THE DETERMINANTS OF CORRUPTION

A Literature Survey and New Evidence

Harry Seldadyo*
Jakob de Haan†

Abstract

This paper examines 70 eonomic and non-economic determinants of corruption. Using Factor Analysis technique, we generate five new indexes on the basis of these determinants. Using Extreme Bound Analysis we examine the robustness of the determinants as well as the new indexes. We find that one of the generated-indexes, namely regulatory capacity, is the most robust variable in explaining corruption.

Keywords : corruption, Factor Analysis, Extreme Bounds Analysis
JEL code : P16, D73

Paper Prepared for the 2006 EPCS Conference, Turku, Finland, 20-23 April 2006

*The Presenter. This paper is not to compete the Knut-Wicksell Prize.

†Address of Corresponding Author: Jakob de Haan, Faculty of Economics, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands, tel. 31-50-3633706, fax 31-50-3633720, email: Jakob.de.Haan@rug.nl

"Let's not mince words ...
We need to deal with the causes of corruption"
James Wolfenson
(*The Economist*, 4 June 2005, p. 66.)

1 Introduction

Corruption is the misuse of entrusted authority for private benefit. This phenomenon is usually found in the public sector as it primarily involves government officials.[1] In the words of Nye (1967, p. 417), it is "endemic in all governments". Thus, as widely recognized, corruption is probably as old as government itself. According to Glynn et al. (1997, p. 7) "... no region, and hardly any country, has been immune."

Corruption affects almost all parts of society. Like a cancer, as argued by Amundsen (1999, p. 1), corruption "eats into the cultural, political and economic fabric of society, and destroys the functioning of vital organs". Transparency International (TI) regards corruption as "... one of the greatest challenges of the contemporary world. It undermines good government, fundamentally distorts public policy, leads to the misallocation of resources, harms the private sector and private sector development and particularly hurts the poor."[2] The World Bank (WB) has identified corruption as "the single greatest obstacle to economic and social development. It undermines development by distorting the rule of law and weakening the institutional foundation on which economic growth depends."[3]

Corruption has also attracted attention in the academic arena; not only in economics, but also in sociology, political science, law, etc. In the words of Andvig (1991, p. 58), it is "a meeting place for research from the various disciplines of the social sciences and history". Thus, research in this subject is basically multi- and interdisciplinary and includes detailed descriptions of corruption scandals, country cases, and cross-country studies. It also ranges from theoretical models to empirical investigations.

During the last two decades, various organizations have collected and published data on corruption. They are drawn from two forms of data sources: poll-based data (primary source) and poll-of-polls-based data (secondary source). Data on corruption are usually expressed on some scale reflecting the perception of respondents, therefore most corruption indicators are not about the actual level of corruption, but about perceived corruption.

A good example of the first type of data —poll-based data— is the International

[1] There are, however, also various forms of corruption in the private sector. Bowles (2000) lists some of them including insider trading, collusion in asset valuation, and 'information brokerage'.

[2] www.transparency.org/speeches/pe_carter_address.html.

[3] www1.worldbank.org/publicsector/anticorrupt/index.cfm.

Country Risk Guide data set covering almost 150 countries since the beginning of the 1980s. Meanwhile, the most popular poll-of-polls-based data —the second type of data— is perhaps the Transparency International data set, calculated on the basis of indexes drawn from various corruption surveys around the globe done by a number of organizations (Lambsdorff, 2000, 2001a, 2002, 2003, and 2004a).[4] Recently, the World Bank has also produced corruption data as part of a governance index also using data collected from various international polls (Kaufmann and Kray, 2002a and 2002b; Kaufmann, et al., 1999, 2000, 2003, and 2005; Kaufmann, et al., 1999).[5]

Updating the surveys of Andvig et al. (2000) and Jain (2001), this paper first reviews and then extends research on the causes of corruption. The reason is straightforward. Since corruption deteriorates the performance of nations (Mauro, 1995; Tanzi and Davoodi, 1997; Gupta et al., 1998; Lambsdorff, 2001b), the determinants of corruption are of considerable importance. Many possible causes of corruption have been suggested in the literature and this paper critically examines these causes. The rest of this paper is constructed as follows. In section 2 we discuss the concept of corruption and how to measure it, while in section 3 we review research on the causes of corruption. In section 4 we present the outcomes of our factor analysis, while in section 5 we outline our methodology and present new evidence. The last section offers some concluding comments.

[4]The complete series is available at www.icgg.org/corruption.index.html.

[5]A series of papers by Daniel Kaufmann and his colleagues at the World Bank since 1999 explains the methodological construction of the index. The complete list is provided in www.worldbank.org/wbi/governance/wp-governance.html.

2 Corruption: Concept and Measurement

In the *Oxford Advanced Learner's Dictionary* (2000, p. 281) corruption is described as: [1] dishonest or illegal behaviour, especially of people in authority, [2] the act or effect of making somebody change from moral to immoral standards of behaviour. Thus, corruption includes three important elements, namely morality, behaviour, and authority. Gould (1991, p. 468) explicitly defines corruption as a moral problem, i.e., corruption is "an immoral and unethical phenomenon that contains a set of moral aberrations from moral standards of society, causing loss of respect for and confidence in duly constituted authority".

However, viewing corruption merely as problems of morality and behaviour tends to individualize a social phenomenon and to simplify it as only 'good' or 'bad' phenomenon; thus it ignores the socio-political context of corruption. To exist, corruption should be supported by discretionary power, economic rents, and a weak judicial system (Jain, 2001). Discretionary power relates to authority to design and administer regulations, which, in turn, is accompanied by the presence of economic rents associated with power. Meanwhile, a weak judicial system refers to low probability of detection and penalty. Even in the absence of a moral problem, the combination of rent, power, and a weak (or even failure of the) judicial system is enough for corruption to exist.

How can corruption be measured? Even though there are numerous journalistic accounts of corruption[6] it is still difficult to precisely estimate the extent of corruption. However, some researchers have tried to estimate corruption.[7] In their studies, corruption is calculated on the basis of micro level data, like data on infrastructure projects or data drawn from firm-level surveys. Unfortunately, these data do not enable a comparative analysis. For this purpose, other type of data are available.

There are two basic approaches to measure corruption at the macro level, namely (1) general or target-group perception and (2) incidence of corruptive activities (also referred to as proxy method). The first type of measures reflects the feeling of the public or a specific group of respondents (sometimes called experts) concerning the

[6]For instance in *The Guardian* (March 26, 2004) Charlotte Denny lists 10 country leaders indicating how much money they made with corruption. In the list there are Mohammed Suharto (Indonesia, 1967-1998) with $15bn-35bn, Ferdinand Marcos (Philippines, 1972-1986) $5bn-10bn, Mobutu Sese Seko (Zaire, 1965-1997) $5bn, Sani Abacha (Nigeria, 1993-1998) $2bn-5bn, Slobodan Milosevic (Serbia, 1972-1986) $1bn, Jean-Claude Duvalier (Haiti, 1971-1986) $300m-800m, Alberto Fujimori (Peru, 1990-2000) $600m, Pavlo Lazarenko (Ukraine, 1996-1997) $114m-200m, Arnoldo Alemn (Nicaragua, 1997-2002) $100m, and Joseph Estrada (Philippines, 1998-2001) with $78m-80m. Source: http://www.guardian.co.uk/indonesia/Story/0,2763,1178382,00.html.

[7]See, for instance, Wade (1982) for the case of India, Murray-Rust and van der Valde (1994) for Pakistan, Manzetti and Blake (1996) for Latin America; more recent publications include Svensson (2003) for Uganda, Kuncoro (2004), and Henderson and Kuncoro (2004) for Indonesia, and Golden and Picci (2005) for Italy.

'lack of justice' in public transactions. Therefore, corruption perception is an indirect measure of the actual level of corruption. The incidence-based approach is based on surveys among those who potentially bribe and those whom bribes are offered. Through this approach, a researcher can get information on how frequently corruption occurs in various types of transactions (The Hungarian Gallup Institute, 2000).[8]

Golden and Picci (2005) criticize survey-based measures of corruption as they have at least two intrinsic weaknesses. First, the reliability of survey information about corruption is largely unknown. Respondents directly involved in corruption may have incentives to underreport such involvement, and those not involved typically lack accurate information. Secondly, the reliability of the index may deteriorate over time. There is a danger that respondents report what they believe based on the highly publicized results of the most index rather than how much 'real' corruption exists.

In terms of representative sampling, surveys among the general public may be better. However, various respondents may have no experience with corruption. Their perception may not be very stable over time, since it is highly depending on how much attention corruption receives in the media. Meanwhile, using specific target groups as the source of corruption perception can yield maximum information about corruption although not necessarily honestly expressed. The drawback is that these groups may not be fully representative, being a corruption-prone sub sample of the general public.

Kaufmann and Kraay (2002b) point out that the advantage of a survey among experts is that it is explicitly designed for cross-country comparability. The disadvantage is that such a poll is typically based on the opinions of a few experts per country, and its quality is highly depending on the knowledge of these expert on the countries they assess. The advantage of surveys among the general public or (foreign) business people is that they reflect the opinions of a larger number of people closely connected with the countries they are assessing. There are also disadvantages of surveys among business people or citizens. First, survey questions can be interpreted in culture-specific ways. A question on 'improper practices', for example, is certainly coloured by country-specific perceptions of what is meant by 'improper'. Second, such approaches are costly resulting in a much smaller set of countries than poll of experts. Furthermore, foreign business people are not accustomed to the local customs and language and tend to oversee the ways how issues are settled locally. As a consequence, their evaluation may be biased.

Table 1 summarizes the approaches of various organizations that publish information on corruption. A quick look at the Table shows that the approaches differ along five dimensions. First, corruption is defined in various forms, ranging from bureau-

[8]We categorize poll- or survey-based measures as 'macro' level analysis instead of 'micro' level of analysis because of two reasons. First, they are usually directed to generate corruption indexes on country-wide basis, not, say, firm-level basis; and these indexes are used for country-level comparison. Second, aggregation methods are usually applied to measure corruption drawn from these polls or surveys.

cratic corruption to political corruption. Second, the indexes fall into two categories: poll-based and poll-of-polls-based indexes. The former generates the indexes from direct surveys (perception or proxy method), while the latter combines or aggregates data from direct surveys into a single index of corruption. Third, the indexes use different scales of measurement, ranging from qualitative statements to quantitative rating systems. Fourth, some organizations focus on particular regions only —like the ones found in Afrobarometer, Asian Intelligence, and Latinobarometro— while others have a wider coverage of countries, like the surveys done by Political Risk Services-International Country Risk Guide (PRS-ICRG), Standard and Poor's, etc. Fifth, the institutions that publish information on corruption are private firms, multilateral organizations, and non-governmental organizations. As a consequence, some indexes are only provided on a commercial basis, while others are supplied for free.[9] In the following, instead of discussing all indexes, we will focus on the mostly used indexes.

[9]Especially for the comercial-survey-based indexes, choice of countries certainly depends on the attractiveness of the countries in terms of investment, business climate, geopolitical influence, etc. These factors are important for international economic and political decisions.

Table 1: Surveys on Corruption

Institution	Survey Name	Rating Range		First Published	Last Published	Country Coverage	Subject Measured
		Corrupt	Clean				
Institute for Democracy in South Africa, Ghana Centre Democratic Development, Michigan State University	Afrobarometer Survey	Percentage fairly-very common	Percentage fairly-very rare	1999	2004	12	1. How common corruption among public officials 2. Whether or not corruption worse under the previous regime
Business Environmental Risk Intelligence	Political Risk Index	0	100	1970s	2003	50	How frequently corruption required in business
Columbia University	State Capacity	Severe	Low	1990	2003	95	Severity of corruption
Standard and Poor's DRI	Country Risk Review	0	10	1996		106	Immediate and secondary risk events
Economist Intelligence Unit	Country Risk	1	10	1980	2004	115	Pervasiveness of corruption
Freedom House	Nations in Transition	7	1	1995	2003	27	Level of corruption
Institute for Management Development	World Competitiveness Yearbook	0	10	1987	2005	60	Bribing and corruption in the public sphere Bribing and corruption in the economy
Impulse	Exporter Bribery Index	10	0	1994		103	Proportion of deal involved corrupt payments
Information International	Survey of Middle Eastern Business				2004	31	How common bribes How costly they for doing

Continued on next page

Continued from previous page

Institution	Survey Name	Rating Range		First Published	Last Published	Country Coverage	Subject Measured
		Corrupt	Clean				
							business How frequently public contracts awarded to friends and relatives
International Country Risk Guide	Political Risk	0	6	1982	2004	144	Gov. off. to demand special payments Illegal payments generally expected
International Working Group	Crime Victim Survey			1989		58	Gov. off. to ask to pay a bribe for his service
Latinobarometro	Latinobarometro Survey	100	0	1988	2003	17	1. Corruption 'increased a lot', 'a little'; 'decreased a lot' 'a little'; or 'remained the same' the last 12 months 2. Direct experience of corruption 3. Proportion of corrupt civil servants
Merchant International Group	Grey Area Dynamics	100	0	1990s	2004	155	Range from bribery of gov. ministers to inducements payable to the 'humblest clerk'
Multilateral Development Bank					2002	47	widespread the incidence of corruption
Political and Economic Risk Consultancy	Asian Intelligence Issue	10	0	1980s	2005	12	1. Extent of corruption 2. Rate of corruption in terms of its quality contribution to the overall living/working environment
Price Waterhouse Coopers	Opacity Index	150	0	2001	2004	35	Frequency of corruption
Transparency International	Perception Index	0	10	1995	2005	146	Composite index

Continued on next page

Continued from previous page

Institution	Survey Name	Rating Range		First Published	Last Published	Country Coverage	Subject Measured
		Corrupt	Clean				
Transparency International	Bribe Payer Index	0	10	1999	2002	15	How common bribes are
Transparency International	Global Corruption Barometer	5	1	2003	2004	64	Perception and experience of corruption
World Bank	Kaufmann Index	-2.5	2.5	1996	2004	199	Composite index
World Bank	Country Policy and Institutional Assessments	1	6	1970s	2004	77	Executive accountable for use of funds
World Bank European Bank for Reconstruction and Development	Business Environment Survey	4	1	1996	2004	22	Additional payment
World Economic Forum Harvard Institute for International Development	Global Competitiveness Report	0	10	1979	2005	104	1. Undocumented extra payments 2. Payments favorable regulations and judicial decisions
World Economic Forum	Africa Competitiveness Report	1	4	1998	2004	25	1. How problematic corruption is 2. Irregular, additional payments
World Markets Research Centre Global Insight	Risk Ratings				2004	196	Likelihood of encountering corrupt officials

We start with the corruption index constructed by PRS-ICRG that has been produced since the beginning of the 1980s covering almost 150 developed and developing countries. The PRS-ICRG data consists of political, economic, and financial indexes, each is rated within a specified range.[10] Corruption is one of the 12 political risk components, on a scale of 0-6[11], with higher score means better performance. It is captured from statements like 'high government officials are likely to demand special payments' and 'illegal payments are generally expected throughout lower levels of government' in the forms of 'bribes connected with import and export licences, exchange rate controls, tax assessment, police protection, or loans' (Tanzi and Davoodi, 1997). It also places weight on actual or potential corruption in the form of excessive patronage, nepotism, job reservations, 'favor-for-favors', secret party funding, and suspiciously close ties between politics and business.

The corruption index developed by Kaufmann of the World Bank index is also part of a broader index, the so-called governance index.[12] Published every two years since 1996, this poll-of-polls-based index covers almost 200 countries and is computed on the basis of some hundred individual variables on perception of corruption, drawn from about 40 data sources produced by more than 30 different organizations. From these sources the definition of corruption ranges from the frequency of additional payments to get things done, to the effects of corruption on the business environment, to measuring grand corruption in the political arena or in the tendency of elite forms to engage in state capture (Kaufmann, et al., 2005).

To combine the various corruption indicators into a single index, the following formula is used (Kaufmann, et al., 1999, 2002). The observed score of country i on indicator j, namely $Y_{i,j}$, is treated as a linear function of an unobserved index of corruption C_j and a disturbance term $\epsilon_{i,j}$:

$$Y_{i,j} = \alpha_j + \beta_j[C_i + \epsilon_{i,j}] \tag{1}$$

where α_j and β_j are unknown parameters mapping the latent variable of corruption (C_i) into the observed corruption $Y_{i,j}$.[13] The unobserved C_i is composed of a cluster of $j = 1, ..., J$ indicators, each one providing a numerical rating of some aspect of corruption in each of the $i = 1, ..., I_j$ countries covered by the indicator. Meanwhile,

[10]Different from economic and financial risk indexes that are computed on the basis of objective quantitative data or combinations of this with qualitative data, the political risk index is entirely based on the subjective analysis of the PRS-ICRG staff of the available information.

