.

The most important implications of the regression estimates are the following: two variables for social inequality – *political power distributed by social groups* and *social group equality in respect for civil rights*, are statistically significant at level .001 in the last model, which introduces country fixed effects and the method PCSE. One unit change in the two explanatory variables leads to increase in the levels of "freedom and fairness" of elections of .382 and .363 respectively. The produced change is noteworthy since the variation scores in the dependent variable are between -2.2 and 2.4 (Table 2). These findings support the hypotheses, thus suggesting a strong and positive relationship between electoral misconduct and fraud. *Hypothesis 4 Instances of electoral misconduct are more frequent in societies with inequality based on social group,* is corroborated.

However, the two variables with emphasis on socioeconomic position do not hold statistical significance persistently across the models; and in different models have an opposite (negative) sign to the one we expected. This means that as their values increase which marks more equality, the level of fraud grows too, unlike what we predicted in the previous sections. The effect of both variables is also quantitatively small with a coefficient of only -.002 in the last model. Therefore, hypothesis 3 *Instances of electoral misconduct are more frequent in societies with inequality based on socio-economic position,* is rejected.

R-squared, presented in Table 3, tells us that for the first model 35% of the variation in electoral misconduct is explained by the five V-Dem variables for inequality. The introduction of the LDV adds 50% to the explanatory power of the model, which means that its inclusion serves as a proxy for the omitted variables as expected. The following models maintain the percent of explained variance very high with values around 80%, which can be considered as very substantial.

In conclusion, the regression estimates show that inequality based on social differences increases the likelihood of electoral manipulations. This means that, for example, when one religious group holds in practice the monopoly of political power, actors will be more likely to engage in unfair tactics. The same is true, when certain social groups are denied civil rights, for example if the freedom of movement is restricted for a certain ethno-linguistic group. However, the empirical tests do not support the assumption that economic inequality leads to more instances of electoral misconduct.

4.1.1 Robustness checks

The comparison of my theory with alternative explanations shows that three of the control variables Gini coefficient, GDP per capita and ethnic fractionalization, do not show statistically significant estimates consistently through the models and produce quantitatively small effect on the dependent variable. The variable for ethnic fractionalization is omitted in the regression models introducing country dummies because that measure is constant over time, and therefore, is collinear with the fixed effects. Gini coefficient is statistically significant in the last model at .5 level but has a negative sign, opposite to my theory. That result in practice shows that a wide income gap between rich and poor is not enough to trigger greater likelihood to rig elections.

The variable for type of electoral system, however, receives significant results in models 4 and 5. The regression estimates show that moving from a majoritarian electoral system (re-coded as 1) to a mixed one (re-coded as 2), and finally to a proportional (re-coded as 3) increases the overall score for “clean” elections with .188 point estimates for each step

(Model 5). It is worth noting that obtaining statistically significant results for some of the control variables, previously tested in the literature, is in practice strengthening the results from my analysis for two main reasons. First, because despite the strong effect of electoral institutions on electoral competitors' behaviour, two of my main explanatory variables maintain significant results. Second, the claim for validity of the data and the new measure created for electoral misconduct is thus supported.

The regime type in a state is the last alternative explanation controlled for. Excluding full autocracies and democracies from the analysis keeps essentially the results obtained so far (Table 4 below). As in the other control models, the beta coefficients are slightly changed. The variable *Political power distributed by socio-economic position* is showing statistically significant results in the first three models. Nevertheless, since in the other main and control models this variable has very high standard errors compared to the beta coefficients, I do not have enough evidence to claim that it has a consistent and independent effect on electoral misconduct.

Another important aspect of the regression results is that the number of observations included in the regression drops drastically with the inclusion of the control variables[21]. This makes it difficult to compare the first models with the rest since N is decreased more than four times (from 2148 to 510). The drop is explained by the fact that the control variables are missing information for the early years of the 20th century, and some countries (for example Malaysia) included in the V-Dem data collection are not coded for some of the control variables.

To make sure the results I have obtained are not biased by the countries included in the last models, I re-estimated the first three models on the same 510 observations as in the last ones (Table A.8 in the appendix). Since the regression estimates are not differing substantially from the initial models in terms of beta coefficients and significance levels, we can in all likelihood rule out that there is bias as result of the decrease of the number of observations.

[21] Table A.10 in the Appendix lists the countries included in regression Models 4 and 5.

Table 4. Regression estimates of the effect of Societal Inequality on Electoral Misconduct excluding full democracies and autocracies

	Model 6	Model 7	Model 8	Model 9	Model 10
	V-Dem IV	LDV	LDV PCSE FE	LDV, CV, PCSE	LDV, CV, PCSE, FE
Power distributed by socioeconomic position	0.231***	0.096**	0.152**	0.068	-0.017
	(0.05)	(0.04)	(0.05)	(0.05)	(0.09)
Power distributed by social group	0.496***	0.189***	0.301***	0.052	0.318**
	(0.06)	(0.04)	(0.08)	(0.09)	(0.12)
Social group equality in respect for civil liberties	0.038	-0.005	0.195***	0.008	0.418***
	(0.05)	(0.03)	(0.05)	(0.06)	(0.11)
Social class equality in respect for civil liberty	0.117*	0.05	0.165*	0.068	-0.037
	(0.05)	(0.03)	(0.08)	(0.05)	(0.16)
Electoral misconduct lagged		0.717***	0.381***	0.631***	0.248***
		(0.03)	(0.05)	(0.06)	(0.06)
Ethnic fractionalization				0.139	.
				(0.09)	.
Gini coefficient				0	-0.009
				(0)	(0.01)
GDP per Capita, PPP (Constant International USD)				0.000**	0
				(0)	(0)
Electoral system				0.094**	0.217**
				(0.03)	(0.07)
Constant	-0.273***	-0.03	-0.284	-0.225	-0.191
	(0.05)	(0.03)	(0.22)	(0.21)	(0.32)
Observations	662	642	642	293	293
ll	-699.03	-382.84			
aic	1408.07	777.67	.	.	.
r2	0.287	0.718	0.815	0.695	0.822

Legend: * p<0.05; ** p<0.01; *** p<0.001.Robust standard errors within parantheses.

To further compare the regression models with and without the control variables, I perform likelihood ratio and Wald tests[22]. Both tests estimate that adding the four control measures as predictor variables results in a statistically significant improvement in the fit of the model and do not create bias (Chen 2003, Chapter 3). The output of the tests is in the Appendix, Table A.9.

Table A.7 in the Appendix presents the regression estimates when the predictor variables are lagged with one year to test for the theory that the decision to employ fraud is influenced largely by the information for inequality from the year before the elections were

[22] The two tests are used to evaluate the difference in the fit of models to data. This is done by nesting one of the models in the second by forcing restrictions on the parameters to make them match in terms of observations number. The tests then answer the question whether the restriction to those parameters significantly reduces the fit of the model to the data (Chen, 2003, Chapter 3).

held. Methodologically, the lagging of the explanatory variables is a suitable check for reversed causality. The regression estimates show that the lag of the independent variables does not change the regression estimates considerably. The beta coefficients have slightly lower values compared to the main models, but the two variables for social inequality remain statistically significant at .001 level. The same results are obtained when the lag of the dependent variable on the right hand side of the equation is excluded (Table A.7).

How good the applied models are depends also on how well they predict the outcome variable. The scatter plot of the observed versus the predicted values of electoral misconduct shows that the last model (Model 6) is good at predicting the occurrence of electoral misconduct, judging by the visible 45 degree distribution pattern of the observations (Chen et al, 2003, Chapter 2).[23] Furthermore, to estimate the goodness of the models, I performed checks for the normality of the residuals. The plots in Graphs A.3-5 in the Appendix show the goodness of fit of the data when all my independent variables are included. Ideally, the residuals should be normally distributed in a bell-shape form in Graph A.3, and be plotted straight on the lines in Graphs A.4 and A.5. Despite the visible deviations, it can be concluded that the residuals are close to normal, and the assumed PCSE method has solved any occurring issues (Chen 2003, Chapter 2). The same is true for the assumption for homogeneity of the residuals variance. The plot of residuals against the predicted values (Graph A.6) shows a pattern which would not be the case in a perfectly fit model. However, it does not give strong enough evidence to disregard the model, especially because of the specifications added to the original model.

Finally, I checked for observations with extreme values in my data that might have affected the regression line unproportionally. Table A.11 presents the nine cases estimated to be outliers with excessive influence. To identify unusual observations I performed standard checks[24] for outliers (observations with large residuals), high leverage (data points with extreme observations on the independent variables), and finally, observations with

[23] Plotting the predicted values from the other models showed similar results.

[24] The checks flagged observations with absolute value of the studentized residuals higher than 2; data points with leverage higher than (2k + 2)/n, where k is the number of independent variables, and n is the number of observations in total. I also included two tests combining information for the residuals and leverage: Cook's distance with a cut-off point for observations higher than 4/n, and DFITS cut-off point is for observation with absolute value higher than 2*sqrt(k/n) (Chen et al 2003, Chapter 2).

extraordinary influence on the results, and test if the removal of them changes the regression line substantially. However, as Table A.12 in the Appendix shows re-estimating the main regression models excluding the nine extreme observations did not change the results significantly so we can draw the conclusion that the regression coefficients are not driven by unusual observations.

