On studying the unstudyable

A quantitative literature review of proposed causes of corruption
and how these have been studied

Kristian Blindheim

Master’s Thesis
Department of Comparative Politics
University of Bergen
Abstract

This thesis is a quantitative review of research on the causes of corruption. The aim is twofold: firstly to establish the degree of agreement between researchers regarding the causal effect of seven different explanatory dimensions on corruption. These explanatory dimensions are: economic liberalization, level of democracy, regime transitions, the strength of democratic traditions, the presence of a free media, level of economic development and natural resource dependency. I find that there is high agreement that economic liberalizations and regime transitions lead to more corruption. There is also high consensus that the presence of a free media and high economic development, respectively, lead to less. Concerning the effect of level of democracy, democratic traditions and natural resource dependency the findings are more diverse. About half of prior research concludes that level of democracy and stronger democratic traditions lead to less corruption. Also about half conclude that natural resource dependency leads to more corruption.

The second aim of this thesis is to discover whether the disagreement between researchers can be explained by the data they have applied as operationalizations of corruption as the dependent variable. This is done through a series of logistic regressions, where the explanatory variables are the different data sources on corruption. These fall into the categories “perceptions based cross-national data”, “experience based cross national data” and “country-or region-specific studies”. The results show that applying experience based cross national data gives significantly lower likelihood for concluding that higher economic development leads to less corruption. It also makes it less likely to conclude that natural resources lead to more corruption. Applying perceptions based cross national data in some instances provide higher likelihood for observing significant effects of the explanatory dimensions treated in this thesis and sometimes lower likelihood. An interesting finding is that the Transparency International Corruption Perceptions Index and studies conducted with country- or region specific foci “agree” on the effects on corruption of all of the treated explanatory dimension where comparison is possible. This is counter to what is assumed based in theory.
Acknowledgements

On completing this project there are several people to whom I owe gratitude. First I must thank my tutor at the Department of comparative Politics, Professor Lise Rakner, for substantial and sound advice. Associate Professor Michael Alvarez and Post Doctor Leiv Marsteinstredet also deserve thanks for good advice and discussions regarding methodological choices.

I also would like to thank the Department’s research groups on democracy and development and methods, respectively, for sound and honest advice on several occasions. In particular, thanks to Professor Tor Midtbø for encouraging me about the feasibility of the project.

My friend, class mate and flat mate, Iver M. Fiksdal, also deserves my gratitude. He has read drafts of my thesis on two occasions and has readily answered almost any query I have had regarding theory, methods and the Universe in general.

My fellow students at Sofie Lindstrøms Hus have provided a good environment for both learning and amusement. Thank you all.

Lastly I would like to thank my girlfriend, Tonje Vollset, for her patience and support.

Kristian Blindheim
Bergen, June 2011
Contents:

Tables: ........................................................................................................................................ vii
Figures: ..........................................................................................................................................viii
Chapter 1: Introduction .................................................................................................................. 1
  1.1 Studying research .................................................................................................................. 3
  1.2 Why study corruption? .......................................................................................................... 4
  1.3 The structure and approach of this thesis .............................................................................. 6
Chapter 2: Corruption: what it is and how it may be explained ............................................... 9
  2.1.1 Defining corruption ......................................................................................................... 9
  2.1.2 Types of corruption ......................................................................................................... 11
  2.2 The scope of the study: seven dependent variables ............................................................... 12
      2.2.1 The effect of economic liberalization on corruption ................................................... 14
      2.2.2 The effect of democracy- versus autocracy on corruption .......................................... 16
      2.2.3 The effect of regime transitions on corruption ............................................................. 18
      2.2.4 The effect of strong versus weak democratic traditions ............................................. 20
      2.2.5 The effect of media presence on corruption ................................................................. 21
      2.2.6 The effect of level of economic development on corruption ....................................... 23
      2.2.7 The resource curse ....................................................................................................... 25
  2.3 Concluding remarks ............................................................................................................. 26
Chapter 3: Independent variables: data sources on corruption ............................................... 27
  3.1 Qualitative versus quantitative methods ................................................................................ 28
  3.2 Data and research on corruption ............................................................................................ 31
      3.2.1 Perceptions based measures ......................................................................................... 31
      3.2.2 Experience based measures ......................................................................................... 34
      3.2.3 Data sources available for country- and region-specific studies .................................. 35
  3.3 Measurement errors: formulating hypotheses ......................................................................... 40
      3.3.1 Possible classical measurement errors in corruption data ........................................... 41
      3.3.2 Possible divergence between experience based and perceptions based data .............. 45
      3.3.3 Non-classical measurement error .................................................................................. 46
      3.3.4 Possible divergence between experience based and perceptions based data .............. 48
  3.4 Concluding remarks ............................................................................................................. 50
Chapter 4: Choices of methods and data for this thesis ............................................................ 52
  4.1 The choice of a quantitative approach ................................................................................... 52
4.2 Reviewing literature quantitatively ................................................................. 53
4.3 Operationalizations of dependent and independent variables ....................... 54
4.4 Binary logistic regression analysis .................................................................... 56
  4.4.1 Interpretation ................................................................................................. 58
  4.4.2 Test statistics ............................................................................................... 59
  4.4.3 The assumptions of logistic regression. ....................................................... 60
4.5 Creating a sample: ............................................................................................. 62
4.6 Concluding remarks ......................................................................................... 64

Chapter 5: Analyses ............................................................................................... 65
  5.1.1 The effect of economic liberalization on corruption ..................................... 65
  5.1.2 The effect of democracy on corruption ..................................................... 67
  5.1.3 The effect of regime transition on corruption: ......................................... 73
  5.1.4 The effect of stronger democratic tradition on corruption ....................... 74
  5.1.5 The effect of presence of a free media on corruption ............................... 77
  5.1.6 The effects of level of economic development on corruption ................... 80
  5.1.7 The resource curse ..................................................................................... 83
  5.2 Discussion of findings ...................................................................................... 86

6. Concluding remarks ........................................................................................... 92
  6.1 Summary of approach and findings .............................................................. 92
  6.2 Implications of findings .................................................................................. 93
    6.2.1 Economic liberalization ............................................................................ 93
    6.2.2 Democracy: transition to-, presence of-, and experience with ............... 93
    6.2.3 Presence of a strong and independent media .......................................... 94
    6.2.4 Level of economic development ............................................................ 95
    6.2.5 Natural resource dependency ............................................................... 95
    6.2.6 The data issue .......................................................................................... 96

7. Literature ............................................................................................................. 97

8. Appendix ............................................................................................................. 118
Tables:

TABLE 1: DEPENDENT VARIABLES OF EACH OF THE SUB-CHAPTERS ........................................ 14
TABLE 2: HYPOTHESES PRESENTED IN THIS CHAPTER .......................................................... 41
TABLE 3: CORRELATIONS BETWEEN DIFFERENT YEARS OF THE CPI .................................. 44
TABLE 4: CORRELATIONS BETWEEN DIFFERENT YEARS OF THE CCI .................................. 44
TABLE 5: CORRELATIONS BETWEEN DIFFERENT CORRUPTION INDEXES ............................. 44
TABLE 6: CORRELATIONS BETWEEN DIFFERENT CORRUPTION INDEXES ............................. 44
TABLE 7: SUMMING UP THE GOALS OF THIS THESIS .............................................................. 52
TABLE 8: OPERATIONALIZATIONS OF EXPLANATORY VARIABLES ......................................... 55
TABLE 9: THE EFFECT OF ECONOMIC LIBERALIZATION ON CORRUPTION: SUMMARY OF DATA SOURCES AND CONCLUSIONS ................................................................. 66
TABLE 10: THE EFFECT OF DEMOCRACY VERSUS AUTOCRACY ON CORRUPTION: SUMMARY OF DATA SOURCES AND CONCLUSIONS ................................................................. 68
TABLE 11: DEPENDENT VARIABLE: CONCLUSIONS THE EFFECT OF DEMOCRACY VERSUS AUTOCRACY ON CORRUPTION .......................................................... 70
TABLE 12: MODEL SUMMARIES .................................................................................................. 70
TABLE 13: RESULTS IN LIGHT OF HYPOTHESES OF CHAPTER 3 ........................................... 73
TABLE 14: THE EFFECT OF REGIME TRANSITION ON CORRUPTION: DATA SOURCES AND CONCLUSIONS ........................................................................................................... 73
TABLE 15: THE EFFECT OF AGE OF DEMOCRACY ON CORRUPTION: DATA SOURCES AND CONCLUSIONS ........................................................................................................... 74
TABLE 16: DEPENDENT VARIABLE: CONCLUSIONS ON EFFECT OF STRENGTH OF DEMOCRATIC REGIMES ON CORRUPTION ......................................................... 75
TABLE 17: MODEL SUMMARIES .................................................................................................. 75
TABLE 18: SUMMING RESULTS IN LIGHT OF HYPOTHESES ..................................................... 77
TABLE 19: THE EFFECT OF PRESS FREEDOM ON CORRUPTION: SUMMARY OF DATA SOURCES AND CONCLUSIONS ........................................................................................................... 77
TABLE 20: DEPENDENT VARIABLE: CONCLUSIONS ON EFFECT OF PRESS FREEDOM ON CORRUPTION ........................................................................................................... 79
TABLE 21: MODEL SUMMARIES .................................................................................................. 79
TABLE 22: RESULTS IN LIGHT OF HYPOTHESES OF CHAPTER 3* .......................................... 80
TABLE 23: THE EFFECTS OF LEVEL OF ECONOMIC DEVELOPMENT ON CORRUPTION: DATA SOURCES AND CONCLUSIONS ........................................................................................................... 80
TABLE 24: DEPENDENT VARIABLE: CONCLUSIONS ON EFFECT OF ECONOMIC DEVELOPMENT ON CORRUPTION ........................................................................................................... 81
TABLE 25: MODEL SUMMARIES .................................................................................................................. 81
TABLE 26: SUMMING UP RESULTS IN LIGHT OF HYPOTHESES OF CHAPTER 3 ......................... 83
TABLE 27: THE EFFECTS OF NATURAL RESOURCE DEPENDENCY: SUMMARY OF DATA SOURCES
AND CONCLUSIONS .................................................................................................................................. 84
TABLE 28: DEPENDENT VARIABLE: CONCLUSIONS ON THE EFFECT OF NATURAL .................. 85
TABLE 29: MODEL SUMMARIES .................................................................................................................. 85
TABLE 30: SUMMING UP FINDINGS IN LIGHT OF HYPOTHESES ......................................................... 86
TABLE 31: SUMMARY OF ALL FINDINGS* ................................................................................................. 88
TABLE 32: SUMMING ALL FINDINGS IN LIGHT OF HYPOTHESES ....................................................... 90
TABLE 33: OVERVIEW OF PUBLICATIONS AND RESULTS INCLUDED IN THE DATA SET ............. 118

Figures:

FIGURE 1: CAUSAL DIAGRAM OF TREISMAN'S SUGGESTION ............................................................... 28
FIGURE 2: CAUSAL DIAGRAM OF ALL HYPOTHESES ............................................................................ 50
Chapter 1: Introduction

Corruption has been studied theoretically and empirically, quantitatively both within single countries and cross-nationally with large samples, it has been studied qualitatively, and over the last ten years a small number of experimental studies on the phenomenon has emerged. They all have their weaknesses. Early modern social scientists studies of corruption were either discussions of issues of definition (like Friedrich 1972), theoretical discussions of causes and consequences of corruption (like Huntington 1968) or empirical analyses applying qualitative data (like Myrdal 1968). Reviewing the scholarly literature on causes of corruption up until 1980, Deysine concluded that “authors do not recognize the same causes of corruption. Emphasis is put either on overregulation and discretion or on the relation between politics and the business world. But most authors recognize that corruption can be found in political parties, in Congress, [and] in the courts.” (Deysine 1980:42). Disagreement on what causes corruption still seems to be the situation (Treisman 2007; You 2010:916).

In this thesis I will review and discuss how corruption, more specifically high level corruption, has been studied empirically, and which strengths and weaknesses the different approaches face. I will seek to establish the degree to which there is agreement on a series of suggested causes of corruption. Based in this I will seek to review whether disagreements may have their roots in methodological choices made by researchers. The differing data sources on corruption will be central in the analyses and will serve as departures for developing explanatory variables on which I will regress the results provided in prior research.

In the beginning of the 1980's the first cross national data on corruption were made available, the Business International Index (Alesina and Weder 2002) and the International Country Risk Guide (PRS-Group 2011). Application of quantitative cross national data when studying corruption was a welcome contribution. With the first publication of the Transparency International Corruption Perceptions Index (CPI) in 1995 (Transparency International 2011) yet another measure was added to the arsenal of tools, which lead to an increase in publications on the topic (Andersson and Bergman 2009:45). Today this is probably, together with the World Banks Control of Corruption Index (CCI) (World Bank 2010), the most commonly applied cross national measure of corruption. Perceptions of corruption “should never be confused with reality”, admits Lambsdorff (2007:20), whom created the CPI. However he also claims that “the given consensus provides some confidence that the perceptions gathered are informative on actual levels of corruption” (2007:20).
However, not all agree that the discussion on measurement of corruption resembles one of any considerable degree of consensus. The debate about the applicability of the cross national survey measures of corruption has been quite intense with several researchers concluding that the cross national measures of corruption simply are not strong enough to help us evaluate the theories presented on the causes of corruption (i.e. Gingerich 2010:364; Treisman 2007). Philip (2006) is one of those think that such measures should be abandoned altogether and that researchers in stead should focus on single countries. However, the measures are still applied by a range of researchers.

The CCI and the CPI have that in common that they are based on expert perceptions of corruption levels in different countries. Treisman suggests that there may be a causal relationship between operationalizing corruption as a dependent variable with the perceptions based measures and the conclusions arrived at; comparing the results applying these measures with measures based on actual experiences with corruption Treisman finds that the likelihood of observing significant effects increase dramatically when applying perceptions based measures. He suggests that there is a causal chain going from what respondents of perceptions based measures believe causes corruption via their responses to the strong explanatory power of regression models applying the perceptions based measures; i.e. if a researchers recognizes the much discussed theory that democracy leads to less corruption, his responses to a corruption perceptions survey may, since actual knowledge of corruption levels may be weak, rather reflect his knowledge of the degree to which the country in question is a democracy. A consequence would be that democracy is shown to have explanatory power on corruption in a regression analysis. If this is true, many of the causes of corruption concluded on when applying perceptions based measures of corruption as dependent variable could be spurious. This claim by Treisman is perhaps the most profound inspiration of this thesis. In addition to comparing perceptions based and experience based cross national data, I will compare the results derived with these to results derived when studying corruption in single countries or regions.

Based on Deysine’s claim that there used to be high disagreement on what causes corruption and Treisman’s indication that an eventual disagreement of what causes corruption may be caused by the data sources applied in the research, I formulate the following two-fold research question:
To what degree is there agreement on causes of high-level corruption in the field? Is there a causal relationship between the data sources researchers have applied and the results they obtain?

1.1 Studying research
Scientific controversy may spring from a number of issues; from the arguments presented above on the data on corruption, it seems that we are facing a scientific controversy that has its roots in data, something that has inspired the latter of the two research questions. Thus I am going to turn a field of study into the matter of inquiry. This is the core idea of the field of research known as the sociology of scientific knowledge (Ashmore 2001). Perhaps the most influential work in the creation of the field was Thomas Kuhn's (Kuhn 1962) *The Structure of Scientific Revolutions*. While Popper (i.e. Popper 1963) who claimed that all scientific findings, while not possible to prove to be true, are possible to falsify with almost complete certainty, Kuhn argued that both true and false are to some extent depending on socially determined views of science (Collins 2001). Lakatos, Worrall and Zahar (1976) showed empirically in their study of the research done on Euler's theorem that scientific discovery is only one factor affecting what researchers consider true. The works of these scientists are crucial to the birth of the sociology of scientific knowledge (Collins 2001). Thus the idea that it may be more than truth itself that determine what we regard as true is not new.

Many aspects of the sociology of scientific knowledge could have been elaborated on. However, the focus of this study will be methodological aspects, in particular concerns about data quality and consequences thereof. Every researcher has his or her values or meta-methodology (Bevir 2008:48). They be explicit or not, they may affect his or her definition of research problems and research questions, and which methodological solutions they choose which again may affect how they conclude (Bevir 2008:48). These thoughts may shed light on the data-debate amongst researchers on corruption.

It is commonly recognized amongst social scientists that research fields should operate as accumulative (Mahoney 2003); generation of new knowledge should be founded upon the findings presented in prior research and the research communities’ understanding as a whole

---

1 Central topics of the field are effects of age (i.e. Bourdieu 1999 [1975]), gender (see i.e. Gilligan 1982), how researchers may be institutionalized into certain theoretical and methodological practices (see i.e. Ziman 1984; or Sismondo 2010), how organization of research may affect the research itself (see i.e. Merton and Storer 1973) etc. For a good overview of science and technology studies, see Sismondo (2010).
should grow. This demands literature reviews (Wolf 1986), something that according to some has been missing in social science research: “Research activity in the social and behavioral sciences has been observed to be “sadly dilapidated” and facing a “crisis” situation” (Wolf 1986). Wolf claims that the natural, biological and physical sciences have come much further in the strength of the accumulative nature of their research fields than the social sciences. This clearly has something to do with the social sciences having less standardized approaches concerning i.e. data collection and methodological opportunities and choices (Wolf 1986).

As a field of research grows larger it clearly is impossible for researchers to sum up the whole field and comment upon every single article in every publication. The task of putting the vast amount of primary research into some sort of order lies with those who review a field of study, through i.e. writing review articles. Ideally these should be unbiased presentations of the findings of the central findings and controversies, presenting the evidence and arguments and letting the reader be the judge, but also commenting on which findings should be seen as the road to follow in the future and which should be left to the past. Only this is a quite delicate task that requires good deals of discretion and scientific judgment (Ziman 1984). However, Ziman also asks whether literature reviews should not take it a step further, and not only review fields of research in the sense of summing up the findings; reviewers might also do good in evaluating and criticizing the research field in question. The task of this thesis is to establish an overview of the degree of agreement on central topics and, with that as a departure, seek to discover whether eventual disagreements have their root in how causes of corruption are studied.

1.2 Why study corruption?
As social scientists we should seek to contribute to the field of study in which we are engaged, by seeking to resolve the major concerns of the field, i.e. by contributing to the theories of the field or the methodological understandings of how the topics of the field could or should be studied (Mahoney 2003). Entering the data-controversy in corruption research, and also seeking to describe degree of agreement of what causes corruption, answers this call. That studying causes of corruption, also is of importance for society in general seems clear. Recall a couple of incidents:

«Bettino, take these too, Bettino, take these too!» a crowd with money in their hands roared at former Italian prime minister Bettino Craxi as he came out of Hotel Raphael
where he was residing (Abruzzoitalia.it 2010). As prime minister of Italy he had been the 
centre of a corruption scandal of ridiculous proportions the investigation of which became 
known as the “mani pulite”- (clean hands) processes. It started out with the arrest of city 
councilor of Milan, Mario Chiesa, for receiving a minor bribe. When Chiesa started to 
cooperate with the investigators, he set off a chain of events that would reveal one of the 
largest corruption scandals in history: «[m]ore than five hundred former parliamentarians 
were implicated, many former ministers, five former premiers, thousands of local 
administrators and public functionaries, the army, the customs service (responsible of 
investigating financial crimes in general), the main publicly owned companies and even 
sectors of the magistracy itself» (Della Porta and Vannucci 1999).

The scandal surrounding French state owned oil company, Elf Aquitaine, may remind 
seemingly unaffected Western-Europeans that corruption is not a distant in time and space-
phenomenon. In 1995 Prime Minister Edouard Balladur appointed Philippe Jaffre new 
director of the company. He was hoping that the new boss, a former finance inspector, would 
uncover corruption in the company that would draw president Francois Mitterrand and the 
socialist party down in the dirt. However, Philippe Jaffre did a better job than Balladur 
expected. In a series of audits initiated by the Jaffre, large scale corrupt activity was revealed 
embroiling a number of high-ranking politicians, both in France and elsewhere (Heilbrunn 
2005), i.e. former foreign minister Roland Dumas, (NTBtekst 2001), former French President 
Francois Mitterrand, former German president Helmut Kohl, Gabonese president Omar 
Bongo, and Congolese president Denis Sassou-Nguesso. During the trials it was revealed that 
Elf had been used an important foreign policy tool for the French state. In the years 1989-
1993 at least 183 million Euros were used as bribes to land Elf contracts in Germany, Spain, 
Russia, South America and Africa (DW-world.de 18.03.2003). Among those finally convicted 
was former chairman of Elf, Loik Le Floch-Prigent. He claimed that all presidents of France 
had been aware of the illicit dealings of the company. A second function of Elf was as slush 
fund for financing political party activity. Le Floch-Prigent claimed during the trials that 
former French president and socialist Francois Mitterrand had ordered him to spread the 
money evenly between the influential parties rather than giving the bulk to Jacques Chirac's 
centre right party which up until then had been the norm (Henley 13.11.2003).

However, the consequences of high-level corruption go much further than simply 
scandalizing and possibly delegitimizing politics. It is now quite widely recognized that
corruption increases poverty and economic inequality (Blackburn and Forgues-Puccio 2007; Fjelde 2009; Apergis et al. 2010), reduces tax revenues (Attila et al. 2009; Baker 2005), incline governments to down-prioritize spending on health and education (Baraldi 2008; Baker 2005), leads to less foreign direct investment (Gani 2007; Baker 2005), slows entrepreneurship (Anokhin and Schulze 2009; Mitchell and Campbell 2009; Baker 2005), increase public spending on defense material and other sectors with many large scale investments (Baker 2005), increase debts of developing countries (Baker 2005), decreases economic growth (Blackburn and Forgues-Puccio 2007; Evrensel 2010), lead to more human trafficking (Studnicka 2010), severely decreases exports (Musila and Sigue 2010), delegitimize politics (Chang and Chu 2006), may threaten political stability (Wong 2009) and lead to more violations of human rights (United Nations 2004).

It has, it must be pointed out, been claimed that corruption may affect many of these dimensions in opposite ways, being socially desirable; however, there seems to be agreement that the conditions under which corruption is socially desirable are extremely rare (Beblavý 2007). In comparison the socially negative consequences of absence of democracy are negligible compared to the negative consequences of corruption (Rothstein 2010).

Considering that corruption generally is on the rise in Africa, and in post-communist Europe (Volejníková 2007) underlines the severity of the consequences of corruption and the importance of seeking to understand the causes of corruption.

1.3 The structure and approach of this thesis

In the next chapter I will start with defining corruption and set the more specific limits of this thesis. I will also describe seven central dimensions that have been commonly applied to seek to explain variance in corruption; economic liberalization, level of democracy, regime transitions, strength of democratic traditions, presence of a strong and independent media, level of economic development and natural resource dependency. These common explanatory dimensions will form the basis for seven dependent variables seeking to capture whether researchers have concluded that i.e. democracy leads to less corruption.

In Chapter 3 I will present what seems to be the central conflict of the field: which data sources on corruption are appropriate when seeking to uncover causes of high-level corruption? This involves presenting in some detail the different sources available. These may broadly be categorized into cross national perceptions based survey data, cross national
experience based survey data, and the range of data sources available for quantitative studies focusing on single- or a few countries or qualitative approaches; this category of studies I will hereafter denote country-or region specific studies. Whether these data sources have been applied in the different studies will serve as explanatory variables.

An important inspiration of this thesis is Treisman (Treisman 2007). He conducts several cross national analyses applying a range of different cross national corruption measures finding that there is a clear pattern that when applying perceptions base measures of corruption as dependent variable more and stronger causal relationships are established compared to when applying experience based measures. He suggests that this may be because the respondents to perceptions based measures let the knowledge of the “status” of the theoretically plausible explanatory variables in the country in question may lead respondents to rate countries in accordance with their knowledge of the explanatory variables; the effect may be, according to Treisman, very high explanatory power when applying these measures as dependent variables. An important difference between my study and Treisman’s is that I also include studies that have not applied cross national survey data on corruption, thus providing an interesting and fruitful comparison. Departing from a thorough discussion of the reliability and validity of the different data sources I, in Chapter 3, formulate a series of hypotheses of causal relationships between different operationalizations of corruption as dependent variables and conclusions of whether the different explanatory variables lead to corruption.

Chapter 4 will present the method of this thesis; I will go through a large amount of previous research preparing a data set which will found the basis for the analyses. After a search for “corruption” in ISI Web of Knowledge I go through 3411 abstracts; data is extracted from 285 publications providing a data set of 1218 units, the units being the results of the relevant explanatory variables on corruption.