[11]The other components are government stability with 0-12 scale, socio-economic conditions (0-12), investment profile (0-12), internal conflict (0-12), external conflict (0-12), military in politics (0-6), religion in politics (0-6), law and order (0-6), ethnic tensions (0-6), democratic accountability (0-6), and bureaucracy quality with a scale of 0-4.

[12]The governance index consists of six elements, namely voice and accountability, political instability and violence, government effectiveness, regulatory quality, rule of law, and control of corruption.

[13]The properties of this model are provided in Kaufmann et al. (1999).

the disturbance term $\epsilon_{i,j}$ captures perception errors, sampling variation, and imperfect measurement of corruption represented by indicator j.

Given the estimates of the model's parameters α_j, β_j, and σ_j, the estimate of corruption for a country produced by this model is the mean of the distribution of unobserved corruption conditional on the J_i observed data points for that country:

$$E[C_i|Y_{i,1},...,Y_{i,J_i}] = \sum_{j=1}^{J_i} \left[\frac{\sigma_{\epsilon_j}^{-2}}{1 + \sum_{j=1}^{J_i} \sigma_{\epsilon_j}^{-2}} \right] \left[\frac{Y_{i,j} - \alpha_j}{\beta_j} \right] \tag{2}$$

In other words, the estimate of corruption is given by the weighted average of (re-scaled) scores of each of the component indicators, where the weights are expressed as the first term of the right hand side of equation [2]. This model allows one to compute the variance of this disturbance term, which is a measure of how informative the index is. The variance of this conditional distribution provides an estimate of the precision of the corruption indicator for each country. The point estimate of corruption is the mean of the conditional distribution given the observed data and ranges between -2.5 (most corrupt) and +2.5 (least corrupt).

The third, perhaps best known, index is the corruption perception index (CPI) computed by Lambsdorff on behalf of the TI since 1995. Constructed as a poll-of-polls-based index, the CPI is designed to capture the perception of well-informed people[14] on corruption which are scored on a range of 0 (high) - 10 (low). The index aggregates the perceptions of respondents with regard to the extent of corruption —defined as the abuse of public power for private benefit. Here the extent of corruption reflects the frequency of corrupt payments and the resulting obstacles imposed on businesses (Lambsdorff, 2004b). The CPI index is available for fewer countries than the ICRG index as there must be at least three primary surveys or sources for corruption available for particular country to be included in the index.

To construct the index, some standardization techniques are needed because every primary survey has its own scaling system and data distribution. The first technique is normal standardization. Weighted equally, every source is standardized by the following formula:

$$Y_{i,j}^s = \left(Y_{i,j}^o - \overline{Y_j^o}\right) \frac{\sigma_{C_{t-1}}^2}{\sigma_{Y_j^o}^2} + \overline{C}_{t-1} \tag{3}$$

where $Y_{i,j}^s$ is the standardized score, $Y_{i,j}^o$ is the orginial score provided by source i-th for country j-th, C_{t-1} is the last year corruption perception index (CPI), σ^2 is the standard deviation, and the bar indicates the mean value of the variable concerned. Applied for all sources and countries this technique basically aims at ensuring that the inclusion of a (new) source —consisting of a certain subset of countries— "should not

[14]The sampling frames of the supporting sources consist of samples ranging from residents living within the countries surveyed, foreigners, to samples of high to mid-level business people.

change the mean and standard deviation of this subset in the CPI" (Lambsdorff, 1998, p. 6).

Since this approach heavily depends on the distribution of the data, an alternative approach —the matching percentile technique— is used, especially if the sources have different distribution from that of the CPI. For this technique the rank of a country is used. An example can illustrate the technique. Firstly, say, there are two sources of data (i.e., source j_t and CPI_{t-1}) composed of a subset of countries. In the year t, the source j assessed country i_1, i_2, i_3, i_4, and i_5 with ordered values of 4.5, 3.5, 3.0, 2.0, and 1.5 respectively on a scale of 1-5. In the year *t-1*, CPI also assessed these five countries respectively with values of 8.0, 9.5, 3.5, 4.5, and 2.5 on a scale of 0-10. Matching percentile techniques thus reorders the scores and assigns them to the countries i_1, i_2, i_3, i_4, and i_5 with values of 9.5, 8.0, 4.5, 3.5, and 2.5 to follow the country rank ordered by source j_t. This procedure is applied to all sources, and the index is calculated from the simple average of the standardized values (Lambsdorff, 2002, 2003, 2004a).

There are, however, two problems with these approaches. First, compared to the previous year indexes, the across-countries standard deviation of the current index calculated via the two approaches tends to be smaller. Second, especially for the normal standardization, there is no guarantee that the score will be in the range of 0-10. Thus, in the computation of CPI a β-transformation is also used for two obvious purposes, i.e., [1] to keep all scores within the desired range of 0-10, and [2] to avoid a decreasing across-countries standard deviation especially if compared to the previous years. To do this, each score (Y) is transformed according to the following function (Lambsdorff, 2002, 2003, 2004a):

$$10 \int_0^1 \left(\frac{Y}{10}\right)^{\alpha-1} \left(1 - \frac{Y}{10}\right)^{\beta-1} dY \tag{4}$$

with the task to find α and β so that the resulting mean and standard deviation of the index have the desired values. In other words, in this transformation once scores of 0 or 10 have been reached, they are not further decreased or increased, respectively. This β-transformation is thus applied to all values that have been standardized via the normal standardization technique or the matching percentile technique. Afterwards, the average of these are calculated to determine the index of every country under review.

However, these techniques are not always applied to construct the whole series of CPI. For the 1995 and the historical data (1980-1985, 1988-1992), the index was constructed by taking simple averages after transforming the various different scales —drawn from different data sources— into the scale of 0-10. The normal standardization technique was introduced in 1996 but stopped in 2001. The matching percentile technique and the β-transformation were introduced in 2002 and applied consistently

since then.[15] As a consequence, the CPI is not a consistent time series. In Lambsdorff's words, "... year-to-year changes may not only result from a changing performance of a country ... changes can result from the different methodologies ... not necessarily from actual changes." (Lambsdorff, 2000, p. 4). Apart from changes in the methodology, a change of the CPI for a particular country may also reflect a change in the number of primary sources available for this country (Johnston 2001b).

Other criticisms have been raised as well. Galtung (2005) argues that the definition of corruption in the CPI does not explicitly distinguish between corruption in different branches of civil service nor corruption in political party financing. Likewise, Johnston (2001b) notes that the definition is skewed to a form of bribery. Andvig (2005) questions whether the different sources to form the CPI cover the same phenomenon.[16]

Since it relies heavily on independently conducted surveys and expert polls, the CPI is not available for a significant number of countries. Its reliance on other sources implies that countries may drop out of the index if the required minimum number of sources is missing. Galtung (2005) therefore concludes that CPI does not measure trends. "The CPI's principal flaw is that it is a defective and misleading benchmark of trends" (p. 12).

Despite all these criticisms, it must be recognized that perception-based indexes have opened the possibility to study corruption empirically as they have made the immeasurable concept measurable. As a result, numerous studies have employed such indexes. In the following section we will systematically review empirical studies on the causes of corruption.

[15] We received this information from personal communication with Lambsdorff, since there is no technical explanation for the indexes produced before 1998 and no clear explanation found in the Lambsdorff's *Framework Document* series.

[16] Likewise, Soreide (2003, p. 7) argues that "Most of the polls and surveys ask for a general impression of the magnitude of the problem, which actually means people's subjective intuitions of the extent of a hidden activity. For the TI index, only one source asks for people's personal experiences with corruption. The quantification of the problem is highly ambiguous. It is not clear to what extent the level of corruption reflects the frequency of corrupt acts, the severity to society, the size of the bribes or the benefits obtained. Most of the surveys do not specify what they mean by the word corruption. It can thus be quite difficult for the respondents to answer when asked about a quantification of 'the misuse of public office for private or political party gain' or when encouraged to rate 'the severity of corruption within the state'."

3 Empirical Determinants of Corruption

Many studies have searched for empirical regularities between corruption and a variety of economic and non-economic determinants. Unfortunately, there is no commonly-agreed-upon theory on which to base an empirical model of the causes of corruption (Alt and Lassen, 2003). At the same time, numerous regression models incorporating a wide variety of explanatory variables have been specified to explain corruption and to find the 'true' determinants. It is often found that, however, a variable is significant in a particular specification of the model, but loses its significance when some other variables are incorporated. In other words, claims concerning the determinants of corruption are conditional, and the robustness of the findings is open to question.

While other categorizations are possible, we identify four broad classes of underlying causes of corruption, namely (1) economic and economic institutions, (2) political, (3) judicial and bureaucratic, and (4) religious and geo-cultural factors. Tables 2-5 summarize all studies that we are aware of, indicating the main results concerning the significance of the variables belonging to the classes of variables that we distinguish.

3.1 Economic Determinants

Economic factors consist of a wide range of economic variables like income or economic policy variables; included here are also demographic variables and economic institutions. To start with, we observe that income is a commonly-used variable to explain corruption (Damania et al., 2004; Persson et al., 2003; and van Rijckeghem and Weder, 1997; among others). Mostly proxied by GDP per capita, income is used to control for structural differences as economic development progresses. It can be generally concluded that a country's wealth is a significant predictor of corruption, even though Kaufmann et al. (1999) and Hall and Jones (1999) question the causal relationship between corruption and income. Two studies with panel data (Braun and Di Tella, 2004; Frechette, 2001) deviate from this main result, finding that income increases corruption, especially when they impose fixed effects.

Income distribution is also argued to affect corruption. As Paldam (2002) puts it, "A skew income distribution may increase the temptation to make illicit gains". Proxied by the Gini coefficient, he claims that income disparity significantly increases corruption. However, using the income share of top 20% of the population under a different specification, Park (2003) does not find a statistically significant relationship. Similarly, Brown et al. (2005) find no evidence that greater income inequality increase corruption.

The size of government is also an important source of corruption. If countries exploit economies of scale in the provision of public services —thus have a low ratio of public service outlets per capita— those who demand the services might be tempted to bribe, e.g., 'to get ahead of the queue'. However, a large government sector may also create

opportunities for corruption; that is, the larger the relative size of the public sector, the greater the likelihood of corrupt behaviour. Thus, there is no consensus among authors on the theoretical relationship between government size and corruption. This is also reflected in the empirical studies of Fisman and Gatti (2002) and Bonaglia et al. (2001) that end up with a different conclusion than the ones of Ali and Isse (2003). Whereas the former finds the negative impact of government spending on corruption, the latter reports the positive impact.

Another variable that according to various authors also explains corruption is import share. Herzfeld and Weiss (2003) and Treisman (2000) report that a higher import share leads to less corruption. A high import share implies lower tariff and non-tariff import restrictions. The presence of such restrictions —like the necessary licenses to import, for example— offers an opportunity to bribe. Likewise, a high export share of raw materials, such as fuel, mineral, and ore, increases the probability of corruption to occur. Since such endowments create rents, this thus exhibits the phenomenon of rents-related corruption which is, according to Tornell and Lane (1998), commonly found in natural-resource-abundant countries.

In line with the above-mentioned argument, restrictions on foreign trade, foreign investment, and capital markets stimulate corruption; see, for example, Knack and Azfar (2003), and Frechette (2001). Likewise, economic freedom —measured by the indexes of the Heritage Foundation/Wall Street Journal and the Fraser Institute— is also found to lessen corruption. Proponents of this view are Gurgur and Shah (2005), Park (2003), and Treisman (2000), but Lederman et al. (2005) and Paldam (2001) find more mixed results. Broadman and Recanatini (2000, 2002) show the existence of a positive relationship between entry barriers and corruption; that is, the greater the barriers to entry and exit faced by firms, and therefore the greater the distortions existing in the competitive environment, the more widespread is corruption.

Finally, we turn to socio-demographic factors associated with corruption. These include schooling, population, and the labour force. Economies with high human capital have low levels of corruption as found in Ali and Isse (2003), Brunetti and Weder (2003), and van Rijckeghem and Weder (1997). However, a counter-intuitive finding is found in Frechette (2001). Using panel data models with fixed effects, he finds that schooling is positive in all regressions explaining corruption. Similar conflicting evidence is found for a country's population. Knack and Azfar (2003) show that as population increases, corruption also rises, while Tavares (2003) reports that population negatively affects corruption.

Another interesting demographic variable is the percentage of female population in the labour force. Swamy et al. (2001) indicate that a higher female labour participation leads to less corruption. Combined with two other gender variables, namely proportion of women in parliament and in government, they find that more influence of women leads to less corruption. Following Gottfredson and Hirshi (1990) and Paternoster and

Simpson (1996), Swamy et al. provide four arguments to explain this finding. First, "women may be brought up to be more honest or more risk averse than men, or even feel there is a greater probability of being caught." Second, "women, who are typically more involved in raising children, may find they have to practice honesty in order to teach their children the appropriate values." Third, "women may feel more than men -the physically stronger sex, that laws exist to protect them and therefore be more willing to follow rules." Lastly, "girls may be brought up to have higher levels of self-control than boys which affects their propensity to indulge in criminal behaviour."

Table 2: Economic Determinants of Corruption*

Variable	Positive-Significant by	Negative-Significant by
Economic Factors		
Income	Braun-Di Tella (2004), Frechette (2001)	Brown, etal. (2005), Kunicova-R.Ackerman (2005), Lederman et al. (2005), Braun-Di Tella (2004), Chang-Golden (2004), Damania et al. (2004), Dreher et al. (2004), Alt-Lassen (2003), Brunetti-Weder (2003), Graeff-Mehlkop (2003), Herzfeld-Weiss (2003), Knack-Azfar (2003), Persson et al. (2003), Tavares (2003), Fisman-Gatti (2002), Paldam (2002-01), Frechette (2001), Bonanglia et al. (2001), Swamy et al. (2001), Abed-Davoodi (2000), Rauch-Evan (2000), Treisman (2000), Wei (2000), Ades-Di Tella (1999), Goldsmith (1999-97), van Rijckeghem-Weder (1997)
Income distribution	Paldam (2002)	
Government expenditure	Ali-Isse (2003)	Fisman-Gatti (2002), Bonaglia et al. (2001)
Government revenue	Lederman etal. (2005), Alt-Lassen (2003)	
Govt. transfer to lower level		Lederman et al. (2005)
Black market premium		Brunetti-Weder (2003), van Rijckeghem-Weder (1997)
Inflation, Inflation vars.	Braun-Di Tella (2004), Paldam (2002-01)	
Economic Institutions		
Foreign aid	Ali-Isse (2003)	Tavares (2003)
Import share		Herzfeld-Weiss (2003),

Continued on next page

Continued from previous page

Variable	Positive-Significant by	Negative-Significant by
		Fisman-Gatti (2002), Frechette (2001), Treisman (2000), Ades-Di Tella (1999)
Raw material export	Herzfeld-Weiss (2003), Tavares (2003), Bonaglia et al. (2001), Frechette (2001)	Frechette (2001)
Trade opennes		Gurgur-Shah (2005), Brunetti-Weder (2003), Knack-Azfar (2003), Persson et al. (2003), Fisman-Gatti (2002), Bonaglia et al. (2001), Frechette (2001), Wei (2000), Ades-Di Tella (1999), Laffont and N'Guessan (1999), Leite-Weidmann (1997)
Economic freedom	Graeff-Mehlkop (2003), Paldam (2001)	Kunicova-R.Ackerman (2005), Gurgur-Shah (2005), Ali-Isse (2003), Graeff-Mehlkop (2003), Park (2003), Treisman (2000), Goldsmith (1999)
Entry barriers, Competitiveness	Broadman-Recanatini (2002-00)	Gurgur-Shah (2005), Suphacahlasai (2005)
Structural reform		Abed-Davoodi (2000)
Infrastructure		Broadman-Recanatini (2000)
Budget constraint		Broadman-Recanatini (2002-00),
Demographic Factors		
Schooling	Frechette (2001)	Ali-Isse (2003), Alt-Lassen (2003), Brunetti-Weder (2003), Persson et al. (2003), Evan-Rauch (2000), Ades-Di Tella (1999-97), van Rijckeghem-Weder (1997)
Population	Damania et al. (2004), Alt-Lassen (2003),	Tavares (2003)

Continued on next page

Continued from previous page		
Variable	Positive-Significant by	Negative-Significant by
	Knack-Azfar (2003), Fisman-Gatti (2002)	
Female labour force		Swamy et al. (2001)

Note: *] Corruption is measured by various indexes; higher score, more corrupt. Significant at conventional levels.