4.2 Discussing linearity vs. inverted "U" effect

Up to this point in the empirical analysis we have only investigated linear relationships between the explanatory and outcome variables. In order to test Hypothesis 2, predicting a relationship that looks like an inverted "U" I used a quadratic specification for the main explanatory variables measuring inequality. Table 4 presents the result of the regression estimates.

The coefficients on the squared specifications are not robust in the model introducing all control variables, country dummies and PCSE method, suggesting there is no evidence for inverted "U" relationship.

Table 5. Regression estimates testing curve-linear relationship between societal inequality and electoral misconduct

	Model 11 V-Dem IV, FE, PCSE	Model 12 All IV, FE, PCSE
Electoral misconduct lagged	0.698***	0.275***
	(0.03)	(0.06)
Power distributed by socioeconomic position	-0.041	-0.014
	(0.02)	(0.09)
Power distributed by social group	0.191***	0.371*
	(0.03)	(0.15)
Social group equality in respect for civil liberties	0.125***	0.401***
	(0.02)	(0.11)
Social class equality in respect for civil liberty	0.043	0.003
	(0.03)	(0.16)
Power distributed by socioeconomic position, squared	0.027	0.013
	(0.02)	(0.05)
Power distributed by social group, squared	0.060*	0.005
	(0.05)	(0.08)
Social group equality in respect for civil liberties, squared	0.042*	-0.084

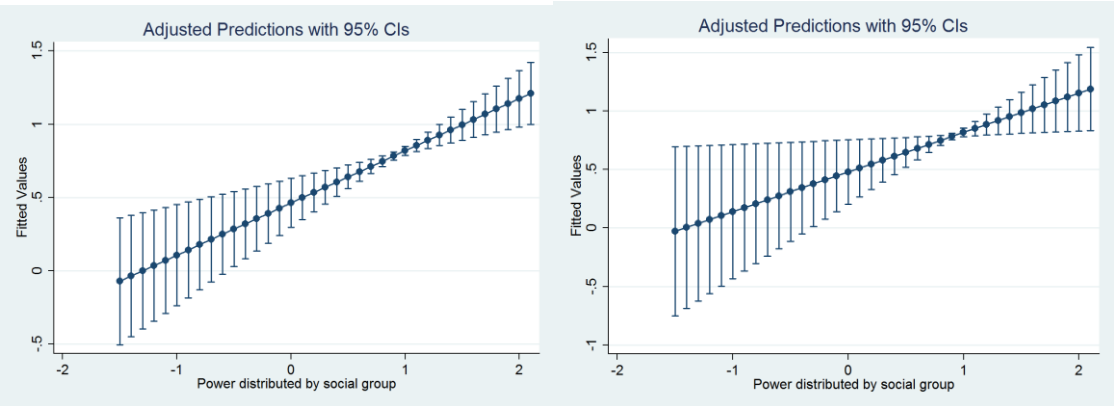
	(0.05)	(0.06)
Social class equality in respect for civil liberty , squared	-0.027	-0.003
	(0.02)	(0.06)
Ethnic fractionalization		-1.858***
		(0.46)
Gini coefficient		(0.011**
		(0)
GDP per Capita, PPP (Constant International USD)		0.000***
		(0)
Electoral system		0.191***
		(0.05)
Constant	-0.380***	-0.368
	(0.07)	(0.23)
R squared	0.872	0.883
Observations	2043	510

Legend: * p<0.05; ** p<0.01; *** p<0.001.Robust standard errors within parentheses.

Graph 3 helps to visualize the effect of inequality on electoral misconduct, presenting the predicted values for the dependent variable at different levels of one of the robust variables measuring inequality – namely, *political power distributed by social groups*, while the other variables are held at their mean[25]. The graphs on the left present the linear terms, and on the right – the quadratic terms, both are contrasted with 95% confidence intervals.

Graph 2. Fitted values of electoral misconduct plotted at different levels of *Power distributed by social group.*

Linear specification Quadratic specification



[25] A similar comparison is used in Ansell and Samuels to investigate linearity of the relationship between economic inequality and the probability for democratic transition (2010).

My theory predicted that countries with middle levels of inequality have the highest levels of electoral misconduct, while full democracies and autocracies should have comparatively lower levels. However, the graphs show that the predicted probability for "cleanest" elections increases when the political power distribution improves in terms of becoming more inclusive for social groups. Thus, at the highest level of the explanatory variable – point estimate 2, marking that social group distinctions are not relevant to the political life, the overall quality of elections is estimated to be very high. Notice also that the confidence intervals are narrower in the linear specification. An explanation for the latter is that the inclusion of the squared terms has increased the value of the standard errors.[26]

In sum, the empirical tests conducted support the idea that societal inequality has a independent powerful effect on electoral behavior and the decision to engage in fraudalent tactics as levels of inequality changes within countries over time. These findings corraborate Hypothesis 1, introduced in this paper, suggesting a positive relationship between inequality and electoral misconduct, and reject Hypothesis 2, arguing for a curve-linear relationship. Regardless of other country-specific conditions, more often than not, different types of electoral manipulations are employed in states in which social group distinctions define the political power distribution and civil liberties access, in conjunction with the prediction of Hypothesis 4.

4.3 Limitations

To test my theoretical assumptions I developed a quantitative study using statistical methods to analyze data with a large number of observations. This design allows estimating the average effect of a set of factors on a phenomenon that we want to explain. The key advantage of using statistical analysis is that its robustness will not be dramatically influenced by omitting minor variables, modest changes in the population or if not every

[26]Graph A.7 in the Appendix presents the predicted values for electoral misconduct at different levels of the other robust explanatory variable – *civil liberties equality by social group*. Again we find no evidence of an inverted U relationship, but for a linear one.

included case is explained (Mahoney, Goertz, 2006:236). In practice this means that the research design of large-N quantitative studies allows generalizing the reached conclusions for big geographical and time span.

In statistical analysis all observations are equal, since the overall pattern of fit is important. However, the main disadvantage of this design is that substantively important cases or cases that do not follow the general pattern are not given special attention unlike in qualitative research (ibid: 231). For this reason findings from qualitative studies tend to be more revealing when moving from a big subset to a particular case.

Mixed research designs combine the mentioned advantages of quantitative and qualitative methods. My research could be supplemented with qualitative case studies in a particular country that fits my theoretical assumption, and one that does not. Analysing and comparing the historical background and changes of the levels of the main variables in the two countries could help illustrate the logic behind the theory developed in this thesis. However, due to time constraints, I was not able to include such analysis in my thesis.

5 Conclusion

Using the newly collected V-Dem data on democracy for 139 countries in the world starting from 1900 until 2012, this thesis evaluates the effect societal inequality has on the occurrence of electoral misconduct. The advantages of that new data allowed designing new single measure that encompasses various legal and illegal tactics attempted to distort the outcome of elections. By using multivariate framework for the explanatory variable, the study explores also what type of inequality aggravates the situation the most.

Both societal inequality and cheating in elections remain topical development and governance problems in many countries. It is intuitive that electoral misconduct is corroding democracy as it is stealing the purpose of elections by replacing the will of the people. The less obvious relationship discussed in this thesis is drawing attention on how

societal inequality can hinder or undermine the democratic development in a country by affecting one of the key tools of democracy – elections.

The empirical findings from this paper suggest that more often than not, inequality based on *social group* differences is associated with electoral misconduct unlike inequality based on economic grounds. To understand the findings better, it is worth repeating that social group is defined as a group of individuals that identify themselves as having common ethnicity, caste, language, race, religion etc., and based on this characteristic distinguish themselves from the other members of the society.

When compared to previous research done on the topic, my results corrobarote and extend further the findings from the single case study on Germany done by Ziblatt (2009) suggesting that structural factors can hinder the transformative effect elections have on institutions. Put in the perspective of the larger topic about democracy and redistribution, my results correspond more to Ansell and Samuels' (2010) skepticism, rather than the findings of Boix (2003), Acemoglu and Robinson (2006) who argue that economic equality as such brings about better environment for democratic transition. The essential inference from my analysis is that indeed equality matters for democracy, but its effect is corrosive only when the underlying differences are translated into the political life or affect the distribution of civil liberties.

Analyzing and understanding the effect economic, social and political inequality can have on electoral manipulations has a number of practical implications too. First, predicting ballot-rigging, election violence and other forms of electoral irregularities can inform policies of international organizations, NGOs and other relevant actors on how to build strategies to prevent them. Second, the theoretical views presented in this paper underline why the different forms of inequality can be dangerous to societies' development and should be addressed. In essence, failing to mitigate the inequality gaps can create a polarized society prepared to break the norms in order to defend their interests better. This perspective on the reasons for fraud, calls for focusing the attention on the long-term trends in a society's development since it argues that immediate tactics employed in a specific election are

largely influenced by preexisting societal power distribution differences. Therefore, attempts to democratize institutionally can be undermined by the conflicting interests of different groups. Examining further how these interests exhibit themselves in the process of democratization promises to be an important and interesting topic for future research.