In the analyses chapter (Chapter 5) I first analyze the data descriptively establishing the degree to which there is agreement on the findings. I find that there is high degree of agreement that economic liberalizations and regime transitions lead to more corruption, while there is high agreement that presence of a free media and high economic development lead to less. About the half of prior research conclude that level of democracy and stronger democratic traditions lead to less corruption. Also about half conclude that natural resource dependency leads to more corruption.
Secondly I analyze the data in a series of logistic regressions seeking to uncover whether there are causal relationships between the data on corruption and the results. The data source, if any, that stands out as providing diverging results is experience based cross national data. The CCI also shows some divergence. However, considering the temperature of the debate on data quality in research on causes of corruption, the degree of divergence between results derived applying different sources is surprisingly small.
Chapter 2: Corruption: what it is and how it may be explained

In this chapter I will first define corruption, and set the thematic limits to this thesis to high-level corruption and what will be defined as “demand-side” causes thereof. Then I will present some commonly suggested and analyzed causes of corruption; economic liberalization, presence of democracy, the presence of a free media, regime transition, strength of democratic traditions, economic development and natural resource dependency. These explanatory dimensions, or more specifically whether researchers have concluded that these explanatory dimensions have effects on high-level corruption will found the basis for the dependent variables of this thesis.

2.1.1 Defining corruption

"Corruption is behavior which deviates from the normal duties of a public role because of private-regarding (personal, close family, private clique) pecuniary or status gains; or violates rules against certain types of private-regarding influence” (Nye 1967:997).

This has become a widely cited definition of corruption. In the last 20 years definitions of corruption have usually been along the lines of this definition by Nye (Baker 2005). Many other definitions have been posed i.e. to clarify what are to be considered “normal duties of a public role”. Gardiner (1993) provides a threefold categorization of normative foundations definitions of corruption may apply: the public interest, laws or the public opinion.

A legal definition:

“if an official’s act is prohibited by laws established by the government, it is corrupt; if it is not prohibited, it is not corrupt even if it is abusive or unethical” (Gardiner, 1993:29).

This definition give us a much clearer view than Nye’s definition of what it is people are not supposed to do. However, it poses a few challenges. Firstly, such a definition makes it difficult to compare corruption across time and space, since nations have different laws and
might change their laws (Heidenheimer 2002:764). This may even be the case between
different jurisdictions of the same country (Noonan in Heidenheimer, 2002:764). I.e. many
acts that today would be considered violation of Norwegian law, would not be considered
corrupt before the revision of the criminal code effective in 2003 (Hjellum 2007). Secondly,
legal definitions pose a problem to the degree laws are challenged by other ethical aspects of
a society; would it be corrupt if a Chinese oppositional political movement broke Chinese law
to finance its activities? Are Chinese laws regarding corruption in concurrence with what one
might consider to be the public interest for instance from a perspective of liberal democracy?

A public interest-definition:

“if an act is harmful to the public interest, it is corrupt even if it is
legal; if it is beneficial to the public, it is not corrupt even if it violates
the law” (Gardiner, 1993:32).

Here the focus is more on the effect of the act rather than whether it is in accordance with a
law (Gardiner, 1993:32). This category of definitions obviously poses the challenge that there
may be different opinions regarding what is beneficial for the public and what is not. Many
have of course suggested definitions of what the public interest is. Even though the public
interest has been defined by many, a common understanding of it may be impossible to
achieve (Johnston 1986:460). The public interest as a departure for empirical research on
corruption seems to be of little value (Gardiner-1992:34).

Public opinion-definitions. Since the beginning of the 70s several studies aiming to
map what the public defines as corrupt have been conducted. Surveys have been conducted
amongst i.e. legislators, other officials or ordinary people who i.e. have been asked to judge a
number of hypothetical situations as corrupt or not corrupt or whether and to what extent
people see corruption as something undesirable (Gardiner 1993). A problem with such
measures is that people’s opinions are unstable and fluctuates (Johnston, 1982 in Gardiner,
1993:33). Another problem is how to decide whom to ask; who is the public? (Erskin 1973 in
Gardiner 1993:33).

In this study I will adhere to the definition posed by Nye; corruption as the misuse of
public office for private gain. Since my aim is wide, geographically, narrowing the definition
down further regarding which norms of conduct of public office should be considered, i.e.
laws, public interest or public opinion, would be impractical. Also since I aim to evaluate how different sources of data affect results and since different sources of data depart from different definitions of corruption it would not make sense to narrow it down further in this regard.

In addition I must discuss which type of corruption I will be discussing in this thesis.

2.1.2 Types of corruption
Regardless of which normative departure definitions of corruption are founded upon, one must also discuss types of corruption; at which level and in which way is a public office abused for private benefit? Corruption is often, particularly in common speech, talked of in more specific ways than “misuse of public office for private gain”. The term bribe is perhaps the specification of corruption most commonly applied in everyday language. Deriving from the discussions above, we can conclude that bribing definitely is a form of corruption, but that corruption clearly also is a lot more i.e. “cronyism, nepotism, patronage, graft, and embezzlement” (Pajunen 2008:654). Since corruption often is studied in its “completeness”, i.e. researchers very often do not study only bribery or nepotism or embezzlement, but the whole concept of corruption, I too will be concerned with the whole concept of corruption. This has been criticized by some since we very well may imagine that different aspects of the concept of corruption may have differing causes and consequences (Lancaster and Montinola 2001). However, since this thesis aims to review literature, and since this is how the phenomenon is most frequently understood in empirical research, this is how I will understand it.

I will, however, narrow the scope of the study down to high-level corruption, also called graft or grand corruption (Andvig et al. 2000). This is a choice that is made for the sake of comparability. For the sake of comparability with studies relying on quantitative cross-national data, I will limit this thesis to country-or region specific studies that have countries or central state institutions such as the executive or legislative bodies as their units of analysis. It seems impractical to compare national level corruption as which the cross national perceptions based measures are often understood with corruption i.e. in emergency departments of Iranian Hospitals (Mirhosseini and Fattahi 2010), were micro level explanatory dimensions may be in focus. An objection to this choice is that many of the
quantitative cross-national seek to capture overall corruption levels in countries, also low-
level corruption. However, it appears to me that corruption within central institutions may be
the corruption the most reflective of the overall level. In addition many of the explanatory
dimensions applied to explain high level corruption are not the same as those applied to
explain low-level corruption, something that further underlines the need to make this choice
to fix the limits of the study.

Further I will narrow the discussion of what may be causes of corruption down to so
called “demand-side” causes of corruption, as opposed to supply-side causes (Beets 2005).
The latter concerns characteristics of those who i.e. pay a bribe; supply-side causes of
corruption may i.e. lie within multinational corporations. Demand side causes typically
concern characteristics of countries, which may affect, in this case, the inclination of high-
level officials to extract money from their office i.e. through fraud or demanding bribes.

2.2 The scope of the study: seven dependent variables
To be entirely sure: the dependent variables of this study aim to capture whether researchers
have found effects of the explanatory variables described in this section on corruption. In
other words, their explanatory variables found basis for my dependent variables. In this
section I will clarify and justify which explanatory dimensions applied in research on causes
of high-level corruption will form basis for my dependent variables.

I established in the preceding sections that I am concerned with high-level corruption
and the causes of high-level corruption that may be understood as “demand-side” causes. A
vast amount of possible causes have been suggested and studied; I cannot cover them all. I
must seek to study the possible causes of corruption that may help me shed light on my
research question; I am concerned with the degree of agreement amongst researchers and
whether eventual disagreement may be caused by the choice of data on corruption. This
implies seeking to establish a general pattern of conclusions within the field’s different
explanatory dimensions, and since more facts make it easier to generalize (King et al.
1994:46; Lijphart 1971; Landman 2002), I should focus on explanatory dimensions that have
been studied by many. This selection has been done partly based on suggestions of previous
extensive reviews of the literature, and partly pragmatically when constructing the data set
(described in Chapter 4), by gradually discovering the extent to which the different
explanatory dimensions have been studied.
Andvig, Fjeldstad, Amundsen, Sissener and Søreide (2000:50-91) describe level of economic development, regime type (i.e. democracy versus autocracy), democratization, degree of federalism and decentralization, the level of public sector salaries and the countries’ openness to international trade as topics that have been common foci of studies seeking to explain corruption levels. Treisman (2007) in a review of efforts to explain corruption cross-nationally, in addition to many of the topics pointed to by Andvig et al., presents the strength of democratic traditions, unlike merely the present degree of democracy, the degree of presence of a free media, degree of female representation in political bodies, economic dependency on natural resources, the degree of regulation of business and the predictability of inflation as central topics in the field. After economists showed increased interest in research on corruption, Hopkin argues, degree of state presence in the economy also has become a much studied topic (Hopkin 2002). Asfar (2001) points to economic liberalization as a topic that has been central in the literature, something that also has been an important anti-corruption policy.

These indications of prior reviewers of the literature form important departures in deciding where to lay the focus of this study. I started collecting data on all of these explanatory dimensions and realized that insufficient amounts of prior work had been conducted on the relationship between high-level corruption and public sector salaries, countries’ openness to international trade, the degree of female representation in political bodies, and the predictability of inflation. Degree of federalism and decentralization conflicts with the focus of this thesis since I am concerned with corruption in the central political bodies of the state. Neither will the degree of regulation be considered in this thesis because it seems more relevant for bureaucratic corruption (Hopkin and Rodriguez-Pose 2007). Degree of state interference has been studied quite extensively; however, I considered that the different aspects of this dimension differed so greatly, i.e. government expenditure (like Gerring and Thacker 2005), taxation (like Hefeker 2010), and state ownership (like Quinn 2008) Sung, 2002 #1060:150}, that they cannot easily be merged into one dependent variable. Neither did they seem to be sufficiently covered in the literature to be treated by themselves in light of the aims of this thesis. Thus the dependent variables that will be applied in this thesis.

These paragraphs present topics that seem to have been deemed central by researchers who have review extensive amounts of the literature. Many more explanatory dimensions have been suggested and studied: A topic that has been given some attention is how different electoral systems may provide incentives for politicians and/or parties that may promote or deter corruption (Golden and Chang 2001; Johnston 2002; Cox 1998; Carey 1995; Persson 2003; Gingerich 2009). Different ethnic
are as summed up in Table 1. However, these variables will also be subject to extensive descriptive analyses, both to form a foundation for the causal analyses, but also because it has a value of interest in itself to discover the extent to which researchers agree on the effects of these dimensions.

Table 1: Dependent variables of each of the sub-chapters

<table>
<thead>
<tr>
<th>Dependent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Whether researchers have concluded that economic liberalization leads to more corruption.</td>
</tr>
<tr>
<td>2. Whether researchers have concluded that democracy leads to less corruption and/or that autocracy leads to more corruption.</td>
</tr>
<tr>
<td>3. Whether researchers have concluded that regime transitions lead to more corruption.</td>
</tr>
<tr>
<td>4. Whether researchers have concluded that stronger democratic traditions lead to less corruption.</td>
</tr>
<tr>
<td>5. Whether researchers have concluded that higher degree of media freedom and presence leads to less corruption.</td>
</tr>
<tr>
<td>6. Whether researchers have concluded that higher economic development leads to less corruption.</td>
</tr>
<tr>
<td>7. Whether researchers have concluded that higher degree of natural resource dependency leads to more corruption.</td>
</tr>
</tbody>
</table>

2.2.1 The effect of economic liberalization on corruption

A central aspect of political economy is state activity in the economy. A central theoretical school that treats this topic is that of neo-liberalism (Stilwell 2006). Neo-liberalism is both an ideology and an analytical tool based in these ideas (Gilpin and Gilpin 2001:14). In this context, of course, it is the analytical tool and not the normative aspects that is of interest. The market, in the neoliberal perspective, is “basically self-regulating, producing optimal (...) outcomes. (Stilwell 2006:261).

An attempt to define neo-liberal theory seems appropriate. From Przeworski (2001:15, 17) I extract the following definition: “The central claim of this perspective (...) is that the market allocates resources to all uses more efficiently than political institutions. (...) [A]ny government action that makes the equilibrium outcome diverge from the competitive equilibrium constitutes a transfer of income. (...) Transfers of income necessarily cause
inefficiency (…) [the consequence of which is] that the society as a whole suffers net losses”. An important mechanism through which government activity causes inefficiencies is government spending which again stimulates activity were people seek to influence governments to provide rents (Przeworski 2001:18). In other words neo-liberalism and the process of liberalization are two related, but clearly separate analytical concepts. My understanding of economic liberalization is the process of transforming an economic system towards being more like what an economic system should be like in the neo-liberal sense. Privatization is a highly related term: “the act of reducing the role of government, or increasing the role of the private sector in an activity or in the ownership of assets” (Savas, 1987:3 Samson 1994). The terms seem to be applied somewhat overlapping.

The conventional wisdom was for years, in accordance with the argument of the neo-liberalists, that “marketization necessarily reduces corruption” (Gong 1997:286; White 1996:150). Adamolekun, Kulemeka and Laleye (1997) describes how curbing corruption was one of the goals when liberalization was initiated in Malawi from 1991. An argument for cutbacks in the public sector is that a smaller public sector may be easier to control (Theobald 1994). According to Bardhan (Bardhan 1997b:332) the argument that liberal societies experience less corruption is valid; but if the neo-liberals are right, why has corruption exacerbated in i.e. Russia and China after the liberalization processes in the 80’s and 90’s? (Bardhan 1997b:332); It is not clear, however, that there is a contradiction between the neoliberal argument and empirical evidence that liberalization processes are associated with more corruption. It is, as Weyland points out (Weyland 1998), the state that must initiate and drive a liberalization process, so it may be that in the process of liberalization the state may actually have an increased role in the market; the rents derived through the opportunities that state intervention formerly provided, may next be provided by accumulating resources from the state assets undergoing privatization i.e. through sale of information to private actors or other manipulation of the process by which liberalization happens (Meny 1996). I’ll give a few examples of how the dimension of economic liberalization has been treated by researchers in practice:

Koyuncu, Ozturkler and Yilmaz (Koyuncu et al. 2010) conduct a panel analysis where they operationalize economic liberalization as “private sector share in total employment”, ”ratio of employment in publicly owned enterprises to total employment”
and "private sector share in GDP", respectively. Since this is a panel analysis it is the changes in these indicators over time that are interesting, and thus it is talk of liberalization processes.

Rossabi (2009) does a qualitative assessment of, amongst other dimensions, the process of economic liberalization in Mongolia and the effects thereof. Cheloukhine and King look at what they call the “wild privatization” of Russian state properties not the least of valuable natural resources (2007; also Black et al. 2000).

Findings are somewhat contradictory. Supporting the assumption of the neo-liberals, Schoenman (Schoenman 2005) finds that in Poland the economic liberalization has indeed contributed to decrease corruption levels, however, in companionship with simultaneous development of a sound democratic system. Koyuncu, Ozturkler and Yilmaz (2010) also find support to the neo-liberal claim studying a sample of transition economies.

Colazingari and Rose-Ackerman claim that states that liberalize their economy without securing an according transition towards political liberalism may experience an increase in corruption. Root (1996) finds that liberalization has lead to more corruption in China, and that this may be due to lack of control with the liberalization process.

Horowitz (2005) claims that economic liberalization processes in Latin-America have neither decreased nor increased the level of corruption. Kohl (Kohl 2004) makes similar findings studying Bolivia (Kohl 2004).

The first dependent variable of this thesis will be concerned with which effect researchers find that economic liberalization has on corruption. This will also be the foundation of descriptive analyses regarding the degree of agreement between researchers.

Dependent variable 1: Whether researchers have concluded that economic liberalization leads to more corruption.

2.2.2 The effect of democracy- versus autocracy on corruption
Dahl defines a democratic regime (polyarchic in his terms) as a regime which (1) is liberal, meaning that the chances of public contestation of policy are high, and where (2) inclusiveness is high, meaning that the opportunities to participate in political life are high. Autocratic regimes have lower degrees of public contestation and participation (Dahl 1971:7). The mechanisms needed to establish a democracy in this sense are in Dahl’s opinion the freedoms of organization and expression, the right to vote and run for office in free and fair elections, right of politicians to compete for support, the availability of sources of information
that are not those of the state, and, finally, the presence of institutions that make government decisions depend on election results and other expressions of preferences (Dahl 1971:3). In light of Dahl’s thoughts we may imagine that such a situation may raise the cost of corruption since the risk of consequences will be higher than in a system where contestation is low. Based in these arguments we may imagine that corruption may be vagrant in autocratic political systems compared to democratic systems. Amongst contemporary researchers these arguments are supported i.e. by Drury, Krieckhaus and Lusztig (Drury et al. 2006) and Pellegrini and Gerlagh (Pellegrini 2007). According to Rose-Ackerman given that the electorate disapproves of corrupt activity, politicians want for reelection will prevent them from committing corrupt acts (Rose-Ackerman 1996; Quah 2004:75).

However, applying a principal-agent model on representative democracies Groenendijk (1997) shows that corruption very well may occur at high levels also in representative democracies. This is partly due to the high cost of monitoring activities of politicians. In addition running for office is a process that demands money (i.e. Damania 2005; Easterly and Levine 1997), and fundraising may easily turn into an illicit process (de Sousa 2005; van Biezen and Kopecky 2007). Citizens have to consider the costs of inspecting politicians and preventing them from being corrupt to the prospect of achieving results. Neither is there any guarantee that the effort will bear results, and results may depend on the efforts of many people, which adds to the uncertainty of a payoff for the effort (Groenendijk 1997). Thus the problem bears similarities to a collective action problem in Olson’s (1971) sense.

Democracy operationalized in ways that comprise the essences Dahl’s definition are of interest in this thesis. An example of how this dimension is commonly operationalized is the Freedom House Indexes of political rights and civil liberties. Researchers apply one (like Lee et al. 2010; or Damania et al. 2004) or both of them added (like Sandholtz and Koetzle 2000). The Vanhanen index of party competition and voter turnout is also frequently applied (i.e. Xin and Rudel 2004; or Chowdhury 2004), as are the Polity III- (i.e. Adsera et al. 2003) and Polity IV (i.e. Calderon and Chong 2006) measures of democracy.

Ateljevic and Budak (Ateljevic and Budak 2010) do a qualitative assessment of the status of democratic institutions and processes and consequences thereof in Croatia. Sun and Johnston (2009) look at the effects of democratic institutions and culture on corruption partially drawing on conclusions from previous research regarding the status of democratic
ideas and institutions India and China, respectively. They also consider levels of democracy in the two countries based on the Polity IV data.

The conclusions regarding the effect of regime type on corruption are diverging. According to Deysine (1980) the attitude that democracy leads to less corruption stood strong in early research on causes of corruption. However, it had become clear back then too, that this hypothesis could not explain the presence or absence of corruption everywhere. Reports from newspapers, trials, public investigations and so on, bore proof of the fact that corruption also could be a severe problem in advanced western democracies (i.e. Pinto-Duschinsky 1976b in Deysine 1980). This theme has not claimed less attention over the years.

Sung (2004), applying a large sample of countries, find that democracy does indeed decrease the amount of corruption. The same do Dalgaard and Olsson (Dalgaard and Olsson 2008).

An example that democracy not necessarily checks corruption is India. Sun argues about India that it is the presence of democracy while there is absence of high economic development that makes the Indian democracy vulnerable to corruption. More economic resources would have helped develop institutions giving them more possibilities for protection against corruption; authoritarian regimes with stronger institutions may even perform better (Sun and Johnston 2009).

Singapore is also an example of opposition to the hypothesis that democracy is better than autocracy regarding the diminishing of corruption; Singapore is an authoritarian state, but one of the least corrupt countries in the world (Low 2001). Another clear opposition to the hypothesis is given by Clarke and Xu (Clarke and Xu 2004) who claim that firms are more likely to pay bribes in countries that are more democratic.

The second dependent variable of this thesis will be concerned with which effect researchers find that different levels of democracy have on corruption. This will also be the foundation of descriptive analyses regarding the degree of agreement between researchers.

Dependent variable 2: Whether researchers have concluded that democracy leads to less corruption and/or that autocracy leads to more corruption.

2.2.3 The effect of regime transitions on corruption
A regime transition is “the interval between one political regime to another [and is] delimited, on the one side, by the launching of the process of dissolution of an authoritarian regime, and
on the other, by the installation of some form of democracy, the return to some form of authoritarian rule, or the emergence of a revolutionary alternative” (O'Donnell and Schmitter 1986:6). Departing from Dahl (1971), we may imagine that processes of transition may cause unpredictability in the political environment, leading to a decrease in the cost of appropriating state resources since the rules of how offices should be conducted may be unclear or because monitoring mechanisms may not be in place. According to Sung incentives for corruption in democratizing states have their roots in the high cost of electoral campaigns, the openness of the state institutions to individual ambitions and the attractiveness of public assets (2004:181).

I will present some examples of how this explanatory dimension has been observed in the literature:

In his study of the effect of transitions in the Czech and Slovak republics, Reed partially takes the transition processes as givens (Reed 1995), which presumably is quite uncontroversial. He relies on previous research to describe the pre-transition status-quo (i.e. Simecka 1984 in Reed 1995) and moves on to describe the Czechoslovak and later Czech and Slovak efforts to establish distinct political and market spheres (Reed 1995:331).

Yu, Chen, Juang, and Hu (2008) do something along the same lines in Taiwan; they describe how Taiwan was ruled by an authoritarian regime from 1949 to 1986 when the opposition managed to form the Democratic Progressive Party which got the democratization process rolling the year after. They rely i.e. on Rigger in describing the processes of transition (2004 in Yu et al. 2008).

Also within this explanatory dimension diverging findings may be presented. Munslow, studying Angola, argues that corruption increased after transition towards multiparty democracy because there where more politicians demanding bribes (1999:558). De Sardan (de Sardan 1999) claims this has been the case in many African states after transition. Democratization has not made corruption recede; rather it has made it rise. Weyland presents a similar argument; the dispersal of power between many in stead of a few, which a typical transition from authoritarian to democratic rule implies, provides more actors needing to consent to a decision, and thus more people to bribe (Weyland 1998).

Similar findings as those of Munslow and de Sardan provide from African states, Sung provides from Argentina, the Philippines and Russia and their democratization processes in the 80’s and 90’s (Sung 2004). These countries have become “prototypes of poor governance”, claims Sung (2004).
Grzymala-Busse (2006) provides results diverging from those mentioned. Studying the recent transitions in Eastern Europe she finds that that increased political competition caused a decrease in rent-seeking. I.e. “[i]n Hungary and in Poland, attempts at rent-seeking were constrained both by their opponents’ immediate airing of any dirty laundry and by the credible threat of replacement posed by the opposition” (Grzymala-Busse 2006:432). This situation provided incentive for all parties to pass governance reforms to prevent opponents from misusing their offices. In the Czech Republic and in Romania the transitions were less successful (Grzymala-Busse, 2003 in Grzymala-Busse 2006:432).

The third dependent variable of this thesis will be concerned with which effect researchers find that regime transitions have on corruption. This will also be the foundation of descriptive analyses regarding the degree of agreement between researchers.

Dependent variable 3: Whether researchers have concluded that regime transitions lead to more corruption.

### 2.2.4 The effect of strong versus weak democratic traditions

A second dimension highly related to democracy, is the strength of and experience with democratic traditions. It has been suggested that even if the transition from an autocratic regime itself should lead to more corruption, that does not mean that democratic institutions and values in the long run may lead to less corruption. Such claims were made i.e. by renowned researchers such as Scott (1969a), Huntington (1968) and Heidenheimer (1970). Thus, even if the overall conclusion of the above section should be that transitions are bad, that does not necessarily prescribe giving up on democracy as promoting good governance (Rose-Ackerman 1999). I’ll provide some examples of how this dimension has been treated empirically:

A common operationalization of this explanatory dimension is to in some manner account for the time-span for which countries have had experience with democracy. I.e. Fan (2009) codes this explanatory dimension as a dummy variable for whether countries have been democracies for all years between 1950 and 2000. Gerring (2005) measures the cumulative years of democratic rule in the twentieth century.

Pietrzyk-Reeves (2006) studies this dimension claiming that establishing strong democratic institutions and control mechanisms is of course important, but that “a cultural and social foundation” is just as important, something the post-communist countries lack due to
their communist heritage; thus he is concerned with the absence of a democratic tradition, rather than the presence of it.

Again there are diverging findings. Studying Uruguay Calderon and Chong (2007) find that corruption decreases with democratic longevity; later democratic regimes in Uruguay have been less troubled with corruption than earlier democratic regimes. Applying a variable of how many years countries have been democratic between 1900 and 1995 Gerring and Thacker (2004) conducting a cross-national study find that it does indeed decrease the level of corruption.

Contradicting these findings, Fan, Lin and Treisman (2009) controlling for a wide range of control variables find that strong democratic traditions has no significant effect on the degree of corruption they experience. Studying India, Singh (1997) provides another example weakening the hypothesis; India has been democratic for most of the time since liberation, but is still plagued with corruption.

