Corruption and governance around the world

Seldadyo, H.

Chapter 3

On the Sensitivity of Corruption Determinants

"Let's not mince words ...
We need to deal with the causes of corruption."

James Wolfenson (2005)

3.1 Introduction

As corruption is generally believed to have negative welfare effects[1] it is important to find out what determines corruption. Many studies have searched for empirical correlations between corruption and a variety of economic and non-economic determinants. Two questions are usually addressed in these studies. First, what variables might explain the cross-country variation in corruption? Second, how robust are these variables in explaining cross-country differences in corruption? Unfortunately, there is no consensus in the literature on both issues. In fact, theory offers little guidance in spec-

Earlier versions of this chapter—joint work with Jakob de Haan—were presented at the European Public Choice Society meetings in 2005 and 2006

[1] A series of studies investigating the impact of corruption include Kaufmann et al. (1999) and Gupta et al. (2001) on income, infant and child mortality rates, low-birth-weight babies, and dropout rates in primary schools, Mauro (1995) on growth, Li et al. (2000) on distribution of income, Tanzi and Davoodi (1997) on the quality of public infrastructure, and Lambsdorff (2003a) on productivity.

ifying a proper regression model for corruption. As a consequence, various specifications incorporating a wide range of explanatory variables have been used to explain corruption and to find the ‘true’ determinants. It is commonly found, however, that a particular variable is significant in a particular model, but becomes insignificant when some other variables are incorporated. Therefore, it is still not clear what variables are really driving corruption.

We employ around 45 variables that in previous studies have been found to be correlated with corruption and examine whether they are robustly related to corruption. We use two sets of corruption measures: aggregated and individual indexes. The indexes of Kaufmann et al. of the World Bank (WB) and Lambsdorff of Transparency International (TI) are aggregated indexes, while those of the International Country Risk Guide (ICRG) and the International Management Development World Competitiveness Yearbook (IMD-WCY) are individual indexes. In addition, we generate an aggregated corruption index using the indicators of the World Economic Forum Global Competitiveness Report (WEF-GCR).

Using a variant of the Sensitivity Analysis (SA) as proposed by Sala-i-Martin (1997) and checking for sample sensitivity, we find that only two variables are consistently related to corruption, namely government effectiveness and rule of law. Other variables like income, regulatory quality, protestant fraction, absolute latitude, economic freedom, or even democracy— commonly argued to be significant determinants— are not robustly correlated with corruption. The remainder of this chapter is organized as follows. Section 3.2 describes our methodology and data in some detail, while sections 3.3 and 3.4 present our evidence. The final section offers some concluding comments.

3.2 The Setup

Studies on the determinants of corruption are marked by three characteristics. First, different authors employ different specifications with different variables to result in a model uncertainty of corruption determinants. The setups of Treisman (2000), Paldam (2002), Ali and Isse (2003), and Park (2003), for example, are not the same even though they all search for a common underlying factors of corruption. Second, it is also quite common that investigators limit their attention to one or a small number of variables of interest. For instance, Graeff and Mehlkop (2003) investigate the link between economic freedom and corruption. Likewise, Brunetti and Weder (2003) inspect the relationship between press freedom and corruption, while Swamy et al. (2001) examine the link between corruption and gender. Finally, to test for the reliability of their variable of interest, researchers usually vary their specifications and examine how sensitive this particular variable is to a certain range of control variables. However, most of the time, the varying specifications are still limited to few variables and a narrow range of variation in model specification. Hence, one may question, how much confidence should be given to the conclusions of previous studies.

Using the SA, we consider a large number of variables that have been claimed to be related to corruption in previous studies. We prefer this tool of analysis because it offers a systematic way to scrutinize what variables are robustly related to corruption. It also allows us to explore a wide range of possible model specifications. The SA departs from a simple setup as follow:

$$C = \alpha_j + \boldsymbol{\beta}_{fj}\boldsymbol{F} + \beta_{xj}X + \boldsymbol{\beta}_{zj}\boldsymbol{Z} + \epsilon \tag{3.1}$$

where C is some corruption index, $\boldsymbol{F}$ is a vector of 'fixed' explanatory variables that are always included in the regression, but which may also be zero; X is the variable of interest; $\boldsymbol{Z}$ is a vector of combinations of up-to-M possible additional explanatory variables drawn from a pool of variables which, according to the literature, may be related to the dependent variable; and ϵ is an error term. It should be noted, however, that the number of variables in $\boldsymbol{F}$ and $\boldsymbol{Z}$ that can be plugged into the model is constrained by the degree of freedoms of the regression as well as by potential multicollinearity problems. Levine and Renelt (1992) include three variables in $\boldsymbol{F}$ and all possible combinations of up-to-three variables in $\boldsymbol{Z}$.