3.2 Political Determinants

Empirical studies on the political causes of corruption can be divided into two broad groups, namely those investigating the impact of political-civil liberty and those examining the effect of decentralization on corruption. Meanwhile, other factors that also have been suggested to affect corruption are the electoral system (Persson et al., 2003; Kunicova and Rose-Ackerman, 2005), governmental administration (Brown et al., 2005; Chang and Golden, 2004), and political instability (Park, 2003).

Although various proxies like civil liberty, political freedom, political rights, length of democratic regime, etc, have been used, there is a consensus that democracy reduces corruption. This conclusion is confirmed if corruption is related to other democracy-related variables, like freedom of the press. This variable is found to be significantly correlated with corruption (Brunetti and Weder, 2003).

The main reason why political liberty tends to reduce corruption is that political liberty imposes transparency and provides checks and balances within the political system. Political participation, political competition, and constraints on the chief executive increase the ability of the population to monitor and legally limit politicians from engaging in corrupt behaviour (Kunicova and Rose-Ackerman, 2005). In addition, it is often found that democratic systems are politically more stable. It is therefore not surprising that authors like Lederman et al. (2005), Park (2003), and Leite and Weidmann (1997) find that that corruption increases in unstable polities.[17]

Some aspects of democratic elections may, however, create opportunities for corruption. Selecting politicians through party lists, for example, can obscure the direct link between voters and politicians, thus degrading the ability of voters to hold politicians accountable (Kunicova and Rose-Ackerman, 2005; Persson and Tabellini, 2003). Chang and Golden (2004), in their study on electoral systems and corruption, find that under open-list proportional representation increases in district size leads to more corruption. Meanwhile, under closed-list proportional representation arrangements, political corruption becomes less prevalent as district magnitude increases. Similar results are also found by Persson et al. (2003).

Decentralization or federalism has also been argued to be crucial to combat corruption, but the empirical evidence is mixed. Measuring decentralization as transfers from central government to other levels of national government as a percentage of GDP, Lederman et al. (2005) find that this variable reduces corruption significantly. Likewise, taking a binary variable of centralized unitary states and decentralized federal systems, Ali and Isse (2003) report that decentralized government lowers corruption. Gurgur and Shah (2005) use the ratio of employment in non-central government administration to general civilian government employment and show that corruption is lower in

[17] Another explanation can also be found in Shleifer and Vishny (1993) who pose that the ephemeral nature of public positions in unstable systems makes officials irresponsible and get them involved in illicit rent-seeking behaviour.

both decentralized unitary and federal states but the impact is higher in decentralized unitary system. Fisman and Gatti (2002) measure decentralization as the sub-national share of total government spending. The numerator is the total expenditure of sub national (state and local) governments, while the denominator is total spending by all levels (state, local, and central) of government. They find the negative effect of fiscal decentralization on corruption, even after controlling for potential joint endogeneity.

In contrast, Kunicova and Rose-Ackerman (2005) using a simple dummy for autonomous regions with extensive taxing, spending and regulatory authority argue that federalism increases corruption, holding other factors constant. Likewise, using a dummy variable for the presence of a federal constitution, Damania et al. (2004) and Treisman (2000) find that a federal structure is more conducive to corruption. "As the political pie is divided between a greater number of geographic entities, opportunities to generate political rents increase" (Brown et al., 2005, p. 12). Similarly, Goldsmith (1999) also demonstrates that federalism is associated with more perceived corruption.

Table 3: Political and Political Institution Determinants of Corruption*

Variable	Positive-Significant by	Negative-Significant by
Democracy, civil liberty		Kunicova-R.Ackerman (2005), Lederman et al. (2005), Gurgur-Shah (2005), Braun-Di Tella (2004), Chang-Golden (2004), Damania et al. (2004), Herzfeld-Weiss (2003), Knack-Azfar (2003), Broadman-Recanatini (2002-00), Paldam (2002), Bonaglia et al. (2001), Frechette (2001), Swamy etal. (2001), Treisman (2000), Wei (2000), Ades-Di Tella (1999-97), Leite-Weidmann (1997), Goldsmith (1999), van Rijckeghem-Weder (1997)
Press freedom, Media		Lederman et al. (2005), Suphacahlasai (2005), Brunetti-Weder (2003)
Decentralization, federalism	Brown et al. (2005), Kunicova-R.Ackerman (2005), Damania et al. (2004), Treisman (2000), Goldsmith (1999)	Gurgur-Shah (2005), Lederman etal. (2005), Fisman-Gatti (2002), Ali-Isse (2003), Wei (2000)
District maginute		Chang-Golden (2004)
Closed list system	Kunicova-R.Ackerman (2005), Persson-Tabellini (2003), Persson et al (2003),	Lederman et al. (2005), Chang-Golden (2004)
Presidentialism	Brown, et al. (2005), Kunicova-R.Ackerman (2005), Lederman et al. (2005), Chang-Golden (2004)	
Number of party	Chang-Golden (2004)	
Political instability	Park (2003), Leite-Weidmann (1999)	
Ideological Polarization		Brown, et al. (2005),

Continued on next page

Continued from previous page		
Variable	Positive-Significant by	Negative-Significant by
Majoritarian plurality		Kunicova-R.Ackerman (2005),
Central planning		Abed-Davoodi (2000),
Women in public position		Swamy et al. (2001)

Note: *] Corruption is measured by various indexes; higher score, more corrupt. Significant at conventional levels.

3.3 Bureaucratic and Regulatory Determinants

The judicial system and the quality of bureaucracy are crucial factors influencing corruption. In this context, the wage level of civil servants may be important, since —as argued by van Rijckeghem and Weder (1997)— public sector wages are highly correlated with the measures of the rule of law and the quality of the bureaucracy, and therefore may have an effect on corruption. In developing economies bureaucrats receive wages that are so low to entice corrupt behaviour. At the same time, low income economies suffer from the lack of institutions for detecting corruption. Measured as the relative magnitude of wage to GDP, Herzfeld and Weiss (2003) identify that an increase in wages significantly lessens corruption. Similary, van Rijckeghem and Weder (1997) claim that government wages as the ratio to manufacturing wages significantly reduces corruption. The influence wage on corruption is also highlighted by Alt and Lassen (2003) and Rauch and Evans (2000). However, other studies reveal that this relationship is not always to be statistically significant (Gurgur and Shah, 2005; Treisman, 2000).

Gurgur and Shah (2005), Brunetti and Weder (2003), and van Rijckeghem and Weder (1997) report that the higher the quality of bureaucracy, the lower the probability for corruption to occur. Along with this finding, it is also interesting to see that the lack of meritocratic recruitment and promotion and the absence of professional training in the bureaucracy are also found to be associated with high corruption (Rauch and Evans, 1997).

Finally, various studies suggest that the rule of law, proxied by various measures, is relevant in explaining corruption. Damania et al. (2004) use the rule of law index of Kaufmann et al. (1999a,b) that takes several indicators into account to measure the extent to which economic agents abide by the rules of society, perceptions of the effectiveness and predictability of the judiciary, and the enforceability of contracts. A group of authors (Brunetti and Weder, 2004; Ali and Isse, 2003; Herzfeld and Weiss, 2003; Park 2003; and Leite and Weidmann, 1999) use the ICRG index to reflect the degree to which the citizens of a country are willing to accept the established institutions to make and implement laws and adjudicate disputes. This index also measures the extent to which countries have sound political institutions, strong courts, and orderly succession of power. All these authors claim that a strong rule of law reduces the likelihood of corruption to take place. This result is significant under various regression specifications.

Table 4: Judicial and Bueraucratic Determinants of Corruption*

Variable	Positive-Significant by	Negative-Significant by
Government wage		Alt-Lassen (2003), Herzfeld-Weiss (2003), Rauch-Evan (2000), van Rijckeghem-Weder (1997),
Quality of bureaucracy		Gurgur-Shah (2005), Brunetti-Weder (2003), van Rijckeghem-Weder (1997)
Merit system		Rauch-Evan (2000)
Rule of law		Damania et al. (2004), Ali-Isse (2003), Brunetti-Weder (2003), Herzfeld-Weiss (2003), Park (2003), Broadman-Recanatini (2000), Leite-Weidmann (1997), Ades-Di Tella (1997),

Note: *] Corruption is measured by various indexes; higher score, more corrupt. Significant at conventional levels.

3.4 Geographical, Cultural, and Religious Determinants

Religion, culture, and geography may also matter for explaining corruption. Countries with many Protestants tend to have lower corruption levels (Chang and Golden, 2004; Bonaglia et al., 2001; Treisman, 2000; La Porta et al., 1999). Paldam (2001) reports that countries dominated by two religions, namely Reform Christianity (e.g., Protestant and Anglican) and Tribal religions, tend to have lower levels of corruption compared to countries in which other religions dominate.

As to cultural variables, many authors find that ethno-linguistic homogeneity tends to reduce corruption (Lederman et al., 2005; La Porta et al., 1999). This finding is explained in terms of the increased difficulties that bureaucrats encounter in extracting bribes from ethnic groups to which he does not belong. The domination of an ethnic group in a country generates an unequal access to power. Minorities with less political access thus collude with bureaucrats for levelling the political and economic landscape. In ethnically diverse communities, a bureaucrat behaves sequentially: first to his close kin, to his ethnic group, and then maybe to his country (Ali and Isse, 2003). As a result, highly fragmented communities are likely to be more corrupt than homogenous societies.

Another cultural variable used to explain corruption is colonial heritage that captures 'command and control habits and institutions and the divisive nature of the society left behind by colonial masters' (Gurgur and Shah, 2005, p. 18). The evidence on the relevance of this variable is, however, mixed. Countries that have been colonialized tend to suffer from corruption (Gurgur and Shah, 2005; Tavares, 2003). Herzfeld and Weiss (2003), on the other hand, find that former British colonies have lower levels of corruption. Persson et al. (2003) measure the influence of colonial history by partitioning all former colonies into three groups, namely British, Spanish-Portuguese, and other colonial origin, and define three binary indicator variables for these groups. They find that former British colonies tend to have a lower current propensity for corruption.

Table 5: Cultural and Geographical Determinants of Corruption*

Variable	Positive-Significant by	Negative-Significant by
Pop. with particular religious affiliation	Paldam (2001), La Porta et al (1999)	Chang-Golden (2004), Herzfeld-Weiss (2003), Persson et al. (2003), Bonaglia et al. (2001), Paldam (2001), Treisman (2000), La Porta et al (1999)
Ethnic heterogeneity	Lederman et al (2005), Suphachalasai (2005), Herzfeld-Weiss (2003), Treisman (2000), La Porta et al (1999)	Bonaglia et al.(2001)
Colonial past	Gurgur-Shah (2005) Tavares (2003)	Herzfeld-Weiss (2003), Persson et al. (2003), Swamy et al. (2001), Treisman (2000)
Distance to large exporter	Ades-Di Tella (1999)	Bonaglia et al.(2001),
Legal origin	Gatti (1999), La Porta et al (1999)	Suphachalasai (2005),
Area wide		Bonaglia et al.(2001),
Latitude		La Porta et al (1999),
Mascullinity	Park (2003)	
Natural resources	Leite-Weidmann (1997)	

Note: *] Corruption is measured by various indexes; higher score, more corrupt. Significant at conventional levels.

4 Data Imputation and Factor Analysis

The main objective of this paper is to rexamine the claims of the above-mentioned authors on the significance of the corruption determinants in various corruption regressions. We do regression analysis on these determinants by taking Kaufmann corruption index 2004 (corka04) as the dependent variable. As the literature suggests a long list of variables causing corruption, we have collected as many variables as possible that have been suggested determining corruption. Table 6 presents the variables we have collected and their sources, while table 7 provides the statistical summary of them.

Since the total number of explanatory variables is huge, no doubt multicollinearity will become a problem in our regression analysis. Some variables, however, can possibly be clustered into some groups representing a particular phenomena. The second problem is that not all data are available for the same set of countries. We have only a few variables capturing all 193 country samples, namely GDP per capita, population density, and country area; for the other variables the number of observations varies from 52 to 191. This implies that we have missing data problem. To deal with the first problem, we use Exploratory Factor Analysis (EFA) to reduce the number of explanatory variables. However, we first solve the missing data problem.

The question of how to treat incomplete data is among the most complicated problems faced by policy analysts. Because of the lack of data, the degree of uncertainty increases with the level of data aggregation and influences the ability to draw accurate conclusions. We minimize the degree of uncertainty using the data imputation technique of Expectation-Maximization (EM) as suggested by Dempster, et al. (1977) and Ruud (1991). The EM algorithm is basically an iterative method that can be divided into two stages. First, in the 'expectation' stage, we form a log-likelihood function for the latent data as if they were observed and taking its expectation. Second, in the 'maximization' stage, the resulting expected log-likelihood is maximized.

Prior to the imputation we transform all variables to improve the distributional characteristics of the data.[18] A more normal (symmetric) distribution implies that the majority of data fall within the two standard deviations of the mean and extreme values occur with small probability. If the observed minimum of the variable is negative, we add a constant such that the transformation of negative values can be computed (see Table 7).

[18]Before the transformation, variables like economic freedom of Heritage Foundation and press freedom are rescaled to give them the same interpretation as the other variables. Thus higher values mean more freedom. The same rescaling is applied for corruption and the other Kaufmann indexes of governance.

Table 6: The Variables

No	Var.	Definition	Year	Source
1	corka04	control of corruption (-2.5: corrupt; +2.5: clean); Reoriented	2004	worldbank.org/wbi/governance/pdf/2004kkdata.xls
2	gdpcap	per capita gdp at constant prices in US dollars	2000	gdp: unstats.un.org/unsd/snaama/downloads/GDPconstantUS-countries.xls
			2000	population: devdata.worldbank.org/edstats/query/default.htm
3	gini	gini index	1993-2000	Human Development Indicator 2005
4	rich10	the richest 10% of population	1993-2000	Human Development Indicator 2005
5	rich20	the richest 20% of population	1993-2000	Human Development Indicator 2005
6	govcon	share of general government final consumption expenditure	2000	unstats.un.org/unsd/snaama/downloads/Shares-countries.xls
7	aidcap	aid per capita (current US$)	2000	devdata.worldbank.org/data-query/
8	debtgni	total debt per gni	2000	rrojasdatabank.org/dev0000.htm
9	totdebt	totdebt (000000)	2000	nationmaster.com/graph-T/eco_deb_ext_cap
10	debtgdp	debt per gdp (00)	2000	nationmaster.com/graph-T/eco_deb_ext_gdp
11	mpor	share of imports of goods and services		unstats.un.org/unsd/snaama/downloads/Shares-countries.xls
12	xpor	share of exports of goods and services		unstats.un.org/unsd/snaama/downloads/Shares-countries.xls
13	xfuel	export of fuel (% total merchandise export)	2000	WDI 2002
14	xmetal	export of ores and metal (% total merchandise export)	2000	WDI 2002
15	area1	size of government, fraser, (0: lowest freedom; 10: highest freedom)	2000	freetheworld.com./2005/2005Dataset.xls
16	area2ab	judicial independence-impartial court, fraser, (0: lowest freedom; 10: highest freedom)	2000	freetheworld.com./2005/2005Dataset.xls
17	area3	sound money, fraser, (0: lowest freedom; 10: highest freedom)	2000	freetheworld.com./2005/2005Dataset.xls
18	area4b	freedom to trade internationally, fraser, (0: lowest freedom; 10: highest freedom)	2000	freetheworld.com./2005/2005Dataset.xls
19	area5b	labor regulation, fraser, (0: lowest freedom; 10: highest freedom)	2000	freetheworld.com./2005/2005Dataset.xls
20	efherit	econ. freedom of heritage found.(1: highest freedom; 5: lowest freedom): Reoriented	2000	heritage.org/research/features/index/downloads.cfm
21	open	share of import + exports of goods and services	2000	unstats.un.org/unsd/snaama/downloads/Shares-countries.xls
22	enrolp	gross enrollment rate (%, primary school)	2000	devdata.worldbank.org/edstats/query/default.htm
23	enrols	gross enrollment rate (%, secondary school)	2000	devdata.worldbank.org/edstats/query/default.htm
24	enrolt	gross enrollment rate (%, tertiary school)	2000	devdata.worldbank.org/edstats/query/default.htm
25	illi	Estimated illiteracy rate and illiterate population aged 15 years and older	2000	uis.unesco.org/TEMPLATE/html/Exceltables/education/ View_Table_Literacy_Country_Age15+.xls
26	popden	population per km sq. area	2000	
27	femlab	female labor force (% of total)	2000	devdata.worldbank.org/edstats/query/default.htm
28	pr	political right (1: highest freedom; 7: lowest freedom)	2000	freedomhouse.org/ratings/index.htm
29	cl	civil liberty (1: highest freedom; 7: lowest freedom)	2000	freedomhouse.org/ratings/index.htm
30	poli	index autocracy-democracy (-10: autocratic; +10: democratic)	2000	cidcm.umd.edu/inscr/polity/polreg.htm
31	demac	democratic accountability (1-6: higher score better performance)	2000	icrg
32	voic	voice accountability (-2.5 - +2.5: higher score better performance)	2000	worldbank.org/wbi/governance/pdf/2004kkdata.xls
33	press	press freedom (0: highest freedom; 100: lowest freedom); Reoriented	2000	freedomhouse.org/research/pressurvey.htm
34	snexg	Sub-National Government Expenditure (% gdp)		government financial statistics
35	snreg	Sub-National Government Revenue (% gdp)		government financial statistics
36	snexe	Sub-National Government Expenditure (% total govt. expend.)		government financial statistics
37	snrer	Sub-National Government Revenue (% total govt. revenue)		government financial statistics
38	distmag	mean district magnitude (House): average no. of legislators elected to the lower house from each district	2000	siteresources.worldbank.org/INTRES/Resources/ DPI2004-no_formula_no_macro.xls
39	presiden	Direct Presidential (0); strong president elected by assembly (1); Parliamentary (2)	2000	siteresources.worldbank.org/INTRES/Resources/ DPI2004-no_formula_no_macro.xls
40	psta	political stability (-2.5 - +2.5: higher score better performance)	2000	worldbank.org/wbi/governance/pdf/2004kkdata.xls
41	gsta	government stability (1-12: higher score better performance)	2000	icrg
42	inco	internal conflict (1-12: higher score better performance)	2000	icrg
43	exco	external conflict (1-12: higher score better performance)	2000	icrg
44	eten	ethnic tension (1-6: higher score better performance)	2000	icrg
45	polariz	Maximum polarization between the executive party and the four principle parties of the legislature;	2000	siteresources.worldbank.org/INTRES/Resources/