The fourth dependent variable of this thesis will be concerned with which effect researchers find that democratic traditions have on corruption. This will also be the foundation of descriptive analyses regarding the degree of agreement between researchers regarding this dimension.

Dependent variable 4: **Whether researchers have concluded that stronger democratic traditions lead to less corruption.**

### 2.2.5 The effect of media presence on corruption

A concept which is closely related to that of democracy, and may be seen as a sub-concept of and a vital part of democracy (Thompson 1995:238), is freedom of press and the degree of presence of a strong press. The effects on corruption of press freedom (Larmour and Barcham 2006) and/or the degree to which the press operates as a critical voice in society (Lindstedt and Naurin 2010) are often treated as an independent analytical phenomenon and is therefore also treated as such in this thesis. Drawing on liberal thinkers such as Jeremy Bentham, James Mill and John Stuart Mill, Thompson points to the importance of a free press as an organ promoting public enlightenment and through which the public may be informed of abuses of public power (Thompson 1995:238). My review will comprise studies were freedom of press is understood along such lines.
The idea of the press as anti-corruption mechanism depends on the media’s self-interest in uncovering corruption, which again is dependent on their freedom to express whatever they want and need to sell i.e. newspapers (Garoupa 1999). Allow me to provide some examples of how this dimension has been treated empirically.

An example of how the degree of presence of a free press is operationalized is the degree of newspaper circulation measured through a count of newspapers per thousand citizens (Lindstedt and Naurin 2010). Tian (2009) applies the Press Freedom Survey from Freedom House which seeks to capture the amount of influences, political and economical, on the way in which the media conduct their work. Studying Indonesia in a qualitative study, Robertson-Snape (1999) argues that the different historical regimes have had different laws regarding whether the press has been allowed to operate freely, i.e. under Suharto they couldn’t. Robertson-Snape also presents the facts that journalists have lacked experience of critical journalism and even have been inclined to receive bribes, as weaknesses of the media as a critical voice in society (2009:601).

Regarding the effect of the media, findings are contradictory. Taiwan may be an example of a country where the media has had a role of helping to decrease corruption. Fell (Fell 2005) found that the liberalization of the media may have helped reduce corruption levels as opposition parties took advantage of this new tool of uncovering corrupt activities within the ruling Chinese Nationalist Party and thus challenging and changing the norms of governing. Brunetti and Weder (Brunetti and Weder 2003) find, in a cross-national study applying the Freedom House measure of press freedom, that higher degree of press freedom does decrease corruption.

One of the most infamous corruption scandals in our time is what later came to be known as the “mani pulite” (clean hands) process. The commercialization of the new media in the years prior to the incidents probably played a crucial role. What triggered the revelations of corruption that was to delegitimize the whole post-war Italian political establishment was triggered by the media’s uncovering of a corruption incident involving local politicians in Milan (Giglioli 1996). However, despite the ousting of numbers of corrupt politicians in the “mani pulite” processes, a non-corrupt Italy proved to be much less than reality. According to Della Porta and Vannucci (Della Porta and Vannucci 2007) the “mani pulite” processes provided a window of opportunity, but politicians were either unable or unwilling to seize the moment. Italy has experienced stable, high levels or corruption. In other words the media,
even though being able to uncover corruption, did not cause a long term decrease in corruption.

In South Africa too, the media has played a clear role when it comes to uncovering corruption through investigative journalistic work. Government members like former president Mbeki (1999-2008) claimed that these scandals were a result of a press dominated by whites rather than being a reflection of real corruption incidents. Hyslop (Hyslop 2005) shows that this not the case within those news papers that actually are engaged in uncovering corruption. Thus the media have probably played a “correct” role which has resulted in revelations based journalistic work not racially motivated. Thus despite what Hyslop describes as a “relatively lively media” (Hyslop 2005:775) that does not look mildly on corruption, corruption has been “ever-growing” in South Africa (Gevisser 2009 in Lemon 2009).

The fifth dependent variable of this thesis will be concerned with which effect researchers find that a stronger and freer press has on corruption. This will also be the foundation of descriptive analyses regarding the degree of agreement between researchers regarding this dimension.

Dependent variable 6: Whether researchers have concluded that higher degree of media freedom and presence leads to less corruption.

2.2.6 The effect of level of economic development on corruption

The relation between level of economic development and democracy has claimed much attention among so called “modernization theorists”. Lipset stated of this relationship that “the more well-to-do a nation, the greater the chances that it will sustain democracy” (Lipset 1959:75). A reason for this may be seen in the argument of Brehm and Wendy that scarcity leads to distrust in other people (Brehm and Rahn 1997:1009) or lower chances for developing a middle class of a critical size or in a less educated population (Lipset 1959:102; Cheng 1989). Similar arguments have been posed concerning corruption. Scott in addition claims that less developed societies may experience more corruption because there is a low level of feeling of political efficacy in the eyes of the citizens, and thus demand expressions are weak. Low development also results in a low degree of interest aggregation (Scott 1969b:325-326), which in a collective action perspective could have been important to deter corruption. Another mechanism suggested is that the richer countries get the more resources
they may spend seeking to improve their institutions i.e. by fighting corruption. Thus the
incentives for adhering to rules may be stronger relative to the incentives for not adhering to
rules. In addition citizens may be more capable of monitoring politicians (Paldam 2002).

This dimension is very commonly operationalized as, or along such lines as, national
income per capita, in fact, it will be shown in the analyses chapter that this is the explanatory
variable most commonly applied in the literature. Level of economic development understood
along such lines is of interest in this thesis. It is also considered a highly important variable to
use as a control variable, and is thus commonly applied as such. Lambsdorff strongly draws in
doubt the finding of Sung (2004) of a curve-linear relationship between democracy and
corruption because he fails to control for income per capita (Lambsdorff 2007:40).

Serra (Serra 2006) suggests that poorer officials are more inclined to demand bribes.
This may be a consequence of poorer countries being less able to pay officials decently (Azfar
and Nelson 2007). Judge, McNatt and Xu (Judge et al. 2011) conducting a review of the
quantitative cross national studies of causes of corruption concludes that 38 out of 42 studies
in their sample have concluded that higher development is associated with less corruption.

In addition to possibly affecting corruption through its effect on democracy, economic
development may help deter corruption through leading to a more educated population and
thus higher degree of literacy, and leading to depersonalized professional relationships. These
features will increase the chance that corruption is exposed and acted against (i.e. Treisman
2000; Paldam 2002).

The weaker ability of people with limited economic resources at their disposal to
uncover and act upon corruption has also been discussed. They also may have fewer
opportunities to organize. In addition the middle class is likely to be smaller in poor countries.
Thus people will be easier to abuse (You and Khagram 2005).

However, not all find that higher economic development is associated with lower
degrees of corruption. Tian and Lo (Tian and Lo 2009) present models with diverging results
regarding economic development, some models present that there is no effect while some
provide that the effect is negative. Lambsdorff (2007) too reviews a range of studies
concluding that some of them do not observe significant relationships between level of
economic development and corruption.

The sixth dependent variable of this thesis will be concerned with which effect
researchers find that level of economic development has on corruption. This will also be the
foundation of descriptive analyses regarding the degree of agreement between researchers on
this dimension.

Dependent variable 6: Whether researchers have concluded that higher economic
development leads to less corruption.

2.2.7 The resource curse
The resource curse thesis was first laid forth by Richard Auty. He claimed that excess of
natural resources, in contrast to past assumptions, were associated with weak economic
growth (Auty 1993). Later the expression has also been associated with weaker democratic
development (Ross 2001). Natural resource dependency has also shown to positively affect
levels of corruption (Sala-i-Martin and Subramanian 2003); “[T]he richer the nation with
respect to natural resources, the poorer the quality of management (…)”, claims Goldman
(2008). The main argument is that presence of natural resources, i.e. oil, diamonds or ore, will
increase the amount of money easily available and thus increase the inclination to demand
bribes (i.e. Ades and Di Tella 1999). Natural resource dependency will in this thesis be
understood as whether and degree to which countries depend economically on such resources
described by Ades and Di Tella.

Montinola and Jackman show that states where substantial amounts of total GDP come
from natural resources corruption are more likely to be states experience a lot of corruption.
(2002). Bhattacharyya and Hodler (2010) find natural resources to increase when found
together with weak democratic institutions, but where democratic institutions are strong, large
natural resource endowments is no problem. The mechanism is presumed to be as follows:
politicians in power wish to appear clean to the electorate. For unclean politicians to appear
clean is much easier in a poor democratic system than in a strong one. In other words strong
democratic institutions will not only be an incentive to look clean, but also to be clean.
(Bhattacharyya and Hodler 2010:1).

The resource curse hypothesis has also been disputed. Some researchers have found no
relation to corruption. Despite expectations of the opposite, You and Khagram (You and
Khagram 2005) found no effect of natural resource abundance on corruption. Moller and
Skaaning (2010), studying post-communist countries, find somewhat diverging results
between different models applying different controls.
The seventh dependent variable of this thesis will be concerned with which effect researchers find that level of natural resource dependency has on corruption. This will also be the foundation of descriptive analyses regarding the degree of agreement between researchers on this dimension.

Dependent variable 8: Whether researchers have concluded that higher degree of natural resource dependency leads to more corruption.

2.3 Concluding remarks
In this chapter I have defined the concept of corruption and narrowed the scope of this thesis to high-level corruption and causes of such that may be found amongst features of the societies in which it occurs, so called demand-side causes. Further, I have presented the central arguments regarding seven features of societies that have been much discussed as possible explanations of degree of presence of corruption. I have also formulated seven dependent variables based on these; my dependent variables will be whether researchers have found that these explanatory dimensions have effect on degree of corruption or not.

In the next chapter I will in closer detail review the way in which corruption has been studied, focusing primarily on the data sources that have been applied as operationalizations of corruption as dependent variable; these data sources will found the basis of the independent variables of this thesis.
Chapter 3: Independent variables: data sources on corruption

A profound inspiration for this thesis is Treisman (Treisman 2007). When comparing results from regression analyses based on cross-national perceptions based measures of corruption, and cross-national experience based measures of corruption, respectively, Treisman shows that effects found when applying perceptions based measures where not found when applying experience based measures. He suggests the following:

“It is possible that the experience-based measures are noisier and less reliable or are measuring a different phenomenon (petty as opposed to grand corruption, perhaps). But it could also be that the widely used subjective indexes are capturing not observations of the frequency of corruption but inferences made by experts and survey respondents on the basis of conventional understandings of corruption’s causes.” (Treisman 2007:213).

In other words Treisman suggests a causal chain going from what amongst respondents to corruption perceptions surveys is considered “common knowledge” about what causes corruption, via the perceptions measure, and to the conclusions of research where these measures are applied as dependent variables. The causal chain that Treisman suggests is summed up in Figure 1.

I want to see how results derived applying these two categories of measures compare to each other and how the different perceptions based measures compare to each other. I also want to see how these results compare to empirical research on causes of corruption that does not rely on neither perceptions based nor experience based cross national measures, but rather has a narrower in depth focus; these I will denote country- or region-specific studies. The bulk of this “other” research it would be correct to term qualitative studies, however, there are also some studies that have narrow geographical foci, but fully or to some extent apply quantitative approaches.

In the following section I will present central arguments regarding methods in the social sciences, discussing the controversy we know well in social science in general between advocates of quantitative methods and qualitative methods respectively. More importantly, I will seek to describe how the data and methods controversy of corruption research fits into the general methods-debate and that it diverges slightly in that the main line of conflict seems to be between quantitative cross-national studies and “the rest” and not between quantitative and
qualitative methods per se. Then I will present and describe the data sources available for research on causes on corruption, starting with the different perceptions based measures, then the different experience based measures and lastly the range of data sources available for country- or region-specific studies. Thereafter I will present central arguments regarding the strengths and weaknesses of the different data sources, focusing on possibilities of unsystematic and systematic measurement errors and effects of such; this will form the basis for formulating seven hypotheses of possible effects application of the different data sources may have on the results derived when applying them.

Figure 1: Causal diagram of Treisman's suggestion

3.1 Qualitative versus quantitative methods
Discussions of applications of different methods are well known the social sciences; “The content” of social science research “is the method” claim King Keohane and Verba (1994). Still, there is great disagreement regarding what this content should be in terms of which methods to apply. Methodologically, the social sciences may be seen as two traditions, the quantitative tradition and the qualitative tradition (King et al. 1994:3). The debate between the different methodological traditions of the social sciences has at times been almost warlike (King et al. 1994:3-4). Through quantitative methods, researchers commonly seek to gather and analyze information about many units, while applying the qualitative approaches involves gathering in depth information about one or a few units (King et al. 1994:3). By definition, the wide range of approaches we associate with qualitative methods usually does not involve quantifications (King et al. 1994:4). The goal of both approaches is scientific inference, in other words “using the facts we know to learn about facts we do not know” (King et al.
That qualitative methods are better when it comes to “digging deep”, is widely recognized (i.e. McKeown 1999; George and Bennett 2005) and also by King et al. (King et al. 1994). Still, quantitative methods have been developed to become highly flexible, so that most issues may be formulated in quantitative language or answered by analyses through quantitative methods (Porter 1995:5-6). A central conflict is, however, if one of the approaches are better when seeking to draw inferences. This is easier to achieve applying quantitative methods claim i.e. King et al. (1994:229), since more facts make it easier to generalize (King et al. 1994:46; also Lijphart 1971; and Landman 2002).

Opponents to this stand claim the attempt to achieve inferences based i.e. on high correlations between phenomena is done at the cost of necessary depth (Pierson 2003). Quantitative methods also, in spite of the higher sophistication they have achieved in later years, provide less opportunities to incorporate context specific explanatory dimensions (or Skocpol 2003; i.e. McKeown 1999). The importance of the latter argument is underlined by Elster’s “plea for mechanisms” (Elster 1998: the title). He argues for studying phenomena at as low levels as possible, so to see their individual components and their immediate causes and consequences and that these mechanisms may ”form building blocks” in explanations of greater complexity (Elster 1998:47). He even claims that there is no room for “law like generalizations” such as those proclaimed by the quantitative macro-camp (Elster 1998:45). An example of this way of studying social phenomena is the method of process tracing, where one seeks to identify causal chains between predictor- and outcome variables with a focus on the details in these processes that i.e. is good for checking spuriousness (George and Bennett 2005:206, 223).

However, the debate also has been influenced by those who want the best from both traditions. Johnson (2004) names this approach “Mixed methods research”. The main argument is that there are many ways of seeking causal inference, and the method that can provide the strongest inferences is a combination of several (Mahoney 2003; Johnson 2004.; Mahoney 1999; Lieberman 2005; Onwuegbuzie 2006).

Compared to what may be seen as the dividing line of conflict between the different “camps of methodologies” in the social sciences in general, in corruption research the dividing line seems not to be between quantitative an qualitative methods but rather quantitative methods applied cross nationally on one side and qualitative and quantitative methods applied in country- and region specific studies on the other (following i.e. Philip
However, it also clearly has much in common with the general debate, since the appearance of cross national measures must be seen in relation to a want to study the phenomenon cross nationally and quantitatively.

Of the early research on corruption, most studies were case studies, which was the “motif in political science studies of corruption” (Gerring and Thacker 2004:299). Researchers, in particular economists, also applied formal modeling (Gerring and Thacker 2004:299). Commenting on early research on corruption Liu (1983:152) claimed that the research was an inconsistent body of studies only loosely attached to theory, and that this was underlined by the insufficient availability of data about changes in amounts of corruption. The lack of usable sources of data on corruption was according to some of the supporters of quantitative methods an enormous problem, and some even suggest it was the reason why the topic was neglected for so long by American political scientists (Williams 1987 in Sandholtz and Koetzle 2000:33). The reason why quantitative data on corruption were so hard to obtain is easy to grasp; no one involved in an act of corruption has any interest in revealing it to the public. Furthermore, the victims of corruption may not even know that they are victims (Sandholtz and Koetzle 2000:33; Goetz 2007:93). The research community, at least segments of it, seemed to believe that the problem was solved with the arrival of the cross national surveys of “corruption perceptions”. Ades and diTella claimed that “The availability of subjective data on corruption has finally provided the field of the economics of corruption with the empirical discipline that is essential to turn its fertile theorizing into policy recommendations” (Ades and DiTella 1997b:514). A considerable amount of studies were published where these measures were applied (Gerring and Thacker 2004), and they’re still coming. Thus one should believe that the need of data would provide opportunities to study causes of corruption through a range of methods including quantitative cross-national approaches, as prescribed by those claiming quantitative approaches or mixed-methods approaches may yield stronger inferences. However, as will be discussed below, not all agree.

The remainder of this chapter will depart from delineation between cross-national quantitative studies, on the one hand, and country- or region specific studies on the other.
3.2 Data and research on corruption

In the following I will start with descriptions of the three most commonly applied perceptions based cross-national data. Then I will describe two less commonly applied perceptions based measures. Following, I will present three experience based cross national measures of corruption. Thereafter, I will present a range of sources available for country- and region specific studies; media-reports, public or internal inquiries, judicial documents, questionnaires and interviews, prior research and source triangulations. Finally I will discuss objections to validity and reliability and possible measurement errors which will form the basis of formulating a series of hypotheses of relationships between measurement error and results when applying these measures as dependent variables.

3.2.1 Perceptions based measures

*International Country Risk Guide (ICRG)*

The first measure to be relatively widely applied as the dependent variable in cross national studies of the causes of corruption was the *International Country Risk Guide* (ICRG), which has been produced since the early 80's. The measure is published by the Political Risk Services Group, and provides a cross country measure of the perceived level of corruption within a country «over a moving average of two-year periods” (PRS-group in Escaleras et al. 2010:440-441). The judgment as to how corrupt the country is made by resident and non-resident experts on the situation in each country. In constructing the numbers for 2009 the following sources were used: «Asian Development Bank, African Development Bank, Bertelsmann Transformation Index, Country Policy and Institutional Assessment, Economist Intelligence Unit, Freedom House, Global Insight, and Merchant International Group (…), Political and Economic Risk Consultancy and the World Economic Forum.” (Escaleras et al. 2010:440-441). The index is a scale from 0 to 6 where higher values indicate less corruption (Dreher and Schneider 2010:221). The index is intended to capture the degree to which “high

---

1 In describing the ICRG the access to information is somewhat limited, since complete access must be bought for more than I can afford. Hence, I do not have access to primary sources of the methodology of the measure, and must rely on what researchers have written about the measure. The ICRG corruption measure must not be confused with ICRGs Quality of Governance Index, of which the corruption index is only one of three components (Brautigam and Knack 2004). After email correspondence with the PRS-Group it has been confirmed that details of the methodology and sources are not publicly available.
government officials are likely to demand special payments,” and “illegal payments are generally expected throughout lower levels of government” in the form of “bribes connected with import and export licenses, exchange controls, tax assessment, police protection, or loans.” (ICRG in Brautigam and Knack 2004:279).

An advantage of this measure is that it has been measured annually since 1982 and thus allows for panel studies (Alesina and Weder 2002). The backside of it is that it is only available to paying customers.

The Corruption Perceptions Index

The Corruption Perceptions Index (CPI) published by Transparency International (TI) is a poll of polls made up of several surveys from several different institutions (Escaleras et al. 2010). The index “ranks countries in terms of the degree to which corruption is perceived to exist among public officials and politicians.” (Transparency International 2010a:1). The index is composed of a series of surveys conducted by a series of well reputed organizations. The argument for applying this wide array of surveys in constructing the index is that it strengthens the reliability. The index ranges from 0 (most corruption) to 10 (least corruption) (Transparency International 2010a).

A criterion for being included in the index is that the surveys seek to measure perceptions of the “overall extent of corruption”, both regarding the size of transactions and how often such events occur amongst politicians and state officials, and that the different countries are given a rating. The surveys provide either the perceptions of business people or groups of experts i.e. academics. Business people responses have been averaged over two years to increase the stability of the index, this has not been considered necessary when dealing with expert assessments. For a country to be included in the survey, at least three independent surveys must be available for that country. Each survey must have been conducted in more than one country with the same methodology in each country (Transparency International 2010a). The methodology of the index has been adjusted twice since it was first published (Transparency International 2010a), something that makes it unsuitable for use in panel studies (Lambsdorff 2007).

How many sources are included in the index varies with their availability (Transparency International 2010a), i.e. in 1998 it consisted of 12 surveys from 7 institutions (Persson et al. 2003).
The Control of Corruption Index

The Control of corruption index (CCI) was developed by The World Bank’s researchers whom were highly critical to the quality of the TI-measure. The main difference in what these researchers did from the TI-researchers was including more sub-indexes from “reliable commercial country risk assessment organizations and non-governmental organizations.” (Andvig, Fjeldstad, Amundsen, Sissener and Søreide 2000:46). Like the CPI, the CCI is a poll of polls, building on a series of surveys. Several of them are also applied in the CPI i.e. surveys from the Asian Development Bank, the African Development Bank, Bertelsmann Transformation Index, Freedom House, Global Insight Global Risk Service Expert\(^4\) The aim of the CCI is similar to that of the CPI: to capture the «perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as "capture" of the state by elites and private interests.» (Kaufmann 2010:4).

This index is one of the World Banks World Governance Indicators and was published biannually from 1996 and annually from 2002 (World Bank 2010d). The index ranges from -2.5 to 2.5 with higher values indicating less corruption (World Bank 2010c).

World competitiveness report

The corruption index of the World Competitiveness Report (WCR) measures the “extent to which improper practices (such as bribing and corruption) prevail in the public sector” (Fisman and Gatti 2002:343). The measure ranges from 0 (less corruption) to 100 (more corruption) (EMF foundation in Fisman and Gatti 2002). Respondents of the survey are top and middle managers in the countries in question (Ades and Di Tella 1999). According to Ades and DiTella (1999; 1997a) it was published in 1989, 1990, 1991 and 1992. Based on the information I have on the measure it would be correct to classify this as a perceptions based measure. The wording changed slightly between the years from “extent to which the country prevents corruption” (1989) to “extent to which government regulations prevent improper practices in the public sphere” (1990) and the extent to which “improper practices (such as bribing and corruption) prevail in the public sphere” (1991-1992) (Ades and DiTella 1997a:1031). This may have had an effect on the results between the years, however, it would still be right to consider the measure a perceptions based measure.

\(^4\) For a complete list of surveys included in the CCI see the website of the Worldwide Governance Indicators (World Bank 2010c).
Global competitiveness survey

The Global Competitiveness Survey (GCS) is published by the World Economic Forum. The respondents are the top leadership from approximately 3000 companies. They are asked to rate “their perception of presence of irregular, additional payments connected with import and export permits, business licenses, exchange controls, tax assessments, police protection or loan applications in their countries” on a scale from 1 (more) to 7 (less) (Fisman and Gatti 2002). In other words, this is a perceptions based measure.

3.2.2 Experience based measures

German exporter index

The German Exporter Index (GEI) is an index based on interviews with German exporters. On average ten, and a minimum of three, respondents with business experience from the country in question are interviewed for each country included in the survey, and are asked to judge the "[t]otal proportion of deals involving kickbacks" (Neumann 1994 in Fisman and Gatti 2002:342). The index ranges from 0 (less) to 10 (more) (Fisman and Gatti 2002:342). The respondents were given a promise of confidentiality (Ades and DiTella 1997a). A strength of the index is that all the respondents come from the same country and thus may to a larger degree have a common understanding of what corruption is.

The Business International measure

Business International Corporation (BI), which was later to be merged with The Economist Intelligence Unit, in the 1980’s an early 1990’s published an experience based measure of corruption. BI surveyed business people stationed abroad (I presume that means outside the US) concerning corruption, amongst other issues (Quinn 2008). Since I have not managed to encounter the original methodology, I have to rely on secondary references: “The Business International index of corruption is based on asking businessmen in different countries, on an ordinal scale, the extent of bribery they face.” (Laffont and N'Guessan 1999:289). Ades and Di Tella present the survey slightly different: "The degrees to which business transactions involve corruption or questionable payments" (Ades and Di Tella 1999). Not viewing the exact wording of the original questions, it still seems appropriate to consider this an
experience base measure. The measure rates countries on a scale from 0 (less) to 10 (more). The Index is available for 68 countries (Fisman and Gatti 2002).

Like The German Exporter Index this index has the advantage that the people surveyed come from the same country (the US), and thus may provide slightly more comparable responses (Ades and DiTella 1997b:498).