For each model j, X is explored using all possible linear combinations of $\boldsymbol{Z}$ given $\boldsymbol{F}$. The 'robust' correlation between X and C might be found via Extreme Bound Analysis (EBA) of Leamer (1983, 1985) and Levine and Renelt (1992) by constructing the upper and lower bounds for β_{xj} defined as

$$ {}^{max}_{min} \beta_{xj} \pm 2\sigma^2_{\beta_{xj}}. \tag{3.2} $$

If within this range β_{xj} does not change sign, the variable is considered robust. If one finds that the sign of β_{xj} changes within this range, the corresponding variable is regarded as 'fragile'.

Certainly, this approach is not without cost. As the bounds are really extreme, Sala-i-Martin (1997) argues that this test is too strong for any variable to pass it, thus "...nothing can be learned ..." (p. 179) from such an approach. He suggests analyzing the entire distribution of the estimates of the parameter of interest (β_{xj}) and examining the fraction of the cumulative density function (CDF) lying on each side of zero.[2] As zero divides the area under the CDF into two, the larger of the two areas is regarded as CDF(0), no matter whether it is above or below zero. Thus, the CDF(0) is always a number between 0.5 and 1.0.

We use equation 3.1, but due to the absence of a commonly-agreed theoretical framework, we set $\boldsymbol{F}$ to be zero. We experiment with a series of models where $\boldsymbol{Z}$ contains several variables ($m = 3, 4, 5$) drawn from a pool

[2] Sala-i-Martin et al. (2004) also propose another technique called Bayesian Averaging of Classical Estimates (BACE) to check the robustness of different explanatory variables in growth regressions. This approach builds upon the approach as suggested by Sala-i-Martin (1997), in the sense that different specifications are estimated to check the sensitivity of the coefficient estimate of the variable of interest. The major innovation of BACE as compared to the Sala-i-Martin's approach is that there is no set of fixed variables included and the number of explanatory variables in the specifications is flexible. The biggest disadvantages of the BACE approach are the need of having a balanced dataset, i.e., an equal number of observations for all regressions (due to the chosen weighting scheme), the restriction of limiting the list of potential variables to be less than the number of observations and the computational burden.

of 45 selected variables. To check the robustness of a variable, we advocate three decision rules. First, at least 95 per cent of the CDF(0) lies on one side of zero; or simply CDF(0)$>$ 0.95. This is different from Sala-i-Martin who sets the CDF(0)$>$ 0.90. We consider his benchmark too low given the one-sidedness of the test. In addition, following Sturm and de Haan (2005), for a variable to be considered robust, it should be significant at the 5 per cent level in at least 90 per cent of all regressions. Finally, a variable must pass the previous two benchmarks for three of the five corruption indexes used in the regressions. To have a complete picture, however, we also show the outcomes under the approaches suggested by Levine and Renelt as well as Sala-i-Martin.

Two issues, however, remain, namely multicolinearity and simultaneity. To deal with the former, we apply the following strategy. First, we exclude individual variables produced by particular sources if they have been included into another variable by other sources. For example, variables such as the Polity IV index, political rights, civil liberty, press freedom, freedom of speech, have been dropped, because Kaufmann et al. (2007) have integrated them into a new variable called 'voice and accountability'. The same applies to the other variables of Kaufmann et al., representing regulatory quality, government effectiveness, rule of law, and political stability.

Second, we chose a variable with the highest number of observations in case we have more than one variable at hand proxying the same concept. For instance, we use the index of economic freedom of the Heritage Foundation instead of that of the Fraser institute in view of its coverage of countries. Third, for some cases we construct indexes for variables explaining the same concept using Principal Component Analysis (PCA). This is done for variables capturing, decentralization (for instance, expenditures and revenues of sub-national governments), human capital (like schooling levels and literacy rate), income inequality (like the Gini coefficient, 10-20 per cent rich population), and women's participation (women in economic, social, and political arenas, female labor force).