Continued on next page

Continued from previous page

No	Var.	Definition	Year	Source
		Also, the maximum difference between the chief executives partys value and the values of the three largest government parties and the largest opposition party		DPI2004-no_formula_no_macro.xls
46	pluralty	Electoral rule: Plurality? (1 if yes, O if no)	2000	siteresources.worldbank.org/INTRES/Resources/ DPI2004-no_formula_no_macro.xls
47	wage	Total Central gov't wage bill (% of GDP)	1996-2000	www1.worldbank.org/publicsector/civilservice/development.htm
48	buqua	bureaucratic quality (1-4: higher score better performance)	2000	icrg
49	geff	government effectiveness (-2.5 - +2.5: higher score better performance)	2000	worldbank.org/wbi/governance/pdf/2004kkdata.xls
50	rlaw	rule of law (-2.5 - +2.5: higher score better performance)	2000	worldbank.org/wbi/governance/pdf/2004kkdata.xls
51	regqua	regulatory quality (-2.5 - +2.5: higher score better performance)	2000	worldbank.org/wbi/governance/pdf/2004kkdata.xls
52	lawor	law and order (1-4: higher score better performance)	2000	icrg
53	budha	% population	2005	worldchristiandatabase.org/
54	hindu	% population	2005	worldchristiandatabase.org/
55	muslim	% population	2005	worldchristiandatabase.org/
56	nonrelig	% population	2005	worldchristiandatabase.org/
57	anglic	% population	2005	worldchristiandatabase.org/
58	cathol	% population	2005	worldchristiandatabase.org/
59	indepen	% population	2005	worldchristiandatabase.org/
60	marginal	% population	2005	worldchristiandatabase.org/
61	orthodox	% population	2005	worldchristiandatabase.org/
62	protest	% population	2005	worldchristiandatabase.org/
63	ethnoa	the probability that two randomly selected individuals from the country in question will not belong to the same ethnic group. Higher value reflects a greater degree of fractionalization.	1960-80	Annett, Anthony. 2001. Social Fractionalization, Political Instability, and the Size of Government. IMF 48(3). imf.org/External/Pubs/FT/staffp/2001/03/annett.htm.; humandevelopment.bu.edu/use_exsisting_index/ show_aggregate.cfm?index_id=234&data_type=1
64	ethnob	Average value of five different indices of ethonolinguistic fractionalization. Its value ranges from 0 to 1. The five component indices are: (1) index of ethnolinguistic fractionalization in 1960, which measures the probability that two randomly selected people from a given country will not belong to the same ethnolinguistic group (the index is based on the number and size of population groups as distinguished by their ethnic and linguistic status); (2) probability of two randomly selected individuals speaking different languages; (3) probability of two randomly selected individuals do not speak the same language; (4) percent of the population not speaking the official language; and (5) percent of the population not speaking the most widely used language.		La Porta et.al (1999) = La Porta et.al (1998)
65	englis	legal origin		La Porta et.al (1999) = La Porta et.al (1998)
66	social	legal origin		La Porta et.al (1999) = La Porta et.al (1998)
67	french	legal origin		La Porta et.al (1999) = La Porta et.al (1998)
68	german	legal origin		La Porta et.al (1999) = La Porta et.al (1998)
69	scandi	legal origin		La Porta et.al (1999) = La Porta et.al (1998)
70	latitu	latitude		La Porta et.al (1999) = La Porta et.al (1998)
71	areakm	land and water area (km sq)		cia.gov/cia/publications/factbook/

Table 7: Transformation and Imputation

No.	Variable	Before Imputation					Transf.	Const.	After Imputation				
		Obs.	Mean	Std. Dev.	Min.	Max			Obs.	Mean	Std. Dev.	Min.	Max.
Corruption													
1	corka04	193	0.02	1.02	-2.50	1.65	sq.	12.50	193	0.05	1.02	-2.50	1.68
Economic Determinants													
2	gdpcap	193	6403.10	10651.61	82.01	77319.59	log	0	193	6403.10	10651.61	82.01	77319.59
3	gini	124	40.18	10.40	24.70	70.70	sr.	0	193	40.48	8.60	24.70	70.70
4	rich10	122	19.03	17.68	4.50	87.20	inv.	0	193	16.76	14.42	4.50	87.18
5	rich20	124	10.40	9.18	3.40	57.60	in.sq.rt.	0	193	9.76	7.47	3.40	57.60
6	govcon	193	0.18	0.13	0.03	1.07	log	0	193	0.18	0.13	0.03	1.07
7	aidcap	159	59.47	133.18	0.00	1105.00	log	0	193	53.02	125.46	0.57	1105.00
8	debtgni	128	83.79	86.40	2.84	726.13	log	0	193	69.55	75.66	2.84	726.13
9	totdebt	142	27513.04	84538.45	10.00	868500.00	log	0	193	28982.92	87381.10	10.00	868497.10
10	debtgdp	141	27.50	31.02	0.19	215.15	log	0	193	24.91	27.01	0.19	215.15
11	mpor	188	0.50	0.38	0.00	3.62	sq.rt.	0	193	0.50	0.38	0.00	3.62
12	xpor	191	0.46	0.40	0.00	4.15	sq.rt.	0	193	0.46	0.39	0.00	4.15
13	xfuel	106	17.36	27.55	0.00	100.00	sq.rt.	0	193	14.05	21.24	0.00	100.00
14	xmetal	109	6.17	11.76	0.00	67.00	sq.rt.	0	193	5.35	9.05	0.00	67.00
15	area1	122	5.75	1.50	2.29	9.19	ident.	0	193	5.80	1.21	2.29	9.19
16	area2ab	122	5.83	1.91	1.98	9.62	log	0	193	5.54	1.76	1.98	9.62
17	area3	122	7.61	1.87	1.25	9.83	sq.	0	193	7.52	1.58	1.25	9.83
18	area4b	121	6.91	1.28	1.66	9.78	sq.	0	193	6.75	1.14	1.66	9.78
19	area5b	122	5.88	1.06	2.71	8.35	sq.	0	193	5.83	0.94	2.71	8.35
20	efherit	146	2.81	0.83	1.00	4.60	sq.rt.	0	193	2.78	0.73	1.00	4.60
21	open	188	0.97	0.73	0.00	6.74	sq.rt.	0	193	0.97	0.72	0.00	6.74
22	enrolp	170	100.09	18.66	15.00	151.00	sq.	0	193	100.33	17.63	15.00	151.00
23	enrols	150	71.39	32.75	6.00	161.00	sq.	0	193	68.78	31.80	6.00	161.00
24	enrolt	118	28.77	21.81	1.00	85.00	ident.	0	193	28.23	17.45	1.00	85.00
25	illi	136	21.91	20.79	0.20	84.00	log	0	193	18.19	19.03	0.20	84.00
26	popden	193	275.74	1359.33	1.66	16981.81	log	0	193	275.74	1359.33	1.66	16981.84

Continued on next page

Continued from previous page

No.	Variable	Before Imputation					Transf.	Const.	After Imputation				
		Obs.	Mean	Std. Dev.	Min.	Max			Obs.	Mean	Std. Dev.	Min.	Max.
27	femlab	171	39.73	8.31	4.00	52.00	sq.	0	193	39.74	7.83	4.00	52.00
Political Determinants													
28	pr	184	4.52	2.24	1.00	7.00	log	0	193	4.60	2.23	1.00	8.25
29	cl	184	4.42	1.82	1.00	7.00	sq.rt.	0	193	4.51	1.83	1.00	7.72
30	poli	155	2.92	6.64	-10.00	10.00	sq.rt.	20	193	2.64	6.05	-10.00	10.00
31	demac	138	3.74	1.83	0.00	6.00	ident.	0	193	3.63	1.59	0.00	6.00
32	voic	187	-0.02	1.00	-2.12	1.64	inv.cub.	12.12	193	0.02	1.01	-2.12	2.07
33	press	183	55.67	23.43	10.00	95.00	sq.rt.	0	193	57.06	23.70	10.00	101.37
34	snexg	53	10.58	7.91	0.00	32.78	sq.rt.	0	193	5.73	5.69	0.00	32.78
35	snreg	52	6.64	5.54	0.00	23.60	sq.rt.	0	193	3.79	3.76	0.00	23.60
36	snexe	57	24.07	14.46	2.08	58.56	sq.rt.	0	193	18.30	10.28	1.47	58.56
37	snrer	56	18.81	14.37	0.88	61.95	sq.rt.	0	193	14.05	9.51	0.62	61.95
38	distmag	148	13.84	27.79	0.70	150.00	log	0	193	11.58	24.67	0.70	150.00
39	presiden	168	0.76	0.93	0.00	2.00	ident.	0	193	0.82	0.91	0.00	2.00
40	psta	164	0.00	1.00	-2.50	1.73	ident.	12.50	193	-12.38	0.99	-15.00	-10.20
41	gsta	139	9.87	0.94	6.58	12.00	cub.	0	193	9.84	0.80	6.58	12.00
42	inco	139	8.76	2.20	0.42	12.00	ident.	0	193	8.86	2.04	0.42	12.87
43	exco	139	9.79	1.57	4.25	12.00	cub.	0	193	9.91	1.39	4.25	12.00
44	eten	138	4.10	1.43	0.00	6.00	ident.	0	193	4.14	1.25	0.00	6.00
45	polariz	152	0.54	0.84	0.00	2.00	ident.	0	193	0.51	0.78	0.00	2.00
46	pluralty	148	0.66	0.48	0.00	1.00	ident.	0	193	0.66	0.47	0.00	1.00
Bureaucratic and Judiciary Determinants													
47	wage	118	6.94	4.41	1.00	23.30	log	0	193	6.96	3.93	1.00	23.30
48	buqua	136	2.21	1.13	0.00	4.00	ident.	0	193	2.07	1.09	0.00	4.00
49	geff	184	-0.01	1.00	-2.50	2.44	ident.	12.50	193	-12.46	1.01	-15.00	-10.06
50	rlaw	185	-0.01	1.00	-2.31	2.20	ident.	12.31	193	0.05	1.03	-2.31	2.20
51	regqua	185	0.00	0.97	-2.50	2.31	ident.	12.50	193	-12.45	0.97	-15.00	-10.19
52	lawor	138	3.93	1.38	1.00	6.00	log	0	193	3.84	1.28	1.00	6.44
Geo-Cultural Determinants													
53	budha	191	4.27	16.71	0.00	93.39	ident.	0	193	4.31	16.62	0.00	93.39
54	hindu	191	2.29	9.61	0.00	80.47	ident.	0	193	2.32	9.57	0.00	80.47

Continued on next page

Continued from previous page

No.	Variable	Before Imputation					Transf.	Const.	After Imputation				
		Obs.	Mean	Std. Dev.	Min.	Max			Obs.	Mean	Std. Dev.	Min.	Max.
55	muslim	191	25.60	36.54	0.00	99.67	ident.	0	193	25.39	36.40	0.00	99.67
56	nonrelig	191	6.20	11.36	0.01	94.09	log	0	193	6.18	11.31	0.01	94.09
57	anglic	191	2.72	7.68	0.00	52.41	ident.	0	193	2.75	7.65	0.00	52.41
58	cathol	191	30.59	32.78	0.00	97.15	sq.	0	193	30.90	32.75	0.00	97.15
59	indepen	191	6.26	10.36	0.01	69.00	log	0	193	6.21	10.32	0.01	69.00
60	marginal	191	1.30	3.64	0.00	36.64	ident.	0	193	1.31	3.62	0.00	36.64
61	orthodox	191	5.80	17.45	0.00	89.77	ident.	0	193	5.81	17.36	0.00	89.77
62	protest	191	14.96	20.54	0.00	90.66	sq.rt.	0	193	14.88	20.45	0.00	90.66
63	ethnoa	148	0.48	0.27	0.00	0.95	sq.rt.	0	193	0.46	0.25	0.00	0.95
64	ethnob	154	0.33	0.30	0.00	1.00	sq.rt.	0	193	0.31	0.28	0.00	1.00
65	englis	191	0.33	0.47	0.00	1.00	ident.	0	193	0.33	0.47	0.00	1.00
66	social	191	0.18	0.39	0.00	1.00	ident.	0	193	0.18	0.39	0.00	1.00
67	french	191	0.43	0.50	0.00	1.00	ident.	0	193	0.42	0.50	0.00	1.00
68	german	191	0.03	0.17	0.00	1.00	ident.	0	193	0.03	0.17	0.00	1.00
69	scandi	191	0.03	0.16	0.00	1.00	ident.	0	193	0.03	0.16	0.00	1.00
70	latitu	191	0.28	0.19	0.00	0.72	sq.rt.	0	193	0.28	0.19	0.00	0.72
71	areakm	193	690010.80	1921548.00	25.40	17100000.00	log	0	193	690138.90	1922649.00	25.40	17100000.00

In doing the EM iteration, we use GDP per Capita, population density, and country area as the predictors, since in terms of the number of observations these are the most complete variables we have at hand.[19] After the imputation, we transform the variables back to the original scales. Table 7 compares the data before and after the imputation.

Now we have a complete data set. The next step is to generate z-scores from the imputed data that have been transformed back to their original scales. This is to have a new data set with mean zero and unit variance. We drop seven categorical variables in this stage because these variables appear as binary dummy variables; thus the z-scores are not appropriate to be applied for binary values. They are political polarization, plurality, and five legal origins, namely english, socialism, french, german, and scandinavia.

Having the z-score at hand, we do EFA to uncover the latent structure (called also dimensions or factors) of our data set and to reduce attribute space from a larger number of variables to a smaller number of factors. Moreover, to have a clearer structure and an easier interpretation of the factors, we rotate the loadings by employing the varimax method. The varimax searches for an orthogonal rotation (i.e., a linear combination) of the original factors such that the variance of the loadings is maximized. Each factor will tend to have either large or small loadings of any particular variable.