World Business Environment Survey

The World Business Environment Survey (WBES) is a survey of managers and owners of over 10,000 companies in 80 countries and one territory that was conducted in 1999-2000. It is based on face-to-face interviews (World Bank 2011c). The survey presents two relevant measures that concern the frequency and amount of bribes paid to government officials (Lee et al. 2010). One of them measures the incident of corruption among business respondents, while the other measures the total amount of money spent on bribes. Since both are operationalizations of corruption, I will include both of them in my analysis. The respondents are presented with one statement and one question: "It Is Common for Firms in My Line of Business to Have to Pay Some Irregular 'Additional Payments' to Get Things Done"; "On Average, What Percentage of Revenues Do Firms Like Yours Typically Pay Per Annum in Unofficial Payments to Public Officials?" (World Bank 2000). The formulations “firms like yours” and “in my line of business” are meant to reduce possible effects of self-incrimination but at the same time aims at capturing the respondents personal experience (Treisman 2007:214).

3.2.3 Data sources available for country- and region-specific studies

Tanzi (Tanzi 1998) lists a series of data sources available for narrow-focus studies: Media reports, internal or public inquiries, different sorts of judicial documents, such as reported crimes or convictions, and finally questionnaire or interview based data. Many researchers also seem to rely on publications by other researchers. These data sources will in turn be elaborated on in the following paragraphs.


6 Response alternatives: " (1) 0%, (2) greater than 0 and less than 1%, (3) 1–1.99%, (4) 2–9.99%, (5) 10–12%, (6) 13–25%, (7) over 25%" (Fan et al. 2009:22).
Media-reports

Corruption reports from well reputed newspapers and magazines, such as Le Monde, the Economist, The Financial Times, and The New York Times etc. may be valuable in uncovering single episodes or the extent of corruption (Tanzi 1998). A great advantage of journalistic sources is the chance journalists may have in many countries to protect their sources (Kvam 1995:21-22). Police or varying control bodies do not to the same extent have this option (Giertsen 2008:212). This advantage of confidentiality may also be used by people whom may know of offenders to inform the public without personal risk. Another strength of the media is their great freedom to choose how they work; there are few procedural demands regarding how information is acquired (Kvam 1995:22). Allow me to provide some examples from the literature:

Eker (Eker 1981:180) surveys reports of corruption in daily newspapers in Nigeria from 1970 to 1979. He finds that the number of reported cases varies along with important changes in society: the reconstruction years following the civil war (1970-1973) with decreasing reported corruption, a period of high increase in government expenditure and inflation hard times in the agriculture etc. (1974-1976), with increasing reported corruption, and finally a period with political stability and relative control on expenditure and inflation, with decreasing reported corruption. It must not be pointed out that these data say nothing about whether the people accused were ever convicted or if they had misused their offices at all, nor do they say anything about the amount of corruption not uncovered.

Schwartz’ also primarily relies on reports of corruption published in the press in the Soviet Union (Schwartz 1979:426). Of course, as Schwartz also points out, relying on media reports from the Soviet era is very problematic due to censorship.

Public or internal inquiries

One may apply data from inquiries in single organizations such as tax administrations or other public organizations or bodies, or public inquires i.e. at state level. The problem regarding internal inquiries is that they may not be publicly available (Tanzi 1998:577). Public inquires typically are publicly available. These are probably more commonly applied in corruption research, thus the focus in these paragraphs will be on such.

---

7. It must be said that Eker also relies on other sources. I only wish to give an example of how media based data have been applied.
The perhaps foremost example of a corruption-related public inquiry is the work done by the Parliamentary Committee of Inquiry on the Mafia (1993) which was related to the “Mani pulite” processes in Italy. These documents were one of the data sources for Della Porta (2001:8-9). I.e. the committee was able to draw very intimate information from insiders like ex-Mafioso Leonardo Messina: “There are magistrates very close to Cosa Nostra [the Mafioso name of the Mafia]. In my own province I have never heard of any magistrates who actually belong to Cosa Nostra, but there are magistrates who are very close to it” (Parliamentary Committee of Inquiry on the Mafia 1993:56 in Della Porta 2001:8-9). Information of how many judges may have had relations to the mafia and how close to the mafia they might have been, if it is true, is information that probably could not be captured by any cross national measure of corruption, at least not any of those that exist today. However, to gather detailed information of such circumstances in a whole country, not to speak of comparable data for many countries, may never be possible.

James (2010:437), studying Thailand, provides an impression of the levels of corruption by referring to reports of the Counter Corruption Commission (now the National Counter-Corruption Commission). These are based on investigations of possible corruption incidents initiated after being brought to attention by government departments or individuals. Based on these investigations the Commission provides an estimate of how much money has been misused. Their numbers indicate that there was an increase in misused money from 317.9 million baht ($13 million) in 1986 to 1.26 billion baht ($50 million) in 1992.” (James 2010:437). James compares these numbers to numbers of the Office of the Auditor-General which has the more routine task of checking ministry accounts. Here, there is an increase in misuse “between 1987 and 1990 from 108.7 million baht ($4.3 million) to 175.5million baht ($7 million)” (James 2010:437). In other words, both sources show an increase, however, the sizes of both the initial numbers and the respective increases, are clearly diverging.

Judicial documents

The source of data that may strike one as most relevant is how many or how severe corruption-related crimes have been committed. Thus data on convictions may be thought to be a good source. However, researchers have realized that such data primarily may be applied for studies within single countries (Goel and Nelson 1998).
Flanary and Watt (Flanary and Watt 1999:527), in a study of corruption in Uganda, apply data from the National Fraud Squad. They present exact numbers of corruption, embezzlement and bribery offences reported, charged and leading to conviction, respectively. Presenting such nuanced data; not only conviction or reported incident may give a more complete image. However, still we know little of how large proportion of the actual amount of corruption is reported or leads to conviction.

Johnston (Johnston 1983:30) does something similar. He counts the number of corruption conviction in 85 judicial districts in the US 1976-1978. Wedeman (Wedeman 2008:10) does something of the same in China; he reports the number of economic crimes filed.

Kneen (Kneen 2000:350) analyses amount of reported bribery cases in a Moscow court, and showing that the number increased from 421 in 1998 to 493 in 1999. An objection here is to the content validity; bribery is a much narrower concept than corruption.

Again the “Mani pulite” processes provides an extraordinary case regarding the amount of information available. By 1998 4000 people had been investigated, the prosecutor had requested that 2970 were charged, the investigative judge had charged 1063, and 438 had been convicted (Della Porta 2001:12). Still, these numbers clearly illustrates the weaknesses of relying on data from the judicial. Whether you rely on data from investigations, prosecutions or convictions will affect the number of cases, and presumably also the severity of the cases a researcher operates with. In addition, different prosecuting authorities may have different traditions regarding how many they want to prosecute, and different investigative judges may have different norms regarding when they press charges, even if, as we may hope, the number of convictions would be pretty much the same between different judges and jurisdictions. However, comparison between countries is probably out of the question.

Questionnaires and interviews
Tanzi also mentions questionnaires as an option for collecting data for country- or region-specific research (Tanzi 1998:577). These can be related to single organizations or political bodies or to whole countries. The same can be said about interviews. However, a challenge in these to regards is that often corruption cannot be studied without at the same time being revealed, and thus is clearly unattractive for most who are involved in illegal activities to
respond to such queries (Blundo 2007:30). Interviewees also may provide problems with inaccurate or selective memory (Treisman 2007:216).

Prior research

Many researchers also seem to rely heavily on prior research and conclusions drawn by colleagues. They do not always seem to consider the validity or reliability of these data. Let me provide some examples:

Matti (Matti 2010:403) states about the Mobutu regime in the Democratic Republic of Congo that “[t]his network was permeated by with corruption at all levels. It is estimated that while in power Mobutu and his close associates plundered between $4 billion and $10 billion of the country’s wealth.” This conclusion is derived by referring to Dunn which has concluded that “[i]t has been estimated that Mobutu and his close friends pillaged between U.S.$4 billion and U.S. $10 billion of the country's wealth” (Dunn 2002:53). This conclusion is drawn with referral to Collins (Collins 1997:278) who claims that “As recently as mid-1996, ex-Zaire was seemingly mired in an endlessly prolonged transition to democracy, yet another chapter in Mobutu's adept maneuvering to retain the privileges and stolen billions (variously estimated at between $4bn and $10bn) he and his close cronies had pilfered over 32 years of dictatorship” without referring to anyone. Thus, our possibilities of controlling these claims are somewhat reduced. In addition it must be pointed out that these numbers are highly diverging with a difference of 6 billion. Such long reference chains clearly are difficult and time-consuming to evaluate the quality of.

Another example is Sun who claims that “[e]conomic scarcities rooted in China’s central planning did create incentives for bureaucratic abuse in the Spartan era of Mao Zedong. But corruption was relatively insignificant compared with the reform era” (Sun and Johnston 2009:8). This is done by referring to a book section by Julie Kwong (1997), and an article by Alan Liu (1983). The former I did not gain access, the latter has collected primary data based media reports. Thus, even if relying on media reports is involved with issues of reliability and validity, it is easier to assess these issues in the case of Suns article, where we have the chance to find the primary source, as opposed to the article of Matti referred to in the previous paragraph.

---

8 Matti also relies on several other data sources. This source is presented only to illustrate this type of data source.
Anyhow, my point is to illustrate that researchers often rely on the work of other researchers as data for their own work.

Source triangulations

In country- or region-specific studies it is probably more common than not to rely on a series of data sources. An example of this is the UN-report “Report of the Panel of Experts on the Illegal Exploitation of Natural Resources and Other Forms of Wealth of the Democratic Republic of Congo” (UN 2001). This report has relied on a series of sources: Primary data collection: “Official documentation from ministries and other institutions as well as recorded minutes of meetings involving various relevant actors”; Secondary sources: “Reports, workshop proceedings, published and unpublished literature”; Interviews: “Structured, semi-structured and open interviews as well as interviews resulting from various network referrals.” (UN 2001: paragraph 9). Clearly, this gives the researchers much greater opportunities to consider the reliability and validity of a range of sources, and thus they may be able to draw stronger conclusions.

3.3 Measurement errors: formulating hypotheses

Several problems can be identified with all of these approaches. In the following sections I will present some criticisms. The “methods debate” in research on causes of corruption seems to be much focused on the issue of which data are appropriate to rely on to draw inferences (following i.e. Philip 2006). In this section I will present the main arguments for and against application of the different categories of data. Measurement error is a central issue. This discussion will form the basis of the formulation of seven hypotheses.

Measurement errors are commonly categorized into unsystematic measurement errors and systematic measurement errors (King et al. 1994), or classical and non-classical measurement errors (Chen 2007), or white noise and black noise (Bertrand and Mullainathan 2001). Unsystematic measurement errors are much less problematic that systematic measurement errors; the consequence of the former is only less efficient estimates (King et al. 1994:158; Midtbø 2007a:88), while the latter may provide us with estimates that simply are wrong (King et al. 1994:156). Bound states that the effects of measurement errors may range from simply providing less efficient estimates to ”situations where (…) observed data exhibit relationships that are not present in the error free data; and (…) even the signs of the
estimated coefficients are reversed.” (Bound et al. 2001:1-2). In addition it may be very hard or impossible to determine the sign of systematic measurement errors (Drosg 2007:15). In the following sections I will first discuss problems of unsystematic measurement errors in the different data sources. Then I will move on to discuss possibilities of the presence of systematic measurement errors. These discussions will found the basis of the formulation of seven hypotheses which are summed up in table 2.

Table 2: Hypotheses presented in this chapter

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Operationalization of corruption</th>
<th>Expected effect on relationship between corruption and variables of Chapter 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a:</td>
<td>Cross-national survey measure</td>
<td>Lower likelihood for observing effects</td>
</tr>
<tr>
<td>H1b:</td>
<td>Cross-national survey measure</td>
<td>Not lower likelihood for observing effects</td>
</tr>
<tr>
<td>H2:</td>
<td>Perceptions based measure</td>
<td>Lower likelihood for observing effects</td>
</tr>
<tr>
<td>H3:</td>
<td>Experience based measure</td>
<td>Lower likelihood for observing effects</td>
</tr>
<tr>
<td>H4:</td>
<td>Cross-national survey measure</td>
<td>Higher likelihood for observing effects</td>
</tr>
<tr>
<td>H5:</td>
<td>Perceptions based measure</td>
<td>Higher likelihood for observing effects</td>
</tr>
<tr>
<td>H6:</td>
<td>Experience based measure</td>
<td>Higher likelihood for observing effects</td>
</tr>
</tbody>
</table>

3.3.1 Possible classical measurement errors in corruption data

The classical measurement error-assumption is that measurement errors of a variable is not related to the true value of the variable, nor are they related to observed or true values of the other variables that may be of interest in our analyses, nor the measurement errors of these; in other words that they are unsystematic (Chen 2007:2). Such measurement errors even each other out; one respondent may over-report, but someone else may under-report and thus bring balance to the measure. This is relatively unproblematic; it would just provide us with less precise answers.

A problem of the cross national measures is how the concept of corruption is commonly defined amongst these measures; they are so vague and general that they become empirically ungraspable. I.e. the amount of corruption observed, depends on how “amount” is defined; does it mean many transactions or large transactions? (i.e. Bardhan 1997a)⁹. These two features may differ quite substantially (Nye 1967). There are in addition many different types of corruption, and a vague formulation clearly opens up for personal interpretations. This may be a source of noise in the data.

⁹ Here, the World Business Environment survey provides an exception since it presents one measure of incident and one of amount (World Bank 2000).
In comparison, the data applied in area-specific studies may provide greater chances to secure the content validity, thus the content validity of these sources may be stronger than that of cross-national measures. Firstly, when studying single countries researchers are not forced to rely on general definitions of corruption that may not capture the content of the concept in country-specific contexts (Tsai 2009). This may give them less erred data. According to many this is “the way” to study corruption, because trends and understandings of corruption are country specific. Some claim it could be studied within regions (Becker et al. 2009). Greater inclination of considering different sources of information is clearly important when considering the quality of the data in country- or region-specific studies.

Based in the classical measurement assumption and the argument that area-specific studies are a way around this I formulate H1a:

H1a: Operationalizing the dependent variable, corruption, as a cross-national survey measure gives lower likelihood for observing effects compared to country- or region-specific studies.

However, there may also be reason to assume that data sources applied in area-specific studies are not much better. When it comes to trying to say something about the amount of corruption, other than claiming that it exists, these data sources may come into problems similar as those of the cross national measures. One cannot presume that media reports, conviction rates or public inquiries may present numbers representative of real level corruption. According to Lambsdorff the cross-national perceptions based measures perform better in this regard; media reports, conviction data etc, say nothing of the overall prevalence of corruption; the judgment of experts may (Lambsdorff 2006).

Documentation of convictions, investigations and such clearly may provide correct claims about the corruption incidents they concern, but they do clearly not provide information about other corruption incidents and may thus to a larger extent reflect the capacity of the police and courts and such measures may suffer from severe bias (Charron and Lapuente 2010; Ades and Tella 1996:6).

An objection to the issues of lack of clarity of what corruption is when applying the cross-national measures is presented by Andvig and Moene (1990 in Andvig et al. 2000). They construct a multiple equilibrium model suggesting that bribe sizes and bribe frequency will be highly correlated, so this problem may be less serious for the perceptions indexes than
many have suggested. Lambsdorff (1999) from this considers it likely that the CPI measures both concepts and that the index captures the most important sides of corruption even if concepts and questions are vaguely formulated (in Andvig et al. 2000). If this is valid for the CPI, it would be valid also for the other perceptions based measures since they in this sense are very similar.

When it comes to classical measurement errors a solution is repeated measurement (Hu 2007:2), which may both help us determine whether there is a measurement problem, and provide more reliable results with the greater amount of data. Comparing data collected at different points in time, checking stability (following Grønmo 2004), and comparing different data based on the same theoretical concept, checking the equivalence (following Grønmo 2004), are common ways of testing the quality of the data. Tests of the stabilities of the CPI and the CCI are presented in tables 3 and 4 respectively. These show that the reliability of both measures is high. For the remaining measures data are not available for such test. However, I may test the equivalence by referring to such tests done by other researchers. These are presented in tables 5 and 6 respectively. All these correlations give reason to argue that the data have high degree of reliability, with the lowest correlation being between the CCI and the WBES (74.4 percent), and with most closing on or being above 90 percent. The lowest correlation between two perceptions based measures is between the WCR and the ICRG at 82 percent\(^{10}\). This seems quite convincing and many do agree on such a conclusion (i.e. Treisman 2000:411; and Arvate et al. 2010). A conclusion like Arvate’s that i.e. “the choice of TI [CPI] should not be a major problem” (Arvate et al. 2010:1015) seems justified.

The data from the ICRG are not publicly available, so I can’t run stability tests myself. Alesina reports that there is very little variation over time in the ICRG index (2002), something i.e. Escaleras, Lin and Register (2010) takes as a sign of high stability.

Data sources available for country-or region-specific studies also provide clear problems regarding validity and reliability. Media reports may be affected by journalists perceptions of the severity of issues with corruption, their interest in covering the topic; that media tend to jump onto cases if they have been covered by other journalists, is not an unknown phenomenon (Midtbø 2007b). This justifies skepticism towards the degree to which media reports may provide a measure that is stable over time. If a corruption scandal should

\(^{10}\) Since I do not have access to primary sources of many of these measures, I must rely on correlations presented by others. I.e. the data from the ICRG is not publicly available so I may i.e. not test the correlation between the CCI and the ICRG and have no secondary source on this correlation.
occur this may affect the attention the topic receives and thus give a rise in media reports both in numbers and in the severity they report without necessarily reflecting an increase in overall incident or severity of incidents. A change in focus from politicians, i.e. in rhetoric or through passing of legislation, may also affect media reports without being reflections of actual changes (Scarrow 2007:201-202). The same skepticism is to some extent also justified towards documents from public inquiries; however, the problem here is rather that they are likely to cover certain incidents or shorter periods of time; thus data based on media reports or public inquiries may not be applicable for establishing general tendencies of corruption, perhaps not even within the same country.

<table>
<thead>
<tr>
<th>Table 3: Correlations between different years of the CPI</th>
<th>Table 4: Correlations between different years of the CCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td></td>
</tr>
<tr>
<td>2001 0.949</td>
<td></td>
</tr>
<tr>
<td>2002 0.949 0.991</td>
<td></td>
</tr>
<tr>
<td>2003 0.952 0.990 0.992</td>
<td></td>
</tr>
<tr>
<td>2004 0.954 0.981 0.980 0.990</td>
<td></td>
</tr>
<tr>
<td>2005 0.952 0.978 0.975 0.986 0.994</td>
<td></td>
</tr>
<tr>
<td>2006 0.956 0.966 0.962 0.975 0.984 0.991</td>
<td></td>
</tr>
<tr>
<td>2007 0.961 0.963 0.960 0.967 0.978 0.958 0.988</td>
<td></td>
</tr>
</tbody>
</table>

Source: Norris (2009)

Table 5: Correlations between different corruption indexes

<table>
<thead>
<tr>
<th>ICRG</th>
<th>WCR</th>
<th>GEI</th>
<th>CPI</th>
<th>BI</th>
<th>GCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICRG</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WCR</td>
<td>.82</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GEI</td>
<td>.88</td>
<td>.88</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPI</td>
<td>.95</td>
<td>.89</td>
<td>.92</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td>.86</td>
<td>.83</td>
<td>.89</td>
<td>.95</td>
<td>1</td>
</tr>
<tr>
<td>GCS</td>
<td>.91</td>
<td>.93</td>
<td>.91</td>
<td>.95</td>
<td>.88</td>
</tr>
</tbody>
</table>


Data drawn from the judicial sector may be less problematic than both media reports and data from inquiries as far as stability over time within the same country is concerned. This claim, however, presupposes that the capacity and ability of police, prosecutors and courts are stable over time, something that very well may not be the case (Butt 2009). If there are changes in the capacities or foci of the police, the prosecuting authority or in the courts this clearly may affect the stability of the data. However, it would still be hard to determine whether the observed corruption levels are representative of the actual levels of corruption.
Based in the arguments that data applied in area-specific studies may be just as problematic as cross-national data, I formulate hypothesis 2:

H1b: *Operationalizing the dependent variable, corruption, as a cross-national survey measure does not give lower likelihood for observing effects compared to country- or region-specific studies.*

### 3.3.2 Possible divergence between experience based and perceptions based data

Even if the different quantitative cross national data are highly correlated with each other, Treisman (2007), as already mentioned does find diverging results between the two. Thus it seems that those 10-30 percentages that measures are *not* correlated may be more severe than it may appear at first sight.

An important difference between the perceptions based cross-national data and the experience based cross-national data are that the respondents are not asked to provide an image of their own personal experience; perceptions may not capture the underlying phenomenon (i.e. Bardhan 2006; or Andvig et al. 2000). Also one of the “fathers” of the CCI admits that this may be a justified concern with perceptions based data (Kaufmann 2010). It seems appropriate to ask whether it is possible to know the answer to questions about the extent of something that per definition is hidden. The experience based measures may be better in this sense, since a personal experience may be considered closer to reality than a mere perception (Russell 1948).

Based in these arguments it seems appropriate to question whether the perceptions based measures may be noisier than the experience based measures, and thus formulate Hypothesis 2:

H2: *Operationalizing the dependent variable, corruption, as a cross-national perceptions based measure gives lower likelihood for observing effects compared to country- or region-specific studies and operationalization through an experience based cross-national measure.*

However, some additional arguments may be posed countering this hypothesis. The fact that the most commonly applied perceptions based measures are polls of polls presumably should contribute to evening out the measurement errors, thus making measurement errors a smaller problem. In addition the World Bank researchers have changed the weighting of their various sources from just taking the simple average so that the sources that provide results diverging
from the rest are given less weight in the construction of the index. This is done to minimize the effects of measurement errors (Kaufmann et al. 1999 and 1999a in Andvig et al. 2000:47). The CPI has gone through similar changes (Lambsdorff 2004 in Connelly and Ones 2008). How specifically the ICRG is constructed, again, is unfortunately not public.

Rubin argues that any attempt at constructing an objective measure, in other words a measure based on actual experiences, will face problems of non-response and thus be as problematic as the perceptions based measures (Rubin and Whitford 2008). Treisman (2007) shows that analyzing causes of corruption applying perceptions based measures of corruption provide clearly more significant effects than when running the same models with experience based measures as the dependent variables. As will be elaborated below, and as Treisman himself argues (2007:228), this may indicate systematic measurement errors in the perceptions based data. However, it could also be that Treisman’s findings occur due to unsystematic measurement errors in the experience based measures.

These arguments, and the empirical evidence provided may point in the direction that the perceptions based measures may be stronger than the experience based measures.

Formulating Hypothesis 3 seems justified.

**H3:** Operationalizing the dependent variable, corruption, as a cross-national experience based measure gives lower likelihood for observing effects compared to operationalization through perceptions based cross-national data and country- or region-specific studies, respectively.

### 3.3.3 Non-classical measurement error

Most frequently, measurement error in a given variable is assumed to be independent of the true level of that and all other variables in the model including measurement errors of those (Bound et al. 2001:1-2). Results achieved when applying cross national survey measures, both experience based and perceptions based are indeed convenient if high explanatory power is the goal, something it often but unjustly seems to be (Midtbø 2007a:104). Treisman presents a number of models with a range of operationalizations of the dependent variable; most of them achieve explanatory power of between 55- and 90 percent and some of them even above 90 percent. This does not, of course, in itself mean that measurement errors are systematic, but it may justify some suspicion. When it comes to unsystematic measurement error the only solution is to compare to more reliable data that we may collect from subsamples of our main samples, i.e. we may go into deeper detail on a subsample and collect data that we with
greater certainty may assume are correct (Chen et al. 2005). Such auxiliary data are unfortunately rarely available (Hu 2007:2-3), and corruption data is no exception. Thus we may not be sure whether the measurement errors we may have are unsystematic. However, there are several arguments having us suspect that the measurement errors in cross national corruption data are systematic.

Bertrand and Mullainathan (2001:3-4) point to three issues of survey data that may lead to systematic measurement errors. Firstly, cognitive problems may involve that the questions asked may make the respondents elicit certain memories that may affect how they answer. Could it be that the mere fact that the respondents are asked about corruption makes them weight experiences with corruption more than they weight not experiencing such incidents? Or in the case of academics who are asked about their perceptions; may they weight certain “more interesting” reports or certain incidents of media coverage more than reports reporting less corruption or the absence of media attention?

A second issue that may affect survey data towards systematic measurement error is social desirability; respondents want to appear as they believe others want them to appear, or to appear in accordance with their own preferences (Bertrand and Mullainathan 2001:3-4). This may be particularly relevant for business managers who answer questions about corruption; they want to appear more in accordance with whichever norms they see as “correct” in any given situation, and thus underreport. Or, since the exact wording in the so called experience based measures may be interpreted as meaning other companies than the respondents company, they may be inclined to over-report the incident of corruption amongst their competitors. Academics may want to appear knowledgeable, and thus perhaps may be inclined to over-report.

Another problem that may cause systematic measurement errors is the problem that people may not know, but still pretend they do (Bertrand and Mullainathan 2001:68-69); Respondents may think they should have an answer simply because they are asked the question, in stead of just admitting that they are not sure or do not know or that their opinion is that it is impossible to know, if that should be the case.