References:

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

Alvarez, R. Michael; Thad E. Hall and Susan D. Hyde (2008) *Election Fraud: Detecting and Deterring Electoral Manipulation*. Washington, D.C.: Brookings Institution Press

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003) *Fractionalization*. Journal of Economic Growth. 8: 155-194., data accessed through Quality of Government database on 2014-05-16.

Ansell, Ben and David Samuels (2010) *Inequality and Democratization: A Contractarian Approach*. Comparative Political Studies 43 (12): 1543-1574

Baland, Jean-Marie, and James Robinson (2006*) Land and Power:Theory and Evidence from Chile*. National Bureau of Economic Research Working Paper 12517

Beck, Nathaniel (2001) *Time-series-cross-section-data: What have we learned in the past few years?* Annual Review of Political Science. 4: 271-293.

Beck, Nathaniel and Jonathan Katz (1995) *What to do (and not to do) with time-series cross-section data.* American Political Science Review. 89(3):634-647

Bermeo, Nancy (2010) *Interests, Inequality, and Illusion in the Choice for Fair Elections*. Comparative Political Studies 43(8/9) 1119–1147

Birch, Sarah (2010) *Perceptions of Electoral Fairness and Voter Turnout*. Comparative Political Studies 43 (12): 1601-1622

Boix, Carles (2003) *Democracy and redistribution.* Cambridge University Press

Bormann, N-C. & Matt Golder (2013) *Democratic electoral Systems Around the World, 1946-2011.* Electoral Studies. Data accessed through Quality of Government database on 2014-05-16

Breunig, Christian and Goerres, Achim (2011) *Searching for Electoral Irregularities in an Established Democracy: Applying Benford's Law Tests to Bundestag Elections in Unified Germany*. Electoral Studies 30 (3): 534-545

Chen, X., Ender, P., Mitchell, M. and Wells, C. (2003) *Regression with Stata*, from http://www.ats.ucla.edu/stat/stata/webbooks/reg/default.htm accessed 2014-05-10

Collier, Paul and Pedro C. Vicente (2012) *Violence, Bribery, and Fraud: the Political Economy of Elections in Sub-Saharan Africa*. Public Choice 153:117–147

Converse, P. E. (1972) *Change in the American Electorate. In The Human Meaning of Social Change*. ed. Campbell, A. & Converse. P.E. New York: Russell Sage

Coppedge, Michael; Gerring, John; Lindberg, Staffan and Jan Teorrell (2013). *Varieties of Democracy Project Description.*

Cox, Gary W. and Keith T. Poole (2002) *On Measuring Partisanship in Roll Call Voting: The U.S. House of Representatives, 1877-1999.* American Journal of Political Science 46 (3): 477-489

Dahl, Robert A. (1971) *Polyarchy: Participation and Opposition*. Yale University Press

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

Elkit, J., & Svensson, P. (1997) *What makes elections free and fair?* Journal of Democracy. 8(3): 32-46.

Gilens, Martin and Benjamin I. Page (2014) *Testing Theories of American Politics: Elites, Interest Groups, and Average Citizens*. Perspectives on Politics, Forthcoming Fall 2014

Greene, K. (2007) *Why Dominant Parties Lose*. New York: Cambridge University Press

Greene, William H. (2008) *Econometrics Analysis.* 6th edition. Upper Saddle River, N.J. : Prentice Hall

Bormann, N-C. & Matt Golder (2013) *Democratic electoral Systems Around the World, 1946-2011*. Electoral Studies.

Hadenius, A. and Jan Teorell (2005) *Assessing Alternative Indices of Democracy*. C&M Working Papers 6, IPSA, August 2005, data accessed through Quality of Government database on 2014-05-16.

Hafner-Burton, Emelie; Susan D.Hyde; Ryan S Jablonski (2013) *When Do Governments Resort to Election Violence?*. British Journal of Political Science 44:149-179

Hicken, Allen (2007) *How Do Rules and Institutions Encourage Vote-Buying?* Elections for Sale, ed. Fred Schaffer. Boulder, CO: Lynn Rienner 47–60.

Howard, M. M. and P. G. Roessler (2006) *Liberalizing Electoral Outcomes in Competitive Authoritarian Regimes*. American Journal of Political Science 50: 365–381

Jackmon, Simon (2001) *Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking*. Political Analysis 9 (3)

Johnson, Valen E. and James H. Albert (1999) *Ordinal Data Modeling*. Springer-Verlag New York Inc.

Jørgen Møller & Svend-Erik Skaaning (2013) *Autocracies, democracies, and the violation of civil liberties*. Democratization 20(1): 82-106

Keele, Luke and Nathan J. Kelly (2006) *Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables*. Political Analysis. 14(2): 186-205

Kitschelt, Herbert; Steven I. Wilkinson (2007) *Patrons, Clients and Policies Patterns of Democratic Accountability and Political Competition*. Cambridge, UK: Cambridge University Press.

Lehoucq, Fabrice Edouard and Ivan Molina (1999) *Political Competition and Electoral Fraud: A Latin American Case Study.* Journal of Interdisciplinary History 30(2): 199-234

Lehoucq, Fabrice Edouard and Ivan Molina (2002) *Stuffing the ballot box: Fraud, election reform, and democratization in Costa Rica*. Cambridge: Cambridge University Press.

Lindberg, Staffan I. (2006) *Democracy and Elections in Africa*. Baltimore: Johns Hopkins University Press.

Lindberg, Staffan I. (2009) *Democratization by Elections: A Mixed Record*. Journal of Democracy 20(3): 86-92.

Lust-Okar, Ellen (2005) Structuring Conflict in the Arab World: Incumbents, Opponents and Institutions. Cambridge: Cambridge University Press

Magaloni, Beatriz (2006) *Voting for Autocracy: Hegemonic Party Survival and its Demise in Mexico*. New York: Cambridge University Press.

Magaloni, Beatriz (2010) *The Game of Electoral Fraud and the Ousting of Authoritarian Rule*. *American Journal of Political Science* 54 (3*): 751:765*

Mahoney, James and Gary Goertz (2006) *A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research*. Political Analysis 14: 227 – 249

McFoul, Michael (2002) *The Fourth Wave of Democracy and Dictatorship: Non-cooperative Transitions in the Postcommunist World*. World Politics 54 (2): 212-244

Møller, Jørgen and Svend-Erik Skaaning (2013) *Autocracies, Democracies, and the Violation of Civil Liberties.* Democratization*.* 20(1): 82-106

Piketty, Thomas (2014) *Capital in the Twenty-First Century*. Belknap Press

Rosset, Jan; Nathalie Giger and Julian Bernauer (2013) *More Money, Fewer Problems? Cross-Level Effects of Economic Deprivation on Political Representation*. West European Politics 36(4): 817-835

Rummel, R.J. (1967) *Understanding Factor Analysis*. The Journal of Conflict Resolution. 11(4): 444-480

Schedler, Andreas (2002)*The Menu of Manipulation*. Journal of Democracy 13 (2): 36-50

Solt, F. (2008) *The Standardized World Income Inequality Database*. Data accessed through Quality of Government database on 2014-05-16

Stokes, Susan; Przeworski, Adam and Manin Bernard (1999) *Democracy, Accountability, and Representation.* Cambridge University Press

Teorell, Jan, Nicholas Charron, Stefan Dahlberg, Sören Holmberg, Bo Rothstein, Petrus Sundin & Richard Svensson. 2013. The Quality of Government Dataset, version 15May13. University of Gothenburg: The Quality of Government Institute

Tilly, Charles (2007) *Democracy*. Cambridge, UK: Cambridge University Press.