In Table 8 we have selected five out of 43 rotated factors that consist of at least two variables with very high factor loadings (i.e., > 0.710 or < -0.710)[20] In *Factor 1* there are 12 variables clustered together with high factor loadings, namely rule of law, judicial independence and impartial court (area2b of Fraser index), government effectiveness, GDP per capita, political stability, regulatory quality, bureaucratic quality, law and order, labor market regulation (area5b of Fraser index), international trade (area4b of Fraser index), internal conflict, and secondary school enrolment. Clearly, this factor is dominated by variables reflecting the capacity of government to regulate and enforce law. Therefore, we call this factor *regulatory capacity*.

[19]Since we have seven binary dummy variables, we follow a simple rule to maintain their original scale: it takes one if the predicted value is >0.500, and zero otherwise.

[20]In our case, these benchmark make the pattern of the factors are much clearer. At the same time, they are also reasonable as the squares of these values exceed 0.500, implying that more than 50% proportion of each manifest variable is explained by the factor.

Table 8: Selected Rotated Factor Loadings

No	Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	...	Factor 43	Uniqueness
1	rlaw	0.929	0.108	-0.075	0.033	-0.225		0.116	0.015
2	geff	0.893	0.095	-0.091	0.010	-0.214		0.077	0.051
3	area2ab	0.892	0.209	-0.091	0.054	-0.160		-0.048	0.040
4	gdpcap	0.850	0.152	-0.082	0.129	-0.155		-0.104	0.055
5	psta	0.817	-0.007	-0.029	0.115	-0.240		0.011	0.060
6	regqua	0.815	0.009	0.018	-0.012	-0.359		0.046	0.060
7	buqua	0.809	0.160	-0.127	0.059	-0.292		-0.006	0.109
8	lawor	0.776	0.136	-0.152	0.054	-0.009		0.044	0.135
9	area5b	0.766	0.031	0.034	0.115	-0.165		-0.047	0.167
10	area4b	0.743	0.110	-0.043	0.169	-0.134		-0.150	0.180
11	inco	0.728	-0.018	-0.011	0.104	-0.116		0.079	0.139
12	enrols	0.710	0.132	-0.152	0.041	-0.225		-0.023	0.110
13	snexe	0.059	0.954	0.002	-0.086	0.005		0.028	0.008
14	snrer	0.051	0.949	0.033	-0.027	0.025		0.027	0.014
15	snreg	0.323	0.858	-0.046	0.019	-0.076		-0.064	0.016
16	snexg	0.392	0.821	-0.082	-0.019	-0.108		-0.054	0.012
17	rich20	-0.104	-0.011	0.909	-0.008	-0.007		0.036	0.131
18	gini	-0.234	-0.079	0.852	-0.032	0.070		-0.001	0.084
19	rich10	-0.120	0.041	0.832	-0.066	-0.063		-0.046	0.197
20	open	0.126	-0.026	-0.028	0.987	-0.036		0.004	0.000
21	mpor	0.065	-0.102	-0.001	0.925	-0.045		-0.014	0.004
22	xpor	0.150	0.058	-0.051	0.909	-0.037		0.012	0.006
23	pr	0.454	-0.010	0.051	0.033	-0.834		0.014	0.041
24	press	0.521	0.013	-0.037	0.046	-0.768		-0.037	0.058
25	poli	0.289	0.132	0.063	0.029	-0.768		-0.017	0.149
26	cl	0.542	-0.034	0.015	0.036	-0.756		0.054	0.031
27	voic	0.601	0.001	-0.031	0.038	-0.749		0.002	0.009
28	area3	0.699	0.062	0.020	-0.080	-0.067		-0.289	0.268
29	presiden	0.469	-0.029	-0.188	0.079	-0.313		0.019	0.343
30	eten	0.429	-0.045	-0.003	-0.039	0.030		-0.036	0.384
31	exco	0.400	-0.079	0.012	0.155	-0.305		-0.004	0.383
32	latitu	0.376	0.336	-0.299	0.041	-0.135		-0.012	0.224
33	demac	0.366	0.155	-0.127	0.046	-0.693		-0.046	0.186
34	enrolt	0.309	0.049	-0.089	-0.060	-0.044		-0.018	0.583
35	gsta	0.232	0.120	0.096	0.004	0.327		-0.007	0.403
36	efherit	0.219	0.013	-0.096	0.016	0.087		-0.009	0.593
37	enrolp	0.217	-0.022	0.065	0.100	-0.101		0.011	0.427
38	debtcap	0.175	0.238	-0.026	-0.057	-0.098		-0.026	0.433
39	popden	0.171	-0.157	-0.039	0.132	0.021		0.000	0.172
40	protest	0.170	0.059	0.071	-0.002	-0.238		0.002	-0.002

Continued on next page

Continued from previous page									
No	Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	...	Factor 43	Uniqueness
41	cathol	0.134	-0.080	0.126	0.107	-0.271		-0.007	-0.002
42	anglic	0.128	-0.124	-0.013	-0.001	-0.070		-0.001	-0.001
43	nonrelig	0.097	0.209	-0.128	0.011	0.037		-0.001	-0.002
44	wage	0.053	-0.346	0.098	0.041	0.040		0.007	0.390
45	marginal	0.041	-0.104	0.024	-0.005	-0.091		0.000	0.005
46	distmag	0.005	0.110	-0.011	0.211	-0.103		-0.004	0.546
47	hindu	-0.014	-0.009	-0.051	-0.014	-0.057		-0.001	-0.001
48	budha	-0.037	-0.038	-0.066	-0.061	0.086		0.003	-0.002
49	indepen	-0.054	0.044	0.270	0.008	-0.036		0.003	-0.001
50	aidcap	-0.062	-0.136	-0.031	0.074	-0.141		-0.008	0.424
51	govcon	-0.069	0.114	-0.043	0.138	0.112		-0.007	0.480
52	areakm	-0.075	0.580	0.049	0.219	-0.009		0.097	0.226
53	area1	-0.083	-0.168	0.212	0.015	0.013		-0.006	0.328
54	debtgdp	-0.087	-0.159	-0.042	-0.002	0.033		-0.051	0.393
55	orthodox	-0.095	0.044	-0.135	0.083	-0.081		0.003	-0.003
56	xmetal	-0.121	0.054	0.122	-0.072	-0.042		-0.025	0.707
57	femlab	-0.150	0.049	-0.188	0.012	-0.322		-0.017	0.309
58	xfuel	-0.175	0.221	0.028	0.032	0.234		-0.008	0.493
59	muslim	-0.194	-0.005	-0.082	-0.104	0.414		0.000	-0.002
60	ethnob	-0.421	0.046	0.141	-0.003	0.063		-0.052	0.271
61	debt	-0.433	-0.092	0.050	-0.014	-0.021		0.136	0.384
62	ethnoa	-0.455	0.142	0.136	-0.054	0.047		0.054	0.232
63	illi	-0.492	-0.052	0.079	-0.182	0.174		-0.003	0.148

Factor 2 includes all proxies of sub national government expenditures and revenues; thus we call this *federalism*. In *Factor 3* we have three variables measuring income inequality, namely gini ratio, the richest 10%, and the richest 20%). We call this factor *inequality*. In *Factor 4* there are three measures of international trade clustered together, namely export ratio, import ratio, and trade volume; thus we call it *trade*. Finally, *Factor 5* may explain political liberty since it captures five correlated political variables, namely political right, press freedom, polity index, civil librety, and voice. We call this factor *political liberty*.

Now taking only variables with substantial loadings (thus ignoring the remaining variables with minor loadings), we generate five new indexes on the basis of the z-scores. These factor-based indexes are computed as follows:

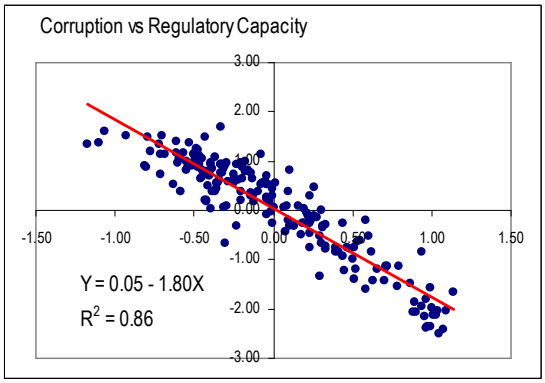
$$\hat{F}_i = \sum_j \frac{\beta_{ij}}{\lambda_i} X_j \tag{5}$$

where $\hat{F}$ is the index we want to construct, X's are the underlying variables (z-score) of the index with high factor loadings, β is the rotated factor loading, and λ is the eigenvalue. This formula is applied for every index we construct.

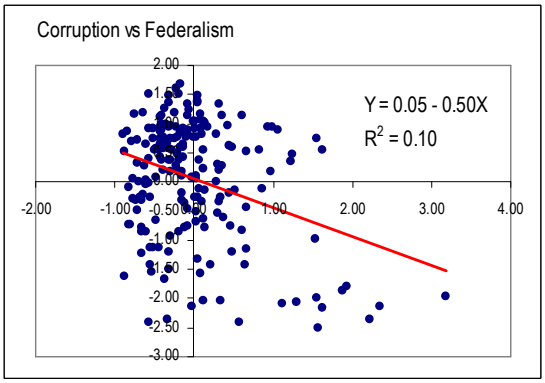
Graphs 1-5 display the plots of corruption against the five new indexes.[21] It is obvious from the Plots that only 'regulatory capacity' and 'political liberty' are related to corruption. The R^2 of the first plot is 0.86 and the fifth's is 0.44, while the rest are far below these values. Although suggestive, these results are only based on bivariate regressions. The next section examines whether these newly generated indexes (and the other non-clustered variables) are robustly correlated with corruption, but first we outline our methodology.

[21] Compared to the other four factor loadings, the loadings of variables falling in *political liberty* are all negative. Hence the interpretation should be reversed: higher scores indicate less political freedom.

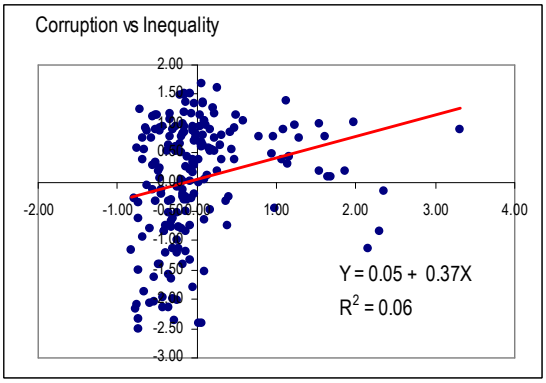
Plot 1:



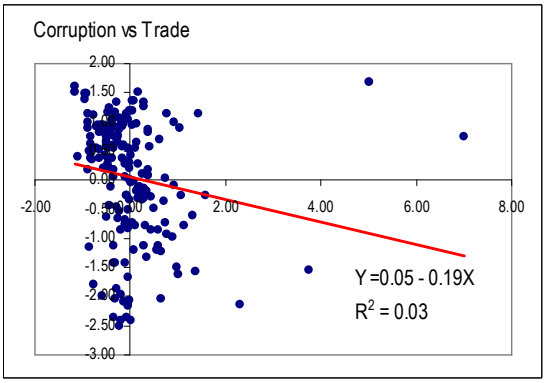
Plot 2:



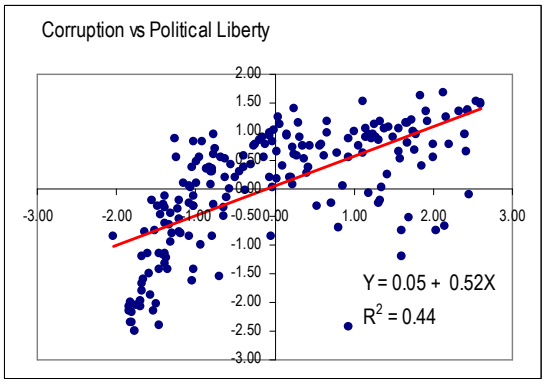
Plot 3:



Plot 4:



Plot 5:



Scatter Plots 1-5: Corruption vs New Indexes

5 Extreme Bounds Analysis

Like in empirical growth models (Temple, 2000), model uncertainty is an important problem in empirical models of the causes of corruption. Researchers usually report a preferred model followed by the results of some diagnostic tests. However, the problem is that "several different models may all seem reasonable given the data, but lead to very different conclusions about the parameters of interest" (Temple, 2000). In such a situation, reporting the result of a single preferred model is misleading, since it underestimates the uncertainty actually present about the parameters. At the same time, there is no strong theoretical framework to help researchers establishing a proper regression model for corruption. This results in a large variety of corruption regression models as well as ways to examine the robustness of variables of interest.

One way to cope with model uncertainty is the extreme bound analysis (EBA).[22] The idea behind EBA is to report an upper and lower bound for parameter estimates, that is to examine the sensitivity of parameters to model specification. In this paper we use the EBA of Levine and Renelt (1992) and Sala-i-Martin's (1997). We include a large number of variables that have been claimed to be related to corruption in previous studies. Our study addresses the question of how much confidence one should have in the conclusion of previous studies.

The EBA can be exemplified as follows (Leamer, 1983; Levine and Renelt, 1992):

$$Y = \boldsymbol{\alpha}_j + \boldsymbol{\beta}_{ij}\boldsymbol{I} + \beta_{mj}M + \boldsymbol{\beta}_{zj}\boldsymbol{Z}_j + u \tag{6}$$

where Y is the dependent variable (Kaufmann's corruption index 2004); $\boldsymbol{I}$ is a vector of "standard" explanatory variables (which may be zero); M is the variable of interest; $\boldsymbol{Z}$ is a vector of up-to-three possible additional explanatory variables, which according to the literature may be related to the dependent variable; and u is an error term. It should be noted that number of variables in $\boldsymbol{I}$ and $\boldsymbol{Z}$ that can be plugged into the model is constrained by the degrees of freedom of the regression as well as the issue of multicollinearity. Levine and Renelt (1992) include three variables (the "fixed-trio") in $\boldsymbol{I}$ and all possible combinations of up-to-three variables (the "flex-trio") in $\boldsymbol{Z}$. However, due to the lack of theoretical guidance and the wide variety of results reported in previous studies, we have decided not to include any variable in the $\boldsymbol{I}$ vector.[23]

The test of extreme bounds for variable M says that if the lower extreme bound for β —i.e., the lowest value for β minus two standard deviations— is negative, while the

[22] There is a growing literature starting with Leamer (1983) and Levine and Renelt (1992), followed by a critique by Sala-i-Martin (1997) and Durlauf and Quah (1999). Recently, Doppelhofer et al. (2005) have proposed Bayesian Averaging of Classical Estimates (BACE) approach to check the robustness of different explanatory variables in growth regressions, while Hendry and Krolzig (2004) suggest the General Unrestricted Model (GUM) which basically needs very few regression to test the robustness of a variable of interest.

[23] Later on, after finding a robust variable(s), we also include this (these) variable(s) in $\boldsymbol{I}$.

upper extreme bound for β —i.e. the highest value for β plus two standard deviations— is positive, the variable M is not robustly related to Y. More formally, the upper and lower bounds are defined as (Levine and Renelt, 1992):

$$\underset{min}{\overset{max}{}}\hat{\beta}_{mj} \pm n\,\sigma_{\hat{\beta}_{mj}} \tag{7}$$

where $n = 2$, and $\sigma_{\hat{\beta}_{mj}}$ is standard deviation of $\hat{\beta}_{mj}$. The robustness thus has two properties. First, the coefficients of β_{mj} in the upper and lower bounds must be consistent. Second, they are significant at the conventional level under various conditioning sets of Z. A violation against these properties makes a variable to be regarded as fragile.

Sala-i-Martin (1997), however, rightly argues that the test applied in the Leamer's as well as the Levine and Renelt's EBA is too strong for any variable to really pass it. If the distribution of the parameter of interest has some positive and some negative supports, one is bound to find one regression for which the estimated coefficient changes sign if enough regressions are run. He therefore suggests analyzing the entire distribution of the estimates of the parameter of interest. Broadly speaking, if the averaged 90% confidence interval of a regression coefficient does not include zero, Sala-i-Martin classifies the corresponding regressor as a variable that is robust.