To decide whether respondents answer honestly, balanced and based on actual knowledge that may conform to reality, is of course hard to determine. However, there seems to be increasing agreement in the literature that the likelihood of systematic measurement
error is higher than previously believed (Chen 2007). That this may occur in survey data has been shown by several researchers (i.e. Bollinger 1998; and Bound and Krueger 1991).

There are also arguments pointing in the direction that the cross national survey data on corruption may be subject to systematic measurement errors. Heywood claims, regarding the perceptions based measures that they may be subject to a problem of endogenity; He considers it likely that the perception of corruption from one year may be affected by the rating the same country has gotten in the index in question or different indexes in previous years (Heywood 2007:698). It is not unthinkable that experience based measures may be subject to similar problems.

Another possible source that may produce bias in the data, and for which the cross national survey data on corruption have been criticized, is sampling. The CPI has been criticized that the share of respondents being business people is biased towards western business people whom may have quite different perceptions of corruption than would local business people or foreign, non-western business people (Philip 2006; Bardhan 2006). I.e. it may be the case that foreigners are forced to pay larger bribes than domestic business people (Bardhan 2006).

That data sources available for country- or region specific studies also may have large problems is clear from the discussion above, however, these problems are likely to be country specific; it is unlikely that all countries experience similar fluctuations or stable biases in the data that are available, i.e. media report or conviction rates. Thus there is little reason to assume that there are measurement errors in these data that are systematic across countries.

Based in these arguments I formulate hypothesis 4:

H4: Operationalizing the dependent variable, corruption, as a cross-national survey measure gives higher likelihood for observing effects compared to country- or region-specific studies.

3.3.4 Possible divergence between experience based and perceptions based data
Researchers have posed arguments pointing in the direction that the perceptions based data have more problems with systematic measurement errors than do experience based cross-national data. Treisman (Treisman 2007) compares the findings applying perceptions based measures and experience based measures, respectively and finds highly diverging patterns of causality, even if as stated above all models achieve high R². Applying perceptions based measures, high development, long traditions with liberal democracy, a free press which is
widely read, a high female representation in central political bodies and a tradition of openness to international trade seem to be factors leading to less corruption. Natural resource dependency, intrusive business regulations and unpredictable inflation seem to lead to more corruption. However, when applying experience based measures, only GDP per capita and, marginally, the time it takes to start a business, have significant effects on the degree of corruption. Treisman further suggests that systematic measurement error in perceptions based data may be caused by the belief among respondents of what causes corruption; the belief that the explanatory variables found to have significant effects have significant effects may have lead respondents of perceptions based measures to rate countries based on the assumed correlates of corruption, and not based on “theoretically independent” judgments of corruption levels (Treisman 2007:228). That there may be systematic measurement errors in perceptions based cross-national data on corruption is also strongly suggested by Hawken and Munck (2011) and also Ko and Samajdar (2010).

Based in this discussion it seems justified to formulate Hypothesis 5:

H5: Operationalizing the dependent variable, corruption, as a cross-national perceptions based measure gives higher likelihood for observing effects compared to operationalization through experience based cross-national data or country- or region-specific studies, respectively.

However, there are also arguments pointing in the direction that experience based cross national data should be troubled with systematic measurement errors. Heywood claims that the fact that there are many business people amongst the respondents i.e. of the CPI may skew the results in favor the opinions of only one relevant segment of people (Heywood 2007:698). If this is the case, the surveys may not give us the whole picture of corruption. Since there is some evidence that major public contracts may be particularly vulnerable to corruption, we may suspect that business people may over-report compared to the real and complete picture (Heywood 2007:698). However, I must partly object to these arguments. Also other experts than business people respond to the surveys. However, for the responses of business people the point may be valid, and thus the final result may be partly skewed. If this indeed is a problem it would be so to a larger extent for the experience based measures than for the measures that are only partly based on responses from business people.

Based in this argument I formulate Hypothesis 6:
H6: Operationalizing the dependent variable, corruption, as a cross-national experience based measure gives higher likelihood for observing effects compared to operationalization through perceptions based cross-national data and country- or region-specific studies, respectively.

**Figure 2: Causal diagram of all hypotheses**

3.4 Concluding remarks
Recall that the idea of discussing the issues explored in this chapter was derived from Treisman’s suggestion that there might be a causal relationship from perceived causes of corruption to perceived levels of corruption, affecting the results when applying these data. Figure 3.1, presented in the beginning of the chapter illustrates this relationship. Figure 3.2 illustrates this relationship and also illustrates the causal relationships presented the hypotheses formulated in this chapter. Note that three possible characteristics and effects thereof have been hypothesized for the cross-national perceptions based and experience based measures, respectively. Two characteristics and effects thereof have been hypothesized regarding data and research that may be considered country- or region-specific; I have not found reason to suspect that measurement errors in these data should be systematic, at least not systematic across countries, which would be of interest in this thesis.

The only clear and simple conclusion that can be drawn based on the discussions of this chapter is that corruption has been studied in many ways, none of which is unproblematic. The advantage for the researcher to dig deeper into the data material, make individual
judgment in the course of the research process, and make qualitative and presumably qualified judgments about reliability and validity as the data are collected and analyzed may strengthen the conclusions drawn from narrow-focus research. However, as King et al. points out researchers of these approaches are not protected from making measurement errors (King et al. 1994:155-156). This should also be clear from the discussions of this chapter. Williams claim may be somewhat pessimistic but nevertheless illustrative: “[i]t seems (...) almost inevitable that the ‘evidence’ used by students of corruption is bound to be fragmentary, biased, anecdotal, potentially misleading, impressionistic and inadequate,” (Williams, 1987:27-28 in Sandholtz and Koetzle 2000:33).

In the next chapter I will move on to elaborating on the mode of analysis of this thesis, logistic regressions. The data sources elaborated on in this chapter will be the explanatory variables, and conclusions of researchers, as discussed in Chapter 2, will be the dependent variables in a series of logistic regressions.
Chapter 4: Choices of methods and data for this thesis

The task of this chapter is to describe how the problems raised in the previous chapters will be answered. Table 7 sums up the dependent variables and hypotheses. The aim of this study is not to determine whether there may be certain scientific works on the causes of corruption that may diverge in its results, but rather (1) to map the general picture of agreement or divergence, and (2) find out whether there are specific causes of such divergence springing from the choices of data on corruption.

I will start out with discussing the choice of a quantitative method for this thesis, then I will discuss how quantitative methods may be applied in reviewing literature; then I will move on to describe how the variables are operationalized, before I describe how I will analyze the data through a series of binary logistic regressions. Finally I will describe how I have constructed the data set.

Table 7: Summing up the goals of this thesis

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Operationalization of corruption</th>
<th>Expected effect on relationship between corruption and variables of Chapter 2 (described in Chapter 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a:</td>
<td>Cross-national survey measure</td>
<td>Lower likelihood for observing effects</td>
</tr>
<tr>
<td>H1b:</td>
<td>Cross-national survey measure</td>
<td>Not lower likelihood for observing effects</td>
</tr>
<tr>
<td>H2:</td>
<td>Perceptions based measure</td>
<td>Lower likelihood for observing effects</td>
</tr>
<tr>
<td>H3:</td>
<td>Experience based measure</td>
<td>Lower likelihood for observing effects</td>
</tr>
<tr>
<td>H4:</td>
<td>Cross-national survey measure</td>
<td>Higher likelihood for observing effects</td>
</tr>
<tr>
<td>H5:</td>
<td>Perceptions based measure</td>
<td>Higher likelihood for observing effects</td>
</tr>
<tr>
<td>H6:</td>
<td>Experience based measure</td>
<td>Higher likelihood for observing effects</td>
</tr>
</tbody>
</table>

4.1 The choice of a quantitative approach

Recall the discussion of the methods debate in the beginning of the previous chapter. Which method would be best to shed light on the queries and hypotheses I have formulated? The complexity of both the phenomenon corruption and the many ways in which it has been studied point in the direction of a qualitative approach to this topic, since qualitative
approaches to a greater degree may study the depths of social phenomena (King et al. 1994; George and Bennett 2005). One may imagine that some of the richness that the qualitative articles, and perhaps also of the quantitative articles, could be reduced applying a quantitative approach. A qualitative approach also would have allowed me to elaborate to a greater extent on finer differences in understandings of the explanatory dimensions applied by different researchers. Choosing a quantitative approach allows me to meaningfully look at a wider set of cases (Landman 2002:897). Quantitative method is superior when it comes to drawing conclusions based on a large population (Mahoney 2003; George and Bennett 2005), and there has been conducted a quite extensive amount of research on this topic; thus there are many cases available. In addition I do not only wish to discover whether there are diverging conclusions regarding the explanatory dimensions I have described; I want to discover the degree to which there is agreement among researchers. This points in the direction of choosing a quantitative approach. Regarding the causal analyses; whether, and the extent to which, results may be determined by data sources of the dependent variable, also is a research question that follows a similar quantitative logic.

4.2 Reviewing literature quantitatively
Wolf (1986) suggests what he calls meta-analysis as a good way of reviewing literature quantitatively. This approach applies the coefficients of former quantitative studies and analyzes them quantitatively. Judge provides an example of this within the field of corruption research (2011). However, I also want to include studies that have not applied quantitative methods. To meet the challenge of making studies applying different approaches comparable for application in quantitative analyses, I first considered applying fuzzy-set analysis to construct the data set; applying such an approach I would started with defining a set i.e. “analyses supporting the hypothesis that democracy leads to less- and autocracy to more corruption”. Then I would have rated the units from 0, “not member of set”, to 1, “full member of set” (following Ragin 2000). This approach is thus merged of both ideas from qualitative and quantitative methods in that you define a qualitative frame and give the units numbers of the degree to which they are members of the set. The data may then i.e. be applied in ordinary linear regression.

A problem, I realized, with this approach is that it would be very hard to determine in such fine detail the degree to which a researchers have found support for the different
hypotheses; in particular this would be true of those who do not provide conclusions expressed in numbers. Even though it would be optimal to describe with numbers the degree to which different researchers have found the effects I am interested in, it seems more plausible to decide whether or not they have found such effects. This is called a *discrete choice* problem (Glasgow 2008:513). More specifically it is talk of an ordered choice situation which, in theory, consists of three categories; The categories fall into a natural order and the simplified expressions of an underlying continuous variable ranging from positive effect, through no effect, to negative effect. In theory the outer limits of a variable are i.e. “democracy explains all of the presence of corruption” and “democracy explains all of the absence of corruption”.

In other words, what I will do is to go through a large amount of relevant scientific publications and categorize whether they find that i.e. higher degree of democracy leads to more corruption (-1), has no effect (0), or leads to less corruption (1). In order to do this I have to choose a cut-off point defining at which point the different analyses should be appointed the different values (Ragin 2008). Regarding the quantitative analyses, statistically significance at the 10 percent level seems as a reasonable cut-off point for negative and positive effects, respectively. Had the samples applied in corruption research been much larger significance at the 5 percent or 1 percent level could have been appropriate. However, most samples, as described in the previous chapter contain only between 50 and 200 units. Non-significant results seem appropriate to assign the value 0. The results not expressed in numbers must be assigned values based on how I interpret the formulations of the researcher. This may be more challenging and must be done with caution and a conscious mind.

Such research has been met with some skepticism. The choices of cut-off points and the process of assigning values may be biased (Ragin 2008), i.e. the researcher’s wish to have his/her hypotheses supported may lead to systematic measurement error in the collection of data (King et al. 1994:155-156). These are issues I must be aware of when collecting the data.

**4.3 Operationalizations of dependent and independent variables**

The dependent variables are operationalized in accordance with Chapter 2 and are summed up in Table 7 above. Considering that the dependent variables defined for this study have three possible values (i.e. -1, 0, 1) it would intuitively be more correct to apply ordered logistic
regression or multinomial logistic regression (Long 1997). However, none of the dependent variables have sufficient amounts of values in all three categories; they may have many on two of them, and few or none on the remaining value. Thus units will be given the appropriate values based on whether they have found the effect described in the statements in Table 7: i.e. if researchers have found that higher level of democracy leads to less corruption I code the articles 1. If they have *not* found this, if they have found that there is no effect or that more democracy gives more corruption, I will code them 0.

The explanatory variables are formulated in accordance with Chapter 3 and are summed up in Table 8. I make one explanatory variable out of each of the three perceptions based measures that are applied by many researchers (the CPI, the CCI and the ICRG measure). The three remaining perceptions based measures and the experience based measures have been merged into “Other perceptions based measure” and “Experience-based measure”, respectively. This has been done in order to reach a sufficient amount of units of each of the values of the explanatory variables, i.e. to overcome problems with discrimination (elaborated below). The former all contains perceptions based measures that are not “polls of polls”; these are less commonly applied and it may be interesting to see if they differ in results from the more commonly applied perceptions based measures. The latter variable is merged for the obvious reason that they are measures of the same kind; experience based.

The variable “Data sources available for country- and region specific studies” is expressed in such a wide category in order to capture the idea that approaches allowing the researcher to base judgments of causal relationships on a wider array of data sources and thus possibly strengthening validity and reliability, which may provide more reliable results.

**Table 8: Operationalizations of explanatory variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Operationalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICRG:</td>
<td>Dummy-variable: If the study in question has applied this measure, I give it the value 1.</td>
</tr>
<tr>
<td>CPI:</td>
<td>Dummy-variable: If the study in question has applied this measure, I give it the value 1.</td>
</tr>
<tr>
<td>CCI:</td>
<td>Dummy-variable: If the study in question has applied this measure, I give it the value 1.</td>
</tr>
<tr>
<td>Other perceptions based measures:</td>
<td>The remaining perceptions based measures have not been applied often enough to form basis of independent explanatory variables and will thus be merged into one variable with the value 1 if such a measure has been applied.</td>
</tr>
<tr>
<td>Experience-based measures:</td>
<td>The experience based measures have not been applied often enough to form basis of independent explanatory variables and will thus be merged into one variable with the value 1 if such a measure has been applied.</td>
</tr>
<tr>
<td>Data sources available for country- and region specific studies:</td>
<td>If an article has applied qualitative method(s), or has applied some quantitative measure developed for one country or region, it will be given the value 1 on this variable.</td>
</tr>
</tbody>
</table>
4.4 Binary logistic regression analysis.
The basic idea of logistic regression analysis is the same as that of ordinary linear regression, but they differ on some details (Skog 2007:352). The reason why we may not apply ordinary linear regression, i.e. Ordinary least squares (OLS) regression, is that the most important preconditions for this kind of analysis are not fulfilled with a dichotomous dependent variable; we will not have linearity, we will experience heteroscedasticity, and lastly we may end up predicting values that are outside the range of the dependent variable, i.e. values higher than 1 (Eikemo and Clausen 2007:83). While OLS is about calculating the estimates so that the error terms are minimized, in logistic regression we calculate the estimates based on the principle of Maximum Likelihood; we want to maximize the likelihood that the dependent and independent variables occur together (Eikemo and Clausen 2007:83). In other words the estimates are calculated so that they to the largest extent possible converge with the observed values of the dependent and independent variables (Eikemo and Clausen 2007:151). OLS is in fact a special case of maximum likelihood. When the error term is normally distributed and the residuals are homoscedastic (as assumed in OLS), the likelihood of observing the exact data set that we have is highest when the sum of squared residuals is as small as possible (Skog 2007:362).

Recall that the mathematical expression for ordinary regression is as follows, a straight line (Skog 2007):

\[ Y = b_1 + b_2 \cdot X \]

When applying logistic regression, however, we cannot assume the same preconditions fulfilled, most importantly because the regression line probably won’t be a straight line. Logistic regression is based on calculations of likelihoods; i.e. the likelihood of ever having tasted an alcoholic beverage may increase slowly from the age of 0 up until the mid-teens, then it will increase dramatically for a few years, and then it will flatten out; it will never reach 100 percent. In other words logistic curves are S-shaped (Skog 2007:354). However, in preparing the data for a logistic regression we will seek to fulfill the linearity assumption we know from OLS. This process consists of two steps, first we need to calculate odds. The odds expresses how much larger the likelihood is for an event to occur than for it not to occur, or vice versa (Skog 2007:363). Based on the odds we calculate the logits (the natural logarithms)
of the odds. The estimation of the model is based on the logits (Long 1997:2). The odds are calculated as follows:

\[
\text{Odds} = \frac{\text{Occurrence}}{\text{Non-occurrence}}
\]

I.e. if there is a 60 percent chance that something will happen, there is a 40 percent chance that it will not happen; the odds is 0.6/0.4=1.5; this means that it is 1.5 times more likely that the event will occur. If the chances for and against occurrence are equal the odds is 1. Specific to the context of this thesis we may state the following:

\[
\text{Odds (conclusion that democracy leads to less and autocracy to more corruption) = } \frac{\text{Number of supporting conclusions}}{\text{Number of non-supporting conclusions}}
\]

In order to analyze the relationships between the different explanatory variables and the dependent variable we need the logits. The value 0 on the logit scale related to the dependent variable is the point in the scale where there are as many units with as without the characteristic of the dependent variable; positive values indicates that more than 50 percent of the units have the characteristic and negative logits indicate that less than half have the characteristic. The advantage of the logit calculation is that it “straightens out” the S-shaped relationship between the variables (Long 1997:2); when the share of values with the characteristic of the dependent variable closes on 100 percent, the logit will approximate infinite positive values and it will approximate infinite negative values when the share closes on 0 percent. In other words we now may observe relationships between the variables that may be expressed approximately as a straight line (Skog 2007:356). We may now formulate the following regression equation:

\[
\hat{Y} = \frac{1}{1+e^{-(b_0+b_1X_1+b_2X_2+\ldots)}} + \epsilon
\]
This equation states the expected value on the dependent variable given the value on the explanatory variable. $\bar{Y}$ indicates the share of units with the characteristic of the dependent variable, or the probability of observing a “1”. The “1”s under and above the bar indicate that this is number for which we want to calculate the probability. The intercept is $b_0$; the value of the dependent variable when the independent variables are 0. $\varepsilon$, the symbol in the right hand end of the equation, I is the error term (Skog 2007:358). The error term in this situation is not normally distributed, and homoscedasticity may not be assumed. These are assumptions of OLS-regression, which are reasons why we may not apply OLS with dichotomous dependent variables (Skog 2007:360).

$b_1$ indicates the increase in the dependent variable when the independent variable ($X_1$) increases by one unit and likewise with $b_2$ in relation to $X_2$ and so on. In logistic regression these numbers inform us of how much the logit changes when the independent variable increases by one unit. The advantage of the recoding into logits is mathematical simplicity, not simple interpretation; a certain change in an independent variable may be expressed with the same relationship to the dependent variable at all levels (Skog 2007:358). In the next section I will describe how we may make these numbers more interpretable.

### 4.4.1 Interpretation
The letter $e$ in the regression equation is the base of a natural logarithm ($2.7182\ldots$). That this number is powered by the expression - $(b_0+b_1\cdot X_1+b_2\cdot X_2\ldots)$ indicates that we want to find the antilogarithm, or the exponential function, of - $(b_0+b_1\cdot X_1+b_2\cdot X_2\ldots)$ (Skog 2007:358); this calculation gives us the number that we in this context call the odds ratio; it expresses (approximately) the relative change in odds; how many times larger or smaller the odds is when the explanatory variable increases by on unit. If we want to express the changes in percent we may, since the explanatory variables also are dichotomous, simply subtract 1 from the odds ratio and multiply this number by 100 (Eikemo and Clausen 2007:92). This is the simplest way of interpreting the regression parameters in logistic regression (Skog 2007:364).

We have discussed possible increases or decreases in odds; but increases or decreases compared to what? We need a reference point in order to determine whether an odds has changed or not. This is done by deciding on an explanatory variable whose odds to which we will compare the odds related to the other explanatory variables. First we calculate the odds of
the reference category; the odds ratios based on the estimates of the other explanatory variables will then describe the *relative change in odds compared to the odds of the reference category* (Long 1997). Technically, in SPSS, this is done by *not* entering the explanatory variable that will serve as reference category into the regression, or in newer versions of SPSS, you may enter all explanatory variables and SPSS excludes one of them. This will make SPSS calculate the odds of the excluded variable and calculate the difference in odds between this and each of the other explanatory variables. It is, of course, important that the data set does not contain variables that shall not be included as explanatory variables (except the reference category) and have units concurring with values on the dependent variable in unfortunate ways, since this could make the reference category wrong and thus mess up the interpretation of the whole model.

4.4.2 Test statistics
Also the test statistics of the logistic model are different from those of OLS-models. When considering the goodness of fit of the whole model we look at the so called -2 log likelihood-value (or just -2LL). This is a measure for how well the data fits the model. Small values indicate that the model fits well (Skog 2007:368). The measure is the equivalent of “sum of squared errors” that we know from OLS-regression (Skog 2007:368). The value is the natural logarithm of the likelihood value multiplied by -2. The value in itself is not very informing, but we can see whether there is a statistically significant change in the value between models and thus determine which model is best (Luke 2004:33-34). We compare the -2LL of a model *without* the explanatory variables (a “reduced model”) with the -2LL of the model with the explanatory variables by conducting a Chi-square test of “goodness of fit” (Mays 2011) which expresses the difference between the two -2LL values (Eikemo and Clausen 2007:88). The chi square of these two values may be tested with a p-value which expresses the likelihood that the difference between the two models is a result of coincidences. In other words if the p-value related to the chi square test of the difference in -2LL values of the reduced model and the full model is 0.1 or lower, we may assume that our explanatory variables contribute significantly to the model (Eikemo and Clausen 2007:88). In other words this forms the equivalent of an F-test in OLS.

In OLS tests of “goodness of fit” and expressions of predictive efficacy, like $R^2$, capture the same substantial notion (Long 1997). In logistic regression, goodness of fit
captures the match between the observed frequencies and the values you would have if the model was the true generator of the data. Predictive efficacy, on the other hand, concerns the ability of a model to correctly predict a case’s value on the dependent variable. We may very well observe a strong goodness of fit, but not high predictive efficacy (Long 1997; also Hosmer and Lemeshow 2000). Several measures of predictive efficacy in logistic regression have been proposed, so called pseudo- $R^2$, i.e. the Cox and Snell $R^2$ and the Nagelkerke $R^2$. However, the meaningfulness of these measures are highly discussed; firstly because they may not express anything meaningful (Hosmer and Lemeshow 2000:164; Skog 2007) and secondly because readers acquainted only with linear regressions may believe they express something along the line of $R^2$ values in OLS-regressions (Hosmer and Lemeshow 2000:164). In fact Hosmer and Lemeshow primarily discuss pseudo- $R^2$ because they believe they have to since these values are reported by many statistics software tools; however, they recommend not reporting them (Hosmer and Lemeshow 2000:167). The attitude that these values should not be reported is also supported by Eikemo and Clausen (Eikemo and Clausen 2007); conclusions of whether our explanatory variables make significant contributions to the model must be based on the Chi-square test of goodness of fit (Eikemo and Clausen 2007:88).

In addition to testing whether the model as a whole significantly contributes to the goodness of fit, we want to see if the individual explanatory variables have significant effects on the dependent variable. In OLS this is done by conducting a t-test and interpreting this directly or looking at the related p-value. In logistic regression we look at the Wald-test or the related p-value. Like the t-test, the Wald test is based on comparing the size of the b-coefficients (here the logits) with the size of the error terms (Skog 2007:368;374). More specifically, by dividing the b-coefficient by that b-coefficients standard error and multiplying that value with itself (King 2008:376). I will interpret the related p-value that SPSS will calculate for me. However, these values must be interpreted with caution since they do not provide the same accuracy as the equivalents of OLS (Skog 2007:375).

4.4.3 The assumptions of logistic regression.
The assumptions of logistic regression analysis are similar to those of ordinary linear regression, only somewhat simpler (Skog 2007:380). Since the dependent variable is dichotomous we don't need to assume the same strict criteria of the distribution of the error term, normal distribution and homoscedasticity (Skog 2007:380).
Firstly, we must assume that (1) the logistic curve gives a correct description of the empirical correlation. This assumption is always fulfilled if the explanatory variables also are dummy-variables. I will include one control variable (elaborated below) which is continuous, thus I need to evaluate whether the assumption is fulfilled. This can be done by conducting a Hosmer-Lemeshow (HL) test; this tests whether the regression curve deviates significantly from the S-curve; if the HL-test is statistically significant the assumption is not fulfilled. This may be solved by adding the quadratic term of the explanatory variable in question to the regression equation or by replacing the variable with its logarithm (Skog 2007:380-381:Chap. 14).

Secondly (2) the observations must be independent of one another. This assumption will always be fulfilled if we operate with a simple probability sample and not longitudinal data (Skog 2007:380); thus this assumption is fulfilled in the analyses of this thesis.