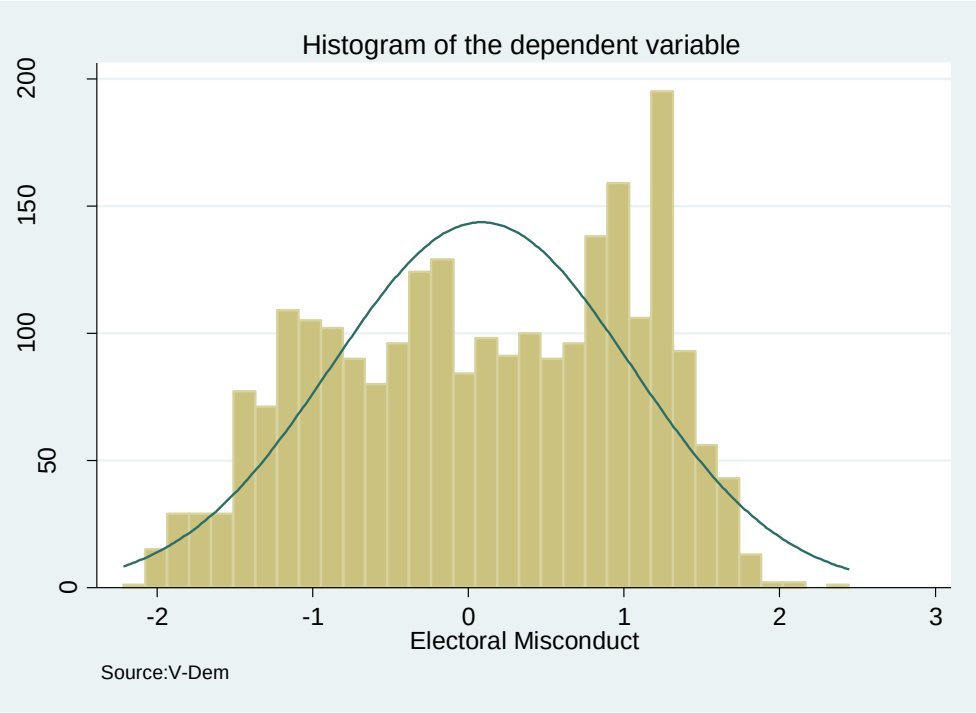
Verba, S., Schlozman, K.L. & Brady, H. (1995) *Voice and equality: Civic voluntarism in American politics.* Cambridge: Harvard University Press

Ziblatt, Daniel (2009) *Shaping Democratic Practice and the Causes of Electoral Fraud: The Case of Nineteenth-Century Germany*. American Political Science Review 103 (1): 1-21

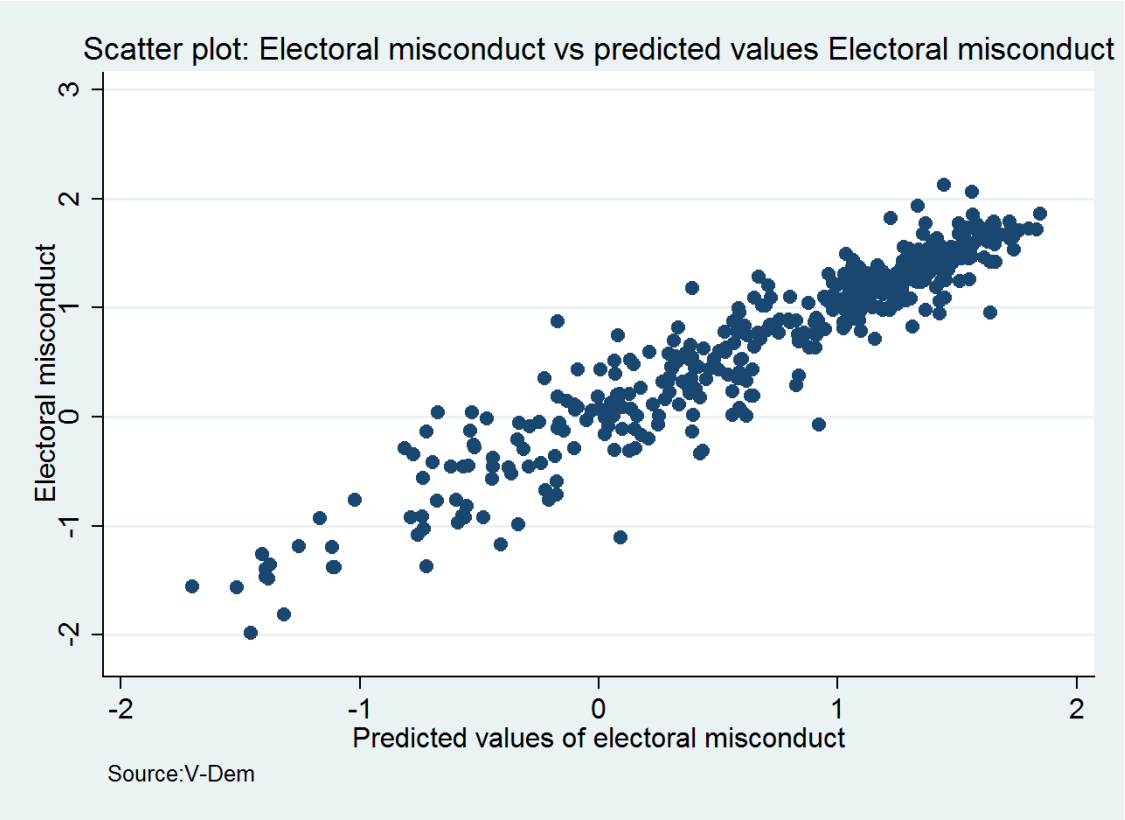
Appendix:

Graphs:

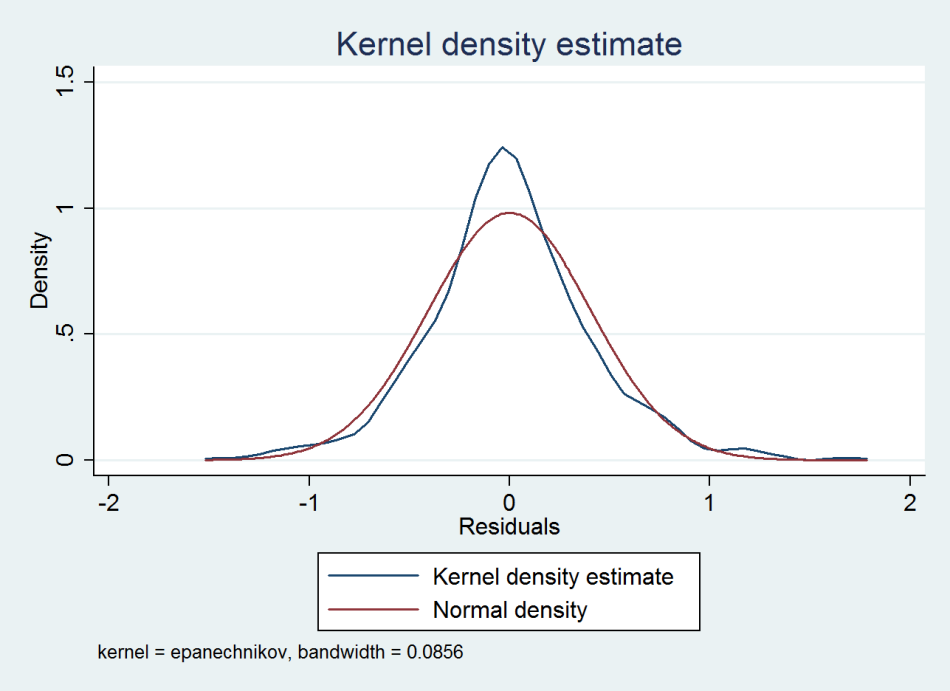
Graph A.1. Distribution of the observations of electoral misconduct



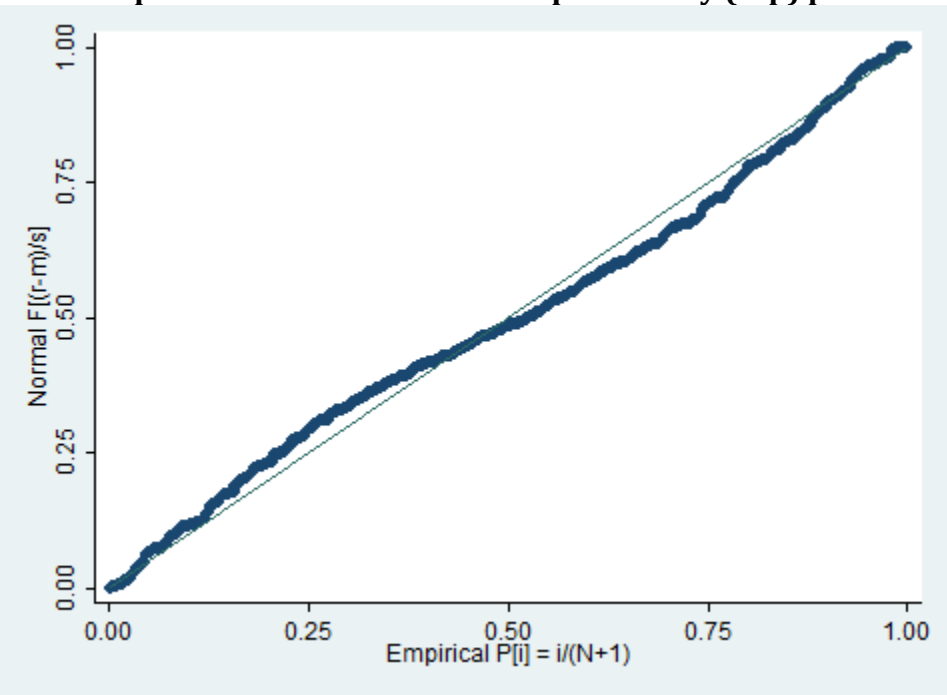
Graph A.2. Scatter plot of the dependent variable versus the predicted values for the dependent variable in Model 8