In our empirical analysis we will use both versions of the EBA. But, since we agree with the critique of Sala-i-Martin (1997) on Leamer's (and Levine and Renelt's) version of the EBA, our conclusion will be based on the Sala-i-Martin variant of the EBA. We not only report the unweighted parameter estimates of β and fraction of regressions significant at 5%, but also the outcomes of the cumulative distribution function (CDF) test. The CDF test is based on the fraction of the cumulative distribution function lying on each side of zero. Since zero divides CDF into two areas (CDF[0] or 1-CDF[0]), no matter whether it is below or above zero, attention is paid only to the largest of the two areas, that is

$$CDF = \max[CDF(0), 1 - CDF(0)] \tag{8}$$

CDF(0) indicates the larger of the areas under the density function either above or below zero, regardless of whether this is CDF(0) or 1-CDF(0). So CDF(0) will always be a number between 0.5 and 1.0. However, instead of using the 90% criterion, we advocate a more stringent criterion, i.e., 95% because of the one-sidedness of the test (Sturm and De Haan, 2005).

Some assumptions must be made to calculate the CDF using the integrated likelihood (L), the point estimate ($\hat{\beta}_{mj}$), and the standard deviation (σ^2_{mj}). First, if $\hat{\beta}_m$ is distributed normally, the weighted mean is obtained by

$$\hat{\beta}_m = \sum_{j=1}^{V} \frac{L_{mj}}{\sum_{i=1}^{V} L_{zi}} \hat{\beta}_{mj}$$

and the weighted mean of variance is

$$\hat{\sigma}_m^2 = \sum_{j=1}^{V} \frac{L_{mj}}{\sum_{i=1}^{V} L_{zi}} \sigma_{mj}^2.$$

Second, if the distribution of $\hat{\beta}_m$ is non-normal, the aggregate CDF can be computed using the individual CDF ($\Phi_{mj}(\hat{\beta}_{mj}, \hat{\sigma}_{mj}^2)$) and the weighted likelihood as before:

$$\Phi_m = \sum_{j=1}^{V} \frac{L_{mj}}{\sum_{i=1}^{V} L_{zi}} \Phi_{mj}(\hat{\beta}_{mj}, \hat{\sigma}_{mj}^2)$$

After knowing its distribution, a variable is labelled robust if 95% of the density function for $\hat{\beta}_m$ lies to the right or left of zero.

We have 48 variables to be used in the EBA, since 27 variables have been replaced by five new indexes.[24] The total number of regressions we run is $\frac{v!}{3!(v-3)!}$; with 48 variables and no $\boldsymbol{I}$ we have 778,320 total regressions or 16,215 regressions per variable we test. In Table 9 columns 3-5 show the outcomes of Leamer's variant of EBA, namely the lower and upper extreme bounds and the fraction of the regressions in which the variable is significantly different from zero. Columns 6-12 show the results of Sala-i-Martin's variant of EBA, namely the estimated coefficients and standard errors as well as normal and non-normal CDF(0). The variables are ordered on the basis of the normal CDF(0).

[24]Now we include the eight categorical variables that were dropped in the factor analysis.

Table 9: Extreme Bounds Analysis: No Variable in 'I'

No	Dependent Variable	Leamer			Sala-i-Martin						
		Lower	Upper	Fraction	Weighted		Unweighted		Normal	Non-Normal CDF	
		Bound	Bound	Signf. at 5%	Beta	St.Dev.	Beta	St.Dev.	CDF	Weighted	Unwght.
1	regcap	-2.242	-1.426	100.000	-1.783	0.061	-1.808	0.065	1.000	1.000	1.000
2	scandi	-3.786	0.488	99.833	-0.685	0.078	-1.872	0.158	1.000	1.000	1.000
3	popden	-0.000	+0.000	99.093	-0.000	0.000	-0.000	0.000	1.000	1.000	0.995
4	eten	-0.448	0.121	97.373	0.072	0.020	-0.190	0.053	1.000	1.000	0.933
5	social	-0.274	3.210	94.573	0.232	0.058	0.588	0.145	1.000	1.000	0.994
6	nonrelig	-0.055	0.031	12.081	0.006	0.002	-0.002	0.007	1.000	0.993	0.561
7	illi	-0.010	0.043	99.741	-0.004	0.001	0.019	0.003	0.999	0.995	0.936
8	wage	-0.071	0.086	18.693	0.015	0.006	0.014	0.015	0.992	0.988	0.759
9	area3	-0.500	0.105	94.345	0.049	0.021	-0.308	0.041	0.991	0.985	0.939
10	latitu	-4.677	0.503	93.827	-0.452	0.215	-1.959	0.373	0.982	0.916	0.978
11	xfuel	-0.008	0.021	74.012	0.002	0.001	0.007	0.003	0.957	0.956	0.954
12	enrolp	-0.020	0.015	71.483	0.003	0.001	-0.006	0.003	0.955	0.952	0.855
13	debt	-0.002	0.009	92.390	-0.001	0.000	0.003	0.001	0.954	0.951	0.933
14	presiden	-0.825	0.029	94.357	-0.056	0.033	-0.482	0.070	0.954	0.952	0.997
15	femlab	-0.041	0.048	11.964	-0.006	0.004	0.005	0.008	0.950	0.914	0.638
16	ethnoa	-0.510	2.454	92.982	-0.174	0.108	1.212	0.244	0.946	0.933	0.942
17	exco	-0.417	0.091	91.644	0.030	0.021	-0.211	0.040	0.927	0.924	0.939
18	aidcap	-0.001	0.003	24.298	0.000	0.000	0.001	0.000	0.924	0.916	0.875
19	area1	-0.167	0.351	31.878	0.027	0.021	0.096	0.058	0.901	0.861	0.901
20	inequal	-0.222	0.848	80.401	0.060	0.047	0.271	0.094	0.901	0.867	0.967
21	french	-1.268	2.925	26.562	-0.064	0.056	0.176	0.130	0.874	0.836	0.788
22	german	-2.684	1.356	92.821	-0.154	0.140	-1.159	0.285	0.863	0.855	0.980
23	marginal	-0.069	0.056	8.967	0.007	0.007	0.000	0.012	0.854	0.844	0.539
24	pollib	-0.071	0.726	93.617	0.033	0.032	0.446	0.045	0.848	0.840	0.992
25	polariz	-0.898	0.087	93.370	-0.036	0.038	-0.472	0.086	0.825	0.799	0.994
26	debtgdp	-0.012	0.011	32.865	0.001	0.001	0.003	0.002	0.805	0.781	0.867
27	indepen	-0.023	0.041	9.127	0.002	0.003	0.006	0.006	0.782	0.772	0.794

Continued on next page

Continued from previous page

No	Dependent Variable	Leamer			Sala-i-Martin						
		Lower	Upper	Fraction	Weighted		Unweighted		Normal	Non-Normal CDF	
		Bound	Bound	Signf. at 5%	Beta	St.Dev.	Beta	St.Dev.	CDF	Weighted	Unwght.
28	budha	-0.014	0.017	1.560	-0.002	0.002	0.002	0.003	0.780	0.773	0.649
29	hindu	-0.025	0.023	7.832	-0.002	0.002	-0.001	0.004	0.744	0.741	0.555
30	debtcap	0.000	0.000	56.238	0.000	0.000	0.000	0.000	0.738	0.719	0.915
31	distmag	-0.010	0.012	1.122	0.000	0.001	-0.001	0.003	0.735	0.681	0.594
32	efherit	-0.596	0.146	71.428	-0.021	0.036	-0.223	0.096	0.722	0.719	0.963
33	federal	-1.074	0.216	90.052	-0.026	0.044	-0.401	0.100	0.719	0.685	0.988
34	demac	-0.466	0.167	87.512	-0.011	0.020	-0.245	0.043	0.710	0.708	0.960
35	trade	-0.585	0.239	0.432	0.019	0.042	-0.130	0.116	0.677	0.679	0.822
36	anglic	-0.050	0.025	32.544	-0.001	0.003	-0.014	0.009	0.653	0.649	0.894
37	gsta	-0.521	0.157	61.813	0.010	0.029	-0.176	0.078	0.644	0.637	0.926
38	orthodox	-0.014	0.023	32.821	0.000	0.001	0.004	0.003	0.631	0.589	0.852
39	cathol	-0.015	0.009	48.930	0.000	0.001	-0.004	0.002	0.604	0.600	0.821
40	protest	-0.027	0.019	73.543	0.000	0.002	-0.010	0.004	0.580	0.575	0.950
41	govcon	-1.347	2.207	10.589	0.036	0.187	0.390	0.380	0.575	0.577	0.785
42	enrolt	-0.028	0.004	92.871	0.000	0.002	-0.012	0.004	0.564	0.563	0.967
43	areakm	0.000	0.000	5.908	0.000	0.000	0.000	0.000	0.561	0.561	0.555
44	ethnob	-0.314	2.275	93.315	0.012	0.106	1.119	0.189	0.545	0.543	0.958
45	xmetal	-0.014	0.032	23.121	0.000	0.003	0.008	0.006	0.538	0.531	0.870
46	pluralty	-0.234	1.062	84.823	0.004	0.062	0.421	0.132	0.523	0.521	0.970
47	englis	-1.620	2.660	26.568	0.003	0.064	-0.182	0.140	0.516	0.513	0.814
48	muslim	-0.011	0.017	73.981	0.000	0.001	0.005	0.002	0.509	0.516	0.870

It is clear from the Table that only one of the five indexes we constructed, namely regulatory capacity, can pass the two tests. In terms of the Leamer's (or the Levine-Renelt's) test, this variable has consistent signs both in the lower and upper extreme bounds, and its coefficients is in 16,215 regressions always significant at the 5% level. In terms of the Sala-i-Martin's test, this variable is normally distributed[25] and all coefficients form one-side CDF. In short, we can conclude that 'regulatory capacity' — consisting of 11 variables— is a robust determinant of corruption. Thus, the message is straight forward: an increase in government regulatory capacity strongly reduces corruption.

We turn now to the rest of robust variables. As we take Sala-i-Martin's test as the benchmark we have 14 other robust variables, namely scandinavian legal origin (negative effect), population density (–)[26], socialism legal origin (+), portion of population with no religion (+), ethnic conflict (+), illiteracy rate (–), government wage (+), sound money (area3 of Fraser index; +), latitude (–), fuel export (+), primary scholl enrollment (+), external debt (–), presidential (–), and portion of female in labor force (–). Two variables are found counter-intuitive, namely illiteracy rate and wage. An increase in illiteracy rate reduces corruption, while an increase in government wage lifts up corruption. The first case is close to Frechette (2001) for schooling variable, but contradicts with Ali and Isse (2003), Alt and Lassen (2003), Brunetti and Weder (2003), Persson et al. (2003), Evan and Rauch (2000), Ades and di Tella (1997; 1999), and van Rijckeghem-Weder (1997). The second case contradicts with Alt and Lassen (2003), Herzfeld and Weiss (2003), Evan and Rauch (2000), and van Rijckeghem-Weder (1997).

Lastly, since we now convincingly have a robust variable (i.e., regulatory capacity) as it passes the two tests, we use this variable as a control variable to be plugged in I. In other words, we run EBA with regulatory capacity as I. After running 713,460 regressions we cannot find any new additional variable to be regarded as a variable robust (Table 10). But now, we find that population density (–), scandinavian legal origin (–), and ethnic conflict (+) pass the two tests.

[25]The rest of the variables also tend to be normal since the correlation between the normal CDF and the non-normal weighted CDF is high.

[26]Although size of the effect is almost zero due to the different unit of measurement between the dependent and independent variables, the coefficient sign is, in fact, negative. This result is counter-intuitive, but it is supported by Damania et al. (2004), Alt and Lassen (2003), Knack and Azfar (2003), and Fisman and Gatti (2002).

Table 10: Extreme Bounds Analysis: Regulatory Capacity in 'I'

No	Dependent Variable	Leamer			Sala-i-Martin						
		Lower	Upper	Fraction	Weighted		Unweighted		Normal	Non-Normal CDF	
		Bound	Bound	Signf. at 5%	Beta	St.Dev.	Beta	St.Dev.	CDF	Weighted	Unwght.
1	popden	-0.000	-0.000	100.000	-0.000	0.000	-0.000	0.000	1.000	1.000	1.000
2	scandi	-1.330	-0.214	100.000	-0.684	0.082	-0.684	0.085	1.000	1.000	1.000
3	eten	0.004	0.126	100.000	0.071	0.020	0.067	0.021	1.000	1.000	0.999
4	social	-0.027	1.018	99.974	0.237	0.060	0.256	0.063	1.000	1.000	1.000
5	nonrelig	-0.003	0.014	93.933	0.005	0.002	0.006	0.002	1.000	0.984	0.993
6	illi	-0.010	0.003	94.473	-0.004	0.001	-0.005	0.001	0.998	0.988	0.996
7	area3	-0.024	0.112	15.949	0.049	0.021	0.037	0.022	0.990	0.986	0.951
8	wage	-0.005	0.036	91.680	0.015	0.006	0.016	0.007	0.989	0.983	0.989
9	latitu	-1.195	0.570	6.680	-0.415	0.216	-0.106	0.182	0.973	0.921	0.662
10	xfuel	-0.002	0.005	1.094	0.002	0.001	0.002	0.001	0.961	0.960	0.901
11	presiden	-0.157	0.039	14.401	-0.055	0.033	-0.060	0.034	0.951	0.949	0.957
12	debt	-0.002	0.000	23.149	-0.001	0.000	-0.001	0.000	0.949	0.945	0.967
13	enrolp	-0.003	0.007	58.676	0.002	0.001	0.003	0.002	0.946	0.941	0.965
14	femlab	-0.018	0.008	9.750	-0.006	0.004	-0.004	0.003	0.942	0.919	0.850
15	ethnoa	-0.592	0.262	1.403	-0.158	0.110	-0.142	0.113	0.925	0.900	0.882
16	aidcap	0.000	0.001	7.839	0.000	0.000	0.000	0.000	0.915	0.906	0.929
17	exco	-0.030	0.096	0.777	0.028	0.021	0.031	0.021	0.908	0.904	0.922
18	inequal	-0.105	0.212	0.079	0.060	0.046	0.032	0.046	0.902	0.880	0.742
19	area1	-0.046	0.105	5.632	0.026	0.021	0.028	0.021	0.895	0.869	0.895
20	german	-0.647	0.779	0.013	-0.157	0.136	-0.077	0.142	0.876	0.865	0.702
21	marginal	-0.015	0.027	0.329	0.007	0.007	0.007	0.007	0.852	0.844	0.841
22	pollib	-0.091	0.188	0.099	0.030	0.033	0.041	0.035	0.818	0.808	0.871
23	french	-0.406	0.785	2.424	-0.052	0.058	-0.057	0.057	0.813	0.783	0.818
24	debtgdp	-0.002	0.004	0.619	0.001	0.001	0.001	0.001	0.809	0.785	0.743
25	budha	-0.008	0.005	0.000	-0.002	0.002	-0.001	0.002	0.804	0.799	0.671
26	indepen	-0.006	0.011	0.000	0.002	0.003	0.002	0.003	0.802	0.792	0.755
27	orthodox	-0.004	0.005	0.257	-0.001	0.001	0.001	0.001	0.771	0.717	0.785

Continued on next page

Continued from previous page

No	Dependent Variable	Leamer			Sala-i-Martin						
		Lower	Upper	Fraction	Weighted		Unweighted		Normal	Non-Normal CDF	
		Bound	Bound	Signf. at 5%	Beta	St.Dev.	Beta	St.Dev.	CDF	Weighted	Unwght.
28	debtcap	0.000	0.000	9.756	0.000	0.000	0.000	0.000	0.771	0.741	0.860
29	efherit	-0.120	0.074	0.000	-0.024	0.035	-0.022	0.037	0.755	0.750	0.723
30	polariz	-0.158	0.101	0.277	-0.026	0.039	-0.053	0.039	0.749	0.724	0.902
31	distmag	-0.002	0.003	0.033	0.000	0.001	0.001	0.001	0.736	0.701	0.774
32	ethnob	-0.328	0.462	0.000	0.068	0.109	-0.031	0.109	0.733	0.713	0.614
33	protest	-0.009	0.007	6.950	0.001	0.002	-0.003	0.002	0.703	0.684	0.891
34	demac	-0.076	0.074	0.000	-0.010	0.019	-0.015	0.021	0.694	0.692	0.752
35	federal	-0.241	0.136	4.124	-0.021	0.044	-0.055	0.046	0.685	0.660	0.864
36	trade	-0.099	0.136	0.000	0.020	0.042	0.029	0.041	0.684	0.683	0.761
37	hindu	-0.008	0.007	0.000	-0.001	0.002	-0.002	0.002	0.672	0.669	0.747
38	englis	-0.411	0.802	0.586	0.023	0.065	-0.009	0.064	0.639	0.592	0.551
39	gsta	-0.090	0.116	0.000	0.009	0.028	0.014	0.033	0.632	0.626	0.656
40	cathol	-0.004	0.004	0.000	0.000	0.001	0.000	0.001	0.619	0.611	0.634
41	anglic	-0.012	0.011	0.000	-0.001	0.003	-0.001	0.003	0.617	0.614	0.589
42	xmetal	-0.006	0.008	0.000	0.001	0.003	0.000	0.002	0.596	0.587	0.518
43	govcon	-0.590	0.597	0.000	0.029	0.188	-0.006	0.187	0.561	0.562	0.512
44	muslim	-0.004	0.004	0.020	0.000	0.001	0.000	0.001	0.543	0.532	0.591
45	enrolt	-0.005	0.005	0.000	0.000	0.002	0.000	0.002	0.543	0.541	0.523
46	pluralty	-0.158	0.211	0.000	0.004	0.061	0.033	0.064	0.523	0.521	0.696
47	areakm	0.000	0.000	0.079	0.000	0.000	0.000	0.000	0.518	0.516	0.503

Up to this point we have succesfully generated five simple factor-based indexes incorporating only variables with high (rotated factor) loadings. Now we turn to another computation technique to generate the indexes. This is to examine whether different method of index (or score) generation result in different impact of the robustness of the variables. To do this, we first do Confirmatory Factor Analysis (CFA) on the basis of the variables that have been rotatedly grouped in the EFA step[27], then we generate the indexes. Different from the previous indexes which are also called factor-based scores or indexes, our indexes are now usually known as factor scores computed on the basis of regression scoring method as follows:

$$\hat{\mathbf{F}} = \mathbf{X}(\mathbf{B}'\mathbf{R}^{-1}) \tag{9}$$

where B is the matrix of factor loadings, X is the matrix of the observed variables, and R is the correlation matrix for the X's.