Thirdly (3), when applying categorical explanatory variables, which I do, we may encounter an issue called discrimination. The source of the problem lies in the way in which the logit and the odds are calculated. Recall that the odds are calculated by dividing occurrence with non-occurrence. If one of these values are zero or close to zero, the quotient will also be zero or close to zero, and calculating the logit will be impossible. A possible solution is to merge one or more explanatory variables (Eikemo and Clausen 2007:129-130).

Fourthly (4) we must seek to make sure that the independent variables we integrate in our models are the variables that actually may explain changes in odds of occurrence of the dependent variables; in other words making sure that there are not other, confounding, variables that we should have included in the model that actually, partly or fully, are the ones explaining the variation in the dependent variables; we need control variables. This assumption is at the same time the one of greatest substantial importance and the one that is hardest to fulfill (Skog 2007:380-381). Recall that in the previous chapter I discussed that many researchers oppose to conducting analyses with very wide geographic scopes because certain causal relationships may be specific to certain context (or Skocpol 2003; i.e. McKeown 1999). In the context of my analyses we may imagine that I may find that i.e. researchers having applied certain data also have studied certain contexts like specific countries, regions, or groups of countries that may have certain features in common. I will seek to control for context by adding control variables for the region in which the analyses of my sample have been conducted. Constructing the data set (elaborated below) I have also
coded which country or group of countries the study has focused on. This has formed the basis for control variables: whether they have focused on communist- or post-communist countries, since corruption by some is seen as having a special place in post-communist (i.e. Kneen 2000) and communist countries (Sun 2008); whether they have studied Sub-Saharan countries; whether they have focused on high-income countries, since effects of the different explanatory dimensions may play very different roles in different economic environments (Huntington 1968); I will also add a control for the year in which the analyses have been conducted, following the notion that social science should be cumulative (Mahoney 2003) and that the accumulation of knowledge may lead researchers to having new and more insight on which they may found stronger analyses and conclusions. I will also apply whether the study has been conducted with OECD countries as samples as control variable. One may argue that this captures much of the same notion as that of high-income countries, however, many new members have been included into the OECD which are not extremely rich. In addition, the organization has long had a goal of promoting democracy and good governance (OECD 2011), thus it is interesting to look at such an effect. I will also in one model include a control for whether the researcher has conducted a panel analysis or not, since the descriptive analyses indicate that this may have affected conclusions.

4.5 Creating a sample:
Selecting the sample I have conducted a search on the word «corruption» in the ISI Web of Knowledge website and done the necessary refinements of the search\(^{11}\). The time span in the search is «all years», meaning that the last article in the search was published in 2010 and the first was published in 1956. This search gave me 3411 matches. Next I went through the abstracts of all 3411 articles in order to determine which seemed relevant. For a few articles it

\(^{11}\) first by ticking «social sciences» and then by excluding philosophy, cardiovascular system and cardiology, demography, medical laboratory technology, linguistics, energy and fuels, surgery, pharmacology and pharmacy, zoology, psychiatry, construction and building technology, agriculture, communication, information science and library science, family studies, computer science, pediatrics, geochemy and geophysics, literature, pathology, water resources, respiratory system, automation and control systems, acoustics, environmental sciences and ecology, plant sciences, film, radio and television, toxicology, geriatrics and gerontology, mathematics, infectious diseases, marine and freshwater biology, instruments and instrumentation, medical informatics, engineering, legal medicine, substance abuse, research and experimental medicine, operations research and management science, telecommunications, history and philosophy of science, food science and technology, reproductive biology, physiology, robotics, sport sciences, archaeology, imaging science and photographic technology, biochemistry and molecular biology, life sciences and biomedicine - other topics, genetics and heredity, forestry, materials science, obstetrics and gynecology, biodiversity and conservation, meteorology and atmospheric sciences, oncology, physics, physical geography, mathematical and computational biology, immunology, anatomy and morphology, optics, neuroscience and neurology, transportation, critical care medicine, radiology, nuclear medicine and medical imaging, biophysics, fisheries, arts and humanities - other topics, geology, mechanics, science and technology - other topics, nutrition and dietetics, ophthalmology and general and internal medicine.
sufficed to read the title, and in a few cases I had to do some deeper reading in the article. Many articles were concerned with i.e. defining corruption or studying consequences of corruption and thus where not relevant. Some articles concerned causes of corruption but were not relevant because they did not focus on whether something has lead to more or less corruption, but i.e. rather takes it as a presupposition of their own study, or that they study the implementation process of anti-corruption policy, the effect of which has not yet been seen. Due to my lingual limitations, only studies in English, Scandinavian languages (except Finnish) or Spanish may be included. After this preliminary reading I ended up with 784 publications which I went through in order to categorize all relevant results and the data with which they had operationalized the dependent variable. In course of the process of going through the articles in more detail I concluded that still many of them were not relevant. I ended up extracting data from 285 of them. Some of them had to be excluded because I could not gain access to them through the UoB library, in all 88 articles. Assuming that these articles would have had the same “fall-out”-percentage in the second round of going through them as the articles to which I had access, I would be missing 32 articles. This is probably not a big problem.

I decided to categorize the results from different analyses as units rather than categorizing different publications as single units if the two should not be the same; i.e. one publication may discuss the effects of economic liberalization in India and China and find diverging results; such a study would provide two units/results. Publications applying cross national quantitative approaches often run several regressions. In such cases I include the models that are “unique” in the sense that they are not repeated several times in the study or repeated with fewer controls. If almost the same model is presented but with other controls than other models in the same publication, this model will be included as a single unit. If identical models are presented, but with differing samples or differing operationalizations of the dependent variable or differing operationalizations of independent variables, their results will also be included as separate units. I have also coded variables as having effect if they are found to have an effect in interactions with a different variable. I ended up having coded 1218 different results from 609 different analyses/models.

---

12 See Appendix for complete list of publications included in the sample.
13 See Appendix for complete list of publications to which I did not have access.
I also categorized the year in which the study was published and its geographic, socio-cultural or socioeconomic scope, i.e. which country or region was studied, or if only high-income countries were studied. These data found the basis for control variables.

4.6 Concluding remarks

I have established that the approach of this thesis is quantitative, and that I will run a series of logistic regressions. The results of previous research will serve as dependent variables, operationalized as dummy-variables. The main explanatory variables will be the data choices made by researchers, also dummy-variables.

In the following chapter I will present the results of the analyses of each of the dependent variables. Starting with descriptive analyses, and then moving on to the causal analyses.
Chapter 5: Analyses
In this chapter I will present the results from the analyses of each of the dependent variables, beginning with descriptive analyses of the degree of agreement between researchers. In the descriptive analyses I will lay emphasis on the outliers, i.e. those that find that democracy leads to more corruption, and try to give an impression of why they have concluded thus. Following the descriptive analyses of each of the dimensions I will, where the data allows for it, regress the conclusions of researchers on the variety of data choices they may have made.

First I will analyze the data I have collected regarding effect of economic liberalization on corruption. These data only allow for descriptive analyses, since the dimension appears to almost solely have been studied in country- or region specific studies, thus there is no foundation for comparison. However, the data allow for an interesting descriptive analysis. Secondly I will present the findings regarding the effect of democracy on corruption. Here I will present some descriptive analyses, and present two logistic models. The effect of press freedom on corruption will be analyzed next. Here I will present three logistic models, one where I exclude a few very influential cases, in addition to descriptive analyses. The data on the effect of regime transition on corruption also only allow for descriptive analyses. Sixth I will present descriptive and causal analyses regarding the effects of level of economic development on corruption-dimension, and finally descriptive and causal analyses of the resource curse dimension, presenting two logistic regression models for these two dimensions also.

5.1.1 The effect of economic liberalization on corruption
The degree of agreement in the field that economic liberalization leads to more corruption is very high, 86.4 percent (Table 9). This clearly casts doubt on the neo-liberal notion that it should do the opposite. Disagreeing results are provided in only eight analyses, of which two conclude that liberalization leads to less corruption and 6 that it has no effect.

This dimension has almost solely been explored in studies with country- or region specific foci. I’ll start with the exception: The CCI is applied as the dependent variable in a panel study of developing countries by Koyuncu, Ozturkler and Yilmaz concluding that liberalization leads to less corruption (2010); this may point in the direction that such an effect may be specific to developing countries. It is however, also interesting that the effect is
observed in a panel analysis; the strength of panel analyses are that they may establish which development occurs first in time; causes clearly must occur before their effects (Skog 2007).

Table 9: The effect of economic liberalization on corruption: summary of data sources and conclusions

<table>
<thead>
<tr>
<th>Effect</th>
<th>CPI</th>
<th>CCI</th>
<th>ICRG</th>
<th>WCR</th>
<th>GCS</th>
<th>GEI</th>
<th>WBES</th>
<th>BI</th>
<th>Country-/reg-spec.</th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>More corruption</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>51</td>
<td>51</td>
<td>86.4 %</td>
</tr>
<tr>
<td>No effect</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>10.2 %</td>
</tr>
<tr>
<td>Less corruption</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3.4 %</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>58</td>
<td>59</td>
<td>100 %</td>
</tr>
</tbody>
</table>

The second study providing results that liberalization processes lead to less corruption is Schoenman (2005). He studies Poland of the 90’s. He also claims that the case of Poland in the 90’s, given the departure as a post-Soviet state, is an “unmatched economic success”. Schoenman highlights the elaborate system of public monitoring of the privatization process and the rapidness with which it was conducted as central aspects of why corruption did not prevail. The conclusion may in other words be diverging because Poland is a divergent case.

A closer look at the six studies concluding that there is no effect of liberalization on levels of corruption reveals that there also may be a geographic pattern in these findings. Four of the analyses focus on Latin-American countries concluding that corruption has not increased as a consequence of economic liberalization (Kohl 2004; Di John 2005; Horowitz 2005; Martimort and Straub 2009). The two remaining focus on China (Sun and Johnston 2009) and Singapore (Low 2001), respectively. The study of China concludes that society has been relatively shielded from an increase in high-level corruption, which is the main concern in this thesis; it has however experienced increased corruption further down in the administrative system. Singapore has long been seen as the exception in Asia (Low 2001); it was not plagued with much corruption before or during liberalization processes in the late 80’s and in the 90’s (Low 2001). This is not to say, of course, that the countries in question, with the notable exception of Singapore, are not troubled with high-level corruption. Only that the liberalization processes did not cause the corruption.

---

The clearest of conclusions regarding this dimension is that there is a relatively overwhelming agreement that economic liberalization processes thus far have lead to increases in corruption. This is not to say, of course, that economically liberalized countries may in the longer run experience better control of corruption. Further insight into these questions is needed.

According to White how people thought about corruption in the 90’s was shaped by a wave of economic liberalization and democratization that started in the 80’s. Not the least has these trends of liberalization and democratization affected the policy agendas of the large worldwide development agencies (White 1996:150) like i.e. The World Bank (Stiglitz 2006) and The International Monetary Fund (Moschella 2009). These thoughts had great political consequences. There is great reason to question whether this was a good idea.

These data obviously do not provide a foundation to compare the results given through country and region specific studies with those of cross-national measures. That so few researchers applying quantitative methods have chosen to study this dimension (only one in this sample) may be that there is a lack of data on whether or not, and when, countries are in a liberalization process. Koyuncu et al. (2010) have solved the problem through a panel analysis applying degree of private versus public activity in the economy which is an operationalization of this explanatory dimension that clearly would not be possible in a cross-sectional analysis. However, it does not at all seem impossible to collect data on whether countries are or have been in a process of economic liberalization at any given point in time. The claim of Bevir (2008) that also theoretical choices may be affected by values may be supported by these findings; perhaps researchers conducting quantitative cross national don’t see the need to study this dimension.

5.1.2 The effect of democracy on corruption

The data I have been able to collect on studies of this dimension, is more extensive and allows also for causal analyses. The main finding is that only a variable of other perceptions based measures (“other” than the ICRG, CPI and CCI) is statistically significant; application of such data makes it less likely to conclude that democracy leads to less corruption. I will start with the descriptive analyses, but will give these somewhat less attention than in the prior section and prioritize the causal analyses since amount of articles is also too large to go through a fair share of them.
The sample on this dimension consists of 237 results (Table 10). Somewhat surprisingly only 54.4 percent of them support the notion that democracy leads to less corruption and/or that autocracy leads to more corruption. 41.8 percent claim that there is no effect, while 3.8 percent claim that democracy leads to more corruption and/or that autocracy leads to less. Let me first try to identify a pattern amongst nine outliers concluding that there is a positive effect.

Table 10: The effect of democracy versus autocracy on corruption: summary of data sources and conclusions

<table>
<thead>
<tr>
<th>More democr. leads to:</th>
<th>CPI</th>
<th>CCI</th>
<th>ICRG</th>
<th>WCR</th>
<th>GCS</th>
<th>GEI</th>
<th>WBES</th>
<th>BI</th>
<th>Country-/reg-spec.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less corruption</td>
<td>28</td>
<td>42</td>
<td>21</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>32</td>
<td>129</td>
</tr>
<tr>
<td>Has no effect</td>
<td>15</td>
<td>16</td>
<td>27</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>18</td>
<td>99</td>
</tr>
<tr>
<td>More corruption</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>44</td>
<td>58</td>
<td>48</td>
<td>18</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>53</td>
<td>237</td>
</tr>
</tbody>
</table>

First Singapore, again, provides an exception (Low 2001). Secondly, applying the WBES as dependent variable and with a sample of transition economies Clarke and Xu (2004) find that level of democracy is positively related with corruption levels; in other words their findings may indicate that such a relationship between democracy and corruption is context specific to transition economies. Italy provides another well recognized exception (Pujas and Rhodes 1999); Italian corruption survives despite democracy. The presence of the mafia and problems with the judicial branch may be factors behind this (Della Porta and Vannucci 2007). Hoetjes makes similar claims in a qualitative study of developing countries from 1976 (Hoetjes 1976). This may indicate another context specific effect, however, one must keep in mind that the data available in 1976 probably were even poorer than they have been in later years. Three exceptions are provided by analyses by Ades and Di Tella (1999). These three analyses apply some of the earliest cross national data available, Business International data (1980-1983) and World Competitiveness Report data (1989-1990). The samples are also quite small 52 and 31 countries, respectively. I am not able to see that any geographic aspects should be at play in the samples. The intuitively most natural conclusion

\[\text{BI: Algeria, Cameroon, Kenya, South Africa, Bangladesh, India, Israel, Japan, Jordan, Korea, Malaysia, Pakistan, Philippines, Singapore, Sri Lanka, Thailand, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, Canada, Dominican Republic, Jamaica, Mexico, Nicaragua, Panama, Trinidad and Tobago, United States, Argentina, Brazil, Chile, Colombia, Ecuador, Peru, Uruguay, Venezuela, Australia, New Zealand, Indonesia. WCR: Hong Kong, India, Japan, Korea, Malaysia, Singapore,}\]
thus seems to be weak data. The final outlier (Sung 2003) is somewhat puzzling, as the author also admits: the CPI (1999-2000) is applied on the whole sample of 99 countries and the explanatory variable of democracy is a dummy of electoral democracy from Freedom House. Sung himself points to research providing (Della Porta and Vannucci 1999; and Rose-Ackerman 1999) that democracy and particularly democratic competition for power may be associated with corruption. However, this does not explain why these findings diverge so highly from the vast majority of researchers concluding that there is either no effect or a negative effect of democracy. The only deviation in his analyses from what most researchers seem to do, as far as I can see, is that he controls for three different measures for female representation in political life. The argument behind female representation, eventually, having negative effect on corruption is that women have higher morale than men (Dollar et al. 2001). However, many have disputed this claim, i.e. that such effects may be culture specific (Alatas et al. 2009), or that it is spurious and only a “myth in the making” (Goetz 2007).

Now I will turn to the causal analyses. In order to avoid problems of discrimination, I have merged the WCR and GCS indexes to a variable I have named other perceptions based measures. The GEI, WBES and the BI have been merged to Experience based measures.

Table 12 shows the reduced model (a model without explanatory variables); the share of conclusions that democracy leads to less corruption. This model forms a basis of comparison to the full model. Let us for the sake of explication pretend that SPSS guessed that all results (constituting my sample) concluded that democracy leads to less corruption; in that sense SPSS was right 54.4 percent of the times in the reduced model. A crucial point when moving on is whether the explanatory variables makes the “guess” correct for a higher share of the units.\textsuperscript{16}

In the full model the percentage of units correctly classified is 66.2. The Chi-square test of the difference -2LL from the reduced model to the full model is statistically significant, indicating that the full model significantly improves the goodness of fit.

Thailand, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, The Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, Canada, Mexico, United States, Brazil, Australia, New Zealand, Indonesia.

\textsuperscript{16} I borrowed this image from Jacob Mays (Mays 2011).
Let’s move on to the different explanatory variables (Table 11). The reference category in this model is *Country- or region specific studies* (odds = 1.5238); the chance of concluding that democracy leads to less corruption or autocracy to more is thus approximately 1.5 times higher than the chance for not doing so when conducting a country- or region-specific study.

In this sample, the odds for concluding that democracy leads to less corruption increases by 14.8 percent when applying the *CPI* as dependent variable compared to the odds for such conclusion for the *country- or region-specific studies*. In other words the analyses, in which the *CPI* has been applied as the dependent variable, are more likely to conclude that democracy leads to corruption; however the effect is not statistically significant, supporting of H1b\(^{18}\).

The *CCI* has a similar effect providing higher likelihood for concluding that democracy leads to less corruption, with an increase in odds of 72.3 percent compared to

---

\(^{17}\) If the countries had a GDP per capita higher than 24000 USD per capita at the time of focus of the study (World Bank 2011a)

\(^{18}\) See table 4.1. on page 52 for summary of hypotheses.
country- or region specific studies; the effect is not statistically significant, also supporting H1b.

In the sample, the analyses applying the ICRG-index are less likely to conclude that democracy leads to more corruption, with a decrease in odds of 49 percent. This effect is statistically significant at the 10 percent level, supporting H2. Lambsdorff (Lambsdorff 2004) claims that the ICRG captures the political risk involved with corruption and not corruption itself and should thus not be expected to give comparable results; these findings may strengthen his claim that the ICRG-measure captures a slightly different concept.

The merged variable of other perceptions based measures shows that studies applying these measures are less likely to conclude that democracy leads to less corruption. The decrease in odds is 96.5 percent. The effect is statistically significant, also supporting H2. These measures have much in common with the CPI and the CCI except they are not polls of polls; these findings may be an indication that aggregating the results of several polls into one index may give stronger measures and thus increase the likelihood of observing effects; the same goes for the procedures applied by the World Bank and Transparency International, respectively, to weight the less diverging polls more. The number of units, one must keep in mind, is not very large with twenty units on this variable.

The merged variable of application of experience based measure shows that applying such a measure, in the sample, decreases the odds with 63.5 percent, however it is marginally insignificant (two-tailed), supporting H1b. However, the number of units with the value 1 on this explanatory variable is small so no strong conclusion may be drawn.

An assumption of logistic regression is that the effects of the explanatory variables are not spurious effects. The way to deal with this issue is to add control variables to the model (Skog 2007:chap. 15). I will seek control for contexts. In light of the descriptive analyses it would make sense to control for Italy and Singapore, however, since a merged variable of these two contexts only gives 3 units this gives problems of discrimination and cannot be done. I have constructed control variables for whether studies have been conducted with focus on Arab countries, communist or post-communist countries, sub-Saharan countries, Latin-American countries, OECD-countries and high-income countries, respectively. The time of writing may also be considered a context for the researcher that may affect the outcome. Thus, I also constructed the variable year published. However, most of these variables prove to have problems with discrimination because units with the value 1 are too few; these cannot be
included. I will include controls for high-income countries, OECD-countries and publication year.

Looking at the p-value of the Chi-square-test we see that the -2LL value has changed significantly from the model without controls (Table 12). Thus the controls contribute significantly to the goodness of fit. The percentage of units correctly classified has increased from 66.2 in the model without controls to 69.6 in the model with controls.

Researchers studying OECD-countries are much more likely to conclude that democracy leads to less corruption; the increase in the odds is 361.4 percent. This may be an expression of the focus of the OECD “to contribute to sound economic expansion in member as well as non-member countries in the process of economic development” (Convention on the OECD 1960: Article 1 (b)), and the fact that both democracy promotion and promotion of good governance are central to OECD (OECD 2011).

Controlling for high-income countries show that researchers focusing on high-income countries are much less likely to conclude that democracy leads to less corruption. The odds decreases by 93.2 percent. This may indicate that when countries become very rich an eventual effect of democracy whither away and becomes less important if important at all.

The effects of the other explanatory variables pretty much remain the same in the model with controls: of the variables with significant effects, the odds ratio of the ICRG further decreases somewhat. The odds ratio of other perceptions based measures is slightly higher, indicating that the likeliness of concluding that democracy leads to less corruption is slightly higher when adding controls, but still much lower than i.e. the reference category. Experience based measures is now very close to being statistically significant at the 10 percent level; this may incline us to conclude that applying experience based measures significantly decreases the likelihood of concluding that democracy leads to less corruption, thus supporting H3. However, this statement is made with caution.

The year published-variable shows us that analyses conducted closer to today are more likely to conclude that democracy leads to less corruption. On average there is an increase in odds of 5.9 percent from one year to the next, which indicates some degree of cumulativeness in the field. However, the degree of agreement on supporting democracy as having a negative effect of corruption is still only 54.4 percent in total. Since the variable publication year is continuous we must test whether the regression line converges to an S-curve; the Hosmer-
Lemeshow (HL) test (Table 12) indicates that there is no statistically significant deviation from an S-shape. Thus, this assumption is fulfilled.

Table 13 sums up the findings of this dimension in light of the hypotheses formulated in Chapter 3.

### Table 13: Results in light of hypotheses of Chapter 3

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Operationalization of corruption</th>
<th>Expected effect on results</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a:</td>
<td>Cross-national survey measure</td>
<td>Lower likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H1b:</td>
<td>Cross-national survey measure</td>
<td>Not lower likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H2:</td>
<td>Perceptions based measure</td>
<td>Lower likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H3:</td>
<td>Experience based measure</td>
<td>Lower likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H4:</td>
<td>Cross-national survey measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H5:</td>
<td>Perceptions based measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H6:</td>
<td>Experience based measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
</tbody>
</table>

### 5.1.3 The effect of regime transition on corruption:

As can be seen in Table 14 this dimension has only been studied in country- or regions-specific studies. This may be due to lack of data measuring whether countries are in transitional situation at any given point in time. However, as a consequence I will only look at this dimension through descriptive analyses.

32 results have been included in the sample. The degree of agreement that transitions lead to more corruption is very high (90.6 percent). 9.4 percent have concluded that there is no effect, while none have concluded that regime transitions lead to less corruption. This is clearly an eye-opener for western governments and organizations promoting democratization around the world; in particular if we take into account the many severe consequences of corruption (elaborated in Chapter 1).

### Table 14: The effect of regime transition on corruption: data sources and conclusions

<table>
<thead>
<tr>
<th>Regime transitions give</th>
<th>CPI</th>
<th>CCI</th>
<th>ICRG</th>
<th>WCR</th>
<th>GCS</th>
<th>GEI</th>
<th>WBES</th>
<th>BI</th>
<th>Country-/reg-spec.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>More corruption</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>90.6%</td>
<td></td>
</tr>
<tr>
<td>Have no effect</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.4%</td>
<td></td>
</tr>
<tr>
<td>Less corruption</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100%</td>
<td></td>
</tr>
</tbody>
</table>

Again, let’s look at the outliers. The conclusion of Moran (1999), studying south Korea, is that rather than being brought about by democratization reforms (1987-1997),
corruption was already there under the previous regimes. A similar situation is described by Lodge (1998) in the transition from Apartheid-South Africa; the democratization process didn’t bring corruption about, however its character changed somewhat. The same is found in Taiwan by Fell (2005), however, the author claims that the political environment shows clear potential of establishing a cleaner political system through the presence of a healthy opposition.

In other words, generally, democracy promoters around the world should expect their efforts to promote democratization of authoritarian states to lead to increases in corruption. It may of course be that corruption will decrease in the long run. However, I have some doubts as to whether this has been the plan of democracy promoters.

5.1.4 The effect of stronger democratic tradition on corruption
A highly related explanatory dimension is posed by the strength of democratic traditions. 124 results have been included in this sample. As seen in Table 15 there is high disagreement regarding whether the age of democracy has an effect on current corruption levels or not. 49.2 percent find that it does indeed decrease extent of corruption, while 50.8 percent disagree. None have concluded that stronger democratic traditions lead to more corruption. There is also relatively wide disagreement between analyses applying the same approaches to reach their conclusions. India again provides an outlier in the one Country- or region specific study that has concluded that there is no effect of the age of democracy on corruption (Singh 1997). India has been democratic for more than 50 years, but still experiences rampant corruption. Amongst the rest of the studies the variation is quite high regarding what they have concluded. Therefore I will move on to the causal analyses.