Under this strategy, we have reduced the number of variables to be considered, from originally around 75 to 45 variables (Appendix 2). These variables can be classified into four broad categories: (1) economic and demographic factors, (2) political institutions, (3) judicial and bureaucratic environment, and (4) geographic and cultural variables.

Meanwhile, to minimize simultaneity problems, corruption is measured in 2005-2006, while the determinants refer to 2000.[3] This also implies that we allow a substantial time for the determinants to have effect on corruption.

We consider a set of corruption indicators as the dependent variable. These variables differ along two dimensions: the methodological construction of the indexes and the number of observations available. The first two indexes—i.e., the WB (scaled from −2.5 to +2.5) and the TI (0-10) corruption indexes, with 201 and 168 observations, respectively—are aggregated indexes based on a variety of individual sources (poll-of-polls indexes). We use also individual indexes including the ICRG (0-6) and the IMD (1-10) indexes, with 140 and 53 observations, respectively. In addition, we exploit PCA to generate an aggregated index that originates from seven indicators of the WEF (scale of 1-7) with 117 observations.[4] The scales of these indexes are adjusted so that a higher score means less corruption.

3.3 Some First Results

Table 3.1 displays the results of our robustness tests in case $\boldsymbol{F}$ is set to be zero, while up-to-three variables are included in $\boldsymbol{Z}$ drawn from a pool of 45 variables[5] using five different corruption indexes as dependent variable. In columns 2-4, we report the two extreme bounds and the fraction of signifi-

[3] Economic growth, however, is measured as the average of the 1990-2000 values, while the income distribution is measured over 1993-2000.

[4] The PCA produces only one component with a high eigenvalue (6.11) that accounts for 87.25 per cent of the combined variance. The 'scoring coefficients' of the indicators range from 0.34 (favoritism) to 0.39 (irregular payments in public contracts), while business costs of corruption as well irregular payments in public utilities, judicial decisions, export-import, and in tax collection have 'scoring coefficients' of 0.38.

[5] Under this setup, for each variable, we run $\frac{(45-1)!}{3!41!} = 13,244$ regressions, or a total of 595,980 regressions for all variables.

cant cases; in columns 5-7, the unweighted CDF(0)[6], the estimate of β, and its standard deviation (σ^2) are displayed. To save space, we only report the variables for which CDF(0)$>$ 0.95.

It follows from Table 3.1 that there are 4-6 variables passing the first rule (i.e., CDF(0)>0.95). However, only half of them pass the second and third rules (90 per cent cases significant at the 5 per cent level and passing the two banchmarks in three of five corruption indexes). The robust variables are government effectiveness and rule of law. The CDF(0) of government effectiveness is very close to one, while in almost all specifications government effectiveness has a significant impact on corruption. For the rule of law variable, the CDF(0) ranges between 0.98-1.00, while the fraction of the regressions in which the coefficient of this variable is significant ranges between 93.92 and 99.98 per cent. Only in the case where the IMD index is the dependent variable, is the rule of law variable slightly below the second benchmark.[7]

Government effectiveness captures "the quality of public service provision, the quality of the bureaucracy, the competence of civil servants, the independence of the civil service from political pressures, and the credibility of the government's commitment to policies". Meanwhile, the rule of law variable denotes mainly "the extent to which agents have confidence in and abide by the rules of society" and "perceptions of the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts." Therefore, rule of law measures "the success of a society in developing an environment in which fair and predictable rules form the basis for economic and social interactions, and importantly, the extent to which property rights are protected" (Kaufmann and Kraay, 2002: 177-178). This may explain why the two variables are robustly correlated with corruption.

[6] Sala-i-Martin (1997) argues that the unweighted average is superior to the weighted average in a situation with a spurious fit.

[7] Notice that the IMD index covers only 53 countries, which is about one-fourth of the WB country coverage.