Using the resulting indexes we run again the two tests of sensitivity analysis as above. The results are displayed in Tables 11 (without I) and 12 (with regulatory capacity as I). The results are not much different. Regulatory capacity is again found robust as it passes the two tests. The rest of variables perform the similar pattern as in the previous. Thus, we can conclude that the different computation technique does not result in different outcomes.

[27]In other words, we use the prior information drawn from the EFA.

Table 11: Extreme Bounds Analysis: No Variable in 'I' (CFA Score)

No	Dependent Variable	Leamer			Sala-i-Martin						
		Lower	Upper	Fraction	Weighted		Unweighted		Normal	Non-Normal CDF	
		Bound	Bound	Signf. at 5%	Beta	St.Dev.	Beta	St.Dev.	CDF	Weighted	Unwght.
1	regcap	-1.160	-0.779	100.000	-0.946	0.031	-0.963	0.033	1.000	1.000	1.000
2	scandi	-3.786	0.488	99.833	-0.641	0.078	-1.866	0.158	1.000	1.000	1.000
3	popden	-0.000	+0.000	98.822	-0.000	0.000	-0.000	0.000	1.000	1.000	0.994
4	eten	-0.448	0.108	97.367	0.065	0.018	-0.190	0.053	1.000	1.000	0.933
5	social	-0.274	3.210	94.450	0.199	0.053	0.585	0.144	1.000	1.000	0.994
6	nonrelig	-0.055	0.031	12.125	0.005	0.001	-0.002	0.007	1.000	0.993	0.560
7	illi	-0.009	0.043	99.698	-0.004	0.001	0.019	0.003	0.998	0.994	0.936
8	wage	-0.072	0.086	19.593	0.016	0.006	0.014	0.015	0.993	0.990	0.763
9	area3	-0.500	0.096	94.357	0.042	0.020	-0.308	0.041	0.983	0.975	0.939
10	ethnoa	-0.521	2.454	93.555	-0.210	0.101	1.205	0.243	0.982	0.974	0.939
11	femlab	-0.041	0.048	12.532	-0.007	0.003	0.005	0.008	0.982	0.961	0.632
12	xfuel	-0.008	0.021	74.419	0.002	0.001	0.007	0.003	0.981	0.980	0.956
13	debt	-0.002	0.009	97.083	-0.001	0.000	0.003	0.001	0.978	0.977	0.932
14	exco	-0.417	0.085	89.948	0.030	0.020	-0.211	0.040	0.934	0.930	0.937
15	enrolp	-0.020	0.015	67.512	0.002	0.001	-0.006	0.003	0.932	0.929	0.853
16	presiden	-0.825	0.036	93.623	-0.045	0.031	-0.480	0.070	0.924	0.921	0.996
17	french	-1.268	2.925	26.253	-0.068	0.052	0.175	0.130	0.905	0.876	0.786
18	aidcap	-0.001	0.003	24.027	0.000	0.000	0.001	0.000	0.875	0.866	0.871
19	debtcap	0.000	0.000	56.516	0.000	0.000	0.000	0.000	0.844	0.811	0.912
20	latitu	-4.677	0.517	93.784	-0.165	0.192	-1.953	0.373	0.805	0.645	0.972
21	marginal	-0.069	0.060	9.405	0.005	0.006	0.000	0.012	0.802	0.794	0.538
22	pollib	-0.984	0.108	93.617	-0.034	0.042	-0.598	0.058	0.793	0.786	0.987
23	inequal	-0.152	0.599	79.920	0.024	0.030	0.189	0.066	0.790	0.747	0.964
24	cathol	-0.015	0.009	48.930	-0.001	0.001	-0.004	0.002	0.770	0.756	0.828
25	indepen	-0.023	0.041	8.819	0.002	0.003	0.006	0.006	0.755	0.748	0.794
26	polariz	-0.898	0.088	93.463	-0.023	0.036	-0.471	0.086	0.738	0.722	0.992
27	budha	-0.014	0.017	1.579	-0.001	0.002	0.002	0.003	0.726	0.722	0.647

Continued on next page

Continued from previous page

No	Dependent Variable	Leamer			Sala-i-Martin						
		Lower	Upper	Fraction	Weighted		Unweighted		Normal	Non-Normal CDF	
		Bound	Bound	Signf. at 5%	Beta	St.Dev.	Beta	St.Dev.	CDF	Weighted	Unwght.
28	anglic	-0.050	0.025	32.495	-0.002	0.003	-0.014	0.009	0.716	0.711	0.893
29	demac	-0.466	0.168	87.512	-0.011	0.019	-0.245	0.042	0.714	0.712	0.960
30	hindu	-0.025	0.023	7.789	-0.001	0.002	-0.001	0.004	0.709	0.706	0.553
31	govcon	-1.347	2.223	10.688	0.094	0.183	0.398	0.381	0.697	0.693	0.798
32	area1	-0.167	0.351	31.853	0.010	0.019	0.096	0.058	0.696	0.659	0.895
33	german	-2.684	1.356	92.908	-0.070	0.137	-1.153	0.285	0.695	0.692	0.967
34	efherit	-0.596	0.146	71.422	-0.016	0.033	-0.222	0.096	0.686	0.682	0.962
35	distmag	-0.010	0.012	1.122	0.000	0.001	-0.001	0.003	0.666	0.632	0.602
36	trade	-0.524	0.209	1.233	0.011	0.033	-0.123	0.103	0.629	0.630	0.838
37	debtgdp	-0.012	0.011	33.068	0.000	0.001	0.003	0.002	0.620	0.601	0.858
38	gsta	-0.521	0.157	61.801	0.008	0.027	-0.176	0.078	0.613	0.607	0.926
39	orthodox	-0.014	0.023	32.747	0.000	0.001	0.004	0.003	0.600	0.585	0.853
40	protest	-0.027	0.021	73.241	0.000	0.002	-0.010	0.004	0.593	0.583	0.941
41	englis	-1.620	2.660	26.778	-0.013	0.059	-0.181	0.140	0.588	0.582	0.808
42	ethnob	-0.312	2.275	93.315	-0.018	0.100	1.114	0.189	0.573	0.571	0.955
43	muslim	-0.011	0.017	75.153	0.000	0.001	0.005	0.002	0.568	0.554	0.875
44	pluralty	-0.233	1.062	84.724	0.008	0.059	0.420	0.132	0.556	0.554	0.971
45	areakm	0.000	0.000	6.136	0.000	0.000	0.000	0.000	0.552	0.539	0.554
46	xmetal	-0.014	0.032	21.745	0.000	0.002	0.008	0.006	0.544	0.548	0.865
47	enrolt	-0.028	0.004	92.858	0.000	0.002	-0.012	0.004	0.526	0.526	0.962
48	federal	-0.726	0.134	90.953	-0.001	0.029	-0.277	0.066	0.515	0.507	0.984

Table 12: Extreme Bounds Analysis: Regulatory Capacity in 'I' (CFA Score)

No	Dependent Variable	Leamer			Sala-i-Martin						
		Lower	Upper	Fraction	Weighted		Unweighted		Normal	Non-Normal CDF	
		Bound	Bound	Signf. at 5%	Beta	St.Dev.	Beta	St.Dev.	CDF	Weighted	Unwght.
1	popden	-0.000	-0.000	100.000	-0.000	0.000	-0.000	0.000	1.000	1.000	1.000
2	scandi	-1.294	-0.191	100.000	-0.643	0.082	-0.638	0.084	1.000	1.000	1.000
3	social	-0.031	0.925	99.868	0.201	0.054	0.217	0.058	1.000	1.000	0.999
4	eten	-0.002	0.113	99.993	0.064	0.018	0.058	0.019	1.000	0.999	0.998
5	nonrelig	-0.003	0.012	93.630	0.005	0.001	0.005	0.002	0.999	0.986	0.992
6	illi	-0.009	0.003	93.076	-0.003	0.001	-0.004	0.001	0.996	0.987	0.994
7	wage	-0.003	0.036	98.215	0.015	0.006	0.017	0.006	0.992	0.987	0.993
8	area3	-0.020	0.101	15.929	0.043	0.020	0.035	0.020	0.986	0.979	0.954
9	xfuel	-0.002	0.005	9.809	0.002	0.001	0.002	0.001	0.983	0.981	0.941
10	femlab	-0.018	0.006	26.166	-0.006	0.003	-0.005	0.003	0.977	0.960	0.947
11	debt	-0.002	0.000	86.047	-0.001	0.000	-0.001	0.000	0.977	0.975	0.983
12	ethnoa	-0.585	0.191	13.801	-0.197	0.102	-0.174	0.106	0.974	0.961	0.941
13	enrolp	-0.003	0.007	5.850	0.002	0.001	0.003	0.002	0.921	0.917	0.945
14	presiden	-0.137	0.045	0.560	-0.044	0.031	-0.048	0.032	0.921	0.916	0.929
15	exco	-0.028	0.090	0.744	0.028	0.020	0.030	0.020	0.920	0.914	0.928
16	french	-0.366	0.721	1.542	-0.064	0.053	-0.053	0.054	0.886	0.849	0.818
17	debtcap	0.000	0.000	16.094	0.000	0.000	0.000	0.000	0.870	0.826	0.912
18	aidcap	0.000	0.001	0.415	0.000	0.000	0.000	0.000	0.867	0.854	0.895
19	latitu	-1.050	0.578	4.895	-0.188	0.197	-0.054	0.174	0.831	0.706	0.580
20	inequal	-0.076	0.135	0.000	0.026	0.030	0.016	0.031	0.806	0.770	0.690
21	marginal	-0.014	0.024	0.000	0.005	0.006	0.006	0.006	0.801	0.792	0.818
22	cathol	-0.004	0.004	0.000	-0.001	0.001	0.000	0.001	0.778	0.759	0.512
23	pollib	-0.212	0.134	0.000	-0.031	0.042	-0.037	0.045	0.773	0.765	0.789
24	indepen	-0.006	0.011	0.000	0.002	0.003	0.002	0.003	0.772	0.763	0.760
25	budha	-0.007	0.005	0.000	-0.001	0.002	-0.001	0.002	0.746	0.741	0.661
26	area1	-0.054	0.088	0.283	0.011	0.019	0.017	0.020	0.724	0.695	0.792
27	demac	-0.074	0.062	0.000	-0.010	0.018	-0.014	0.020	0.713	0.710	0.756

Continued on next page

Continued from previous page

No	Dependent Variable	Leamer			Sala-i-Martin						
		Lower	Upper	Fraction	Weighted		Unweighted		Normal	Non-Normal CDF	
		Bound	Bound	Signf. at 5%	Beta	St.Dev.	Beta	St.Dev.	CDF	Weighted	Unwght.
28	german	-0.527	0.795	0.007	-0.074	0.134	0.002	0.136	0.710	0.704	0.508
29	protest	-0.008	0.007	0.468	0.001	0.002	-0.002	0.002	0.706	0.683	0.848
30	efherit	-0.110	0.072	0.000	-0.017	0.033	-0.019	0.035	0.702	0.696	0.704
31	anglic	-0.012	0.010	0.000	-0.002	0.003	-0.001	0.003	0.694	0.689	0.618
32	govcon	-0.491	0.627	0.000	0.087	0.182	0.054	0.186	0.684	0.678	0.611
33	distmag	-0.002	0.003	0.000	0.000	0.001	0.000	0.001	0.671	0.641	0.701
34	polariz	-0.141	0.101	0.000	-0.016	0.037	-0.041	0.036	0.669	0.656	0.859
35	hindu	-0.008	0.006	0.000	-0.001	0.002	-0.001	0.002	0.647	0.644	0.694
36	trade	-0.083	0.101	0.000	0.012	0.033	0.018	0.032	0.647	0.647	0.715
37	debtgdp	-0.002	0.004	0.112	0.000	0.001	0.000	0.001	0.638	0.614	0.589
38	gsta	-0.077	0.104	0.000	0.007	0.027	0.013	0.031	0.605	0.599	0.656
39	muslim	-0.004	0.004	0.000	0.000	0.001	0.000	0.001	0.601	0.584	0.508
40	ethnob	-0.326	0.433	0.000	0.024	0.103	-0.043	0.103	0.593	0.577	0.661
41	areakm	0.000	0.000	0.026	0.000	0.000	0.000	0.000	0.580	0.569	0.567
42	pluralty	-0.146	0.200	0.000	0.009	0.058	0.035	0.060	0.564	0.562	0.715
43	enrolt	-0.004	0.005	0.000	0.000	0.002	0.000	0.002	0.542	0.541	0.561
44	orthodox	-0.003	0.005	0.830	0.000	0.001	0.001	0.001	0.536	0.530	0.806
45	englis	-0.357	0.747	0.494	-0.004	0.060	-0.003	0.061	0.524	0.519	0.521
46	xmetal	-0.007	0.007	0.000	0.000	0.002	0.000	0.002	0.501	0.504	0.585
47	federal	-0.144	0.101	0.540	0.000	0.029	-0.026	0.029	0.500	0.504	0.789

6 Conclusion

In this paper we survey the literature on the causes of corruption. The literature suggests a long list of variables claimed as statistically significant determinants. We collect as many as variables as possible and examine whether suhc claims are always consistent under various regression specification. We, however, face two problems. First, since the variables are drawn from different sources, these sources do not have the same set of observations. Thus, we have a missing data problem. Second, as we work with a huge number of variables, inter-dependencies among the determinants result in a multicollinearity problem once we run corruption regression.

To cope with the first problem, we do an imputation technique called 'Expectation-Maximization'. Through this techniques we generate a complete set of data consisting of 193 observations and 70 variables. To solve the second problem, we do Exploratory Factor Analysis in which we reduce the dimension of data. Through this technique, 27 variables can be reduced into five new variables, namely 'regulatory capacity', 'federalism', 'inequality', 'trade', and 'political liberty'. For these new variables we generate two types of index, namely factor-based score and factor score.

To examine whether these new variables and the other non-clusterred variables are robust in explaining variation in corruption, we employ two tests of Extreme Bounds Analysis. Using these two tests, we find that one of our new indexes, namely 'regulatory capacity', is the most robust variables. But, using the Sala-i-Martin's test, we find that about 11-14 variables can pass the test. We also find that different index generation still creates the similar result. In short, we convincingly conclude that 'regulatory capacity' is the most robust determinant of corruption. The other robust determinants are population density (–), scandinavian legal origin (–), ethnic tension (+), socialism legal origin (+), portion of population with no religion (+), ethnic conflict (+), illiteracy rate (–), government wage (+), sound money (area3 of Fraser index; +), latitude (–), fuel export (+), primary scholl enrollment (+), external debt (–), presidential (–), and portion of female in labor force (–).