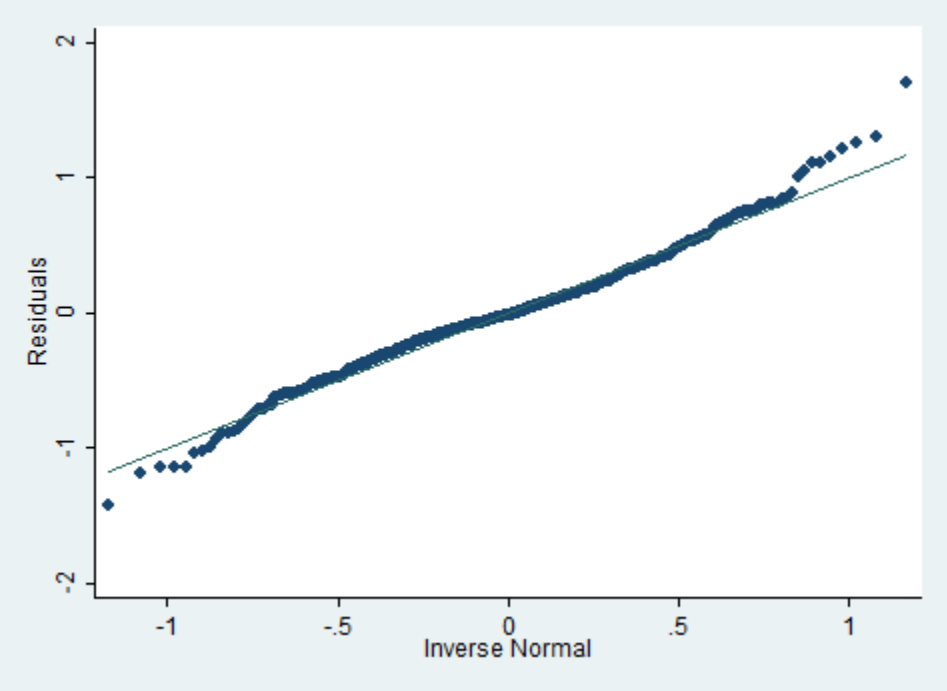
Graph A.3 Density plot of the residuals



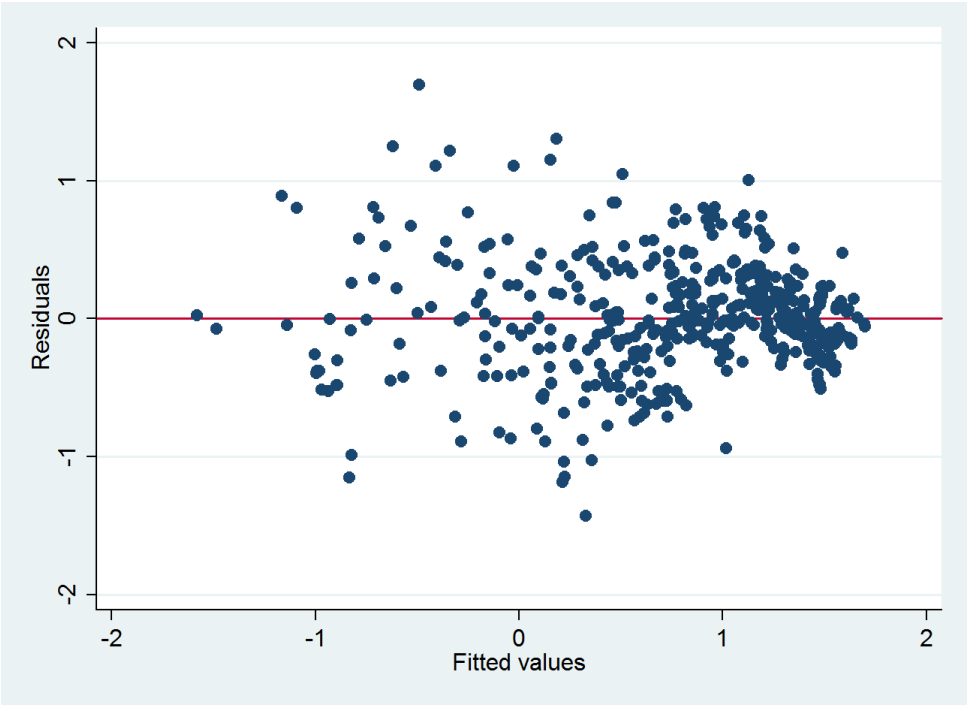
Graph A.4 Standardized normal probability (P-p) plot



Graph A.5 Standardized quantile plot

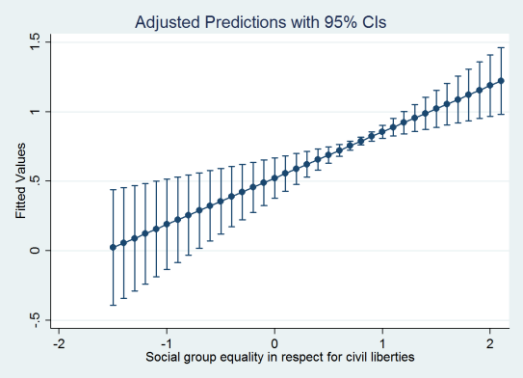


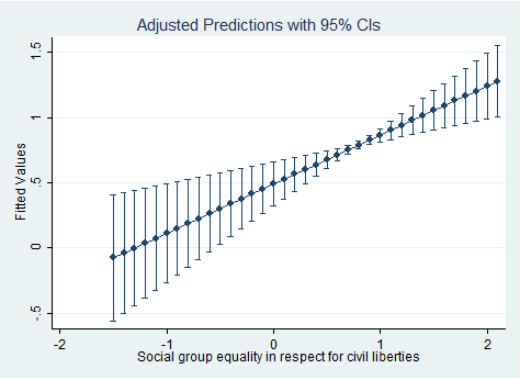
Graph A.6 Plot of the residuals against the predicted values



Graph A.7 Fitted values of electoral misconduct on different levels of *Social group equality in respect for civil liberties.*

Linear specification Quadratic specification





Tables:

Table A.1. Factor analysis output

Factor analysis/correlation — Number of observations = 2453

Method: principal factors — Retained factors = 1

Rotation: (unrotated) — Number of parameters = 10

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	4.65125	3.55422	0.8337	0.8337
Factor2	1.09703	0.73841	0.1966	1.0303
Factor3	0.35862	0.30502	0.0643	1.0945
Factor4	0.05359	0.09108	0.0096	1.1042
Factor5	-0.03749	0.01097	-0.0067	1.0974
Factor6	-0.04846	0.02183	-0.0087	1.0888
Factor7	-0.07029	0.05166	-0.0126	1.0762
Factor8	-0.12195	0.00969	-0.0219	1.0543
Factor9	-0.13164	0.03965	-0.0236	1.0307
Factor10	-0.17129	.	-0.0307	1.0000

LR test: independent vs. saturated: chi2(45) = 1.5e+04 Prob>chi2 = 0.0000

Factor loadings (pattern matrix) and unique variances

Variable	Factor1	Uniqueness
Elections free and fair	0.8784	0.2284
Free campaign media	0.5806	0.6629
Government intimidation	0.8182	0.3306
Other voting irregularities	0.7579	0.4255
Elections multiparty	0.7325	0.4635
Other electoral violence	0.4957	0.7543
Election voter registry	0.7113	0.4940
Eligibility restricted	0.1812	0.9672
Election vote-buying	0.5688	0.6765
EMB autonomy restricted	0.8088	0.3459

Table A. 2. List of countries included in the analysis

1	Afghanistan	41	El Salvador	81	Mauritania	121	Taiwan
2	Albania	42	Estonia	82	Mexico	122	Tajikistan
3	Algeria	43	Ethiopia	83	Moldova	123	Tanzania
4	Angola	44	Finland	84	Mongolia	124	Thailand
5	Argentina	45	France	85	Montenegro	125	Togo
6	Armenia	46	Gabon	86	Morocco	126	Trinidad and Tobago
7	Australia	47	Gambia	87	Mozambique	127	Tunisia
8	Austria	48	Georgia	88	Namibia	128	Turkey
9	Azerbaijan	49	Germany	89	Nepal	129	Turkmenistan
10	Belarus	50	Ghana	90	Netherlands	130	Uganda
11	Belgium	51	Greece	91	New Zealand	131	Ukraine
12	Benin	52	Guinea	92	Nicaragua	132	United Kingdom
13	Bhutan	53	Guinea-Bissau	93	Niger	133	Uruguay
14	Bolivia	54	Haiti	94	Nigeria	134	Uzbekistan
15	Bosnia and Herzegovina	55	Honduras	95	Norway	135	Venezuela
16	Botswana	56	Hungary	96	Pakistan	136	Vietnam, DR
17	Brazil	57	India	97	Palestine/Gaza	137	Vietnam, Republic of
18	Bulgaria	58	Indonesia	98	Palestine/West Bank	138	Zambia
19	Burkina Faso	59	Iran	99	Panama	139	Zimbabwe
20	Burma/Myanmar	60	Ireland	100	Peru		
21	Burundi	61	Italy	101	Philippines		
22	Cambodia	62	Ivory Coast	102	Poland		
23	Cameroon	63	Jamaica	103	Portugal		
24	Canada	64	Japan	104	Romania		
25	Cape Verde	65	Jordan	105	Russia		
26	Central African Republic	66	Kazakhstan	106	Rwanda		
27	Chad	67	Kenya	107	Sao Tome and Principe		
28	Chile	68	Korea, South	108	Senegal		
29	Colombia	69	Kyrgyzstan	109	Serbia		
30	Comoros	70	Laos	110	Sierra Leone		
31	Congo, DR	71	Latvia	111	Slovakia		
32	Congo, Rep. of the	72	Lebanon	112	Somalia		
33	Costa Rica	73	Lesotho	113	Somaliland		
34	Croatia	74	Liberia	114	South Africa		
35	Czech Republic	75	Libya	115	Spain		
36	Denmark	76	Lithuania	116	Sri Lanka		
37	Djibouti	77	Macedonia	117	Sudan		
38	Dominican Republic	78	Madagascar	118	Swaziland		
39	Ecuador	79	Malaysia	119	Switzerland		
40	Egypt	80	Mali	120	Syria		