Table 3.1: Robustness Analysis (No $\boldsymbol{F}$, $\boldsymbol{Z}$ up-to-3)

Variable	Levine-Renelt			Sala-i-Martin		
	Lower Bound	Upper Bound	Fraction Signif.	Unweigt. CDF(0)	Unweigt. β	Unweigt. σ^2
Corruption: WB ($N = 201$)						
Government Effectiveness	0.037	1.615	100.000	1.000	0.930	0.048
Rule of Law	-0.326	1.366	99.977	1.000	0.894	0.051
Ln GDP per capita	-0.419	2.103	90.562	0.978	0.578	0.067
Presidentialism	-0.245	0.998	83.902	0.962	0.290	0.070
Aboslute Latitude	-0.262	1.252	81.297	0.962	0.290	0.070
Voice-Accountability	-0.540	2.009	87.149	0.954	0.602	0.069
Corruption: TI ($N = 168$)						
Government Effectiveness	0.351	4.073	100.000	1.000	2.009	0.122
Rule of Law	-1.424	3.453	99.894	1.000	1.854	0.125
Ln GDP per capita	-1.097	4.794	90.939	0.980	1.227	0.150
Aboslute Latitude	-0.662	2.850	82.520	0.962	0.648	0.159
Corruption: ICRG ($N = 140$)						
Government Effectiveness	-0.684	2.213	99.071	0.999	0.933	0.111
Rule of Law	-1.073	1.934	98.671	0.997	0.883	0.108
Polarization	-0.166	0.921	96.957	0.994	0.354	0.090
Wage Bill per GDP	-2.002	0.576	79.047	0.952	-0.273	0.116

continued on next page ...

Variable	Levine-Renelt			Sala-i-Martin		
	Lower Bound	Upper Bound	Fraction Signif.	Unweigt. CDF(0)	Unweigt. β	Unweigt. σ^2
Corruption: IMD ($N = 53$)						
Government Effectiveness	-3.736	9.398	99.343	0.998	2.681	0.295
Regulatory Quality	-1.610	7.415	95.417	0.992	2.852	0.329
Rule of Law	-5.268	6.200	93.922	0.984	2.391	0.329
Protestant Fraction	-3.067	5.710	91.000	0.983	0.965	0.262
Population	-6.074	4.058	77.076	0.959	-1.075	0.363
Economic Freedom	-1.723	4.752	88.183	0.958	1.562	0.397
Corruption: WEF ($N = 117$)						
Government Effectiveness	-0.628	5.785	99.985	1.000	2.254	0.204
Ln GDP per capita	-1.764	4.285	99.071	0.997	1.679	0.206
Rule of Law	-2.923	3.924	97.584	0.996	2.002	0.208
Regulatory Quality	-2.824	5.001	88.206	0.954	2.087	0.278

There is also another variable passing the three benchmarks: income. Income passes the first two benchmarks as its CDF(0) is about 0.98-0.99 and the lowest fraction significant is 91 per cent. The two benchmarks are achieved when corruption index is WB, TI, and WEF—the three aggregated indexes. Hence, when $\boldsymbol{F}$ is set to be zero and $\boldsymbol{Z}$ is set up-to-three variables, income appears as a robust variable according to the three rules.

The other variables do not consistently pass the three yardsticks. Regulatory quality, for example, only passes the tests when corruption is proxied by the IMD index; the same holds for Protestantism. Absolute latitude has a CDF(0) above 0.95 under the use of the WB and TI indexes, but the fraction of significant regressions is still far from the second benchmark. Democracy—captured by voice and accountability—is significant only in 87 per cent of the regressions, although it passes the CDF test for the use of WB index. Economic freedom has similar figures.

Keeping $\boldsymbol{F}=0$, we now experiment with $\boldsymbol{Z}$ containing up-to-four and up-to-five variables[8] where those passing the first yardstick are presented in Table 3.2. The results reinforce our previous findings: government effectiveness and rule of law turn out to be robust variables. The CDF(0) statistics of these determinants are very close to 1.00. At the same time, we find that in 96-100 per cent of the regressions the impact of government effectiveness is significant, while the corresponding figure for the rule of law variable is about 96-99 per cent.[9] Compared to the previous results, also the standardized impact is very similar. Nevertheless, the results for income are now different. It passes the first two benchmarks only when the WEF index is used, both under up-to-four and five variables in $\boldsymbol{Z}$. For the other corruption indexes, income does not pass the tests. The other variables cannot be regarded as robust variables according to the three rules.

[8] For each variable, we run respectively $\frac{(45-1)!}{4!40!} = 135,751$ and 1,086,008 regressions, or a total of 6,108,795 and 48,870,360 regressions for all variables.

[9] Under the IMD index and $\boldsymbol{Z}$ containing up-to-five variables, the fraction of significant regressions of rule of law is only slightly below the yardstick: 89.6 per cent.