References

[1] Abed, George T. and Hamid R. Davoodi. 2000. "Corruption, Structural Reforms, and Economic Performance in the Transition Economies". IMF Working Paper WP/00/132.

[2] Ades, A. and Di Tella, R. 1997. "The new economics of corruption: a survey and some new results". *Political Studies* XLV: 496-515.

[3] Ades, A. and Di Tella, R. 1999. "Rents, Competition, and Corruption". *American Economic Review* 89(4): 982-92.

[4] Ali, M. Abdiweli and Hodan Said Isse. 2003. "Determinants of Economic Corruption: A Cross-Country Comparison". *Cato Journal* 22(3): 449-466.

[5] Alt, James E. and David Dreyer Lassen. 2003. "The Political Economy of Corruption in American States". *Journal of Theoretical Politics* 15(3): 341-365.

[6] Amundsen, Inge. 1999. "Political Corruption: An Introduction to the Issues". Chr. Michelsen Institute Development Studies and Human Rights Working Paper 7.

[7] Andvig, Jens Chr. 1991. "The Economics of Corruption: A Survey". *Studi Economici* 43(1): 57-94.

[8] ————. 2005. "A House of Straw, Sticks or Bricks? Some Notes on Corruption Empirics". Paper presented at the Fourth Global Forum on Fighting Corruption and Safeguarding Integrity.

[9] Andvig, Jens Chr., Odd-Helge Fjeldstad, Inge Amundsen, Tone Sissener, and Tina Soreide. 2000. "Research on Corruption: A Policy Oriented Survey". Chr. Michelsen Institute and Norwegian Institute of International Affairs.

[10] Bonaglia, Federico, Jorge Braga de Macedo, and Maurizio Bussolo. 2001. "How Globalization Improves Governance". Discussion Paper No. 2992. Paris, France: Centre for Economic Policy Research, Organisation for Economic Co-operation and Development.

[11] Bowles, Roger. 2000. *Corruption. In* Boudewijn Bouckaert and Gerrit de Geest (eds.). *Encyclopedia of Law and Economics: The Economics of Crime and Litigation*, Vol. V. Edward Elgar, Cheltenham, UK.

[12] Braun, Miguel and Rafael Di Tella. 2004. "Inflation, Inflation Variability, and Corruption". *Economics and Politics* 16: 77-100.

[13] Broadman, Harry G. and Francesca Recanatini. 2000. "Seed of Corruption: Do Market Institutions Matter?". The World Bank Policy Research Working Paper No. 2368.

[14] ————————. 2002. "Corruption and Policy: Back to the Roots". *Policy Reform* 5: 37-49

[15] Brown, David S., Michael Touchton, and Andrew B. Whitford. 2005. "Political Polarizationas a Constraint on Government: Evidence from Corruption" on SSRN http://ssrn.com/abstract=782845.

[16] Brunetti, Aymo and Beatrice Weder. 2003. "A Free Press is Bad News for Corruption". *Journal of Public Economics* 87:1801-1824.

[17] Caiden, G. E. and Caiden, N.J. 1977. "Administrative Corruption". *Public Administration Review* 37(3): 301-309.

[18] Chang, Eric CC and Golden, Miriam A. 2004. "Electoral Systems, District Magnitude and Corruption". Paper presented at the 2003 annual meeting of the American Political Science Association, August 28-31, 2003

[19] Damania, Richard, Per Fredriksson, and Muthukumara Mani, 2004. "The Persistence of Corruption and Regulatory Compliance Failures: Theory and Evidence". *Public Choice* 121: 363-390.

[20] Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* 45: 289-309.

[21] Doppelhofer, G., R.I. Miller and X. Sala-i-Martin (2000), "Determinants of Long-term Growth: A Bayesian Averaging of Classical estimates (BACE) Approach", NBER Working Paper No. 7750.

[22] Durlauf, S.N. and D.T. Quah. 1999. "The New Empirics of Economic Growth," in *Handbook of Macroeconomics* Vol. 1, John B. Taylor and M. Woodford (eds), North Holland, Amsterdam.

[23] Fisman, Raymond J. and Roberta Gatti. 2002. "Decentralization and Corruption: Evidence across Countries". *Journal of Public Economics* 83: 325-345.

[24] Frechette, Guillaume R. 2001. "A Panel Data Analysis of the Time-Varying Determinants of Corruption". Paper presented at the EPCS 2001.

[25] Galtung, Fredrick. 2005 (forthcoming). "Measuring the Immeasurable: Boundaries and Functions of (Macro) Corruption Indices". *In* Fredrick Galtung and Charles Sampford (eds.). *Measuring Corruption*. Ashgate, UK.

[26] Gatti, R. (1999), "Corruption and Trade Tariffs, or a Case for Uniform Tariffs", World Bank Policy Research Working Paper No. 2216.

[27] Glynn, Patrick, Stephen J. Kobrin, and Moises Naim. 1997. The Globalization of Corruption. *In* Kimberly Ann Elliott (ed.). *Corruption and the Global Economy*. Institute for International Economics.

[28] Golden, Miriam A. and Lucio Picci. 2005. "Proposal for a New Measure of Corruption: Illustrated with Italian Data". *Economics and Politics* 17(1): 37-75.

[29] Goldsmith, Arthur A. . 1999. "Slapping the Grasping Hand: Correlates of Political Corruption in Emerging Market". *American Journal of Economics and Sociology* 58(4): 865-883.

[30] Gottfredson, Michael, and Travis Hirshi. 1990. *A General Theory of Crime*. Stanford CA: Stanford Univ. Press.

[31] Gould, David J. 1991. "Administrative Corruption: Incidence, Causes and Remedial Strategies". *In* A. Farazmand (ed.). *Handbook of Comparative and Development Public Administration*. Marcel Dekker Inc., New York.

[32] Graeff, P. and Mehlkop, G. 2003. "The Impacts of Economic Freedom on Corruption: Different Patterns for Rich and Poor Countries", *European Journal of Political Economy* 19: 605-620.

[33] Gupta, Sanjeev, Hamid Davoodi, and Rosa Alonso-Terme. 1998 "Does Corruption Affect Income Inequality and Poverty". IMF Working Paper 98/76.

[34] Gurgur, Tugrul and Anwar Shah. 2005. "Localization and Corruption: Panacea or Pandora's Box". World Bank Policy Research Working Paper 3486.

[35] Hall, Robert E. and Jones, Charles I. 1999. "Why do Some Countries Produce so much more Output per Worker than Others". *Quarterly Journal of Economics* 114: 83-116.

[36] Herzfeld, Thomas and Christoph Weiss. 2003. "Corruption and Legal (In)-Effectiveness: An Empirical Investigation". *European Journal of Political Economy* 19: 621-632.

[37] Henderson, J. Vernon and Ari Kuncoro. 2004. "Corruption in Indonesia". NBER Working Paper No. 10674.

[38] Hendry, David F. and Hans-Martin Krolzig. 2004. "We Ran One Regression". Department of Economics, Oxford University.

[39] Jain, Arvind K. 2001. "Corruption: A Review". *Journal of Economic Surveys* 15(1): 71-121.

[40] Jayawickrama, N. 2001. "Transparency International Combating Corruption through Institutional Reform". *In* Annette Y. Lee-Chai and John A. Barg (eds.). *The Use and Abuse of Power: Multiple Perspectives on the Causes of Corruption*. Psychology Press Philadelphia, Florence, and Sussex.

[41] Johnston, Michael. 2001. "Measuring Corruption: Number versus Knowledge versus Understanding". *In* Arvind K. Jain (ed.). *The Political Economy of Corruption*. Routledge, London and New York.

[42] Kaufmann, Daniel and Aart Kraay. 2002a. "Governance Indicators, Aid Allocation and the Millennium Challenge". World Bank mimeo.

[43] ————. 2002b. "Growth Without Governance". World Bank Policy Research Working Paper No. 2928

[44] Kaufmann, Daniel and Aart Kraay, and Massimo Mastruzzi. 1999. "Governance Matters". World Bank Policy Research Working Paper 2196.

[45] ————. 2000. "Governance Matters: From Measurement to Action". *Finance and Development* 37(2): 10-13.

[46] ————. 2003. "Governance Matters III: Governance Indicators for 1996-2002". World Bank mimeo.

[47] ————. 2005. "Governance Matters IV: Governance Indicators for 1996-2004". World Bank mimeo.

[48] Kaufmann, Daniel and Aart Kraay, and Pablo Zoido-Lobatn, P. 1999a. "Aggregating Governance Indicators". World Bank Policy Research Working Paper 2195.

[49] ————. 1999b. "Governance Matters". World Bank Policy Research Working Paper 2195.

[50] Knack, Stephen and Omar Azfar. 2003. "Trade Intensity, Country Size and Corruption". *Economics of Governance* 4: 1-18.

[51] Kuncoro, Ari. 2004. "Bribery in Indonesia: Some Evidence from Micro-Level Data". *Bulletin of Indonesian Economic Studies* 40(3): 329-354.

[52] Kunicova, Jana and Susan Rose-Ackerman. 2005. "Electoral Rules and Constitutional Structures as Constraints on Corruption", *British Journal of Political Science* 35 (4): 573-606.

[53] Laffont, Jean J. and Tchetche N'Guessan. 1999. "Competition and Corruption in an Agency Relationship". *Journal of Development Economics* 60: 271-95.

[54] La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R.W. 1998. "The Quality of Government", *Journal of Law, Economics, and Organization* 15: 222-79.

[55] Lambsdorff, Johann Graf. 1998. "Transparency International (TI) 1998 Corruption Perception Index, Framework Document". Transparency International and University of Passau.

[56] ————. 2000. "Background Paper to the 2000 Corruption Perceptions Index: Framework Document". Transparency International and Göttingen University.

[57] ————. 2001a. "Background Paper to the 2001 Corruption Perceptions Index: Framework Document". Transparency International and Gottingen University.

[58] ————. 2001b. "How Corruption in Government Affects Public Welfare: A Review of Theory". Mimeo, Georg-August-University of Gottingen.

[59] ————. 2002. "Background Paper to the 2002 Corruption Perceptions Index: Framework Document". Transparency International and Gottingen University.

[60] ————. 2003. "Background Paper to the 2003 Corruption Perceptions Index: Framework Document". Transparency International and University of Passau.

[61] ————. 2004a. "Background Paper to the 2004 Corruption Perceptions Index: Framework Document". Transparency International and University of Passau.

[62] ————. 2004b. "Corruption Perceptions Index 2004". *In* Diana Rodriguez, Gerard Waite, and Toby Wolfe (eds.). *Global Corruption Report 2005*. Transparency International-Pluto Press, London.

[63] Leamer, Edward E. 1985. "Sensitivity Analysis Would Help", *American Economic Review* 75, 308-313.

[64] Lederman, Daniel, Norman V. Loayza, and Rodrigo R. Soares. 2005. "Accountability and Corruption: Political Institutions Matter". *Economics and Politics* 17: 1-35.

[65] Leite, Carlos A. and Jens Weidmann. 1999. "Does Mother Nature Corrupt?: Natural Resources, Corruption, and Cconomic Growth". Working paper WP/99/85, International Monetary Fund, Washington, DC.

[66] Levine, Ross and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82(4): 942-963.

[67] Manzetti, Luigi and Charles Blake. 1996. "Market Reforms and Corruption in Latin America." *Review of International Political Economy* 3: 671-682.

[68] Mauro, Paulo. 1995. "Corruption and growth." *Quarterly Journal of Economics* 110(3): 681-712.

[69] Murray-Rust, D. Hammond, and Edward J. Vander Velde. 1994. "Changes in Hydraulic Performance and Comparative Costs of Lining and Desilting of Secondary Canals in Punjab, Pakistan". *Irrigation and Drainage Systems* 8(3): 137-158.

[70] Nye, Joseph S. 1967. "Corruption and Political Development: A Cost-Benefit Analysis". *The American Political Science Review* 61(2): 417-427.

[71] Wehmeier, S (ed.). 2000. *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press.

[72] Paldam, Martin. 2001. "Corruption and Religion: Adding to the Economic Model". *Kyklos* 54: 383-414

[73] Paldam, Martin. 2002. "The Cross-Country Pattern of Corruption: Economics, Culture and the Seesaw Dynamics". *European Journal of Political Economy* 18: 215-240.

[74] Park, Hoon. 2003. "Determinants of Corruption: A Cross-National Analysis" *The Multinational Business Review* 11(2): 29-48.

[75] Paternoster, Raymond and Sally Simpson. 1996. "Sanction Threats and Appeals to Morality: Testing a Rational Choice Model of Corporate Crime". *Law and Society Review* 30(3): 549583.

[76] Persson, Torsten and Guido Tabellini. 2003. *The Economic Effects of Constitutions*. Cambridge, Mass.: MIT Press.

[77] Persson, Torsten, Guido Tabellini, and Francesco Trebbi. 2003. "Electoral Rules and Corruption". *Journal of the European Economic Association* 1(4): 958-989.

[78] Rauch, James and Peter Evans. 2000. "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries". *Journal of Public Economics* 75: 49-71.

[79] Ruud, P. A. 1991. Extentions of Estimation Methods using the EM Algorithm. *Journal of Econometrics* 49: 305-341.

[80] Sala-i-Martin, Xavier. 1997a. "I Just Ran Four Millions Regressions." Mimeo, Columbia University.

[81] ————. 1997b. "I Just Ran Two Millions Regressions." *American Economic Review*, 87(2): 178-183.

[82] Scott, James C. 1972. *Comparative Political Corruption*. Englewood Cliffs, NJ: Prentice-Hall.

[83] Shleifer, Andrei and Robert W. Vishny. 1993. "Corruption". *Quarterly Journal of Economics* 108(3): 599-617.

[84] Sik, Endre. 2002. "The Bad, the Worse and the Worst: Guestimating the Level of Corruption". *In* Kotkin, Stephen and Andras Sajo. (eds.) *Political Corruption in Transition: A Skeptic's Handbook*. Budapest; New York: Central European University Press.

[85] Soreide, Tina. 2003. "Estimating Corruption: Comments on Available Data". Chr. Michelsen Institute. Mimeo.

[86] Spearman, Charles. 1904. "General Intelligence, Objectively Determined and Measured. *American Journal of Psychology* 15: 201-293.

[87] Sturm, Jan-Egbert and Jakob de Haan. Forthcoming. "Determinants of Long-term Growth: New Results Applying Robust Estimation and Extreme Bounds Analysis". *Empirical Economics*.

[88] Suphachalasai, Suphachol. 2005. "Bureaucratic Corruption and Mass Media" Paper presented at the 2005 EPCS

[89] Svensson, Jakob. 2003. "Who Must Pay Bribes and How Much? Evidence From a Cross-Section of Firms". *Quarterly Journal of Economics* 118(1): 207-230.

[90] Swamy, Anand, Stephen Knack, Young Lee, and Omar Azfar. 2001. "Gender and Corruption". *Journal of Development Economics* 64: 25-55.

[91] Temple, Jonathan. 2000. "Growth Regressions and What the Textbooks don't Tell You." *Bulletin of Economic Research* 52(3): 181-205.

[92] Tanzi, Vito and Hamid R. Davoodi. 1997. "Corruption, Public Investment, and Growth". IMF Working Paper 97/139.

[93] Tavares, Jose. 2003. "Does Foreign Aid Corrupt?" *Economic Letters* 79: 99-106.

[94] *The Guardian*, March 26, 2004.

[95] The Hungarian Gallup Institute. 2000. Basic Methodological Aspects of Corruption Measurement: Lessons Learned from the Literature and the Pilot Study. In UNODCCP and UNICRI. Joint Project against Corruption in the Republic of Hungary: Preliminary Assessment and Feedback on the Corruption Pilot Study.

[96] Tornell, Aaron and Philip Lane. 1998. "Voracity and Growth" CEPR Discussion Paper No. 2001.

[97] Treisman, Daniel. 2000. "The Causes of Corruption: A Cross-National Study". *Journal of Public Economics* 76: 399-457.

[98] Van Rijckeghem, Caroline and Beatrice S. Weder. 1997. "Corruption and the Rate of Temptation: Do Low Wages in the Civil Service Cause Corruption?". IMF Working Paper WP/97/73.

[99] Wade, Robert. 1982. "Irrigation Reform in Conditions of Populist Anarchy: An Indian Case". *Journal of Development Economics* 14(3): 285-303.

[100] Wei, Shang-Jin. 2000. "Local Corruption and Global Capital Flows". *Brookings Papers on Economic Activity* 2: 303-352.