Table A.3. Descriptive statistics: Electoral misconduct

Observations	2453
Sum of Wgt.	2453
Mean	0.08
Std. Dev.	0.96
Variance	0.92
Skewness	-0.18
Kurtosis	1.91

Table A.4. Pearson correlation coefficients (*p<0,5)						
	Electoral misconduct	Social class equality in respect for civil liberty	Social group equality in respect for civil liberties	Power distributed by socioeconomic position	Power distributed by social group	Ethnic fractionalization
Electoral misconduct	1					
Social class equality in respect for civil liberty	0.4838*	1				
	0					
Social group equality in respect for civil liberties	0.5166*	0.7556*	1			
	0	0				
Power distributed by socioeconomic position	0.4468*	0.7054*	0.5967*	1		
	0	0	0			
Power distributed by social group	0.5083*	0.6150*	0.5783*	0.6169*	1	
	0	0	0	0		
Ethnic fractionalization	-0.2710*	-0.3312*	-0.2973*	-0.2920*	-0.3910*	1
	0	0	0	0	0	
Gini Coefficient	-0.0849	-0.1034*	-0.0202	-0.0635*	-0.0890*	0.3715*
	0.0004	0	0.4031	0.0092	0.0003	0
GDP per Capita, PPP	0.5633*	0.4392*	0.4497*	0.4182*	0.4090*	-0.5195*
	0	0	0	0	0	0
Freedom House/Imputed Polity	0.8837*	0.4451*	0.4757*	0.4339*	0.6072*	-0.3716*
	0	0	0	0	0	0
Electoral system	0.3984*	0.2377*	0.2040*	0.1208*	0.3261*	-0.1730*
	0	0	0	0	0	0
	Gini coefficient	GDP per Capita, PPP	Freedom House/Imputed Polity	Electoral system		
Gini coefficient	1					
GDP per Capita, PPP	-0.5216*	1				
	0					
Freedom House/Imputed Polity	-0.2625*	0.6002*	1			
	0	0				
Electoral system	-0.1730*	0.2973*	0.4147*	1		
	0	0	0			

Table A.5. Multicollinearity check for the Independent variables

Variable name	VIF	1/VIF
Civil liberties equality for social groups	3.51	0.28
Civil liberties equality for social class	3.44	0.29
Electoral misconduct lagged	3.19	0.31
GDP/PPP	2.65	0.38
Political power by social group	2.19	0.46
Gini coefficient	1.77	0.57
Political power by socioeconomic position	1.72	0.58
Ethnic fractionalization	1.51	0.66
Electoral system	1.37	0.73
Mean VIF	2.42	

Table A.6 Regression estimates step by step addition of the control variables				
	Adding Ethnic Fractionalization	Adding Gini coefficient	Adding Electoral system	Adding GDP per capita
Electoral misconduct lagged	0.825***	0.837***	0.694***	0.680***
	(0.02)	(0.02)	(0.04)	(0.05)
Power distributed by socioeconomic position	0.047	0.050*	0.044	0.029
	(0.02)	(0.02)	(0.03)	(0.04)
Power distributed by social group	0.158***	0.149***	0.111	0.067
	(0.03)	(0.03)	(0.06)	(0.06)
Social group equality in respect for civil liberties	0.063**	0.035	0.043	0.069
	(0.02)	(0.02)	(0.04)	(0.05)
Social class equality in respect for civil liberty	0.008	0.016	0.085*	0.044
	(0.02)	(0.02)	(0.04)	(0.04)
Ethnic fractionalization	0.004	0.018	-0.047	0.016
	(0.05)	(0.05)	(0.07)	(0.08)
Gini coefficient		0.001	0	0.001
		(0)	(0)	(0)
Electoral system			0.065***	0.068**
			(0.02)	(0.02)
GDP per Capita, PPP (Constant International USD)				0.000
				0
Constant	-0.103***	-0.107*	-0.067	-0.126
	(0.03)	(0.05)	(0.07)	(0.16)
N	1426	1185	756	510
ll	-735.998	-540.416	-302.715	-198.058
aic	1485.997	1096.833	623.43	416.116

Legend: * p<0.05; ** p<0.01; *** p<0.001.Robust standard errors within parentheses.

Table A.7 Regression re-estimations, lagged independent variables; no lag on the dependent variable.

	Lagged V-Dem variables	LDV, lagged independent vars	Lagged independent and control variables, PCSE, FE	No lag on the DV, Control variables, PCSE, FE
Electoral misconduct, lagged		0.905***	0.310***	
		(0.01)	(0.08)	
Power distributed by socioeconomic position	0.136***	0.011	0.081	0.072
	(0.04)	(0.02)	(0.06)	(0.06)
Power distributed by social group	0.329***	0.008	0.357**	0.509***
	(0.04)	(0.02)	(0.12)	(0.09)
Social group equality in respect for civil liberties	0.265***	0.019	0.497***	0.477***
	(0.03)	(0.02)	(0.07)	(0.09)
Social class equality in respect for civil liberty	0.064	0.034	-0.009	-0.026
	(0.04)	(0.02)	(0.1)	(0.13)
Ethnic fractionalization			.	.
			.	.
Gini coefficient			-0.007	-0.010*
			(0.01)	(0)
GDP per Capita, PPP (Constant International USD)			0	0
				(0)
Electoral system			0.052	0.222***
			0.06	0.05
Constant	-0.219***	0.011	-0.190	-0.959***
	(0.02)	(0.01)	(0.28)	(0.21)
Observations	2067	2040	458	516
ll	-2511.918	-952.522		
aic	5033.837	1917.043	.	.

Legend: * p<0.05; ** p<0.01; *** p<0.001.Robust standard errors within parantheses.

Table A.8. Re-estimating models 1-3 on the sub-sample of countries			
	Model 1 V-Dem	Model 2 LDV	Model 3 LDV, PCSE, FE
Power distributed by socioeconomic position	0.111*	0.022	-0.009
	(0.056)	(0.038)	(0.063)
Power distributed by social group	0.403***	0.105	0.397***
	(0.086)	(0.061)	(0.095)
Social group equality in respect for civil liberties	0.193**	0.074	0.369***
	(0.073)	(0.049)	(0.096)
Social class equality in respect for civil liberty	0.127*	0.028	-0.005
	(0.059)	(0.047)	(0.114)
Electoral misconduct lagged		0.707***	0.290***
		(0.048)	(0.06)
Constant	-0.057	0.04	-0.2
	(0.066)	(0.044)	(0.121)
Observations	510	510	510
ll	-407.575	-201.713	
aic	827.149	417.425	.

Legend: * p<0.05; ** p<0.01; *** p<0.001.Robust standard errors within parantheses.