Table 3.2: Robustness Analysis ($\boldsymbol{F}$=0, $\boldsymbol{Z}$ up-to-4 and up-to-5)

Variable	$\boldsymbol{F}$=0, $\boldsymbol{Z}$ up-to-4			$\boldsymbol{F}$=0, $\boldsymbol{Z}$ up-to-5		
	Fraction Signif.	Unweigt. CDF(0)	Unweigt. β	Fraction Signif.	Unweigt. CDF(0)	Unweigt. β
Corruption: WB						
Government Effectiveness	100.000	1.000	0.924	100.000	1.000	0.918
Rule of Law	99.912	1.000	0.878	99.782	0.999	0.860
Ln GDP per Capita	86.465	0.963	0.523			
Corruption: TI						
Government Effectiveness	100.000	1.000	1.999	99.999	1.000	1.986
Rule of Law	99.622	0.999	1.805	99.174	0.997	1.751
Ln GDP per Capita	87.930	0.970	1.117	84.734	0.958	1.018
Corruption: ICRG						
Government Effectiveness	97.761	0.996	0.918	95.826	0.993	0.901
Rule of Law	97.778	0.996	0.864	96.712	0.994	0.846
Political Polarization	93.812	0.988	0.305	89.867	0.982	0.268
Corruption: IMD						
Government Effectiveness	98.701	0.996	2.665	97.780	0.993	2.650
Regulatory Quality	93.627	0.988	2.758	91.929	0.984	2.684
Rule of Law	91.929	0.973	2.303	89.591	0.960	2.214
Protestant Fraction	85.291	0.970	0.901	79.823	0.954	0.852

continued on next page ...

Variable	$\boldsymbol{F}$=0, $\boldsymbol{Z}$ up-to-4			$\boldsymbol{F}$=0, $\boldsymbol{Z}$ up-to-5		
	Fraction Signif.	Unweigt. CDF(0)	Unweigt. β	Fraction Signif.	Unweigt. CDF(0)	Unweigt. β
Corruption: WEF						
Government Effectiveness	99.890	1.000	2.231	99.629	0.999	2.208
Ln GDP per Capita	97.626	0.992	1.552	95.398	0.987	1.438
Rule of Law	96.032	0.992	1.930	94.281	0.988	1.861

Having two robust variables at hand, we treat government effectiveness and rule of law as the fixed variables ($\boldsymbol{F}$) and rerun Model 3.1. Applying the same decision rules, we do not find any variable that consistently has a robust relationship to corruption. Some variables are able to pass the CDF test, but most of them fail to pass the second one. None passes the third test. In the following we narratively report the results.

Using the WB index, we find that population, fraction of population belong to Hindu, GDP per capita growth, plurality, government expenditure, and fraction of Buddhists pass the first yardstick, but only population can pass the second benchmark. This is different from the results drawn from the TI index; here, the fraction of Hindu, voice and accountability, and population are able to pass the first test, but only the first can pass the second yardstick. Also, a different result is found when we use the ICRG index as the dependent variable. Now, debt, wage, and polarization pass the first test, while none but debt passes the second benchmark. Meanwhile, variables passing the first yardstick under the IMD index are voice and accountability, fraction of Catholic, export of ores and metal, and human capital. Yet, none passes the second test. Finally, under the WEF index, three variables—voice and accountability, fraction of Catholic, and export of ores and metal—pass the first test, but none passes the second.

3.4 Effect of Observations

Up to this point, $\boldsymbol{Z}$ includes all observations (N). Now we turn to an experiment where N is varied. In this experiment, we order the observations according to their corruption scores and estimate equation 3.1 employing $N = 100, 125, 150, 175$ observations, or respectively 50, 62.5, 75, and 87.5 per cent of the total observations. We also compare these results if the corruption scores are ordered from low to high. In the first ordering the first 50 per cent of the observations are dominated by corrupt nations, while the reverse applies to the second ordering. In this part of the analysis, we use only the WB index as it covers almost all countries over the world. The results are reported in Table 3.3 where $\boldsymbol{Z}$ is up to three variables.

As follows from Table 3.3, government effectiveness and rule of law again appear as the robust variables correlated with corruption since they consistently pass the three benchmarks regardless of the number of observations. Even, when the observations are reduced until 50 per cent, the CDF(0) and fraction of significant are still far above the benchmark. The performance of these variables is stable under both orderings. The other variables do not consistently pass the tests; this holds true even for income and democracy (proxied by voice and accountability) that are commonly argued to be significant in explaining corruption. Such variables are sensitive not only to the existence of other (control) variables, but also to the number and composition of the set of observations.