<table>
<thead>
<tr>
<th>Stronger democratic tradition gives</th>
<th>CPI</th>
<th>CCI</th>
<th>ICRG</th>
<th>WCR</th>
<th>GCS</th>
<th>GEI</th>
<th>WBES</th>
<th>BI</th>
<th>Country-/reg-spec.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less corruption</td>
<td>23</td>
<td>28</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>61</td>
</tr>
<tr>
<td>Has no effect</td>
<td>30</td>
<td>4</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>1</td>
<td>63</td>
</tr>
<tr>
<td>More corruption</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>53</td>
<td>32</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>5</td>
<td>124</td>
</tr>
</tbody>
</table>

Table 15: The effect of age of democracy on corruption: data sources and conclusions
The chi-square-test of the change in the -2LL is statistically significant (Table 17), thus the explanatory variables improve the goodness of fit from the reduced model. The correctly predicted units increase from 57.5 percent to 67.9 percent.

Since there are so few Country- or region specific studies, CPI will serve as reference category (odds: 0.767); applying the CPI the chance is smaller for concluding that stronger democratic traditions lead to less corruption than for not doing so.

The variable experience based measures in this analysis faces problems with discrimination. It cannot meaningfully be merged with any other explanatory variable. It must therefore be excluded from the regression. We may note that all of the units applying the WBES-measure as dependent variable, have concluded that there is no effect of age of democracy on corruption; this may be because respondents to this survey may be less affected by theories claiming that democracies should be less corrupt than many of the respondents to perceptions based measures (following Treisman 2007).

### Table 16: Dependent variable: conclusions on effect of strength of democratic traditions on corruption

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Full model</th>
<th>With controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p-value of</td>
<td>p-value of</td>
</tr>
<tr>
<td></td>
<td>Wald-test</td>
<td>Wald-test</td>
</tr>
<tr>
<td><strong>CPI</strong> (Odds: 0.767)</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>CCI</td>
<td>0.000</td>
<td>9.130</td>
</tr>
<tr>
<td>ICRG</td>
<td>0.676</td>
<td>0.783</td>
</tr>
<tr>
<td>Country- or reg.specif.</td>
<td>0.152</td>
<td>5.217</td>
</tr>
<tr>
<td>Year published</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Constant</td>
<td>0.338</td>
<td>0.767</td>
</tr>
</tbody>
</table>

* Reference category

### Table 17: Model summaries

<table>
<thead>
<tr>
<th></th>
<th>Correct predictions</th>
<th>Chi²-test of -2 LL</th>
<th>p-value of HL-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced model</td>
<td>57.5</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Full model</td>
<td>67.9</td>
<td>21,689 ***</td>
<td>--</td>
</tr>
<tr>
<td>With controls</td>
<td>77.4</td>
<td>23,752 ***</td>
<td>.000</td>
</tr>
</tbody>
</table>

N: 124

***p<0.01, **p<0.05, *p<0.10

Looking at the coefficients we see that applying the CCI increases the likelihood of concluding that longer democratic traditions lead to less corruption; the odds increases by 813.0 percent. This is counter to expectations that these two variables should show similar effects. We must recall however, that in a sample with few units, small changes in the explanatory variables may cause great changes in the coefficients, thus the results must be
interpreted with caution. Neither of the other explanatory variables give significant changes in the odds compared to the CPI-variable. In the sample the odds decreases by 21.7 percent with the ICRG, and increases by 421.7 percent with country- or region-specific studies. Thus it would be correct to state that it is the CCI that deviates from the rest, thus supporting H5 stating that perceptions based measures increase the likelihood for observing effects. However, it is in comparison only to the variable country- and region-specific studies a conclusion of support to H5 can be made and such studies have only been conducted by five researchers thus this conclusion must be drawn with caution. It must also be added that the CPI and the ICRG, also being perceptions based variables oppose H5. Thus no conclusion consistent with one hypothesis can be drawn based on these numbers.

The only control variable with a sufficient amount of units is the year published-variable. When running a model including this variable in the full model as a control, it is not significant. Since this explanatory variable is continuous I run a Hosmer-Lemeshow test to see if the relation between the variables is expressed through a logistic (S-curved) line. The test is statistically significant, indicating that the curve deviates significantly from the logistic curve. Running the model first with the log (10) of the year-variable, and also with a quadratic term together with the original year measure, both give significant Hosmer-Lemeshow tests. Thus this assumption is not and may not be fulfilled for this variable (Skog 2007).

The surprising divergence between the results arrived at when applying the CPI and the CCI respectively made me go through these articles in order to find a pattern; I found that of those concluding with no effect amongst those applying the CPI 16 analyses are provided by Kwok and Tadesse (2006); the effect seems to give in to controls of foreign direct investment, a measure of individualistic versus collectivistic cultures, and measures for rule of law. That may be some of the reason for the divergence between the CCI and the CPI. It also may indicate that these three variables applied by Kwok and Tadesse are explanatory variables that should be investigated further and more commonly should be included be included as controls.

Table 18 sums up the findings of this dimension in light of the hypotheses from Chapter 3.
Table 18: Summing results in light of hypotheses

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Operationalization of corruption</th>
<th>Expected effect on relationship between corruption and variables of Chapter 2</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a:</td>
<td>Cross-national survey measure</td>
<td>Lower likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H1b:</td>
<td>Cross-national survey measure</td>
<td>Not lower likelihood for observing effects</td>
<td>Yes</td>
</tr>
<tr>
<td>H2:</td>
<td>Perceptions based measure</td>
<td>Lower likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H3:</td>
<td>Experience based measure</td>
<td>Lower likelihood for observing effects</td>
<td>(Not tested)</td>
</tr>
<tr>
<td>H4:</td>
<td>Cross-national survey measure</td>
<td>Higher likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H5:</td>
<td>Perceptions based measure</td>
<td>Higher likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H6:</td>
<td>Experience based measure</td>
<td>Higher likelihood for observing effects</td>
<td>(Not tested)</td>
</tr>
</tbody>
</table>

5.1.5 The effect of presence of a free media on corruption

89 of the analyses in the sample present effects on corruption of press freedom or the extent of media presence (Table 19). The overall agreement that a stronger media leads to less corruption is quite large, with 79.8 percent of the analyses concluding that this leads to less corruption. 20.2 percent conclude that there is no effect. None of the analyses in the sample have concluded that press freedom or more media presence lead to more corruption.

However, the results are diverging between the different methodological approaches; the majority of the country- or region specific studies conclude that there is no effect, while very few of the studies applying cross national perceptions based measures conclude likewise. Let’s first look at the outliers among those applying perceptions based measures of corruption.

Table 19: The effect of press freedom on corruption: summary of data sources and conclusions

<table>
<thead>
<tr>
<th>Stronger Press gives</th>
<th>CPI</th>
<th>CCI</th>
<th>ICRG</th>
<th>WCR</th>
<th>GCS</th>
<th>GEI</th>
<th>WBES</th>
<th>BI</th>
<th>Country-/reg-spec.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less corruption</td>
<td>33</td>
<td>15</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>71</td>
<td>79.8 %</td>
</tr>
<tr>
<td>Has no effect</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>18</td>
<td>20.2 %</td>
</tr>
<tr>
<td>More corruption</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 %</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>20</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>89</td>
<td>100%</td>
</tr>
</tbody>
</table>

Five of these have applied the CCI as dependent variable. Three of these analyses are done in Andersen (2009) who studies the effect of degree of e-government on the degree of corruption. He applies a variety of samples and controls and in three models he finds no effect.
of degree of press freedom as measured by Freedom House; the fact that a measure of e-government is included, a variable which is not commonly applied may be the cause of the diverging results. The two remaining analyses with similar conclusions are conducted by Lindstedt and Naurin (Lindstedt and Naurin 2010). These results may be due to the operationalization of media presence as number of radio receivers per capita.

The one diverging result where the WCR has been used is an analysis by Brunetti and Weder (2003). They conduct run several models with other operationalizations of corruption primarily finding that press freedom leads to less corruption. They claim that this diverging finding with the WCR is due to a bias in the WCR sample towards industrialized countries.

Of the country- or region specific studies concluding that there is no effect, nine are from a study by Larmour and Barcham (2006). They conduct a series of case studies of small pacific island states. In other words, most of the conclusions amongst country- or region specific studies that there is no effect of press freedom and outreach may be specific to the contexts of the Pacific island states. For the remainder of the studies both with conclusion of effect and no effect I am not able to find any clear pattern. Let’s turn to the causal analyses. In addition to a model without and with controls, I will also run an additional model where I exclude the units concerned with the small Pacific Islands and the two units operationalizing media presence as number of radio receivers.

In the causal analysis of this dimension I encounter a problem with discrimination on the ICRG-variable. Therefore this variable must be excluded from the analysis, since there does not seem to be any other explanatory variable with which it makes sense to merge the ICRG. However, we may note that all of the 13 analyses applying this measure concluded that press freedom leads to less corruption.

Looking at the p-value of the Chi-square-test of the change in -2LL (Table 21), the full model significantly improves the goodness of fit from the reduced model to the full model. However the increase in correct classifications is only from 71.6 percent to 75.3 percent. Let’s turn to the different explanatory variables.

In the sample, analyses applying the CCI and the CPI are much more likely than country- or region specific studies to conclude that press freedom and outreach lead to less corruption. The effects are statistically significant, thus supporting H5. The Other perceptions based-variable also shows increased likelihood of concluding that media presence leads to less corruption, but is not significant thus opposing H5 and supporting H1b. Running a model
with the control variables that have the sufficient amount of units, shows that the year published-variable has a significant effect on conclusions; there is a decrease in odds of 22.7 percent per year closer to today. With this control there is also a substantial increase in the odds of the CCI and the CPI, by 2274.5 percent and 26220.0 percent respectively.

Table 20: Dependent variable: Conclusions on effect of press freedom on corruption

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Full model</th>
<th>With controls</th>
<th>Without outliers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p-value of Wald-test</td>
<td>Odds ratio</td>
<td>p-value of Wald-test</td>
</tr>
<tr>
<td>CPI</td>
<td>0.001</td>
<td>8,800</td>
<td>0,000</td>
</tr>
<tr>
<td>CCI</td>
<td>0.041</td>
<td>4,000</td>
<td>0,002</td>
</tr>
<tr>
<td>Other percept-based</td>
<td>0.846</td>
<td>1,333</td>
<td>0,653</td>
</tr>
<tr>
<td>* Country-or region specific*</td>
<td>Odds: 0.750</td>
<td>1.000</td>
<td>Odds: 0.750</td>
</tr>
<tr>
<td>Communist- or post com</td>
<td>--</td>
<td>--</td>
<td>0,850</td>
</tr>
<tr>
<td>OECD</td>
<td>--</td>
<td>--</td>
<td>0,521</td>
</tr>
<tr>
<td>Year published</td>
<td>--</td>
<td>--</td>
<td>0,065</td>
</tr>
<tr>
<td>Constant</td>
<td>0.514</td>
<td>0.750</td>
<td>0,065</td>
</tr>
</tbody>
</table>

* Reference category

Table 21: Model summaries

<table>
<thead>
<tr>
<th>Whole models</th>
<th>Correct predictions</th>
<th>Chi²-test of -2LL</th>
<th>p-value of HL-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced model</td>
<td>71.6 %</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Full model</td>
<td>75.3 %</td>
<td>13,116 ***</td>
<td>--</td>
</tr>
<tr>
<td>With controls</td>
<td>82.7 %</td>
<td>22,666 ***</td>
<td>0.586</td>
</tr>
<tr>
<td>Reduced model - outliers</td>
<td>82.9 %</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Full model - outliers</td>
<td>82.9 %</td>
<td>2.059</td>
<td></td>
</tr>
</tbody>
</table>

N: 81 N, without outliers: 70

***p<0.01, **p<0.05, *p<0.10

However, as shown in the descriptive analysis, there may be some units that unjustly affect these results. I will therefore run an additional model where I exclude the units from the study of small pacific island states (Larmour and Barcham 2006), and the units where press freedom has been operationalized as radio receivers per capita (Lindstedt and Naurin 2010). In this analysis there is no significant change between the reduced model and the full model, and none of the explanatory variables are statistically significant. The most conservative conclusion that may be derived from this is that the sample is not satisfactorily large to draw any conclusion. A less conservative conclusions may be that there, when removing units that are outliers either because they are context specific or because researchers may have operationalized media presence unconventionally, are no differences between the choices of data on corruption. Table 22 sums up these findings in light of the hypotheses of Chapter 3.
Table 22: Results in light of hypotheses of Chapter 3*

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Operationalization of corruption</th>
<th>Expected effect on relationship between corruption and variables of Chapter 2</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a:</td>
<td>Cross-national survey measure</td>
<td>Lower likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H1b:</td>
<td>Cross-national survey measure</td>
<td>Not lower likelihood for observing effects</td>
<td>Yes</td>
</tr>
<tr>
<td>H2:</td>
<td>Perceptions based measure</td>
<td>Lower likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H3:</td>
<td>Experience based measure</td>
<td>Lower likelihood for observing effects</td>
<td>(Not tested)</td>
</tr>
<tr>
<td>H4:</td>
<td>Cross-national survey measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H5:</td>
<td>Perceptions based measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H6:</td>
<td>Experience based measure</td>
<td>Higher likelihood for observing effects</td>
<td>(Not tested)</td>
</tr>
</tbody>
</table>

* Based on results of model without outliers

5.1.6 The effects of level of economic development on corruption

Looking at Table 23 we see that 407 analyses have been conducted of the relationship between economic development and corruption, thus being the explanatory dimension studied the most. 78.6 percent of the analyses conclude that higher economic development leads to less corruption 18.2 percent conclude that it has no effect and 3.2 percent conclude that it increases corruption.

Table 23: The effects of level of economic development on corruption: data sources and conclusions

<table>
<thead>
<tr>
<th>Higher economic development</th>
<th>CPI</th>
<th>CCI</th>
<th>ICRG</th>
<th>WCR</th>
<th>GCS</th>
<th>GEI</th>
<th>WBES</th>
<th>BI</th>
<th>Country-/reg-spec.</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less corruption</td>
<td>116</td>
<td>115</td>
<td>54</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>23</td>
<td>4</td>
<td>1</td>
<td>320</td>
</tr>
<tr>
<td>Has no effect</td>
<td>21</td>
<td>12</td>
<td>20</td>
<td>13</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>74</td>
</tr>
<tr>
<td>More corruption</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>137</td>
<td>127</td>
<td>84</td>
<td>19</td>
<td>2</td>
<td>4</td>
<td>27</td>
<td>6</td>
<td>1</td>
<td>407</td>
</tr>
</tbody>
</table>

Of the ten analyses applying the ICRG as dependent variable and concluding that higher economic development leads to more corruption seven analyses with different samples are provided by Baksi, Bose and Pandey (2009), and as they also point out, these results are surprising. They control for fuel exports and import as share of GDP and in addition they are panel analyses; this combination may be the reason why these results are obtained. The three remaining are done by Dreher and Schneider (2010) with samples of only high-income countries, something that might indicate that of rich countries, the richest are more corrupt.
The three studies applying the WCR-data concluding that higher economic development leads to more corruption is done by Ades and DiTella (1997a) applying a very small sample of only 32 countries which seems to be somewhat biased towards high-income countries, thus expressing a similar sample bias as the study by Dreher and Schneider.

The data material is too large to discuss in detail all the 74 analyses finding no effect. I will move on to the causal analysis. Again the values for positive effect and no effect will be merged. The *Country- or region specific studies* must be omitted due to discrimination, which is unsurprising since there is only one unit. The *CPI* will serve as reference category (odds: 5.5238). The chance of concluding that higher economic development leads to less corruption is approximately five and a half times higher when applying the CPI than for not doing so. Since, as discussed, some outliers have applied panel analysis, I will include a dummy for whether such analysis has been applied.

### Table 24: Dependent variable: conclusions on effect of economic development on corruption

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Full model</th>
<th></th>
<th>With controls</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p-value of Wald-test</td>
<td>Odds ratio</td>
<td>p-value of Wald-test</td>
<td>Odds ratio</td>
</tr>
<tr>
<td><strong>CPI</strong>*</td>
<td>Odds: 5.5238</td>
<td>1.000</td>
<td>Odds: 5.5238</td>
<td>1.000</td>
</tr>
<tr>
<td>CCI</td>
<td>.198</td>
<td>1.647</td>
<td>.092</td>
<td>1.952</td>
</tr>
<tr>
<td>ICRG</td>
<td>.001</td>
<td>.327</td>
<td>.000</td>
<td>.013</td>
</tr>
<tr>
<td>Other percept.-based</td>
<td>.000</td>
<td>.056</td>
<td>.359</td>
<td>.638</td>
</tr>
<tr>
<td>Experience based</td>
<td>.256</td>
<td>.585</td>
<td>.000</td>
<td>.282</td>
</tr>
<tr>
<td>Panel/time-series</td>
<td>.085</td>
<td>.534</td>
<td>.546</td>
<td>.762</td>
</tr>
<tr>
<td>Årstall</td>
<td></td>
<td></td>
<td>.000</td>
<td>.833</td>
</tr>
<tr>
<td>GDPstørre24000USD</td>
<td>.495</td>
<td></td>
<td>.588</td>
<td></td>
</tr>
<tr>
<td>(Post-)communist</td>
<td>.922</td>
<td></td>
<td>1.083</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>.000</td>
<td>6.193</td>
<td>.000</td>
<td>2.073E+160</td>
</tr>
</tbody>
</table>

* *Reference category*

### Table 25: Model summaries

<table>
<thead>
<tr>
<th></th>
<th>Correct predictions</th>
<th>Chi²-test of -2 LL</th>
<th>p-value of HL-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced model</td>
<td>78.6 %</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Full model</td>
<td>81.3 %</td>
<td>56,740 ***</td>
<td>--</td>
</tr>
<tr>
<td>With controls</td>
<td>84.5 %</td>
<td>74,178 ***</td>
<td>.019</td>
</tr>
</tbody>
</table>

N: 407

**p<0.01, *p<0.05, *p<0.10**

19 Hong Kong, India, Japan, Korea, Malaysia, Singapore, Thailand, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, Canada, Mexico, United States, Brazil, Australia, New Zealand, Indonesia, Hungary.

20 I attempted to run both ordinal and multinomial regression, however, the ordinal regression did not satisfy the assumption of parallel lines and the multinomial regression faced issues with discrimination.
The chi square test of the change in -2LL is significant (Table 25). The change of prediction accuracy is (only) 2.7 percentage points from 78.6 percent to 81.3 percent.

Table 24 shows that applying the CCI does not provide results significantly different from the results obtained applying the CPI; in the sample there is an increase in odds of 64.7 percent. Applying the ICRG gives a significantly lower likelihood of concluding that higher economic development has no effect; the change in odds is 67.3 percent. This again may support Lambsdorffs (Lambsdorff 2004) doubt in the measure’s grasp of the concept of corruption. The variable Other perceptions based measures also significantly decreases the likelihood for concluding that higher economic development leads to less corruption, with a decrease in odds of 94.4 percent. However, there are few units where these measures have been applied, thus hard conclusions should be avoided.

Applying experience based measures, interestingly, and counter to what should be expected from Treisman (2007) and the assumption that these measures to a lesser degree should be victims of measurement bias, there is no statistically significant change in the odds between the CPI and experience based measures; in the sample the odds decreases by 34.4 percent. Thus hypothesis H3 is not supported.

The dummy for panel study is significant, reducing the odds for concluding that higher economic development leads to less corruption by 46.6 percent.

I run a second model including the controls that have sufficient amount of units put together with the dependent variable. The Chi-square-test of the change in -2LL shows that the controls significantly contribute to the explanatory power. The percentage of correct predictions has increased to 84.5 percent.

Looking at the coefficients of the explanatory variables in the model with controls shows that the CCI compared to the CPI now has a significant effect on the likelihood to conclude that higher economic development leads to less corruption, with an increase in odds of 95.2 percent. The difference between the CPI and otheropinion is no longer statistically significant. The experience based measure now gives a statistically significant change in odds compared to that of the CPI; the odds for concluding that higher economic development leads to less corruption has decreased by 71.8 percent compared to the CPI. This supports H3.

With these controls, applying the ICRG-measure has now even lower likelihood for concluding that there is an effect; the decrease in odds is of 98.7 percent compared to that of the CPI. Thus interpretation in light of the hypothesis concluding that perceptions based
measures should diverge, is somewhat ambiguous since the perceptions based measures provide so diverging results. The panel analysis variable has now lost its effect.

The only of the control variables that have significant effect on conclusions is the year in which the analysis have been published; the odds of concluding that higher economic development leads to less corruption has decreased with an average of 16.7 percent for each year later in time. This may indicate that the field is experiencing a slide away from something that there has been a quite extensive agreement on, at least amongst researchers applying the cross-national measures. Note that the Hosmer-Lemeshow test is significant indicating that the relationship between the year published-variable and the dependent variable does not conform to a logistic curve. Running the model again with quadratic term and the log (10) of the variable, respectively, provide the same result. That the test is significant may indicate that this development has not been stable over time.

Table 26 sums up the findings of this section in light of the hypotheses formulated in Chapter 3.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Operationalization of corruption</th>
<th>Expected effect on results</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a:</td>
<td>Cross-national survey measure</td>
<td>Lower likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H1b:</td>
<td>Cross-national survey measure</td>
<td>not lower likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H2:</td>
<td>Perceptions based measure</td>
<td>Lower likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H3:</td>
<td>Experience based measure</td>
<td>Lower likelihood for observing effects</td>
<td>Yes</td>
</tr>
<tr>
<td>H4:</td>
<td>Cross-national survey measure</td>
<td>Higher likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H5:</td>
<td>Perceptions based measure</td>
<td>Higher likelihood for observing effects</td>
<td>Partly</td>
</tr>
<tr>
<td>H6:</td>
<td>Experience based measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
</tbody>
</table>

5.1.7 The resource curse

The sample of this dimension consists of 112 results of which the results are highly diverging (Table 27). 50.9 percent of the analyses have concluded that natural resource dependency leads to more corruption while 48.2 percent have concluded that it has no effect. One study (0.9 percent) has concluded that natural resource dependency leads to less corruption. The dimension has been studied with close to the full width of data sources. The category of studies with the highest degree of agreement is the WCR; these six results come from Ades and di Tella (1999). These analyses have a relatively small sample consisting of only 31
countries\textsuperscript{21}, which again seems biased towards rich western countries, something that may be related to the results.

Table 27: The effects of natural resource dependency: Summary of data sources and conclusions

<table>
<thead>
<tr>
<th>Natural Resource dependency gives</th>
<th>CPI</th>
<th>CCI</th>
<th>ICRG</th>
<th>WCR</th>
<th>GCS</th>
<th>GEI</th>
<th>WBES</th>
<th>BI</th>
<th>Country- /reg-spec.</th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>More corruption</td>
<td>15</td>
<td>18</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>9</td>
<td>57</td>
<td>50.9%</td>
</tr>
<tr>
<td>Has no effect</td>
<td>12</td>
<td>7</td>
<td>11</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td>54</td>
<td>48.2%</td>
</tr>
<tr>
<td>Less corruption</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.9%</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>25</td>
<td>21</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>3</td>
<td>9</td>
<td>112</td>
<td>100%</td>
</tr>
</tbody>
</table>

Between the Country- or region specific studies there also is a high degree of agreement. The one study concluding that natural resources may lead to less corruption is by Brownsberger studying Nigeria (1983). Brownsberger argues that since the country was troubled with poverty and inequality, officials were more prone to nepotism and extraction. The author claims that oil revenues would, and had already started to, reduce corruption in Nigeria through aiding poverty and inequality and thus letting those entering government already have fulfilled their material needs. These conclusions must be seen in light of the fact that the article is written early in the theoretical development of the field. Regarding the data available at the time the author illustratively comments that “I have been tracing shadows in this article” (Brownsberger 1983). Talking of outliers between those concluding that there is no effect and a positive effect, makes little sense since they are almost equal in numbers. I will therefore turn to the causal analyses.

The chi square test of the change in -2LL shows that the explanatory variables significantly contribute to the goodness of fit (Table 29); correct classifications increase from 51.3 percent to 63.7 percent.

The variable other perceptions based measures must be excluded due to problems with discrimination. Country- or region-specific studies will serve as reference category (odds: 9.000): within this category there is nine times higher chance for concluding that natural resource dependency leads to more corruption than for not doing so.