Table A.9. Comparing the regression results between the nested models

Likelihood ratio test results:

. lrtest n1 n2

Likelihood-ratio test LR chi2(3) = 28.68

(Assumption: n1 nested in n2) Prob> chi2 = 0.0000

Wald test output

. test al_ethnicginiwdi_gdpcgol_est_fill

(1) o.al_ethnic = 0

(2) gini = 0

(3) wdi_gdpc = 0

(4) gol_est_fill = 0

Constraint 1 dropped

F(3, 426) = 8.35

Prob> F = 0.0000

Table A.10 List of countries included in Models 4 and 5					
1	Albania	31	Ireland	61	Switzerland
2	Argentina	32	Italy	62	Thailand
3	Armenia	33	Kenya	63	Trinidad and Tobago
4	Australia	34	Korea, South	64	Turkey
5	Austria	35	Kyrgyzstan	65	Uganda
6	Belgium	36	Laos	66	Ukraine
7	Benin	37	Latvia	67	United Kingdom
8	Bolivia	38	Lebanon	68	Venezuela
9	Brazil	39	Lithuania		
10	Bulgaria	40	Macedonia		
11	Burundi	41	Madagascar		
12	Cape Verde	42	Mali		
13	Central African Republic	43	Mauritania		
14	Comoros	44	Moldova		
15	Congo, Republic of the	45	Mongolia		
16	Croatia	46	Nepal		
17	Czech Republic	47	Nicaragua		
18	Denmark	48	Niger		
19	El Salvador	49	Nigeria		
20	Estonia	50	Pakistan		
21	Finland	51	Peru		
22	France	52	Philippines		
23	Georgia	53	Poland		
24	Ghana	54	Portugal		
25	Greece	55	Senegal		
26	Guinea-Bissau	56	Sierra Leone		
27	Honduras	57	Slovakia		
28	Hungary	58	Spain		
29	India	59	Sri Lanka		
30	Indonesia	60	Sudan		

Table A.11 List of outliers and influential observations by country and election year					
Country name	Year	Studentized residuals	Leverage value	Cook's distance	Dfit
Threshold value		absolute(r)> 2	>(2k+2)/n	> 4/n	absolute(Dfit)> 2*sqrt(k/n)
Lebanon	1992	-2.59	0.05	0.03	-0.59
Sudan	1996	-2.59	0.07	0.04	-0.71
Benin	2011	-2.51	0.05	0.03	-0.58
Pakistan	1985	-2.08	0.06	0.02	-0.51
Lebanon	2000	2.89	0.05	0.04	0.66
Burundi	1993	3.49	0.05	0.05	0.80
Sierra Leone	1996	3.75	0.08	0.09	1.08
Hungary	1990	4.15	0.06	0.08	1.01
Mauritania	2006	5.46	0.06	0.15	1.36

Table A.12 Regression re-estimation Model 1-5 on the effect on Societal inequality on Electoral misconduct without outliers and influential observations

	V-Dem IV	LDV	LDV PCSE FE	LDV, CV, PCSE	LDV, CV, PCSE, FE
Power distributed by socioeconomic position	0.074*	-0.01	-0.051*	-0.033	-0.033
Power distributed by social group	0.365***	0.094***	0.186***	0.106*	0.235**
Social group equality in respect for civil liberties	0.229***	0.037*	0.129***	0.001	0.209**
Social class equality in respect for civil liberty	-0.008	0.004	0.032	0.064*	0.15
Electoral misconduct lagged		0.877***	0.737***	0.736***	0.370***
Ethnic fractionalization				0.048	(omitted)
Gini coefficient				0.001	-0.010*
GDP per Capita, PPP (Constant International USD)				0.038*	0.139**
Electoral system				0.000*	0
Constant	-0.247***	-0.031**	-0.315***	-0.476	-0.372
Observations	2141	2012	2012	493	493
ll	-2499.0	-841.6			
chi2			514.8	2483.3	478.9
aic	5010.0	1697.2	.	.	.

Legend: * p<0.05; ** p<0.01; *** p<0.001.

List of indicators used in the analysis:

Indicators included in the measure for electoral misconduct (V-Dem data: Coppedge et al, 2013):

1. Candidate restriction by ethnicity, race, religion, or language (B)

v2 tag: v2elrstrct

Project manager: Pamela Paxton

Question: Is the eligibility of candidates for national executive or legislative office formally restricted (by constitution or statute) by ethnicity, race, religion, or language?

Responses:

0: Yes, there are such statutory restrictions.

1: No, but law requires that candidates must renounce (or hide) their ethnic, racial, religious, or linguistic identity.

2: No, there are no such restrictions.

Scale: Ordinal

2. EMB autonomy (C)

v2 tag: v2elembaut

Project manager: Staffan I. Lindberg

Question: Does the Election Management Body (EMB) have autonomy from government to apply election laws and administrative rules impartially in national elections?

Clarification: The EMB refers to whatever body (or bodies) is charged with administering national elections.

Responses:

0: No. The EMB is controlled by the incumbent government, the military, or other *de facto* ruling body.

1: Somewhat. The EMB has some autonomy on some issues but on critical issues that influence the outcome of elections, the EMB is partial to the *de facto* ruling body.

2: Ambiguous. The EMB has some autonomy but is also partial, and it is unclear to what extent this influences the outcome of the election.

3: Almost. The EMB has autonomy and acts impartially almost all the time. It may be influenced by the *de facto* ruling body in some minor ways that do not influence the outcome of elections.

4: Yes. The EMB is autonomous and impartially applies elections laws and administrative rules.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

3. Elections multiparty (C)

v2 tag: v2elmulpar

Project manager: Staffan I. Lindberg

Question: Was this national election multiparty?

Responses:

0: No. No-party or single-party and there is no meaningful competition (includes situations where a few parties are legal but they are all *de facto* controlled by the dominant party).

1: Not really. No-party or single-party (defined as above) but multiple candidates from the same party and/or independents contest legislative seats or the presidency.

2: Constrained. At least one real opposition party is allowed to contest but competition is highly constrained – legally or informally.

3: Almost. Elections are multiparty in principle but either one main opposition party is prevented (*de jure* or *de facto*) from contesting, or conditions such as civil unrest (excluding natural disasters) prevent competition in a portion of the territory.

4: Yes. Elections are multiparty, even though a few marginal parties may not be permitted to contest (e.g. far-right/left extremist parties, anti-democratic religious or ethnic parties).

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

4. Election voter registry (C)

v2 tag: v2elrgstry

Project manager: Staffan I. Lindberg

Question: In this national election, was there a reasonably accurate voter registry in place and was it used?

Responses:

0: No. There was no registry, or the registry was not used.

1: No. There was a registry but it was fundamentally flawed (meaning 20% or more of eligible voters could have been disenfranchised or the outcome could have been affected significantly by double-voting and impersonation).

2: Uncertain. There was a registry but it is unclear whether potential flaws in the registry had much impact on electoral outcomes.

3: Yes, somewhat. The registry was imperfect but less than 10% of eligible voters may have been disenfranchised, and double-voting and impersonation could not have affected the results significantly.

4: Yes. The voter registry was reasonably accurate (less than 1% of voters were affected by any flaws) and it was applied in a reasonable fashion.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

5. Election vote buying (C)

v2 tag: v2elvotbuy

Project manager: Staffan I. Lindberg

Question: In this national election, was there evidence of vote and/or turnout buying?

Clarification: Vote and turnout buying refers to the distribution of money or gifts to individuals, families, or small groups in order to influence their decision to vote/not vote or whom to vote for. It does not include legislation targeted at specific constituencies, i.e., "porkbarrel" legislation.

Responses:

0: Yes. There was systematic, widespread, and almost nationwide vote/turnout buying by almost all parties and candidates.

1: Yes, some. There were non-systematic but rather common vote-buying efforts, even if only in some parts of the country or by one or a few parties.

2: Restricted. Money and/or personal gifts were distributed by parties or candidates but these offerings were more about meeting an 'entry-ticket' expectation and less about actual vote choice or turnout, even if a smaller number of individuals may also be persuaded.

3: Almost none. There was limited use of money and personal gifts, or these attempts were limited to a few small areas of the country. In all, they probably affected less than a few percent of voters.

4: None. There was no evidence of vote/turnout buying.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

6. Election other voting irregularities (C)

v2 tag: v2elirreg

Project manager: Staffan I. Lindberg

Question: In this national election, was there evidence of other *intentional* irregularities by incumbent and/or opposition parties, and/or vote fraud?

Clarification: Examples include use of double IDs, intentional lack of voting materials, ballot-stuffing, misreporting of votes, and false collation of votes.

This question does *not* refer to lack of access to registration, harassment of opposition parties, manipulations of the voter registry or vote-buying (dealt with in previous questions).

Responses:

0: Yes. There were systematic and almost nationwide other irregularities.

1: Yes, some. There were non-systematic, but rather common other irregularities, even if only in some parts of the country.

2: Sporadic. There were a limited number of sporadic other irregularities, and it is not clear whether they were intentional or disfavored particular groups.

3: Almost none. There were only a limited number of irregularities, and many were probably unintentional or did not disfavor particular groups' access to participation.

4: None. There was no evidence of intentional other irregularities. Unintentional irregularities resulting from human error and/or natural conditions may still have occurred.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

7. Election government intimidation (C)

v2 tag: v2elintim

Project manager: Staffan I. Lindberg

Question: In this national election, were opposition candidates/parties/campaign workers subjected to repression, intimidation, violence, or harassment *by the government, the ruling party, or their agents*?

Clarification: Other types of clearly distinguishable civil violence, even if politically motivated, during the election period should *not* be factored in when scoring this indicator (it is dealt with separately).

Responses:

0: Yes. The repression and intimidation by the government or its agents was so strong that the entire period was quiet.

1: Yes, frequent: There was systematic, frequent and violent harassment and intimidation of the opposition by the government or its agents during the election period.

2: Yes, some. There was periodic, not systematic, but possibly centrally coordinated – harassment and intimidation of the opposition by the government or its agents.

3: Restrained. There were sporadic instances of violent harassment and intimidation by the government or its agents, in at least one part of the country, and directed at only one or two local branches of opposition groups.

4: None. There was no harassment or intimidation of opposition by the government or its agents, during the election campaign period and polling day.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

8. Election other electoral violence (C)

v2 tag: v2elpeace

Project manager: Staffan I. Lindberg

Question: In this national election, was the campaign period, election day, and post-election process free from other types (*not by the government, the ruling party, or their agents*) of violence related to the conduct of the election and the campaigns (but not conducted by the government and its agents)?

Responses:

0: No. There was widespread violence between civilians occurring throughout the election period, or in an intense period of more than a week and in large swaths of the country. It resulted in a large number of deaths or displaced refugees.

1: Not really. There were significant levels of violence but not throughout the election period or beyond limited parts of the country. A few people may have died as a result, and some people may have been forced to move temporarily.

2: Somewhat. There were some outbursts of limited violence for a day or two, and only in a small part of the country. The number of injured and otherwise affected was relatively small.

3: Almost. There were only a few instances of isolated violent acts, involving only a few people; no one died and very few were injured.

4: Peaceful. No election-related violence between civilians occurred.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

9. Election free campaign media (C)

v2 tag: v2elfrcamp

Project manager: Staffan I. Lindberg

Question: Do parties or candidates receive either free or publicly financed access to national broadcast media during national election periods?

Responses:

0: Either no parties or only the governing party receives free access.

1: Some parties in addition to the governing party receive free access.

2: All parties receive free access.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

10. Election free and fair (C)

v2 tag: v2elfrfair

Project manager: Staffan I. Lindberg

Question: Taking all aspects of the pre-election period, election day, and the post-election process into account, would you consider this national election to be free and fair?

Clarification: The only thing that should *not* be considered in coding this is the extent of suffrage (by law). Thus, a free and fair election may occur even if the law excludes significant groups (an issue measured separately).

Responses:

0: No, not at all. The elections were fundamentally flawed and the official results had little if anything to do with the 'will of the people' (i.e., who became president; or who won the legislative majority).

1: Not really. While the elections allowed for some competition, the irregularities in the end affected the outcome of the election (i.e., who became president; or who won the legislative majority).

2: Ambiguous. There was substantial competition and freedom of participation but there were also significant irregularities. It is hard to determine whether the irregularities affected the outcome or not (as defined above).

3: Yes, somewhat. There were deficiencies and some degree of fraud and irregularities but these did not in the end affect the outcome (as defined above).

4: Yes. There was some amount or human error and logistical restrictions but these were largely unintentional and without significant consequences.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

Inequality indicators (V-Dem data, Coppedge et al, 2013):

1. Power distributed by social group (C)

v2 tag: v2pepwrsoc

Project manager: John Gerring

Question: Is political power distributed according to social groups?

Clarification: A social group is differentiated within a country by caste, ethnicity, language, race, region, religion, or some combination thereof. (It does *not* include identities grounded in sexual orientation or socioeconomic status.) Social group identity is contextually defined and is likely to vary across countries and through time. Social group identities are also likely to cross-cut, so that a given person could be defined in multiple ways, i.e., as part of multiple groups. Nonetheless, at any given point in time there are social groups within a society that are understood - by those residing within that society – to be different, in ways that may be politically relevant.

Responses:

0: Political power is monopolized by one social group comprising a minority of the population. This monopoly is institutionalized, i.e., not subject to frequent change.

1: Political power is monopolized by several social groups comprising a minority of the population. This monopoly is institutionalized, i.e., not subject to frequent change.

2: Political power is monopolized by several social groups comprising a majority of the population. This monopoly is institutionalized, i.e., not subject to frequent change.

3: Either all social groups possess some political power, with some groups having more power than others; or different social groups alternate in power, with one group controlling much of the political power for a period of time, followed by another – but all significant groups have a turn at the seat of power.

4: All social groups have roughly equal political power *or* there are no strong ethnic, caste, linguistic, racial, religious, or regional differences to speak of. Social group characteristics are not relevant to politics.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

2. Power distributed by socioeconomic position (C)

v2 tag: v2pepwrses

Project manager: John Gerring

Question: Is political power distributed according to socioeconomic position?

Clarification: All societies are characterized by some degree of economic (wealth and income) inequality. In some societies, income and wealth are distributed in a grossly unequal fashion. In others, the difference between rich and poor is not so great. Here, we are concerned not with the degree of social inequality but rather with the political effects of this inequality. Specifically, we are concerned with the extent to which wealth and income translates into political power.

Responses:

0: Wealthy people enjoy a virtual monopoly on political power. Average and poorer people have almost no influence.

1: Wealthy people enjoy a dominant hold on political power. People of average income have little say. Poorer people have essentially no influence.

2: Wealthy people have a very strong hold on political power. People of average or poorer income have some degree of influence but only on issues that matter less for wealthy people.

3: Wealthy people have more political power than others. But people of average income have almost as much influence and poor people also have a significant degree of political power.

4: Wealthy people have no more political power than those whose economic status is average or poor. Political power is more or less equally distributed across economic groups.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

3. Social class equality in respect for civil liberty (C)

v2 tag: v2clacjust

Project manager: Svend-Erik Skaaning

Question: Do poor people enjoy the same level of civil liberties as rich people do?

Clarification: This question specifies the extent to which the level of civil liberties is generally the same across socioeconomic groups so that people with a low social status are not treated worse than people with high social status. Here, civil liberties are understood to include access to justice, private property rights, freedom of movement, and freedom from forced labor.

Responses:

0: Poor people enjoy much fewer civil liberties than rich people.

1: Poor people enjoy substantially fewer civil liberties than rich people.

2: Poor people enjoy moderately fewer civil liberties than rich people.

3: Poor people enjoy slightly fewer civil liberties than rich people.

4: Poor people enjoy the same level of civil liberties as rich people.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

4. Social group equality in respect for civil liberties (C)

v2 tag: v2clsocgrp

Project manager: Svend-Erik Skaaning

Question: Do all social groups, as distinguished by language, ethnicity, religion, race, region, or caste, enjoy the same level of civil liberties, or are some groups generally in a more favorable position?

Clarification: Here, civil liberties are understood to include access to justice, private property rights, freedom of movement, and freedom from forced labor.

Responses:

0: Members of some social groups enjoy much fewer civil liberties than the general population.

1: Members of some social groups enjoy substantially fewer civil liberties than the general population.

2: Members of some social groups enjoy moderately fewer civil liberties than the general population.

3: Members of some social groups enjoy slightly fewer civil liberties than the general population.

4: Members of all salient social groups enjoy the same level of civil liberties.

Scale: Ordinal, converted to interval by the measurement model.

Cross-coder aggregation: Bayesian item response theory measurement model (see *V-Dem Methodology,* posted at V-Dem.net).

Control variables:

1. Ethnic fractionalization Alesina, Devleeschauwer, Easterly, Kurlat & Wacziarg (Alesina et al 2003)
The definition of ethnicity involves a combination of racial and linguistic characteristics. The variables reflect the probability that two randomly selected people from a given country will not share a certain characteristic, the higher the number the less probability of the two sharing that characteristic.
Years: 1946-2012
N: 189
Note: Country Constant Variable

Dataset accessed through the Quality of Government dataset
(http://www.qog.pol.gu.se/data/datadownloads/ 2014-03-20)

2. GDP per capita, PPP (constant international $) World Bank (World Bank WDI 2013)
GDP per capita based on purchasing power parity (PPP). PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United States. GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources.
Years: 1980-2011
N: 181

Dataset accessed through the Quality of Government dataset
(http://www.qog.pol.gu.se/data/datadownloads/ 2014-03-20)

3. Gini Household Disposable Income (Solt 2008)
Estimate of Gini index of inequality in equivalized (square root scale) household disposable income, using Luxembourg Income Study data as the standard.
The Standardized World Income Inequality Database
A custom missing-data algorithm was used to standardize the United Nations University's World Income Inequality Database; data collected by the Luxembourg Income Study served as the standard.
Years: 1960-2010
N: 169
Dataset accessed through the Quality of Government dataset
(http://www.qog.pol.gu.se/data/datadownloads/ 2014-03-20)

Supplemented by data gathered by Ansell and Samuels (2010).

4. Electoral System Type (Bormann & Golder 2013)
The basic type of electoral system used in the elections.
(1) Majoritarian
(2) Proportional
(3) Mixed
Years: 1946-2011
N: 134

Dataset accessed through the Quality of Government dataset
(http://www.qog.pol.gu.se/data/datadownloads/ 2014-03-20)

5. Democracy (Freedom House/Imputed Polity (Hadenius & Teorell 2005)

Scale ranges from 0-10 where 0 is least democratic and 10 most democratic. Average of Freedom House (fh_pr and fh_cl) is transformed to a scale 0-10 and Polity (p_polity2) is transformed to a scale 0-10. These variables are averaged into fh_polity2.
The imputed version has imputed values for countries where data on Polity is missing by regressing Polity on the average Freedom House measure. Hadenius and Teorell (2005) show that this average index performs better both in terms of validity and reliability than its constituent parts.
Years: 1972-2012
N: 207
Dataset accessed through the Quality of Government dataset
(http://www.qog.pol.gu.se/data/datadownloads/ 2014-03-20)