We again examine the sensitivity of our findings using up-to four and five variables in the $\boldsymbol{Z}$ vector (Tables 3.4-3.5). Apart from government effectiveness and rule of law, no variable is found to be robust according to the three decision rules. Different $\boldsymbol{Z}$, N, or ordering procedure do not change the conclusion that these variables are robust. One, however, may question why government effectiveness and rule of law are always robust in their correlation with corruption in all circumstances. Figure 3.1 clarifies it. As the plots scatter around their 'means', clearly no influencing outlier is found in the figure. The figure shows that none of the nations with low quality of government effectiveness and rule of law appears as a corruption-free nation. Likewise, those regarded as clean countries always perform high quality of government effectiveness and rule of law. This fact holds up using a variety of measures of corruption.

Furthermore, there are three interesting features displayed in Tables 3.1-3.5 with respect to β. First, the CDF(0) increases as N increases, regardless of the way of ordering. The same applies to fraction of significant regressions. This confirms the classical regression issue: a bigger sample gives more convincing results. Second, the impact of government effectiveness is always higher in all circumstances compared to that of rule of law. The same holds true for both corrupt (the first 50 per cent ascending order) and clean (the first 50 per cent descending order) countries.

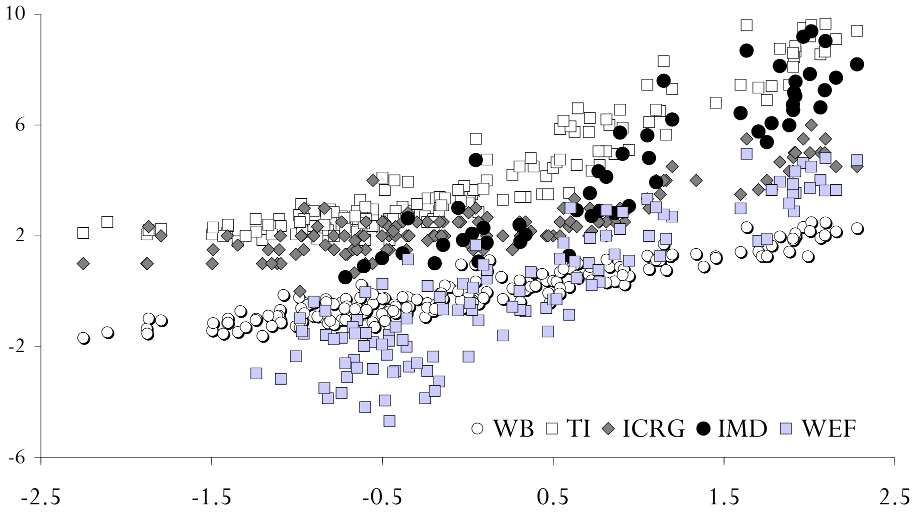
Finally, the regressions based on the ascending-order produce much lower β in the first 50 per cent of the observations, than the descending-order regressions do. The coefficient of government effectiveness is about 0.4-0.5 under the former, and 0.8-0.9 under the latter. Similar results are found for the rule of law variable. This implies that the impact on corruption of government effectiveness and rule of law are considerably different in corrupt and non-corrupt regimes. In corrupt regimes the impact is about 50 per cent lower than in clean regimes, reflecting heterogeneity in the size of the impact of corruption determinants.

Table 3.3: Robustness Analysis ($\boldsymbol{F}$=0, $\boldsymbol{Z}$ up-to-3, Various N)

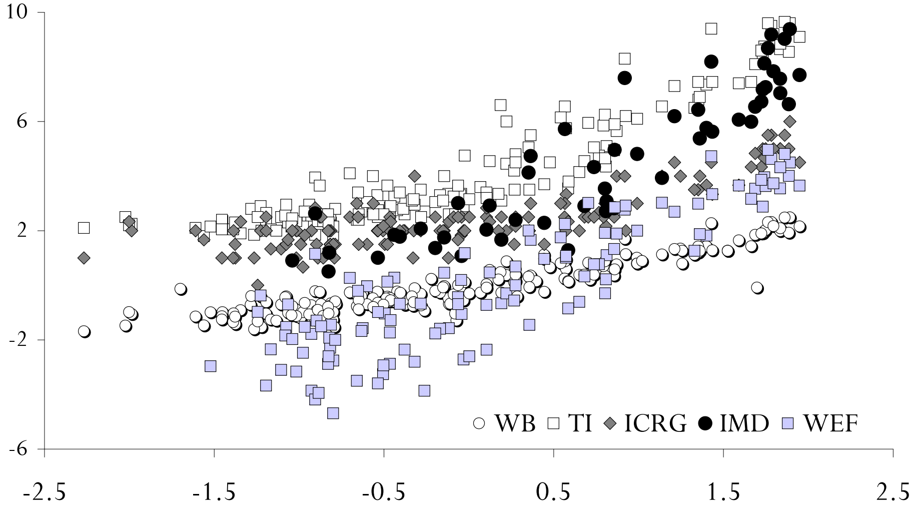
Dependent Variable	Z=up-to-3; Ascending Order			Z=up-to-3; Descending Order		
(Ind. Var: WB Index)	% Sign.	CDF(0)	Beta-U	% Sign.	CDF(0)	Beta-U
$N = 100$						
Government Effectiveness	99.517	0.998	0.464	99.706	1.000	0.872
Rule of Law	97.357	0.995	0.436	98.014	0.997	0.869
Regulatory Quality	96.338	0.994	0.290			
Voice and Accountability	90.818	0.979	0.229	90.169	0.951	0.757
Political Stability				92.170	0.983	0.634
$N = 125$						
Government Effectiveness	99.977	1.000	0.593	99.841	1.000	0.895
Rule of Law	99.607	0.999	0.538	99.857	0.999	0.865
Voice and Accountability	95.296	0.988	0.309			
Regulatory Quality	92.706	0.970	0.353			
Political Stability				90.335	0.965	0.577
Absolute Latitude				79.364	0.954	0.287
$N = 150$						
Government Effectiveness	100.000	1.000	0.727	99.992	1.000	0.918
Rule of Law	99.902	1.000	0.672	99.962	1.000	0.872
GDP per Capita	92.057	0.982	0.325	89.671	0.971	0.579
Voice and Accountability	92.147	0.978	0.379			
Area	79.991	0.965	-0.173			
Regulatory Quality	92.532	0.957	0.482			

continued on next page ...

Dependent Variable	Z=up-to-3; Ascending Order			Z=up-to-3; Descending Order		
(Ind. Var: WB Index)	% Sign.	CDF(0)	Beta-U	% Sign.	CDF(0)	Beta-U
Foreign Debt				86.477	0.968	-0.150
Presidentialism				83.268	0.966	0.255
Absolute Latitude				80.301	0.958	0.299
Political Stability				89.037	0.955	0.590
$N = 175$						
Government Effectiveness	100.000	1.000	0.810	100.000	1.000	0.919
Rule of Law	99.947	1.000	0.761	99.977	1.000	0.868
Area	92.578	0.985	-0.263			
GDP per Capita	92.306	0.984	0.399	90.645	0.973	0.565
Voice and Accountability	90.788	0.974	0.444			
Regulatory Quality	91.838	0.953	0.567			
Presidentialism				92.683	0.985	0.316
Ethnic Division				78.232	0.955	-0.195
Foreign Debt				81.931	0.954	-0.139



(a) Government Effectiveness



(b) Rule of Law

Figure 3.1: Corruption (y-axis) and Two Robust Determinants (x-axis)

Table 3.4: Robustness Analysis ($\boldsymbol{F}$=0, $\boldsymbol{Z}$ up-to-4, Various N)

Dependent Variable	Z=up-to-4; Ascending Order			Z=up-to-4; Descending Order		
(Ind. Var: WB Index)	% Sign.	CDF(0)	Beta-U	% Sign.	CDF(0)	Beta-U
$N = 100$						
Government Effectiveness	98.703	0.994	0.459	99.472	0.999	0.872
Regulatory Quality	93.882	0.988	0.285	97.613	0.996	0.858
Rule of Law	95.571	0.987	0.434			
Voice and Accountability	84.007	0.957	0.210			
Political Stability				88.293	0.972	0.571
$N = 125$						
Government Effectiveness	99.888	1.000	0.594	99.805	1.000	0.894
Rule of Law	99.470	0.999	0.537	99.623	0.999	0.851
Voice and Accountability	90.871	0.976	0.294			
Regulatory Quality	90.289	0.962	0.349			
Political Stability				86.044	0.951	0.522
$N = 150$						
Government Effectiveness	99.993	1.000	0.727	99.987	1.000	0.914
Rule of Law	99.719	0.999	0.668	99.865	0.999	0.851
GDP per Capita	88.520	0.972	0.303	85.563	0.952	0.530
Voice and Accountability	86.947	0.962	0.359			
Foreign Debt				81.840	0.950	-0.145
$N = 175$						
Government Effectiveness	99.998	1.000	0.806	100.000	1.000	0.914

continued on next page ...

Dependent Variable	Z=up-to-4; Ascending Order			Z=up-to-4; Descending Order		
(Ind. Var: WB Index)	% Sign.	CDF(0)	Beta-U	% Sign.	CDF(0)	Beta-U
Rule of Law	99.801	1.000	0.750	99.907	1.000	0.850
GDP per Capita	88.505	0.973	0.370	86.863	0.958	0.515
Area	85.358	0.970	-0.232			
Voice and Accountability	85.261	0.956	0.415			
Presidentialism				86.228	0.970	0.264

Table 3.5: Robustness Analysis ($\boldsymbol{F}$=0, $\boldsymbol{Z}$ up-to-5, Various N)

Dependent Variable	Z=up-to-5; Ascending Order			Z=up-to-5; Descending Order		
(Ind. Var: WB Index)	% Sign.	CDF(0)	Beta-U	% Sign.	CDF(0)	Beta-U
$N = 100$						
Government Effectiveness	97.187	0.986	0.438	99.166	0.999	0.871
Regulatory Quality	90.805	0.977	0.279			
Rule of Law	93.426	0.972	0.433	96.744	0.993	0.845
Political Stability				83.815	0.957	0.514
$N = 125$						
Government Effectiveness	99.713	0.999	0.594	99.762	1.000	0.893
Rule of Law	99.079	0.998	0.535	99.248	0.997	0.835
Voice and Accountability	85.972	0.962	0.281			
Regulatory Quality	87.945	0.955	0.347			
$N = 150$						
Government Effectiveness	99.970	1.000	0.726	99.985	1.000	0.909
Rule of Law	99.441	0.999	0.662	99.675	0.998	0.828
GDP per Capita	84.697	0.959	0.284			
$N = 175$						
Government Effectiveness	99.991	1.000	0.803	100.000	1.000	0.909
Rule of Law	99.555	0.999	0.738	99.763	0.999	0.830
GDP per Capita	84.457	0.959	0.344			
Presidentialism				78.751	0.950	0.221

3.5 Concluding Remarks

Over the last two decades, the number of empirical studies on the determinants of corruption has rapidly increased. These studies employ a variety of models using various economic and non-economic variables. Yet, there is no 'true' model due to the absence of an encompassing theory on the determinants of corruption. Still, many variables have been claimed to be significant in explaining cross-country variation in corruption. Since the 'true' model is far from known, the question which variables are really correlated with corruption remains.

The literature offers various approaches in dealing with this issue including the EBA of Levine and Renelt (1992) and the SA of Sala-i-Martin (1997). Slightly modifying these approaches, we advocate three sequential decision rules to label a variable of interest 'robust' or 'fragile'. We use 45 variables and search all possible linear combinations to check their robustness. We also experiment with up-to-three, four, and five variables in $\boldsymbol{Z}$ but none in $\boldsymbol{F}$.

We find that two variables that are consistently passing the three tests. The two variables are government effectiveness and rule of law, drawn from the governance dataset of Kaufmann et al. (2007). Government effectiveness and rule of law repeatedly appear robust since their CDF(0) is always above 0.95 and in more than 90 per cent of the regressions these variables are found to be significant at the 0.05 level. Also, the robustness of these variable does not change if we use different corruption indexes. Two graphical expositions demonstrating the closeness of the scattered data to their 'means' support this finding. The other variables, however, are either found to pass the test erratically or not pass the test at all across the use of different dependent variables.

On the basis of these findings, we conclude that corruption is always lower in countries governed by high quality officials and where the rule of law is strongly enforced. This conclusion is not sensitive to the change in the number and composition of observations as well as the ordering of observations. Also, these two variables always appear robust regardless of the types of corruption index used in the regressions. The other variables do not consistently pass the tests when different corruption indexes are employed, even though these indexes are highly correlated. In sum, of a bunch of determinants claimed significant in explaining cross-country variation in corruption, only government effectiveness and rule of law can be regarded as robust variables.