\textsuperscript{21} Hong Kong, India, Japan, Korea, Malaysia, Singapore, Thailand, Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, The Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, Canada, Mexico, United States, Brazil, Australia, New Zealand, Indonesia.
Looking at the coefficients (Table 28) we see that the odds of concluding that natural resource dependency leads to more corruption decreases by 12.5 percent when applying the CPI compared to the reference category, *Country- or region specific studies*. Applying the CCI gives increase in odds of 80 percent. Applying the ICRG gives a decrease in odds by 36.4 percent. However, none of these three effects are statistically significant, thus supporting H1b.

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Full model</th>
<th>With controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p-value of</td>
<td>Odds ratio</td>
</tr>
<tr>
<td></td>
<td>Wald-test</td>
<td></td>
</tr>
<tr>
<td>CPI</td>
<td>.831</td>
<td>.875</td>
</tr>
<tr>
<td>CCI</td>
<td>.376</td>
<td>1.800</td>
</tr>
<tr>
<td>ICRG</td>
<td>.493</td>
<td>.636</td>
</tr>
<tr>
<td>Experience-based</td>
<td>.020</td>
<td>.194</td>
</tr>
<tr>
<td><em>Country- or reg.-spec</em></td>
<td>Odds: 9.000</td>
<td>1.000</td>
</tr>
<tr>
<td>High-income</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>(Post)-Communist</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Sub-Saharan countries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year published</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Constant</td>
<td>.469</td>
<td>1.429</td>
</tr>
</tbody>
</table>

* Reference category

<table>
<thead>
<tr>
<th></th>
<th>Correct predictions</th>
<th>Chi²-test of -2LL</th>
<th>p-value of HL-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced model</td>
<td>51.3 %</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Full model</td>
<td>63.7 %</td>
<td>13,644 ***</td>
<td>--</td>
</tr>
<tr>
<td>With controls</td>
<td>69.0 %</td>
<td>26,775 ***</td>
<td>0.000</td>
</tr>
</tbody>
</table>

N: 112

***p<0.01, **p<0.05, *p<0.10

Applying an *experience based measure* as the dependent variable significantly decreases the likelihood of concluding that natural resource dependency leads to more corruption; the decrease in odds is 80.6 percent. This may indicate that natural resource countries are more attractive to business people and thus they let this attractiveness affect their responses, thus not necessarily providing information of corruption, but rather general attractiveness. It may also be that these measures simply are noisier, thus supporting H3.

I now run a model including the controls that have a satisfying number of units. Adding controls significantly improves the goodness of fit of the model; correct classifications increase from 63.7 percent to 69.0 percent (Table 29).

In the model with controls the main findings remain the same regarding which variables are statistically significant and the signs of the effects (Table 28). In the sample,
analyses conducted focusing on countries with a high GDP per capita are more likely to conclude that natural resource dependency leads to more corruption; the effect is not significant.

Interestingly studies conducted with focus on communist- or post-communist countries are significantly more likely to conclude that natural resource dependency leads to more corruption; the odds increases by 755.3 percent.

Focusing on the Sub-Saharan region increases the odds of concluding that natural resource dependency leads to more corruption by 1372.2 percent. The effect is marginally not significant.

The tendency to conclude that natural resource dependency leads to more corruption has increased with an average of 12.9 percent per year. This effect is statistically significant. However, the Hosmer-Lemeshow test is also significant; again the year-variable is the only variable that may be the problem since it is the only continuous variable. Replacing it with the log (10) of the variable does not solve the problem. Neither does including a quadratic calculation of the variable. Thus it must be interpreted as it is, but with caution. One possible answer to this is that the increase in odds is not steady but varies from year to year.

Table 30 sums up the findings of this dimension in light of the hypotheses formulated in Chapter 3.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Operationalization of corruption</th>
<th>Expected effect on results</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a:</td>
<td>Cross-national survey measure</td>
<td>Lower likelihood for observing effects</td>
<td>Partially</td>
</tr>
<tr>
<td>H1b:</td>
<td>Cross-national survey measure</td>
<td>Not lower likelihood for observing effects</td>
<td>Partially</td>
</tr>
<tr>
<td>H2:</td>
<td>Perceptions based measure</td>
<td>Lower likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H3:</td>
<td>Experience based measure</td>
<td>Lower likelihood for observing effects</td>
<td>Yes</td>
</tr>
<tr>
<td>H4:</td>
<td>Cross-national survey measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H5:</td>
<td>Perceptions based measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H6:</td>
<td>Experience based measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
</tbody>
</table>

5.2 Discussion of findings

I have analyzed seven dimensions commonly applied by researchers seeking to explain high-level corruption. Five of these dimensions have provided sufficient data to conduct analyses where I have sought to discover whether the disagreements of the effects on corruption may be caused by which data on corruption that have been applied. The results are somewhat diverging; first, the descriptive analyses have shown that there is high degree of agreement on
some of the explanatory dimensions while on others the disagreement is almost as great as it could be. The causal analyses also show diverging results as to whether data sources systematically determine conclusions. The results of both the descriptive analyses and the causal analyses are summed up and sought put together in Table 31 (below).

There is high agreement that economic liberalizations and regime transitions, respectively, lead to more corruption, while there is high agreement that higher economic development and higher media presence, respectively, are associated with less corruption. There seems to be very high disagreement regarding whether natural resource dependency leads to more corruption or not. There also seems to be high disagreement on whether present time degree of democracy and strength of democratic traditions lead to less corruption. Democracy has been strongly promoted in order to enhance the security and well-being of people in authoritarian regimes; firstly this process in itself is probably highly damaging, secondly not even once democracy has been established is it clear that it helps decreasing corruption, and thirdly there is no guarantee that it will do so in the long run neither. In light of the knowledge we have on the consequences of corruption it is not obvious that this is a right priority (Rothstein 2010), at least not if we cannot find ways to achieve democratization without increasing corruption. That liberalization processes, at least in the short run, have failed in deterring corruption also seems clear, with the potential of similar negative consequences. An interesting question is whether it is possible to achieve economic liberalization and its potential beneficial consequences without at the same time increasing corruption.

A common objection to the claim that higher economic development leads to less corruption is that the direction of causality may go the other way; that corruption leads to lower economic development. If the data had allowed, more panel analyses would have been useful on this matter would have been useful. Per today, even though there are researchers applying i.e. the CPI (Baksi et al. 2009; Chowdhury 2004) in panel studies, the ICRG is the only measure that is suited for panel analyses since it has been published with the same methodology for many years.

Improving the freedom and the presence of the media may be a policy measure that in is highly recommendable, since the agreement that it lowers degree of corruption is relatively high.
When it comes to the suggested causal relationships between the different data sources on corruption and results, I first would disagree with i.e. Philip (2006) that the perceptions based cross-national measures should be disregarded altogether. Looking at Table 31 we may observe that not one significant divergence in results has been found between the CPI and country- and region-specific studies. Determining whether they both are marked by

---

**Table 31: Summary of all findings***

<table>
<thead>
<tr>
<th>Economic liberalization</th>
<th>Democracy</th>
<th>Regime transition</th>
<th>Strong democratic tradition</th>
<th>Media presence</th>
<th>Economic development</th>
<th>Natural resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Majority's conclusion</td>
<td>More</td>
<td>Less</td>
<td>More</td>
<td>No effect</td>
<td>Less</td>
<td>Less</td>
</tr>
<tr>
<td>Degree of agreement</td>
<td>86.4 %</td>
<td>54.4 %</td>
<td>90.6 %</td>
<td>50.8 %</td>
<td>79.8 %</td>
<td>78.6 %</td>
</tr>
</tbody>
</table>

**Divergence in findings caused by data sources**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>yes/no</td>
<td>yes/no</td>
<td>yes/no</td>
<td>yes/no</td>
<td>yes/no</td>
<td>yes/no</td>
<td>yes/no</td>
</tr>
<tr>
<td>CPI</td>
<td>-</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>CCI</td>
<td>-</td>
<td>No</td>
<td>-</td>
<td>Yes/ more likely</td>
<td>Yes/ more likely</td>
<td>No</td>
</tr>
<tr>
<td>ICRG:</td>
<td>-</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>Not determined***</td>
<td>Yes/ less likely</td>
</tr>
<tr>
<td>Other percept.</td>
<td>-</td>
<td>Yes/ less likely</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Experience-based</td>
<td>-</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>Yes/ less likely</td>
<td>Yes/ less likely</td>
</tr>
<tr>
<td>Country- or region specific</td>
<td>-</td>
<td>No</td>
<td>-</td>
<td>No</td>
<td>-</td>
<td>No</td>
</tr>
<tr>
<td>High income</td>
<td>-</td>
<td>Yes/ less likely</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
</tr>
<tr>
<td>(Post)- communist</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>No</td>
<td>Yes/ more likely</td>
</tr>
<tr>
<td>OECD</td>
<td>-</td>
<td>Yes/ more likely</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>-</td>
</tr>
<tr>
<td>Panel analysis</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>-</td>
</tr>
<tr>
<td>Sub-Saharan Year</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
<td>-</td>
</tr>
<tr>
<td>published</td>
<td>-</td>
<td>Yes/ more likely</td>
<td>-</td>
<td>Yes/ less likely</td>
<td>Yes/ less likely</td>
<td>Yes/ more likely</td>
</tr>
</tbody>
</table>

* This table presents results for models with controls.
** For this dimension I here present the results for the model without outliers; with outliers the results are very different.
*** This variable could not be entered into the regression due to discrimination; all 13 analyses applying this measure concluded that stronger media presence gives less corruption

---
correctness or wrongness is a different task, but they are at least marked by some of the same. Within the dimensions where both of these two data sources have been applied, most of the other data sources also provide similar likelihoods of observing the different effects. Thus it would not be out of line suggesting that results derived applying the CPI and country- and region specific data sources, respectively, not only are alike, but that there also is something correct about them both.

One divergence has been found between the CCI and country- and region specific studies; on the democratic traditions dimension. Since the CCI here is the only data source producing diverging data source, it is at least more likely that it is the CCI there is something wrong with than all of the other data sources. Alternatively the divergence may be caused by features of the studies applying the CCI that I have not been able to uncover.

Looking at the “CCI-row” we may observe that the CCI is also denoted as diverging from the rest on the economic development-dimension. However, here the divergence is quite extensive between the other data sources as well. While the CPI served as reference category, the only other category of data sources that agreed with the CPI on this dimension is the other perceptions based measures-variable. The CCI diverges making it more likely to conclude that higher economic development leads to less corruption, while the ICRG and experience base measures makes it less likely. Thus it is hard, on this dimension to point out diverging data sources, since most of them seem to be. However, that the CCI-variable provides findings in the way of “more likely” to conclude in accordance with the expected effects on this dimension and the democratic traditions-dimension may indicate systematic measurement error in accordance with Treisman’s suggestion. However, the CCI-variable also provides non-divergent findings on three of the dimensions.

The ICRG “agrees” with most of the other data sources on all dimensions where it has been included in causal analyses except the already mentioned economic development-dimension. Thus I find it hard to draw this measure strongly in doubt.

The variable other perceptions based measures has been included with relatively few units in all the causal analyses. Thus it is hard to draw solid conclusions. The measure “agrees” with the CPI-variable on the economic development-variable, provides results conforming to the rest on the media presence-dimension, and results diverging from the rest on the democracy-dimension.

The findings indicate that there may be unsystematic measurement errors in the
experience based cross-national data. On the natural resource dimension and the economic development dimension the *experience based measures*-variable makes it less likely to conclude that the dimensions have the expected effects. On the natural resources dimension, the experience based data provide a likelihood of such conclusions that is diverging from all of the others.

Table 32: Summing all findings in light of hypotheses

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Operationalization of corruption</th>
<th>Expected effect on results</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1a:</td>
<td>Cross-national survey measure</td>
<td>Lower likelihood for observing effects</td>
<td>Yes, partly</td>
</tr>
<tr>
<td>H1b:</td>
<td>Cross-national survey measure</td>
<td>Not lower likelihood for observing effects</td>
<td>Yes, partly</td>
</tr>
<tr>
<td>H2:</td>
<td>Perceptions based measure</td>
<td>Lower likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H3:</td>
<td>Experience based measure</td>
<td>Lower likelihood for observing effects</td>
<td>Yes</td>
</tr>
<tr>
<td>H4:</td>
<td>Cross-national survey measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H5:</td>
<td>Perceptions based measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
<tr>
<td>H6:</td>
<td>Experience based measure</td>
<td>Higher likelihood for observing effects</td>
<td>No</td>
</tr>
</tbody>
</table>

Thus the overall picture in light of the hypotheses (see Table 32) is, firstly, that H1a is partly supported since they do not consistently provide lower likelihood for observing effects; however, in three of the five dimensions there is at least one cross national measure supporting H1a. H1b is also partly supported since there also are many instances where the cross national measures are not less likely to provide effects. Thus the conclusion of whether H1a or H1b is supported is somewhat ambiguous.

H2 is not supported, since the overall picture is not that perceptions based measures give lower likelihood for observing effects; rather the findings are slightly diverging between the perceptions based measures. H3 is supported; there seems to be a tendency towards lower likelihood for observing effects applying experience based measures. H4 is not supported; there is no general tendency towards cross national measures in general giving higher likelihood for observing effects. Neither is H5 supported; the likelihood for observing effects is not generally higher applying perceptions based measures: two findings point in the direction of supporting H5, two findings point in the complete opposite direction, while the majority of the findings show that the effects are generally the same across measures, thus *not* supporting H5. Finally, H6 is not supported; in no instances do application of experience based measures give higher likelihood for observing effects.

“[B]uilding governance indicators constitute an ongoing project”, claims Apaza
(Apaza 2009). There seems, in my opinion, to be little reason to conclude that the project of developing such measures resembles a sinking ship as some would claim. These findings to some extent poses a challenge to Treisman’s claim that the perceptions based measures are more problematic than the experience based measures; if one of these two categories of corruption data are more problematic than the rest, the findings of this thesis points in the direction that the experience based measures are more problematic. This may also be said of the CCI.
6. Concluding remarks

6.1 Summary of approach and findings
The research question of this thesis is twofold: *To what degree is there agreement on causes of high-level corruption in the field? Is there a causal relationship between the data sources researchers have applied and the results they obtain?*

I have responded to these questions by delimiting the scope of the study to a few explanatory dimensions commonly suggested and analyzed: based in these dimensions I have conducted an extensive literature search in ISI Web of Knowledge. Reading and categorizing large amounts of publications I have constructed a data set based on the effect researchers have found that these explanatory dimensions have on corruption, which data on corruption they applied, which geographical area they studied and when they did so.

Analyzing these data I have found that there is high degree of agreement that economic liberalization and regime transitions lead to more corruption and high agreement that the presence of a free media leads to less corruption. Concerning whether present level of democracy and long democratic traditions, respectively, lead to less corruption researchers disagree greatly. The same is true of whether natural resource dependency leads to more corruption.

The general picture of whether data sources on corruption explain the divergences in conclusions is that they don’t. There are, however, a few exceptions. Application of the Control of Corruption Index of the World Bank gives higher likelihood of concluding that stronger democratic traditions and higher economic development, respectively, lead to less corruption. Application of the International Country Risk Guide corruption index of the Political Risk Services Group give lower likelihood of concluding that higher economic development leads to less corruption. Experience based measures seem to provide diverging results in more instances (two out of three). Application of these give lower likelihood of concluding that natural resources leads to more and higher economic development leads to less corruption. Thus there is only one dimension, level of democracy, where experience based measures do not diverge. The merged variable of other perceptions based data show diverging results on the democracy dimension.
6.2 Implications of findings

6.2.1 Economic liberalization
Most economic liberalization processes so far have gone astray, leading to higher degree of corruption. It seems uncontroversial to state that these processes, at least in the 80’s and 90’s, gave unintended and possibly severe consequences.

Organizations such as the World Bank and the International Monetary Fund still promote economic liberalization processes. I.e. the World Bank’s many “[t]oolkits for policymakers and reform leaders” prescribe privatization as possible solutions to many political and administrative challenges i.e. for improving waste services, ports, highways, water and sanitation, bus services and telecommunications. However, they now also acknowledge that public service delivery, even public monopoly, may be a solution in some instances i.e. for bus services (World Bank 2011b). Stiglitz describes the privatization promotion of the i.e. The World Bank and The International Monetary Fund as being of a more orderly kind now than it was in the 80’s and 90’s also focusing on how privatization should be done and that state institutions must be developed accordingly (Stiglitz 2006). Some of these policies now also underline that there may be risk of increased corruption and suggests policies for preventing this, i.e. by establishing multi-sectorial regulatory agencies (World Bank 2001:28). The findings of this thesis underline the importance of considering economic liberalization versus non-private solutions with soberness.

These findings also underline the importance of sound regulation and auditing of the private sector, which also is underlined in the United Nations Convention Against Corruption (United Nations 2004). That it also is in the interest of the private sector to engage in anti-corruption work has also been raised to attention. This has been expressed i.e. in resolution 3/2 of the Conference of the States Parties to the United Nations Convention Against Corruption (United Nations 2009). That privatization per definition is positive for social well-being is, however, a notion that seems to have been proven wrong.

6.2.2 Democracy: transition to-, presence of-, and experience with-
It has come to be recognized that the socially negative consequences of absence of democracy are negligible compared to the negative consequences of the presence of corruption (Rothstein 2010). In addition it seems highly questionable whether good governance can be promoted through democratization; at least that has generally not been the case so far. Transitions have
in an overwhelming amount of cases lead to more corruption. There is no guarantee that consolidation of democracy in itself will help that situation. Also knowing that the prospects for improvement in the long-run are not overwhelmingly positive, it is justified to question how we should add this up with the many democratization campaigns of international organizations and western governments; i.e. the democratization-engaged UN-programs like the Development Programme and the Democracy Fund (United Nations 2011), or national programs such as that of the United States Agency for International Development or Norwegian Agency for Development Cooperation (USAID 2010; NORAD 2011).

It seems as a good start to seek to figure out whether corruption prevention and democratization are projects that stand in opposition to one another or if this only is the consequence of the way in which democratization processes have been conducted so far: is it possible to at the same time promote democratization and good governance? Have democratization processes so far lead to more corruption because they were poorly conducted?

The attitude so far seems to be that these are policies that should coexist; i.e. the African Union Convention on Preventing and Combating Corruption departs from- and restates the goal of democratization formulated in the Sixty-fourth Ordinary Session of the Council of Ministers (1996) (African Union 2003), thus indicating that democratization is seen as a tool to prevent corruption, a suggestion to which there seems to be little empirical support. This is not to say, of course, that democracy may not be of great value in itself.

### 6.2.3 Presence of a strong and independent media

The findings in this thesis underline the importance of the presence of a strong media. This is a principle that also has come to be widely recognized in anti-corruption work. For instance this is expressed in The African Union Convention which aims to “[c]reate an enabling environment that will enable civil society and the media to hold governments to the highest levels of transparency and accountability in the management of public affairs” (African Union 2003: Article 12). Intentions which are also duly stated in the Southern African Development Community’s Protocol Against corruption (SADC 2001: Article 4). Establishing rights to access of the public to appropriate information is obviously a crucial part of this, which is i.e. stated in the OECD Anti-Corruption Action Plan for Asia and the Pacific (2001:Pillar 3). The importance of freedom of information and a free media is also stated in The Twenty Guiding
Principles for the Fight Against Corruption of the Council of Europe (Council of Europe: Committee of Ministers 1997). Thus, here international policy seems to be well in line with the findings made by researchers.

6.2.4 Level of economic development
The immediate policy implications of economic development leading to less corruption are not obvious. Also, caution should also be taken regarding the overall conclusion that it does lead to less corruption. Some doubt has been presented as to whether the direction of causality goes from level of economic development to corruption or if it may completely or partly go in the opposite direction (i.e. Guetat 2006). Most studies of this relationship are cross sectional, thus not being able to uncover an eventual opposite direction of causality. However, the findings have also been supported in panel studies.

Other factors related to the level of economic development and whether these may lead to less corruption need to be dug into more deeply; i.e. level of education (Mocan 2008), overall human development including the distribution of material and monetary resources (You and Khagram 2005; Welzel 2002) and the development of self-expression values (Welzel 2002; Kim 2010).

6.2.5 Natural resource dependency
About half of the studies looking into the effect of natural resource dependency on corruption have found that it does indeed increase corruption. No studies from recent years have found that it leads to less. It is clear that efforts must be made to develop institutions that let peoples of resource rich countries benefit in greater accordance with the potential natural resources provide.

The efforts to clean the diamond industry that started with the Kimberley Process Certification Scheme has made quite some development in ridding the industry of corruption, amongst more severe issues. That is not to say that there is not a lot of work left to do (Taylor 2011). Efforts are also made to bring greater transparency to the oil and gas sectors i.e. with the G8 2009 Leaders Declaration (article 43) which restates a decade-long effort of the G8 to promote transparency (Transparency International-USA 2011). However, the issue of natural resources and corruption has been given surprisingly little attention in anti-corruption
conventions and declarations (i.e. United Nations 2004, 2000; OECD 2010; African Union 2003; Southern African Development Community 2001; Organization of American States 1996; Council of Europe 1999b, 1999a). Many countries would probably benefit from increased attention to these matters.

6.2.6 The data issue

“Crossnational measures of corruption are riddled with problems”, claim Hawken and Munck (Hawken 2009:4). As this statement illustrates the debate on how to study corruption has been heated (also i.e. Gingerich 2010:364; Treisman 2007; Philip 2006). This thesis has not comprised the whole range of issues with which these data may be concerned. Some of the most central arguments of the debate have been presented. Foremost I have applied these arguments to present several hypotheses on the nature of the measurement errors with which the data may be troubled and possible consequences of these.

The findings of the effects of data choices on conclusions incline me to suggest, opposing to Treisman (Treisman 2007), that it is the experience based measures that are more troubled with measurement error; whether these are unsystematic and thus give weaker regression models, or if they are systematic in the sense that business people are likely to underreport incidents of corruption, is difficult to determine.

"Having some data, even if of poor quality, is a less grave problem than having no data at all” (Hawken and Munck 2009:21). The findings of this thesis restate this, perhaps on a slightly more positive note. The status on the cross national data is not one of hopelessness. However, the differences between the different perceptions based measures should be looked into; from what do they stem? Even if, in comparison with country- and region specific studies, the CPI seems as the more reliable of the three most commonly applied perceptions based measures, more research should be done, i.e. of the kind in conducted in this thesis to further evaluate the findings also when applying this index.
7. Literature


Council of Europe: Committee of Ministers. 1997 "RESOLUTION (97) 24: ON THE TWENTY GUIDING PRINCIPLES FOR THE FIGHT AGAINST CORRUPTION ": Adopted by the Committee of Ministers on 6 November 1997 at the 101st session of the Committee of Ministers.


Hawken, Angela, Munck, Gerardo L. 2009. "Do You Know Your Data? Measurement Validity in Corruption Research." Authors are professors at Pepperdine University and University of Southern California, respectively.


Henley, Jon. 13.11.2003. "Gigantic sleaze scandal winds up as former Elf oil chiefs are jailed." The Guardian.


Norris, Pippa. 2009 "Cross sectional data set."


Skog, Ole-Jørgen. 2007. Å FORKLAARE SOSIALE FENOMENER. En regresjonsbasert tilnærming. 2nd ed. Norway: Oslo: Gyldendal Norsk Forlag AS.


Treisman, D. 2007. "What have we learned about the causes of corruption from ten years of cross-national empirical research?" *Annual Review of Political Science* 10:211-44.


8. Appendix

Table 33: Overview of publications and results included in the data set

<table>
<thead>
<tr>
<th>Number of units from article</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Author(s)</td>
</tr>
<tr>
<td>---</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Davis, D. E.</td>
</tr>
</tbody>
</table>


1 Economy, E. 2004. "Don't break the engagement." *Foreign Affairs* 83 (3):96-.


<p>| 1 | Harriss-White, B. 1996. &quot;Liberalization and corruption: Resolving the paradox (a discussion based on south Indian material).&quot; <em>Ids Bulletin-Institute of Development Studies</em> 27 (2):31-&amp;. |</p>
<table>
<thead>
<tr>
<th></th>
<th>Author(s)</th>
<th>Year</th>
<th>Title</th>
<th>Journal Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Martinez, G. X.</td>
<td>2006</td>
<td>&quot;The political economy of the Ecuadorian financial crisis.&quot;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Author(s)</td>
<td>Title</td>
<td>Journal</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Author(s)</td>
<td>Year</td>
<td>Title</td>
<td>Journal/Source</td>
</tr>
<tr>
<td>---</td>
<td>-----------</td>
<td>------</td>
<td>-------</td>
<td>----------------</td>
</tr>
</tbody>
</table>


1 Wong, K. Y. 1992. "INFLATION, CORRUPTION, AND INCOME-
Articles that from their headline and/or abstract indicated relevance, but to which I did not have access through the University of Bergen network:


The following article was omitted because the operationalization of the dependent variable is not satisfactorily explicated: